Structure/Function Studies of the P2X Receptor

Helen Digby

Department of Cell Physiology and Pharmacology

University of Leicester

June 2007

Structure/Function Studies of the P2X Receptor

Helen Digby

Abstract

P2X receptors are a novel family of ligand-gated ion channels that open in response to the binding of extracellular ATP. There are no crystal structures available for the P2X receptor family and they share little homology with other ATP-binding proteins. For this reason, experimental evidence has been relied upon to determine the topology of the P2X receptor family and infer function. The family presents numerous drug targets for the treatment of cystic fibrosis, regulation of blood pressure, pain and irritable bowel syndrome to name a few (Khakh & North, 2006). A better understanding of the structure of P2X receptors gained through mutagenesis and more recently bioinformatic techniques will enable better drug design and development.

The amino acid sequence of a protein determines its secondary structure which in turn dictates the tertiary or 3D structure. Protein function is dependent on protein structure and for this reason structural studies can give important insight into protein function. Alanine-replacement of conserved glycine residues in the P2X receptor family has been studied to determine the impact of the flexible nature of glycine in the extracellular segment of the $P2X_1$ receptor. A key functional residue, glycine-250 ($P2X_1$ receptor numbering), was identified and its predicted location at the C-terminal end of an α-helix confirms such a role. However, it is not possible to predict the tertiary structure of a protein based on amino acid sequence alone; the so-called 'protein-folding problem'. Therefore, this thesis has relied upon sequence analysis and protein structure prediction methods as tools for extracting structural information from P2X receptor sequences. In particular, searching for structural templates on which to model both the putative ligand-binding segment of the $P2X_1$ receptor and the newly defined cysteine-rich domain of the extracellular segment.

Published homology models for the rat $P2X_4$ receptor using class II aminoacyl-tRNA synthetases as structural templates exist. Such models have been validated by mutagenesis studies and residues thought to be important in ATP-binding at the $P2X_4$ receptor have been identified. These residues have been aligned to the $hP2X_1$ receptor sequence and the corresponding residues mutated to cysteine. It is clear that this $P2X_4$ receptor model does not directly translate as a model of ATP-binding at the $hP2X_1$ receptor due to inconsistencies in mutagenesis data and the unreliable nature of the original homology model. In contrast, a model of ATP-binding at the $P2X_1$ receptor based on experimental data does provide an interesting insight into those residues involved in ATP-binding at the $P2X_4$ receptor thus enhancing the published homology models and validating the $P2X_1$ receptor models of ATP-binding.

Acknowledgments

To Richard for his commitment to me and my thesis and to getting it right,

to Mike for his enthusiasm,

to Jon for teaching me all I know in the lab,

to Catherine for our TV chats and body conditioning class,

to Milly for being my great friend,

to Phill for helping me all the time,

to Mum and Dad for money,

to Aaron for being Aaron,

Thank you!

Abbreviations

αβ-meATP	αβ-methylene ATP
2D	two dimensional
3D	three dimensional
3DPSSM	three dimensional position specific scoring matrix
AATS	aminoacyl-tRNA synthetases
ADP	adenosine diphosphate
AFM	atomic force microscopy
AMP	adenosine monophosphate
AMPA	alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
Ap_5A	P(1),P(5)-di(adenosine 5')-pentaphosphate
ATP	adenosine triphosphate
BLAST	basic local alignment search tool
BzATP	2',3'-O-(4-benzoyl)-ATP
CNS	central nervous system
cRNA	capped RNA
DIDS	4,4'-diisothiocyanatostilbene-2,2'-disulfonic acid
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DRG	dorsal root ganglion
DSP	dithio*bis*(succinimidylpropionate)
DTSSP	3, 3'-Dithiobis(sulfosuccinimidylpropionate)
ER	endoplasmic reticulum
GABA	gamma-aminobutyric acid neurotransmitter
GPCR	G-protein coupled receptor
HEK	human embryonic kidney cells
HUNK	hormonally upregulated neu-tumor associated kinase
KO	knock out
LTP	long-term potentiaton
mRNA	messenger RNA
MTS	methanethiosulfonate
MTSEA	2-aminoethyl methanethiosulfonate
MTSEA-biotin	N-biotinoylaminoethyl methanethiosulfonate
MTSES	2-sulfonatoethyl methanethiosulfonate

MTSET	[2-(trimethylammonium)-ethyl]-methanethiosulfonate
NANC	non-adrenergic, non-cholinergic
NMDA	N-methyl-D-aspartic acid
NMR	nuclear magnetic resonance
NR	non redundant
P2XR	P2X receptor
PAGE	polyacylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	protein data bank
PKC	protein kinase C
PPADS	pyridoxal-phosphate-6-azophenyl-2',4'-disulfonate
PSI-BLAST	position specific iterated - basic local alignment search tool
RNA	ribonucleic acid
SCAM	substituted cysteine accessibility mutagenesis
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
TM	transmembrane
TM1	transmembrane domain 1
TM2	transmembrane domain 2
TOPITS	threading one dimensional predictions into three dimensional structures
tRNA	transfer RNA
TWIK	the weakly inward rectifying K^+ channel
UTP	uridine triphosphate
WT	wild-type

Chapter One

1.0 Preface

Structural studies give insight into the molecular basis of functional properties and in the case of P2X receptors may be useful for rational drug design. This introduction gives a historical background to ATP as a transmitter and the classification of purinergic receptors. The structure of the P2X receptor family and the evidence for the proposed topology will be discussed in detail giving a current picture of the structure and functional domains.

1.1 Adenosine Triphosphate (ATP)

1.1.1 History

Research on chemical synaptic transmission showed that the stimulation of certain peripheral nerves produced responses in their target organs. With time it became clear that in the majority of cases these responses were mediated by acetylcholine or noradrenaline but there were exceptions and this led to the idea that there must be something else, a neurotransmitter that is non-adrenergic and non-chlinergic (NANC) (Burnstock, 1972).

In 1929 ATP was identified in muscle extracts and in the following years it was recognized as a source of readily available chemical energy for all living cells. Due to its essential intracellular role it seemed unlikely that ATP would be used for any extracellular physiological action as why would cells release such a fundamental molecule? Also the size and charge of ATP could not cross the plasma membrane by simple diffusion (Glynn, 1968; Chaudry, 1982). However the actions of extracellular adenine compounds in the mammalian heart were first reported in 1929 by Drury and Szent-Gyorgyi. They showed that adenosine and AMP extracted from heart muscle had physiological effects including heart block, arterial dilation, lowering of blood pressure and inhibition of intestinal contraction. This led to a number of early investigations into the effects of adenosine and ATP on a variety of tissues which were reviewed in a book by Green and Stoner (1950).

An important advance was made in 1959 by Holton who reported that ATP was released from sensory nerves in rabbit ear arteries (Holton, 1959). Berne (1963) introduced the concept that adenosine might by the physiological regulator of coronary blood flow during reactive hyperaemia (Berne, 1963) and together these advances were the first in a major line of

research concerned with purines as neurotransmitters. It was Burnstock (1972) who suggested that ATP was a transmitter involved in non-adrenergic, non-cholinergic nerve-mediated responses of the smooth muscle in the gastrointestinal tract and bladder (Burnstock, 1972). He followed this with his theory on cotransmission (Burnstock, 1976) where he defined a 'transmitter' to be any substance that is synthesised and stored in nerve cells, released during nerve activity and interacts with specific receptors on the postsynaptic membrane leading to changes in postsynaptic activity and suggested that some species employ multiple transmitters. It is now widely accepted that ATP is an important neurotransmitter or cotransmitter in both the peripheral and central nervous systems (CNS) (Burnstock, 1990; Edwards & Gibb, 1993; Fredholm, 1995). Burnstock went on to distinguish two major types of purinergic receptors; P1 (selective for adenine) and P2 (selective for ATP and ADP). Four subtypes of P1 receptor have been cloned and the subdivision of P2 receptors recognises two familes; P2Y and P2X (Abbracchio & Burnstock, 1994). This classification of purinoceptors will be discussed in detail in section 1.2.

1.1.2 Structure, Synthesis and Release

Structure

ATP is a nucleotide consisting of a purine base (adenine) attached to a ribose sugar by a glycosidic bond. Three phosphate groups are attached at the 5' carbon atom of the ribose sugar (Figure 1.1).

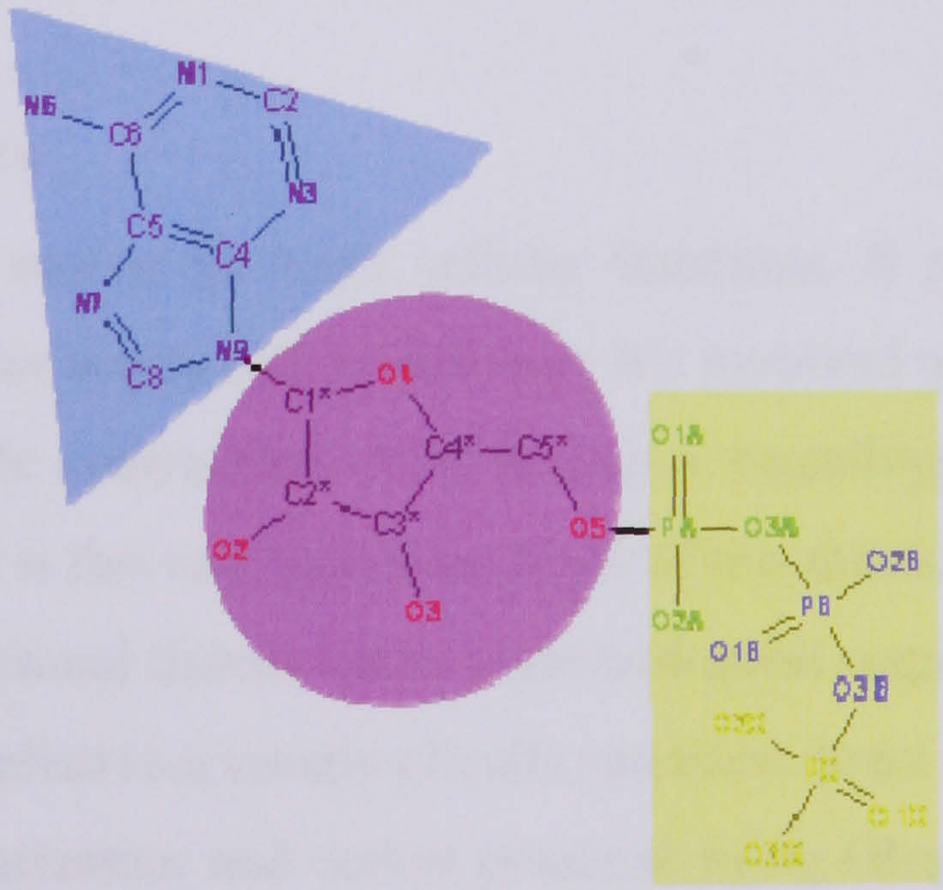


Figure 1.1 Structure of ATP. ATP consists of a purine base (adenine) (blue triangle) attached to a ribose sugar (pink circle) with three phosphate groups (yellow rectangle) attached to the ribose sugar.

Synthesis

Energy is released during hydrolysis of the third phosphate group. ATP can be produced by a number of distinct cellular processes; in eukaryotes the three main pathways used to generate ATP are glycolysis (glucose is metabolized to form pyruvate), the citric acid cycle (Krebs cycle) and oxidative phsphorylation in the mitochondria (ATP generated via the electron transport chain). Adenine nucleotide transporters move ATP out of the internal membrane of the mitochondrion in exchange for ADP.

Release

There are three mechanisms of ATP release:

1 Via ATP-filled vesicles. This is considered to be a major release mechanism in neurons, neuroendocrine cells, mast cells and platelets (Burnstock, 1972). Also, the high concentration of ATP with other neurotransmitters in synaptic vesicles has clarified ATP is co-stored and co-released with other neurotransmitters. Although it is not known how vesicles become loaded with ATP is it known that ATP released in this way is a transmitter signal for nerves in the gut, nodose neurons, in the dorsal root ganglia and in specialized nerves in the brain (Schwiebert & Zsembery, 2003).

2 Through mitochondrial porins (Bodin & Burnstock, 2001). These proteins cross the cell membrane and act as a pore through which molecules can diffuse.

3 By diffusion after sudden rupture of intact cells by tissue injury.

1.1.3 Physiological Roles

ATP is the main energy source for most cellular functions. It plays a critical role in the transport of macromolecules across cell membranes; it's involved in maintaining cell structure and is required for muscle contraction. ATP is also a signalling molecule, recognized by purinergic receptors and it is this role that is the focus of this thesis. Extracellular signalling in both neuronal and non-neuronal tissues exerts wide biological responses. *In vitro* studies have demonstrated that ATP applied to a variety of cells can elicit direct excitatory actions that lead to rapid membrane depolarization and action potential firing (Brake & Julius, 1996). In the central nervous system P2X receptors are widespread and ATP has been shown to be involved in both synaptic transmission and cotransmission with other transmitters such as GABA, glycine and glutamate (Robertson *et al.*, 2001). ATP acting on specific P2X receptors in the CNS has been implicated in thermoregulation where experiments on conscious rats involving

intracerebro-ventricular administration of an ATP analog and P2X antagonists showed that both activation and blockade of central P2X receptors caused changes in body temperature (Gurin *et al.*, 2003). Evidence also exists suggesting that ATP may play a role in Long-Term Potentiation (LTP) which has been implicated in memory formation (Wieraszko, 1996). Such observations as these prove that ATP can act as an excitatory neurotransmitter in the CNS.

In non-neuronal tissue, extracellular ATP is known to act on a variety of cells within the immune system, including mast cells, macrophages and lymphocytes and has been implicated as an initiator of programmed cell death (for a full review see North 2002). ADP and ATP also play a crucial role in platelet activation, and their receptors ($P2Y_1$, $P2Y_{12}$ and $P2X_1$) are potential targets for antithrombotic drugs (for a review see Gachet 2006).

1.2 Purinergic Receptors

'Purinergic receptors' were first formally proposed by Burnstock and classified into two subgroups, P1 purinoceptors at which adenosine is the main ligand and P2 purinoceptors activated by ATP and ADP (Burnstock, 1972). P1 receptors all couple to G proteins and were divided into two subtypes, A1 and A2 (van Calker *et al.*, 1979; Londos *et al.*, 1980). A2 receptors were further divided into types A_{2A} and A_{2B} based on the identification of high (A_{2A}) and low affinity (A_{2B}) binding sites in a human fibroblast cell line (Bruns *et al.*, 1986). The evidence for these distinct receptors came from cloning experiments while recombinant studies revealed distinct pharmacological profiles. The fourth distinct adenosine receptor to be identified was the A3 receptor cloned from rat striatum (Zhou *et al.*, 1992). This was identical to a previously isolated clone from a rat testis cDNA library encoding a G protein-coupled receptor (GPCR) although its ligand had not then been identified (Meyerhof *et al.*, 1991).

In 1985 Burnstock and Kennedy, based on agonist potencies proposed two subtypes of P2-purinoceptors namely P2X receptors and P2Y receptors (Burnstock & Kennedy, 1985). In the following years, a number of other P2-purinoceptors subtypes were proposed including a P_{2T} receptor which is ADP-selective and responsible for platelet aggregation, a P_{2Z} receptor first identified on mast cells and mediates cytotoxity and degranulation (Gordon 1986) and a P_{2U} receptor activated equally by ATP and UTP and widely distributed (O'Connor *et al.*, 1991). In addition P_{2S} (Wiklund & Gustafsson, 1988), P_{2R} (von Kugelgen & Starke, 1990), P_{2D} (Pintor *et al.*, 1993), uridine nucleotide-specific receptors ('pyrimidinoceptors') (Seifert & Schultz, 1989; von Kugelgen & Starke, 1990), P3 (Shinozuka *et al.*, 1988) and P4 (Pintor & Miras-

Portugal, 1995) receptors have also been proposed and of these the P_{2T}, P_{2U}, P_{2Z} and uridine nucleotide-specific receptors have been cloned. However, receptor sub classification based on pharmacological criteria alone is not considered evidence enough so the identity of the other proposed subtypes of P2 receptor remains to be proven.

The first P2 purinoceptor to be cloned was a GPCR isolated from chick brain and termed the $P2Y_1$ receptor (Webb *et al.*, 1993). About the same time, a second GPCR, $P2Y_2$, was isolated from mouse neuroblastoma (Lustig *et al.*, 1993). A year later, Valera *et al* (1994) and Brake *et al* (1994) simultaneously cloned the $P2X_1$ receptor from the rat vas deferens and the $P2X_2$ receptor from rat PC12 cells respectively (Brake *et al.*, 1994; Valera *et al.*, 1994). On the basis of these cloning studies the P2 receptors were classified into a P2Y (G-protein-coupled) family and a P2X (ion-gated channel) family (Abbracchio & Burnstock, 1994) (Figure 1.2).

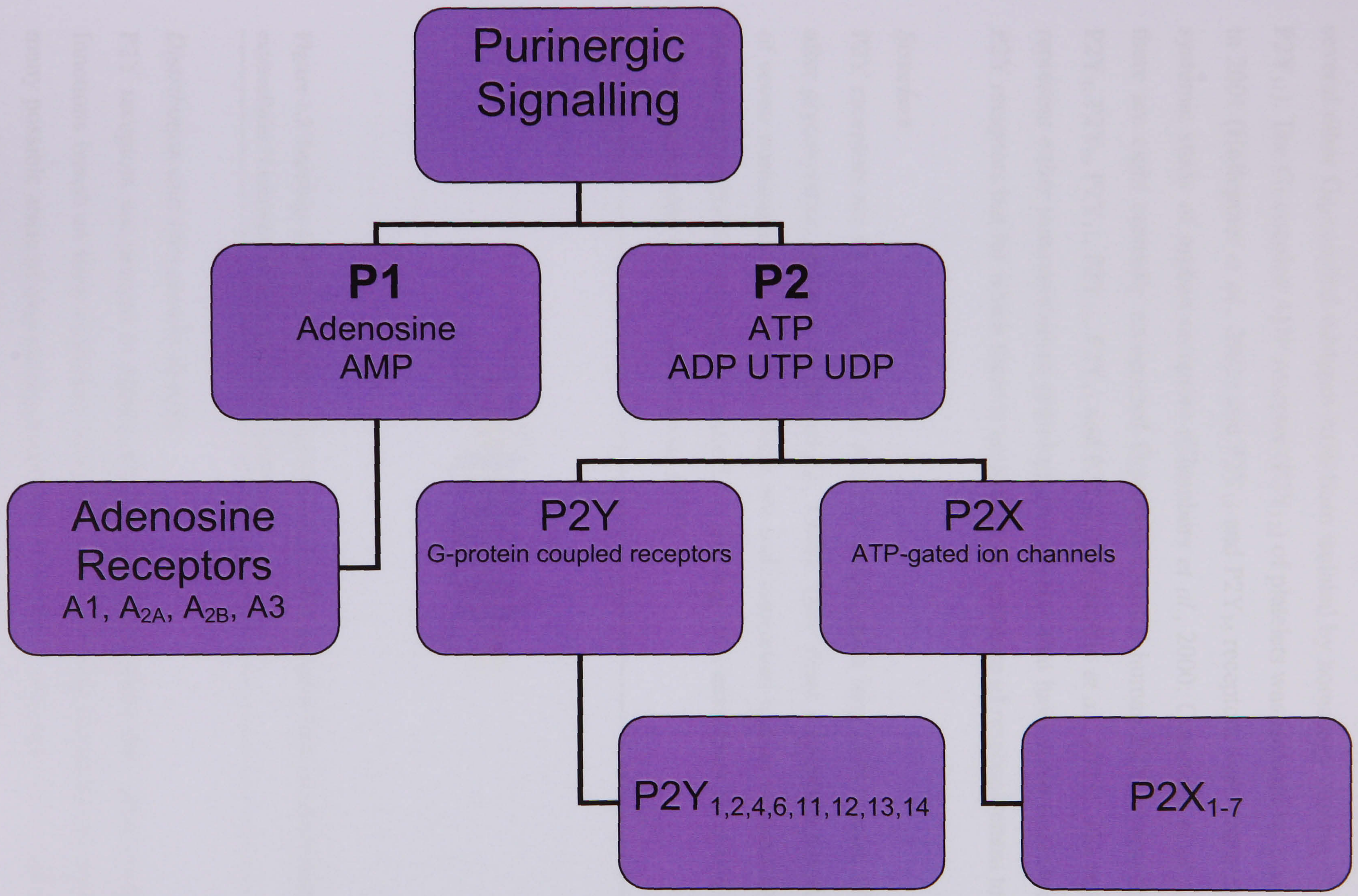


Figure 1.2 Purinergic Signalling Flow Chart. Purinergic receptors are classified into two subgroups; P1 and P2 purinoceptors at which adenosine and ATP are the main ligands respectively.

1.2.1 P2Y Receptors

History

Since the first cloned P2 receptors ($P2Y_1$ and $P2Y_2$) (Lustig *et al.*, 1993; Webb *et al.*, 1993) several other Gq-coupled subtypes have been isolated by homology cloning ($P2Y_4$, $P2Y_6$ and $P2Y_{11}$). The Gi-coupled ADP receptor ($P2Y_{12}$) of platelets was isolated by expression cloning in 2001 (Hollopeter *et al.*, 2001) and $P2Y_{13}$ and $P2Y_{14}$ receptors were characterized during a systemic study of orphan receptors (Chambers *et al.*, 2000; Communi *et al.*, 2001). Today, there are eight currently recognized G-protein coupled human P2Y receptors; $P2Y_1$, $P2Y_2$, $P2Y_4$, $P2Y_6$, $P2Y_{11}$, $P2Y_{12}$, $P2Y_{13}$ and $P2Y_{14}$ (Abbracchio *et al.*, 2003). The missing numbers represent either nonmammalian orthologs or receptors that have some sequence homology to P2Y receptors but for which there is no functional evidence of responsiveness to nucleotides.

Structure

P2Y receptors are all in the range of 308-377 amino acids long with masses of 41 to 53 kDa after glycosylation (Ralevic & Burnstock, 1998). They share a common topology consisting of seven transmembrane domains which are well conserved and an extracellular N-terminus and an intracellular C-terminus (Figure 1.3). Binding sites have been identified on the 6^{th} and 7^{th} transmembrane domains (Burnstock, 1997).

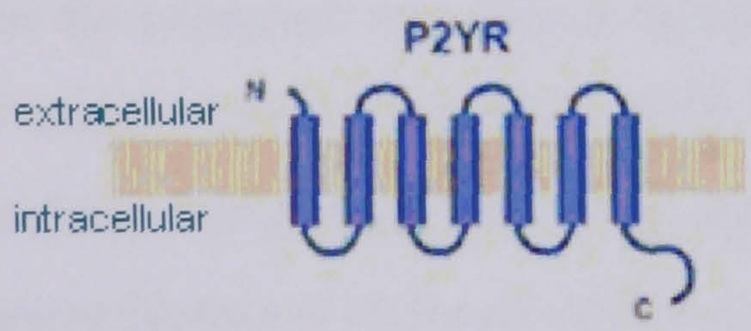


Figure 1.3 Topology of a P2Y receptor. G-protein-coupled P2Y receptors have seven transmembrane domains, extracellular N terminus and intracellular C terminus.

Distribution and Therapeutic Benefit

P2Y receptors are present in almost all human tissues where they exert various biological functions based on their G-protein coupling. The following examples are just a few of the many possible areas of therapeutic benefit the P2Y family represent. For a full review of P2Y receptor distribution and function see Burnstock and Knight 2004.

- Platelets express $P2X_1$, $P2Y_1$ and $P2Y_{12}$ receptors (the role of the $P2X_1$ receptor will be discussed in more detail later) and they are all involved in blood clotting. The presence and function of the $P2Y_1$ receptor was confirmed by the use of selective $P2Y_1$ antagonists (Leon *et al.*, 1997; Hechler *et al.*, 1998) and by studies of $P2Y_1$ knockout (KO) mice showing the bleeding phenotype (Fabre *et al.*, 1999; Leon *et al.*, 1999). Each receptor has a selective role during platelet aggregation (Hechler *et al.*, 2005) and the use of P2Y antagonists as antithrombotic agents in the prevention of recurrent stroke and heart attacks represents a very interesting therapeutic application. $P2Y_{12}$ is the target of the anti-thrombotic drug Clopidogrel. This drug produces a metabolite which acts as an irreversible $P2Y_{12}$ receptor antagonist and stops bleeding (Abbracchio *et al.*, 2006).

- $P2Y_2$ receptors are involved in insulin release from pancreatic islet β cells and may be a good candidate for innovative antidiabetic drugs (Solini *et al.*, 2003).

- All cloned P2Y receptors have been found in both healthy and failing human hearts (Banfi *et al.*, 2005) which supports a functional role in myocardial function (Vassort, 2001).

- P2Y receptors are expressed in all regions of the kidney tubule and glomerulus. $P2Y_{1,4,6}$ receptors in the proximal convoluted tubule and loop of Henle are involved in the reabsorption of water, ions and nutrients. P2Y receptors mediate rennin secretion and P2Y receptor-related compounds are being explored for the treatment of chronic renal failure and transplantation-induced erythrocytosis. In addition, an increase in expression of $P2Y_2$ and $P2Y_6$ receptors has been observed in polycystic kidney disease and therefore P2Y antagonists and inhibitors of ATP release may represent possible therapeutic agents (Turner *et al.*, 2004).

1.2.2 P2X Receptors

The focus of this thesis lies with the P2X receptor family and more specifically the $P2X_1$ receptor.

Seven genes encoding P2X receptors namely $P2X_{1-7}$ were cloned in the mid-1990s. The first two P2X receptor subunits were isolated from rat vas deferens smooth muscle ($rP2X_1$

receptor) (Valera *et al.*, 1994) and PC12 cells (rP2X$_2$ receptor) (Brake *et al.*, 1994) by expression cloning. This is a technique that requires no prior knowledge of DNA sequence or location. cRNA from the tissue of interest expressing the receptor is transfected into cells and then isolated by selection and cloning. In the case of the rP2X$_1$ and rP2X$_2$ receptor, cRNA was injected into oocytes. Five additional subunits of the P2X receptor family, rat P2X$_{3-7}$ receptors were subsequently cloned using either screening of diverse libraries or polymerase chain reaction (PCR) techniques or by a combination of the two (Chen *et al.*, 1995) (Bo *et al.*, 1995; Buell *et al.*, 1996; Collo *et al.*, 1996; Surprenant *et al.*, 1996). The characterization of the P2X receptor family has been further helped by the cloning of various human homologous cDNAS; hP2X$_1$ (Valera *et al.*, 1995), hP2X$_2$ (Lynch *et al.*, 1999), hP2X$_3$ (Garcia-Guzman *et al.*, 1997b), hP2X$_4$ (Garcia-Guzman *et al.*, 1997a), hP2X$_6$ (Urano *et al.*, 1997) and hP2X$_7$ (Rassendren *et al.*, 1997b). However, efforts to amplify a "full-length" cDNA were unsuccessful for the hP2X$_5$ receptor sequence (Le *et al.*, 1997).

The seven P2X subunits range from 384 to 595 amino acids in length and are 40-55% pairwise identical with the P2X$_7$ receptor sequence least like the others (Figure 1.4) (North, 2002). This amino acid conservation across the family hints at important structural features or functional domains.

Hydrophobicity plots provided the initial evidence for the suggested membrane topology. Two significant runs of approximately 20 amino acids, long enough to cross the plasma membrane if an α-helical structure was assumed were identified (Brake *et al.*, 1994). Both of these regions were characteristic of transmembrane domains because they had an excess of positively charged amino acids at their predicted cytoplasmic ends and aromatic residues were at interfacial regions. In addition sequence analysis revealed the absence of a signal peptide sequence after the starting methionine suggesting that P2X receptors have intracellular amino and carboxy termini (Brake *et al.*, 1994; Soto *et al.*, 1997). This provided the basis for the model in Figure 1.5.

```
hP2X1           ----------MARRFQEELAAFLFEYDTPRMVLVRNKKVGVIFRLIQLVVLVYVIGWV 48
hP2X2           MAAAQPKYPAGATARRLARGCWSALWDYETPKVIVVRNRRLGVLYRAVQLLILLYFVWYV 60
hP2X3           ------------------MNCISDFFTYETTKSVVVKSWTIGIINRVVQLLIISYFVGWV 42
hP2X4           ----------MAGCCSALAAFLFEYDTPRIVLIRSRKVGLMNRAVQLLILAYVIGWV 47
hP2X5           ----------MGQAGCKGLCLSLFDYKTEKYVIAKNKKVGLLYRLLQASILAYLVVWV 48
hP2X6           ----------MGSPGAT-TGWGLLDYKTEKYVMTRNWRVGALQRLLQFGIVVYVVGWA 47
hP2X7           ---------------MPACCSCSDVFQYETNKVTRIQSMNYGTIKWFFHVIIFSYVC-FA 44
                               .  *.* :    :.   * :   .:  :. *.  :.

hP2X1           FLYEKGYQTS-SGLISSVSVKLKGLAVTQ-------LPGLGPQVWDVADYVFPAQGDNSF 100
hP2X2           FIVQKSYQESETGPESSIITKVKGITTS----------EHKVWDVEEYVKPPEGGSVF 108
hP2X3           FLHEKAYQVRDTAIESSVVTKVKGSGLY----------ANRVMDVSDYVTPPQGTSVF 90
hP2X4           FVWEKGYQETDSVVS-SVTTKVKGVAVTN-------TSKLGFRIWDVADYVIPAQEENSL 99
hP2X5           FLIKKGYQDVDTSLQSAVITKVKGVAFTN-------TSDLGQRIWDVADYVIPAQGENVF 101
hP2X6           LLAKKGYQERDLEPQFSIITKLKGVSVTQ-------IKELGNRLWDVADFVKPPQGENVF 100
hP2X7           LVSDKLYQRKEPVIS-SVHTKVKGIAEVKEEIVENGVKKLVHSVFDTADYTFPLQG-NSF 102
                :: .* **        :: .*:**                    : *. ::. * :  . :

hP2X1           VVMTNFIVTPKQTQGYCAEHP--EGGICKEDSGCTPGKAKRKAQGIRTGKCVAFNDTVK- 157
hP2X2           SIITRVEATHSQTQGTCPESIRVHNATCLSDADCVAGELDMLGNGLRTGRCVPYYQGPSK 168
hP2X3           VIITKMIVTENQMQGFCPESE--EKYRCVSDSQCGP--EPLPGGGILTGRCVN-YSSVLR 145
hP2X4           FVMTNVILTMNQTQGLCPEIP-DATTVCKSDASCTAGSAGTHSNGVSTGRCVAFNGSVK- 157
hP2X5           FVVTNLIVTPNQRQNVCAENEGIPDGACSKDSDCHAGEAVTAGNGVKTGRCLRRENLARG 161
hP2X6           FLVTNFLVTPAQVQGRCPEHPSVPLANCWVDEDCPEGEGGTHSHGVKTGQCVVFNGTHR- 159
hP2X7           FVMTNFLKTEGQEQRLCPEYP-TRRTLCSSDRGCKKGWMDPQSKGIQTGRCVVHEGNQK- 160
                 ::*..  *  * *  *.*       *  *  *        . *: **:*:

hP2X1           TCEIFGWCPVEVDDDIPRPALLREAENFTLFIKNSISFPRFKVNRRNLVEEVNAAHMKTC 217
hP2X2           TCEVFGWCPVEDGASVSQF-LGTMAPNFTILIKNSIHYPKFHFSKGN-IADRTDGYLKRC 226
hP2X3           TCEIQGWCPTEVDT-VETP-IMMEAENFTIFIKNSIRFPLFNFEKGNLLPNLTARDMKTC 203
hP2X4           TCEVAAWCPVEDDTHVPQPAFLKAAENFTLLVKNNIWYPKFNFSKRNILPNITTTYLKSC 217
hP2X5           TCEIFAWCPLETSSRPEEP-FLKEAEDFTIFIKNHIRFPKFNFSN-NVMDVKDRSFLKSC 219
hP2X6           TCEIWSWCPVESGVVPSRP-LLAQAQNFTLFIKNTVTFSKFNFSKSNALETWDPTYFKHC 218
hP2X7           TCEVSAWCPIEAVEEAPRPALLNSAENFTVLIKNNIDFPGHNYTTRNILPGLN----ITC 216
                ***: .*** *         :   * :**:::** : :. .:    * :          *

hP2X1           LFHKTLHPLCPVFQLGYVVQESGQNFSTLAEKGGVVGITIDWHCDLDWHVRHCRPIYEFH 277
hP2X2           TFHEASDLYCPIFKLGFIVEKAGESFTELAHKGGVIGVIINWDCDLDLPASECNPKYSFR 286
hP2X3           RFHPDKDPFCPILRVGDVVKFAGQDFAKLARTGGVLGIKIGWVCDLDKAWDQCIPKYSFT 263
hP2X4           IYDAKTDPFCPIFRLGKIVENAGHSFQDMAVEGGIMGIQVNWDCNLDRAASLCLPRYSFR 277
hP2X5           HFGPK-NHYCPIFRLGSVIRWAGSDFQDIALEGGVIGINIEWNCDLDKAASECHPHYSFS 278
hP2X6           RYEPQFSPYCPVFRIGDLVAKAGGTFEDLALLGGSVGIRVHWDCDLDTGDSGCWPHYSFQ 278
hP2X7           TFHKTQNPQCPIFRLGDIFRETGDNFSDVAIQGGIMGIEIYWDCNLDRWFHHCRPKYSFR 276
                 :       **::::* :.  :*  *  :*  ** :*: : * *:**     * * *.*

hP2X1           GLYE---EKNLSPGFNFRFARHFVEN-GTNYRHLFKVFGIRFDILVDGKAGKFDIIPTMT 333
hP2X2           RLD--PKHVPASSGYNFRFAKYYKIN-GTTTRTLIKAYGIRIDVIVHGQAGKFSLIPTII 343
hP2X3           RLDSVSEKSSVSPGYNFRFAKYYKMENGSEYRTLLKAFGIRFDVLVYGNAGKFNIIPTII 323
hP2X4           RLDTRDVEHNVSPGYNFRFAKYYRDLAGNEQRTLIKAYGIRFDIIVFGKAGKFDIIPTMI 337
hP2X5           RLDN-KLSKSVSSGYNFRFARYYRDAAGVEFRTLMKAYGIRFDVMVNGKG---------- 327
hP2X6           LQE---------KSYNFRTATHWWEQPGVEARTLLKLYGIRFDILVTGQAGKFGLIPTAV 329
hP2X7           RLDDKTTNVSLYPGYNFRYAKYYKEN-NVEKRTLIKVFGIRFDILVFGTGGKFDIIQLVV 335
                          .:*** * ::    .   * *:* :***:*::* * .

hP2X1           TIGSGIGIFGVATVLCDLLLLHILP----------------KRHYYKQKKFKYAEDMG 375
hP2X2           NLATALTSVGVGSFLCDWILLTFMN----------------KNKVYSHKKFDKVCTPS 385
hP2X3           SSVAAFTSVGVGTVLCDIILLNFLK----------------GADQYKAKKFEEVN--- 362
hP2X4           NIGSGLALLGMATVLCDIIVLYCMK----------------KRLYYREKKYKYVED-- 377
hP2X5           ----------AFFCDLVLIYLIK----------------KREFYRDKKYEEVRGLE 357
hP2X6           TLGTGAAWLGVVTFFCDLLLLYVDR----------------EAHFYWRTKYEEAKAPK 371
hP2X7           YIGSTLSYFGLAAVFIDFLIDTYSSNCCRSHIYPWCKCCQPCVVNEYYYRKKCESIVEPK 395
                          :.: * ::                              *  .* .

hP2X1           PG----------------AAERDLAATSSTLGLQENMRTS------------------ 399
hP2X2           HPSGSWPVTLA-------RVLGQAPPEPGHRSEDQHPSPPSGQEGQQG---------- 426
hP2X3           ------ETTLK-------IAALTNPVYPSDQTTAEKQSTDSG---------------- 391
hP2X4           ------------------YEQGLASELDQ---------------------------- 388
hP2X5           DSSQEAEDEASGLGLSEQLTSGPGLLGMPEQQELQEPPEAKR---------------- 399
hP2X6           ATANSVWRELA-------LASQARLAECLRRSSAPAPTATAA---------------- 406
hP2X7           PTLKYVSFVDESHIR---MVNQQLLGRSLQDVKGQEVPRPAMDFTDLSRLPLALHDTPPI 452

hP2X2           AECGPAFPPLRPCPISAPSEQMVDTPASEPAQASTPTDPKGLAQL-------------- 471
hP2X3           -----AFS------IGH---------------------------------------- 397
hP2X5           --------------GSSSQKGNGSVCP---QLLEPHRST------------------- 421
hP2X6           --------------GSQTQTP-GWPCPSSDTHLPTHSGSL----------------- 431
hP2X7           PGQPEEIQLLRKEATPRSRDSPVWCQCGSCLPSQLPESHRCLEELCCRKKPGACITTSEL 512
                FRKLVLSRHVLQFLLLYQEPLLALDVDSTNSRLRHCAYRCYATWRFGSQDMADFAILPSC 572
                CRWRIRKEFPKSEGQYSGFKSPY                                     595
```

Figure 1.4 Multiple Sequence Alignment human P2X Receptor Family (opposite) generated by ClustalW (Thompson *et al.*, 1994). Consensus symbols: '*' residues are identical in all the sequences in the alignment, ':' conserved substitutions observed and '.' semi-conserved substitutions observed.

Although topology predictions based on amino acid sequence alone can provide some insight into structure they require substantiation and with no 3D structure available for the P2X receptor family, structure predictions have relied upon experimental and mutagenesis techniques to identify residues and domains of importance. The next section (Section 1.3) will examine the topology of the P2X receptor in a simple outside-in approach (Figure 1.5) providing evidence for the predicted structure and the insight this has given to possible function.

1.3 Structure and Function of the P2X Receptor Family

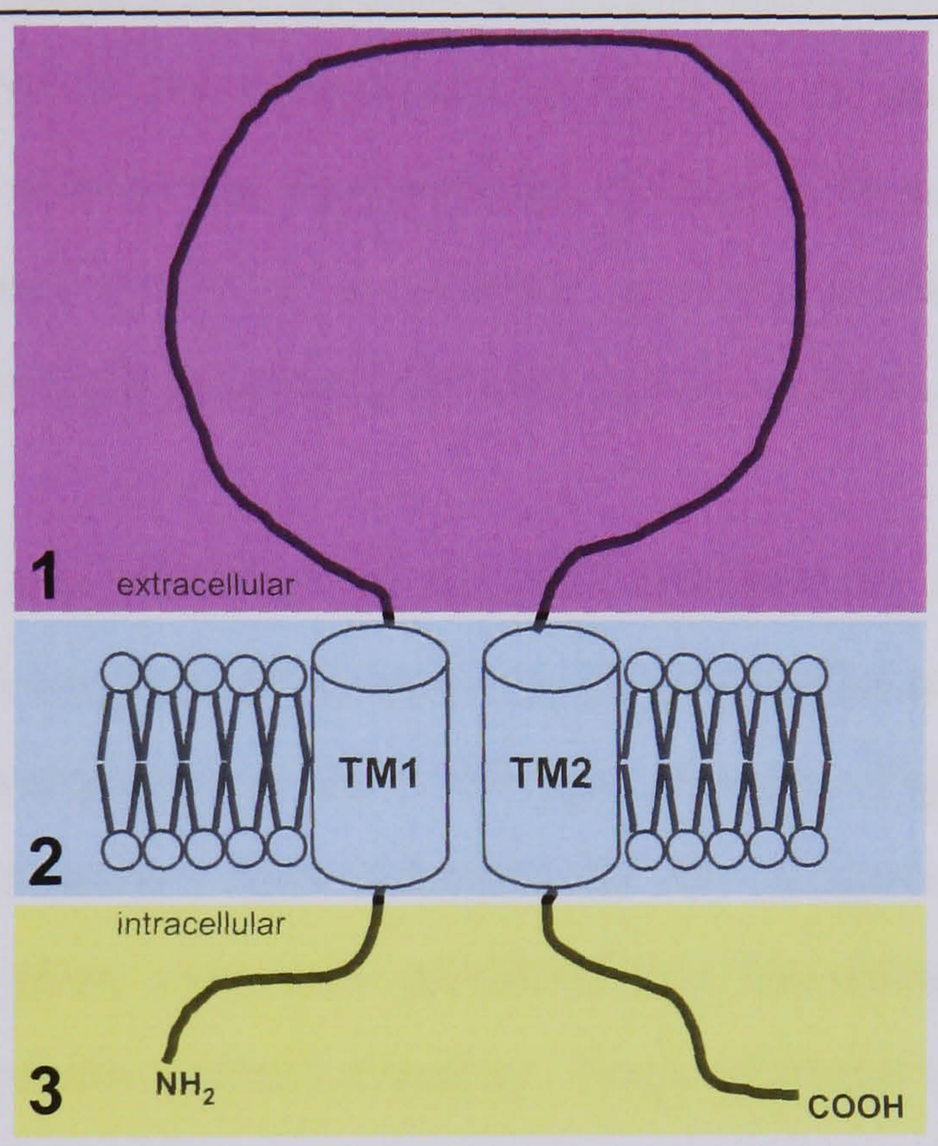


Figure 1.5 Topology of a P2X receptor subunit. The common topology shared by a P2X receptor subunit has a large extracellular domain (1. shaded pink), two transmembrane domains (2. shaded blue) and intracellular amino and carboxy termini (3. shaded yellow).

1. Extracellular domain

Topology

The predicted large extracellular loop of the P2X receptor family (~280 amino acids) constitutes two thirds of the receptor mass. Evidence from the analysis of glycosylation patterns supports this topology. N-linked glycosylation is a post-translational modification, that occurs at asparagine residues and it is the process whereby sugars bind to proteins forming glycoproteins. An asparagine residue can only accept a sugar if it is part of a consensus sequence N-X-S/T and if the sequence is located extracellular. N-glycosylation begins in the endoplasmic reticulum (ER) and is completed within the Golgi complex. Folding of the protein in the golgi membrane means that residues/regions that will be

extracellular are inside the golgi and subject to modification. Therefore glycosylation can only occur at positions that are to be located extracellularly. All P2X receptor subunits have consensus sequences for N-linked glycosylation. By determining the location of new "introduced" glycosylation sites the boundaries of the extracellular loop can be determined. This technique has been used in different studies with different P2X receptor subtypes to gain an overall understanding of those parts of the receptor that lie extracellularly. These individual studies will be considered in chronological order in more detail next.

Valera *et al* (1995) identified 5 potential sites for N-glycosylation in the predicted extracellular region of the human P2X protein and noted that they were conserved in rat and mouse (Valera *et al.*, 1995). *In vitro* translation in the absence and presence of microsomes was used to examine the size of the recombinant human P2X protein. In the presence of microsomes a 60kD product was seen in contrast to 45kD in the absence of microsomes suggesting that the receptor is glycosylated, thus the central section must be extracellular.

This work was confirmed and extended for the $P2X_2$ receptor by Newbolt *et al* (1998) who identified three asparagine residues all located in the extracellular domain at positions 182, 239 and 298 that are glycosylated in the wild-type receptor (Figure 1.6) (Newbolt *et al.*, 1998). Individual point mutations were tolerated by the receptor to produce a functional channel however multiple point mutations produced non-functional channels. Therefore full glycosylation is not required for normal function. Three artificial consensus sites (N-X-S/T) were constructed in the predicted N-terminal region (asparagine at position 9, 16 or 26). None of these sites were glycosylated supporting the idea of an intracellular N-terminus. Similar consensus sites were introduced in the at the C-terminal end of the first transmembrane domain (TM1) (position 62 and 66) and at the N-terminal end of the second transmembrane domain (TM2) (position 324). Residues P62N, I66N and K324N were glycosylated indicating that they lie extracellularly. In the case of the $P2X_2$ receptor these studies agree with the proposed topology of intracellular N and C-termini, two hydrophobic domains and a large extracelullar loop. It is presumed the other homologous members of the family possess this same topology.

The effect of prevention or removal of glycosylation in recombinant $P2X_2$ receptors and its effect on expression and function was also investigated (Torres *et al.*, 1998a). Two approaches were used; tunicamycin treatment to prevent glycosylation (at glycosylation sites located at positions 182, 239 and 298 (Torres *et al.*, 1998b)) and site-directed mutagenesis to

remove the glycosylation sites. Tunicamycin specifically inhibits the first step in the formation of the core oligosaccharide in N-linked glycosylation. HEK293 cells stably transfected with the $P2X_2$ receptor subunit showed little or no response to ATP after tunicamycin treatment and loss of function was observed with the elimination of all three N-linked glycosylation sites from the $P2X_2$ receptor by site-directed mutagenesis. Expression of the non-glycosylated receptors produced by either method was greatly reduced suggesting that the non-glycosylated $P2X_2$ receptors were retained inside the cell. This provided evidence for the significant role of N-linked glycosylation in the cell surface expression of a P2X receptor subunit (Torres *et al.*, 1998a).

Rettinger *et al* (2000) studied the contribution of N-glycosylation to the assembly, surface appearance and ligand recognition of the $P2X_1$ receptor expressed in *Xenopus laevis* oocytes (Rettinger *et al.*, 2000). Analysis of mutants possessing glutamine instead of asparagine at the consensus site revealed that four out of the five putative N-glycosylation sites in rat $P2X_1$ sequence are used, Asn^{284} being the exception due to a proline residue in the +4 position (Nicke *et al.*, 1998). Mutants migrated as homotrimers when resolved by blue native polyacylamide gel electrophoresis (PAGE). If all four sites were eliminated $P2X_1$ receptors did not form and if three sites were eliminated, the formation of the receptor was severely impaired. Therefore at least one N-glycan per subunit is required for the formation of the $P2X_1$ receptor. Elimination of Asn^{210} caused a 3-fold reduction in the potency of ATP suggesting it may contribute to ligand binding.

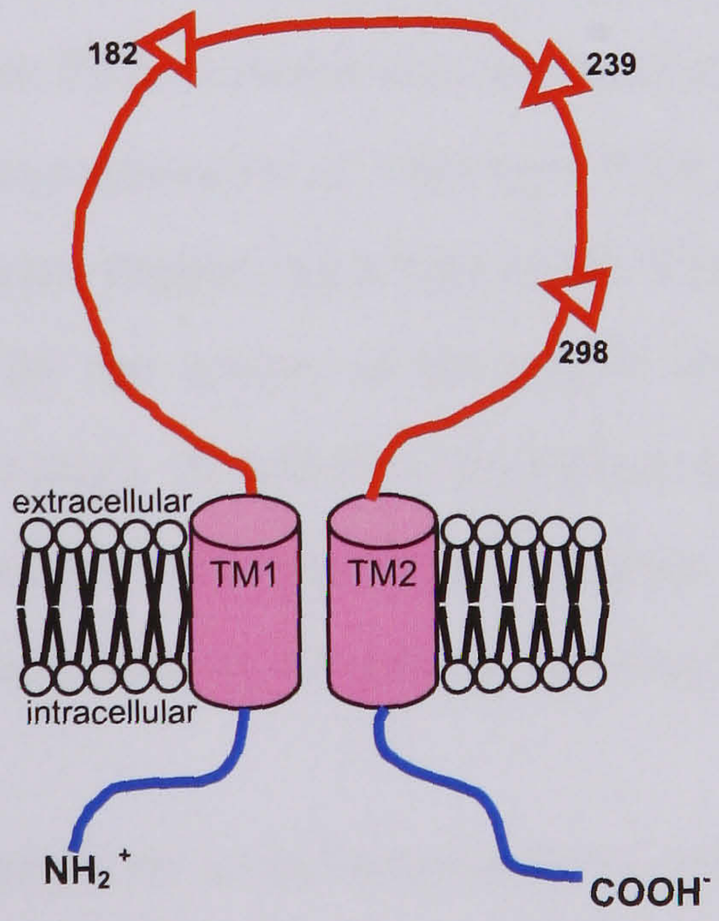


Figure 1.6 Topology of a P2X receptor subunit. The common topology shared by a P2X receptor subunit has intracellular amino and carboxy termini (blue), two transmembrane domains (pink) and a large extracellular domain (red). N-linked glycosylation sites are marked with red triangles (numbering refers to $P2X_2$ receptor).

Interaction of Subunits

The extracellular loop is predicted to contain the ligand binding site of the P2X receptor subunit as extracellular ATP application causes the channel to open. It is thought that P2X channels have three binding sites for ATP (Bean, 1990) but it is not known whether each subunit is required for the binding of one molecule of ATP. Analysis of single channel currents has suggested that sequential binding of ATP shows positive co-operativity and that partially liganded channels do not open (Ding & Sachs, 1999). The question of where the agonist binding site(s) are located is raised and whether it is formed within one subunit or at the interface of two neighbouring subunits. Analysis of many ATP binding proteins shows common consensus ATP binding motifs such as the Walker motif (Traut, 1994). However such motifs are not present in P2X receptors and due to the lack of sequence homology between P2X receptors and other known ATP binding proteins the binding sites have not yet been located.

Studies that address the role of individual P2X subunits in ATP binding have brought further understanding of the location of the binding site. In a recent study, individual point mutations of residues K69 and K308 (rat $P2X_2$ receptor numbering) to alanine in the $P2X_2$ and $P2X_3$ receptors were chosen to investigate the binding site location (Wilkinson *et al.*, 2006) due to previous work which implicated both residues in ATP-binding (Ennion *et al.*, 2000). Currents recorded from HEK293 cells expressing each mutant channel (K63A and K299A $P2X_3$ receptor numbering) in the $P2X_3$ receptor prevented channel function. Coexpression of wild-type $P2X_3$ and mutated $P2X_2$ subunit revealed that the heteromeric $P2X_{2/3}$ channel functioned normally when either lysine in the $P2X_2$ subunit was mutated but not when both lysines were mutated to alanine. Conversely, coexpression of wild-type $P2X_2$ with a mutated $P2X_3$ subunit produced non-functional heteromers suggesting a bias to the $P2X_3$ receptor subunit in channel formation. This was confirmed by the rescue of the single lysine mutant $P2X_2$ subunit by wild-type $P2X_3$ (but not the converse). In addition the failure to rescue function in the $P2X_2$ subunit with both lysines mutated by wild-type $P2X_3$ suggests that these residues from two different subunits interact in agonist binding or channel opening (Wilkinson *et al.*, 2006).

The idea the ATP-binding site might be located between receptor subunits and that residues from different subunits bind ATP was the subject of another mutagenesis study that used a disulfide cross-linking approach to identify pairs of residues that are in close proximity of the predicted location of the ATP binding site of the $P2X_1$ homotrimer (Marquez-Klaka *et al.*, 2007). Eight residues in the extracellular domain of the $P2X_1$ receptor were replaced by

cysteine. All eight residues had previously been shown to be essential for high ATP potency. Every pairwise combination of each single cystine mutant was expressed in *Xenopus laevis* oocytes and the purified receptor complexes were analyzed by nonreducing SDS-PAGE to scan for disulfide formation between P2X subunits. Of all the pairings, a band corresponding in size to a dimer was only observed with the K68C and F291C mutants indicating the close proximity of these residues. A successful cross link of three subunits each containing the double mutant K68C/F291C was achieved and observed at ~150kDa, the predicted size of a $P2X_1$ trimer. Previous mutagensis data suggests that residues K68 (Ennion *et al.*, 2000) and F291 (Roberts & Evans, 2004) are involved in ATP-binding at the $P2X_1$ receptor. If the ATP-binding site is formed between neighbouring subunits as the cross-linking would suggest, binding of ATP should influence or prevent disulfide bond formation. This was tested by expressing the K68C and F291C single-mutant pair and the respective double mutant subunits and treating the receptor complexes during the purification process (while they were still attached to the Ni^{2+}-NTA agarose beads) with β-mercaptoethanol. Elution was then performed under nondenaturing conditions in the presence or absence of varying concentrations of ADP or ATP. The results showed that the specific binding of ATP prevented the close interaction of K68C and F291C either by directly interacting with one or both of these residues or by restraining the receptor in a conformation in which the distance between these residues is constrained (Marquez-Klaka *et al.*, 2007). Taken together this evidence points to an ATP-binding site at the interface of neighbouring P2X subunits.

The Amino Acid Sequence

Mutagenesis studies that have targeted conserved residues have been used to identify particular areas or residues of importance. Conserved residues across the family are considered to be important as they have remained through the evolutionary process. There are 93 conserved amino acids across at least six P2X family members in the extracellular ligand-binding loop which could be important in terms of structure or could have a functional role. Alanine replacement mutagenesis has been used in the past because alanine is a small, hydrophobic residue so side-chain function should be removed without disrupting structure. The theory is if a residue is important for ATP action then replacement of the conserved amino acid with alanine should change the properties of the receptor. A lack of effect on channel properties/function would indicate that a conserved residue was not essential for normal receptor function. Groups of residues that share particular biochemical properties have been targeted across the extracellular loop of the P2X receptor family. These are summarized below:

Contribution of Conserved Cysteine Residues

The structure of the extracellular domain is heavily influenced by ten cysteine residues that are 100% conserved across the P2X receptor family. Cysteine residues have the ability to form disulfide bonds due to a reactive sulfhydryl group and disulfide bonds are known to play an important role in the formation and maintenance of ion channel structure (Ennion & Evans, 2002a). For example in the nicotinic receptor the disulfide bonds are thought to stabilize individual subunits (Brejc *et al.*, 2001). In the TWIK potassium channel an intersubunit disulfide bond is responsible for dimerization of the receptor (Lesage *et al.*, 1997). However, it is not likely that disulfide bonds between cysteine residues in P2X receptors are intersubunit because a variety of denaturing agents that do not break disulfide bonds still dissociate $P2X_1$ multimers into monomeric receptors. This demonstrates that $P2X_1$ receptor multimers are maintained by noncovalent interactions (Nicke *et al.*, 1998). Experiments using the human $P2X_1$ receptor and a series of single and double cysteine-to-alanine mutations have shown that disulfide bonds can form between the conserved cysteine residues. The following disulfide bonds were proposed based on phenotypic comparisons of single and double cysteine mutants: C117-C165, C126-C149, C132-C159, C217-C227. The significance of this is realized when the structure of the receptor is modeled, as such bonding patterns impose certain restraints on the resulting structure. Clyne *et al* (2002) performed a similar study using the $P2X_2$ receptor suggesting disulfide bonds between C113-C164, C258-C267 and C214-C224 and two additional bonds between C124, C130, C147 and C158 (Clyne *et al.*, 2002) although they were unable to assign specific cysteine pairs to these two bonds. The pairings that were specified were identical to those proposed by Ennion *et al* (2002) (Figure 1.7).

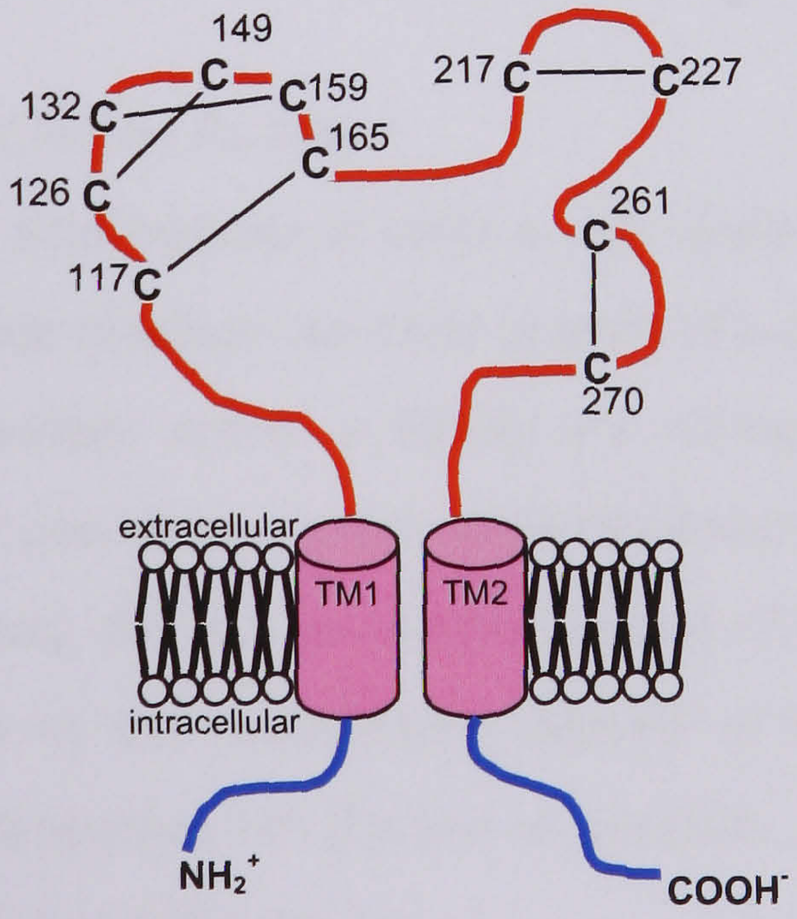


Figure 1.7 Topology of a P2X receptor subunit. The common topology of a P2X receptor subunit has intracellular amino and carboxy termini (blue), two transmembrane domains (pink) and a large extracellular domain (red). Conserved cysteine residues are shown with **C** and the predicted pattern of disulfide bonds are shown with black lines (Ennion & Evans, 2002a).

Contribution of Conserved Proline Residues

In addition to cysteine, proline residues have the ability to influence and constrain the structure and binding/gating movements of the extracellular loop of P2X receptors. Proline residues can influence the secondary structure of proteins by introducing kinks in α-helices due to their unique structure. The side chain of a proline residue is covalently bonded to the main chain nitrogen which results in a ring structure as the main chain amide nitrogen is prohibited from participating in hydrogen bonding. There are four 100% conserved proline residues namely P93, P166, P228 and P272 and three less conserved across the P2X family (P196, P174, P225). Alanine replacement mutagenesis and partial agonists were used with the human $P2X_1$ receptor to determine the role of the conserved prolines.

Partial agonists such as, Ap_5A (P(1),P(5)-di(adenosine 5')-pentaphosphate) or BzATP (2',3'-O-(4-benzoyl)-ATP), activate the $P2X_1$ receptor but can only produce a submaximal response. Such compounds can be used to investigate the reason why a decrease in ATP potency is seen for some mutant $P2X_1$ receptors. When an agonist binds to a receptor it triggers the gating process (the process whereby an ion channel changes state). If a change in ATP potency is observed for a mutant it can be due to a change in binding affinity at the receptor and/or a change in the gating process by which the channel opens. It was found that P93A had a small effect on the action of ATP and experiments involving the partial agonists BzATP and Ap_5A revealed that it is likely P93 is involved in ATP binding. P272A did not form a functional channel and was expressed at low levels in the cell and at the cell surface (Roberts & Evans, 2005).

Contribution of Conserved Glycine Residues

Glycine is a unique amino acid because it lacks a side chain. This gives glycine its flexible nature and as a result, glycine residues can exist in parts of a protein other amino acids would not. Conserved glycine residues across a family are strong indicators of loops/secondary structure breakers and they can allow certain conformational changes or movement within a protein that may be necessary for the successful binding of an agonist. Mutation of the 16 conserved glycine residues of the extracellular domain of the $hP2X_1$ receptor to alanine, revealed that the unique properties of glycine at position 250 are important for channel function (Digby *et al.*, 2005) (see Chapter 3).

Contribution of Positively Charged Residues

There are 11 conserved positive charges in the extracellular loop of P2X receptors. These conserved residues were mutated to alanine in the human $P2X_1$ receptor to determine their role in the binding of the negatively charged phosphate groups of ATP. In particular K68A and K309A were found to dramatically reduce the potency of ATP by >1400-fold while a small change in ATP potency was seen for K70R and R292K/A. These residues are situated close to the predicted position of the two transmembrane domains suggesting that the ATP binding site may form close to the entrance to the pore (Ennion *et al.*, 2000). Jiang *et al* (2000) identified two lysine residues namely K69 and K71 in the rat $P2X_2$ receptor critical for the action of ATP also (Jiang *et al.*, 2000).

Contribution of Negatively Charged Residues

There are seven conserved negatively charged residues within the extracellular loop of P2X receptors and they may contribute to the ligand-binding site as ATP binding generally involves hydrogen bond formation with charged or polar side chains (Jiang *et al.*, 2000). Experiments have been carried out on the human $P2X_1$ receptor whereby the conserved negatively charged residues have been mutated in an attempt to see the effect on the action of ATP. Negatively charged aspartic and glutamic acid residues were neutralised by mutation to alanine but the result had little effect on ATP potency. In conclusion these individual amino acids are not required for ATP recognition at the human $P2X_1$ receptor (Ennion *et al.*, 2001). However, this is not the case in the $P2X_2$ receptor where there is evidence to suggest some of these negatively charged amino acids are involved in ATP binding. Mutant D315A increased the sensitivity to ATP (~10-fold) (Jiang *et al.*, 2000). Mutant D315V reduced the sensitivity to ATP by ~60-fold suggesting that D315 may serve to maintain agonist sensitivity (Nakazawa *et al.*, 1998). This could underlie some of the subtype dependent properties and it may be the interaction with non-conserved residues that is important.

Contribution of Aromatic Residues

In many ATP binding proteins, aromatic amino acids (phenylalanine, tryptophan and tyrosine) coordinate the binding of the adenine group. Conserved aromatic residues, of which there are 20 conserved in the extracellular loop of P2X receptors, have been replaced by alanine in the human $P2X_1$ receptor to determine their effect on the properties of the channel. It was found that the majority of conserved aromatic residues are not essential for the normal functioning of the channel. F195A and W259A did not form functional channels and Western blot analysis revealed that F195A was not expressed in total cellular lysates either because the

receptor is degraded or inefficiently translated suggesting that this residue is essential for the normal folding and expression of the protein. Mutant W259A was detected in total cellular lysates at a reduced level so it was concluded the inefficient trafficking of the protein to the cell surface was responsible for the lack of function from this mutant. Mutants F185A (10-fold) and F291A (~160-fold) result in a decrease in ATP potency suggesting that these are most likely to be involved in ATP action at the receptor. Studies using partial agonists BzATP and Ap_5A to understand whether such decreases in ATP potency at mutant $P2X_1$ receptors was due to changes in agonist binding or gating of the channel revealed that residues K68, F185, F291, R292 and K309 are directly involved in ATP binding to the receptor (Roberts & Evans, 2004).

Contribution of Conserved Polar Glutamine, Asparagine and Threonine Residues

Polar residues can play an important role in the activation of ligand-gated ion channels including hydrogen bonding to the agonist and stabilizing conformational changes. Studies on the extracellular domain of the rat $P2X_2$ receptor suggest that some conserved polar glutamine, asparagine and threonine residues are important in regulation of ATP action at the receptor (Jiang *et al.*, 2000). The extracellular loop of the P2X receptor family has a number of uncharged conserved polar residues. These were mutated to alanine in the $hP2X_1$ receptor and the effect on the responsiveness to ATP observed. For the majority of mutants no effect was seen on ATP potency. Mutants Q95A, Q112A, Q114A and T158A showed changes in the efficacy for the partial agonists BzATP and Ap_5A which suggests that they may contribute to the gating of the channel. A larger decrease in ATP potency was seen for mutants T186A, N204A and N290A (6, 3 and 60-fold decreases respectively) and for T186A and N290A BzATP and Ap5A were no longer agonists but still bind to the receptor as the ability to modulate the response to co-applied ATP was not affected. A model of ATP binding at the $hP2X_1$ receptor where the residues $^{185}FT^{186}$ and $^{290}NFR^{292}$ play a role is likely (Roberts *et al.*, 2006).

A recent model, detailing the key residues identified by mutagenesis studies involved in ligand binding at the $P2X_1$ receptor are summarized in Figure 1.8.

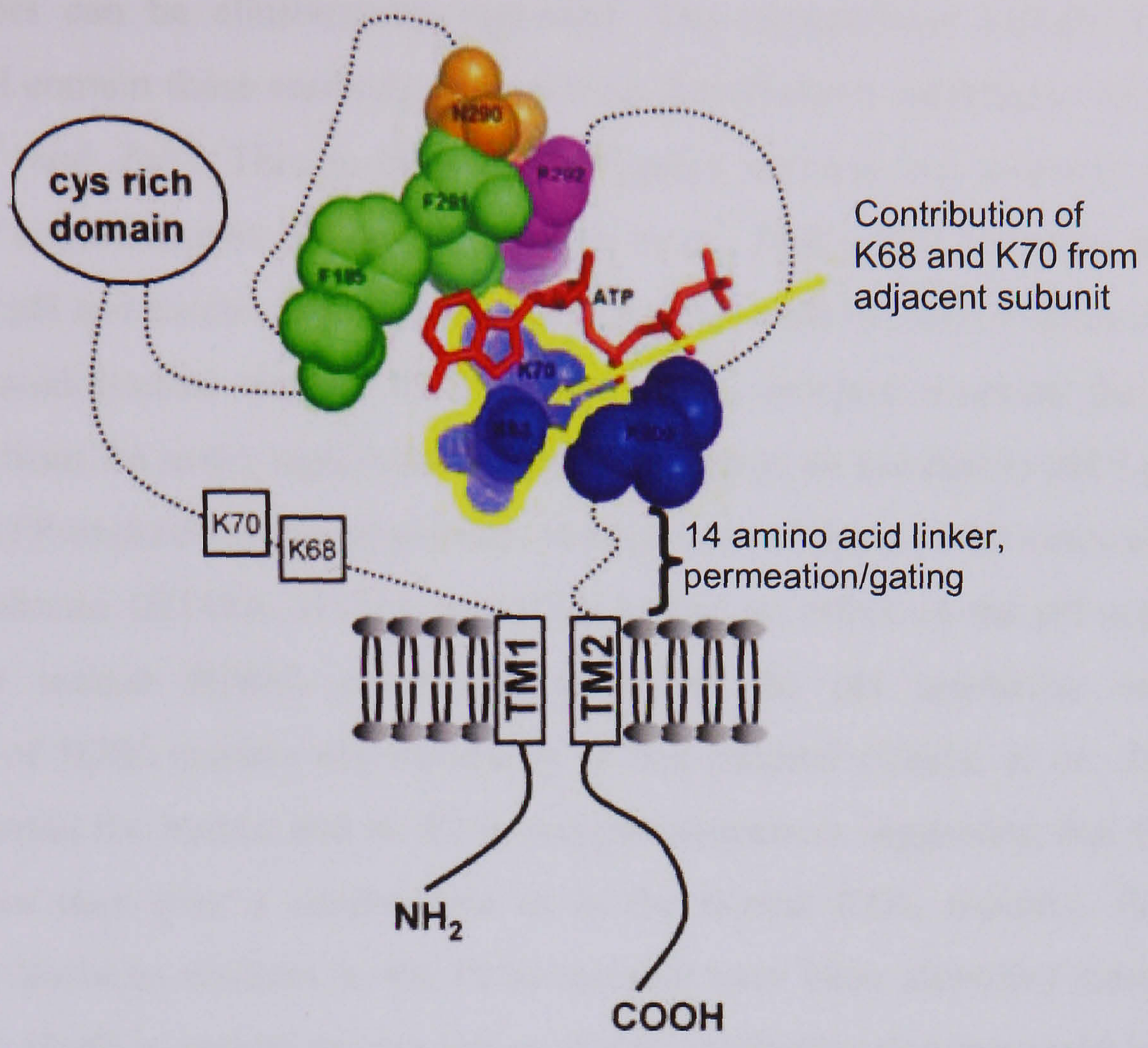


Figure 1.8 Model of ATP Action at $P2X_1$ Receptors. The common topology shared by a P2X receptor subunit is shown; intracellular amino and carboxy termini, two transmembrane domains and a large extracellular loop. Residues thought to be involved in ATP action at the $P2X_1$ receptor are written in black and ATP is shown in red. In addition the ATP-binding site of rat synapsin II [Protein Data Bank code 1i7l (Berman *et al.*, 2000)] has been used to predict the structure of the ATP-binding site in the $P2X_1$ receptor based on sequence analysis and a motif (NWK) similar to that found in all P2X receptor subunits (^{290}NFR^{292}) and predicted by mutagenesis studies to bind ATP (Ennion *et al.*, 2000; Roberts & Evans, 2004; Roberts & Evans, 2006). Positively charged lysine residues, 68, 70 and 309 are shown in blue but a recent study (Marquez-Klaka *et al.*, 2007) suggests that K68 and K70 are likely to come from a neighbouring subunit. The cysteine-rich domain is likely to contain three disulphide bonds; C117-C165, C126-C149 and C132-C159 (Ennion & Evans, 2002a). The ATP-binding region of the receptor is linked to opening of the channel by a 14 amino acid gating linker; mutagenesis of this region regulates the time course of the response with little or no effect on ATP potency. Figure taken from (Roberts *et al.*, 2006).

The Binding Site for Allosteric Modulators

P2X receptors can be allosterically regulated. The extracellular domain is likely to be involved and contain those residues that underlie the allosteric modulation of P2X receptors by H^+, Ca^{2+} and Zn^{2+}. This is because ATP-gated currents are sensitive to changes in extracellular concentrations of these ions (Egan *et al.*, 2004). P2X receptors are sensitive to extracellular pH and much attention has been concerned with the role of extracellular histidine residues in acidification. Studies of the human $P2X_4$ receptor revealed that changing the external pH from 7.4 to 6.5 significantly reduced, whereas an increase to pH 8 potentiated the maximum ATP-evoked current amplitude. Mutagenesis of three of the extracellular histidine residues to alanine (H140A, H172A and H241A) had no effect on the pH sensitivity of the channel but mutant H286A completely abolished the pH sensitivity suggesting that protonation of H286 confers pH sensitivity of this channel (Clarke *et al.*, 2000). H286 is conserved across the human and rat $P2X_4$ receptor sequences suggesting that H286 in the rat $P2X_4$ receptor may play a similar role as in the human $P2X_4$ receptor. Roles for other extracellular histidine residues in the $P2X_4$ receptor have been identified using site-directed mutagenesis. H140 is critical for the action of Cu^{2+} indicating that this residue forms part of the inhibitory metal-binding site of the $P2X_4$ receptor (Coddou *et al.*, 2003). H241 of $P2X_4$ receptors has been suggested to regulate the antagonist (suramin and PPADS) sensitivity of the receptor (Xiong *et al.*, 2004) (Xiong *et al.*, 2005).

2. Transmembrane Domains

Substituted cysteine accessibility mutagenesis (SCAM) has been used to identify residues that line the channel wall. This technique introduces individual point mutations to cysteine residues in a protein which can bind to highly soluble cysteine-reactive compounds. Methanethiosulfonate (MTS) derivatives form covalent (irreversible) disulphide bonds only with free cysteines and they exhibit extremely rapid reactivity which means the reaction rate can be quantified. By measuring how the reaction rate changes under various conditions for example as the size and charge of the reagent is varied, between closed and open states and depending on whether the reagent is applied to the internal or external membrane surface, questions like which residues line the protein surface and where is the channel gate can be answered.

Ion Conduction

Cysteine scanning mutagenesis in the $P2X_2$ receptor revealed that domains near either end of the first transmembrane domain influence ion conduction through the pore of the $P2X_2$

receptor (Haines *et al.*, 2001). The first transmembrane domain has also been implicated in the transduction pathway linking agonist binding to opening of the channel gate in the P2X receptor family (Haines *et al.*, 2001). This theory was developed on chimeric receptors consisting of those receptor subunits sensitive to αβ-methylene ATP ($P2X_{1 \text{ and } 3}$) and those that are not ($P2X_{2, 4, 5, 6 \text{ and } 7}$). Site directed mutagenesis indicated that the C-terminal end of TM1 was important in determining the agonist selectivity of channel gating for the chimeras (Haines *et al.*, 2001). Jiang *et al* (2000) has also carried out cysteine scanning mutagenesis this time on a region near the first transmembrane domain showing that residues in this domain also line the pore (Jiang *et al.*, 2000).

The involvement of TM2 in ion conduction in the $P2X_2$ receptor has been investigated by two groups in an attempt to identify the parts of the molecule that form the ionic pore (Rassendren *et al.*, 1997a; Egan *et al.*, 1998). SCAM analysis and the use of methanethiosulfonates were used in both cases and together they helped to identify the channel pore by mapping residues that lie on either side of the channel gate. Rassendren *et al* (1997) mutated residues in the predicted second hydrophobic domain of the $P2X_2$ receptor to cysteine and examined the effect of 3 MTS compounds (Rassendren *et al.*, 1997a). The results indicated that part of the pore of the P2X receptor is formed by the second hydrophobic domain and that residues L338 and D349 are on either side of the channel gate. In addition three mutants I328C, N333C and T336C inhibited currents evoked by ATP with the application of extracellular charged methanethiosulfonates suggesting that they lie in the outer vestibule of the pore (Rassendren *et al.*, 1997a).

Egan *et al* (1998) carried out similar experiments using SCAM to determine the involvement of the predicted second transmembrane of the $P2X_2$ subunit in ion conduction (Egan *et al.*, 1998). Residue G342 was identified by Egan *et al* 1998 as being part of the gate based on the principle that channels are water-filled pores lined in part by transmembrane-spanning domains and that Ag^+ and $MTSEA^+$ react more readily with ionized thiolates exposed to water (those residues lining the pore) than with nonionized thiols exposed to lipid (residues buried in the lipid membrane). In this way the accessibility of residues helped to predict their position; those that are accessible project into the channel and those inaccessible are those buried in the lipid membrane. G342C was accessible from both sides of the membrane and therefore must lie at or near the gate (Egan *et al.*, 1998). Mutation of the corresponding residue G342 in the $P2X_4$ receptor prevents the transition between gating modes that occurs in lowered extracellular Ca^{2+} (Khakh *et al.*, 1999). Mutation of G342 to alanine had multiple

effects which have been attributed to the fact that glycine has more conformational freedom than alanine due to its lack of side chain. This mutation would disrupt the ability of a possible TM2 gating domain to change conformation (Egan *et al.*, 2004; Li *et al.*, 2004).

The structure of TM2 or at least the outer part has been predicted to be α-helical in nature (Silberberg *et al.*, 2005). Alanine (Li *et al.*, 2004) and tryptophan (Silberberg *et al.*, 2005) scanning mutagenesis also revealed its participation in conformational changes/movement that occurs during receptor activation. Mutagenesis of each amino acid in TM2 to alanine revealed two clusters of alanine sensitive residues (those mutants that affected ATP potency). The first clustered around P329 and the second around G342 ($rP2X_2$ receptor numbering). It was suggested based on agonist potency, channel kinetics and efficacy of the partial agonist BzATP that the first cluster of residues of TM2 may form a linker region that connects the extracellular ATP binding site to the intrapore gate (Egan *et al.*, 2004; Li *et al.*, 2004). Tryptophan mutagenesis of TM2 resulted in 7 mutations located at the outer end of TM2 causing complete loss of channel function despite cell surface biotinylation experiments revealing that all seven non-functional mutants were successfully trafficked to the cell membrane at levels comparable to WT (Silberberg *et al.*, 2005). It was concluded that these mutants were likely to have disrupted ion conduction or gating and a critical functional role was suggested for the outer end of TM2.

Cysteine replacement studies in the rat $P2X_2$ receptor and the use of charged methanethiosulfonates which assess the accessibility of the substituted cysteines revealed that V48 was likely to be located extracellularly and play a role in the gating movements of channel opening (Jiang *et al.*, 2001). HEK293 cells expressing the double mutant V48C/I328C receptor resulted in reduced current amplitude but exhibited large steady inward currents at -60mV. Application of reducing agents dithiothreitol or bismercaptoethanol increased the ATP-induced current and at the same time decreased the persistent inward current. These observations led to the suggestion that V48 and I328 were sufficiently close for a disulphide bond to form. Although it is not possible to conclude whether the bond that forms is inter or intra subunit the fact a bond can form places a restraint on future P2X receptor models (Figure 1.9).

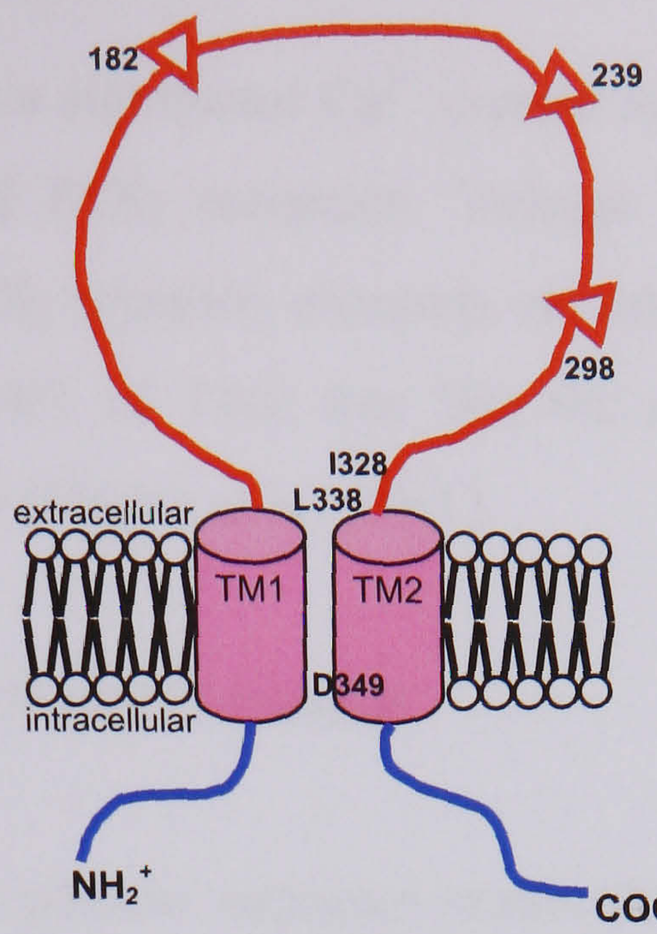


Figure 1.9 Topology of a P2X receptor subunit. The common topology shared by a P2X receptor subunit has intracellular amino and carboxy termini (blue), two transmembrane domains (pink) and a large extracellular domain (red). N-linked glycosylation sites are marked with red triangles (numbering refers to $P2X_2$ receptor). Residues L338 and D349 are thought to be on either side of the channel gate (Rassendren *et al.*, 1997a).

Agonist Binding, Gating and Potency

The involvement of both the transmembrane domains in determining agonist potency and binding/gating was investigated by alanine mutagenesis of each residue in TM1 and TM2 (Egan *et al.*, 2004). Change in channel function was observed as change in ATP potency, BzATP efficacy and deactivation kinetics. TM1 showed a regular pattern of potency-sensitive substitutions across the entire membrane and this is suggestive of a helical structure. However, no regular pattern of hits was seen for TM2. Mutants Y43A and F44A produced larger than normal holding currents in the absence of ATP and this coupled with the significant changes to the efficacy of BzATP of many mutants, suggest that TM1 participates in the opening of the channel.

Subunit Assembly

TM2 has been implicated in P2X subunit assembly. Deletion mutagenesis demonstrated that neither the amino or carboxyl terminal domains appear to be important for assembly so a strategy using chimeras of $P2X_1$ and $P2X_3$ receptors was adopted. Results supported the idea that the extracellular domain alone was not sufficient for assembly of full-sized subunits to form and that a transmembrane domain was involved in inter-subunit interaction. Taken together the deletion mutagenesis and chimera results led to a possible explanation that TM2 is critical for subunit co-assembly (Torres *et al.*, 1999).

Ion Permeability

P2X receptor channels conduct a significant Ca^{2+} current and the molecular basis of this was determined by mutagenesis of $P2X_2$ receptors. Voltage clamp studies of HEK293 cells expressing WT or mutant $P2X_2$ receptor channels revealed that the side chains of polar residues (T336, T338 and S340) of TM2 that line the channel pore influence Ca^{2+} and monovalent cation permeability (Migita *et al.*, 2001).

3. Intracellular Domain: N and C-termini

Structure

Analysis of the P2X receptor protein sequence revealed the absence of a signal peptide sequence after the starting methionine suggesting that P2X receptors have intracellular amino and carboxy termini (Brake *et al.*, 1994; Soto *et al.*, 1997).

N-terminus

The length of the N terminal region of P2X receptors is similar across all seven subtypes (~30 amino acids). This region contains a conserved protein kinase C (PKC) phosphorylation site, ([ST]-X-[RK]) in every P2X receptor subtype (Figure 1.10).

C-terminus

The intracellular C-terminus of the P2X receptor family varies in length between subunits and displays the least sequence similarity between P2X receptor subunits (Buell *et al.*, 1996). The $P2X_7$ receptor has a distinct C-terminus as it is much longer than any other P2X receptor family member. It is unclear as to why this is but it could be due to an evolutionary insertion.

Function

Functional roles have been identified for the N and C-termini. The N-terminus and the C-terminus have been implicated in regulating the time-course of response. The C-terminus has also been implicated in receptor expression.

Time-course

Mutation of the PKC site in $P2X_2$ and $P2X_3$ receptors had a marked effect on the time course of response. $P2X_2$ receptors have relatively slow desensitization responses and mutation of the PKC site resulted in a decreased receptor phosphorylation and a speeding of the time-course of the response (Boue-Grabot *et al.*, 2000). A similar mutation of the PKC site in the $P2X_3$ receptor resulted in a non functional channel (Paukert *et al.*, 2001). Mutation of the PKC site

(T18A $P2X_1$ receptor numbering) has been shown to affect the peak current in the $P2X_1$ receptor and speed up the rate of receptor desensitization by ~10-fold (Ennion & Evans, 2002b). Further studies in the $P2X_1$ receptor revealed that the PKC site $^{18}TPR^{20}$ is phosphorylated following the binding of ligand to the receptor to prevent current rundown but that the involvement of the three PKC sites in a functional trimeric channel differ (Liu *et al.*, 2003).

Desensitization

The role of the C terminus in controlling receptor desensitization has been investigated by monitoring the changes in intracellular free Ca^{2+} concentration in cells stimulated with ATP, the common agonist for all P2X receptor subunits (Koshimizu *et al.*, 1999). An increase in the intracellular Ca^{2+} concentration was observed on activation of P2X receptors and was determined by the rate of desensitization. The differing C termini of the P2X receptor subunits could account for the differences in rates of desensitization.

Expression

Rat $P2X_4$ receptors contain a C-terminal motif ($Y^{378}xxGL^{382}$) that is totally conserved among mammalian $P2X_4$ receptor sequences but is absent from other P2X receptor subtypes. This tyrosine-based sorting signal is necessary for efficient endocytosis of the $P2X_4$ receptor and it may be important for the control of functional expression (Royle, Bobanovic et al. 2002). Mutation of the $Y^{378}xxGL$ motif decreases $P2X_4$ receptor internalization and increases surface expression (Royle, Bobanovic et al. 2002; Royle, Qureshi et al. 2005). Mutation of Y378 to alanine in the $rP2X_4$ receptor significantly increases the amplitude of current compared to the $P2X_4$ WT receptor (Toulme, Soto et al. 2006). In addition, two residues K373 and Y374 in the C-terminus of the $P2X_4$ receptor (located before the endocytosis motif) have been identified by mutagenesis as being key determinants of $P2X_4$ receptor desensitization indicating a role for this region of the receptor in determining the duration of the physiological action of ATP at the $P2X_4$ receptor (Fountain & North, 2006).

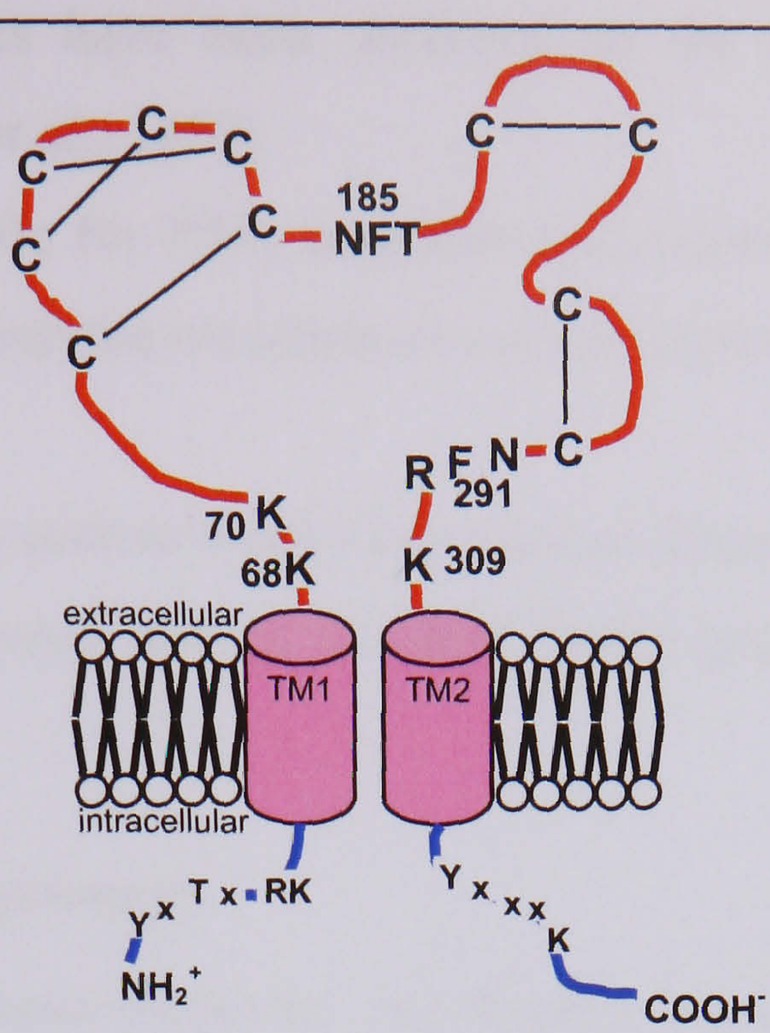


Figure 1.10 Topology of a P2X receptor subunit. The common topology shared by a P2X receptor subunit has intracellular amino and carboxy termini (blue), two transmembrane domains (pink) and a large extracellular domain (red). Conserved cysteine residues are shown with **C** and the predicted pattern of disulfide bonds are shown with black lines (Ennion & Evans, 2002a). Residues thought to be involved in ATP action at the $hP2X_1$ receptor are shown in black (numbering refers to $hP2X_1$ receptor). The highly conserved PKC phosphorylation site is located on the amino terminus and the conserved tyrosine-based sorting motif is located on the carboxyl terminus. The structure of TM1 is likely to be α-helical and so is the outer part of TM2.

1.4 P2X Channel Formation

1.4.1 Multimeric Channels

Given that the membrane topology of P2X receptors is strikingly different to other ligand-gated ion channel families, the organisation of P2X is interesting. Every ligand gated ion channel studied so far forms functional channels by multimeric arrangements and it was hypothesised that P2X receptors would behave similarly. It also seemed unlikely that just two transmembrane domains could form the pore. Further lines of evidence exist supporting a multimeric P2X receptor channel:

1. Bean studied concentration responses to ATP in P2X channels expressed in bullfrog and rat neurons and first suggested that more than one ATP molecule must bind in order for a channel to open (Bean, 1990). This was later confirmed by dose-response curves on recombinant receptors that revealed a Hill co-efficient of greater than one again suggesting that more than one drug molecule is required to bind in order to activate the channel (Brake *et al.*, 1994) indicating a multimeric complex.

2. Novel phenotypes have been observed on the mixing of properties of single subunits (Lewis *et al.*, 1995).
3. Antibodies specific for $P2X_2$ receptors will coimmunoprecipitate $P2X_3$ receptors and vice versa when the two channels are co-expressed (Radford *et al.*, 1997).

This raised further questions such as would they form as homo- or hetero-multimeric channels and how many receptors would form a channel? These questions will be discussed in the following sections.

1.4.2 Homomers and Heteromers

Pharmacological and functional properties are directly determined by subunit composition. Therefore different subunit combinations give rise to different phenotypes. Recombinant expression of single subunits resulted in functional receptors so it was concluded that P2X receptors are capable of forming as homomers.

Evidence for the formation of heteromeric P2X receptor channels was provided by copexpression studies of $P2X_2$ and $P2X_3$ receptor subunits (Lewis *et al.*, 1995). P2X receptors were first described in sensory neurons (Jahr & Jessell, 1983; Krishtal *et al.*, 1983; Abbracchio & Burnstock, 1994; Surprenant *et al.*, 1995) however Lewis *et al* reported the cloning of a complementary DNA ($P2X_3$) from rat dorsal root ganglion (DRG) that had properties dissimilar to those of sensory neurons (fast, αβ-meATP desensitizing currents). To address this, coexpression of $P2X_2$ and $P2X_3$ receptor subtypes were tested to see if a heteromer could produce an ATP-gated current which exhibited minimal desensitization but was fully activated by αβ-meATP. Coexpression of $P2X_2$ and $P2X_3$ produced ATP-activated currents that closely resembled those in sensory neurons. These properties could not be accounted for by addition of the two channels indicating that a new channel had formed by subunit heterpolymerization (Lewis *et al.*, 1995).

Functional studies of native receptors and patterns of subunit gene expression both suggest that heteromeric assembly of subunits in this family is possible. As a consequence of heteromeric assembly an important question is which subunits are capable of forming heteromers and in which combinations? Only by considering this are we able to identify the possible phenotypes that could arise in addition to homomeric assembly. Torres *et al* (1998) addressed this question using epitope-tagged $P2X_{1-7}$ receptor subunits and co-immunoprecipitation assays (Torres *et al.*, 1998b). To test whether different P2X subunits

could co-assemble into heteromers when co-expressed in HEK293 cells different pairs of P2X·FLAG/P2X·HA subunits were co-transfected and their ability to interact with each other examined using the co-immunoprecipitation technique such that cells co-expressed $P2X_1$·FLAG with $P2X_2$·HA or the converse pair $P2X_1$·HA and $P2X_2$·FLAG. The detergent solubilised membranes were immunoprecipitated with the anti-FLAG antibody and the resulting complexes were analyzed by Western blot using the anti-HA antibody. It was found that hetero-oligomeric assembly could occur and that it did so in a preferential fashion. The results of this study are summarised in Table 1.1.

	$P2X_1$	$P2X_2$	$P2X_3$	$P2X_4$	$P2X_5$	$P2X_6$	$P2X_7$
$P2X_1$	+	+	+	-	+	+	-
$P2X_2$		+	+	-	+	+	-
$P2X_3$			+	-	+	-	-
$P2X_4$				+	+	+	-
$P2X_5$					+	+	-
$P2X_6$						-	-
$P2X_7$							+

Table 1.1 Summary of observed P2X receptor subunit co-assembly (data taken from (Torres *et al.*, 1998b))

Note, no protein-protein interactions were found for any combinations containing the $P2X_7$ receptor subunit. The $P2X_6$ receptor subunit, although not able to co-assemble with itself, can co-assemble into heteromeric assemblies with many other subunits. This suggests that $P2X_6$ may act in a regulatory role within hetero-oligometric assemblies. It was also found that lack of a novel phenotype does not necessarily mean that two subunits are not interacting to form a heteromeric channel complex. The results for the heteromeric assembly of the $P2X_2$ and $P2X_3$ receptors are consistent with the findings of Radford *et al* (Radford *et al.*, 1997). The findings also agree with results for the $P2X_4$/$P2X_6$ receptors (Le *et al.*, 1998) and the $P2X_1$/$P2X_5$ receptors (Torres *et al.*, 1998c; Le *et al.*, 1999; Surprenant *et al.*, 2000).

Having established which P2X receptor subtypes can form heteromers and in which combinations, a further question arises as to the rules of subunit association. Torres *et al* (1999) published further findings on receptor subunit assembly where a motif(s) present in the second transmembrane domain of P2X receptor subunits was implicated in the assembly of P2X receptor channels (Torres *et al.*, 1999). This is in contrast to the four transmembrane spanning ligand-gated ion channel, the nicotinic acetyl-choline receptor, in which the amino terminus has been shown to be important for assembly. It is clear that a generic process for ligand-gated ion channel assembly does not exist.

1.4.3 Subunit Assembly

Given that homo- or heteromultimeric P2X channels exist, the question arose concerning how many subunits are required to form functional channels? This was examined by four different technical approaches:

1. Nicke *et al* (1998) used two techniques with $P2X_1$ and $P2X_3$ receptors to investigate the arrangement of P2X receptors (Nicke *et al.*, 1998). The first technique was blue native PAGE, used to estimate the molecular mass of the $P2X_1$ receptor under nondenaturing conditions. P2X receptors migrated as non-covalently linked homo-trimers whereas the positive control $\alpha_2\beta\gamma\delta$ nicotinic acetylcholine receptor (coexpression of α, β, γ and δ subunits) migrated as expected as a pentamer. If n-octylglucoside was used for P2X receptor solubilization, homo-hexamers were seen perhaps due to the covalent bonding of two trimers.

   The second was cross-linking experiments which allowed the purification of N-terminally His-tagged P2X receptors by Ni^{2+} agarose chromatography. Compounds dithio*bis*(succinimidylpropionate) (DSP), 3, 3'-Dithiobis(sulfosuccinimidylpropionate) (DTSSP) and 4,4'-diisothiocyanatostilbene-2,2'-disulfonic acid (DIDS) were used to cross-link His-$P2X_1$ subunits as they can join molecules together by a covalent bond. All three compounds were able to link $P2X_1$ subunits to dimers and trimers. PPADS, a P2X receptor antagonist, has an aldehyde group capable of forming a Schiff base (a compound formed by a condensation reaction between an aromatic amine and an aldehyde or ketone) with a primary amino group. Bifunctional PPADS analogues when added to His-$P2X_1$ cross-linked $P2X_1$ subunits to dimers and trimers. The same was true of $P2X_3$. The relative amounts of dimers and trimers produced depended on the distance between the two reactive groups of the cross-linkers. Increasing the length and flexibility of the spacer led to an increase in the amount of trimers produced. Both these techniques resulted in evidence to suggest P2X receptors can form as trimers or multiples of trimers. Similar results have been reported for the rat $P2X_7$ receptor (Kim *et al.*, 2001).

2. Stoop *et al* (1999) conducted studies on the mutant T336C rat $P2X_2$ receptor (Stoop *et al.*, 1999). This residue (T336) had previously been mutated to cysteine and found to be sensitive to block of current through the expressed channel by cysteine-reactive methanethiosulfonates (MTS compounds) (Rassendren *et al.*, 1997a; Egan *et al.*, 1998). The reaction of an –SH group of a channel cysteine with MTSET ([2-(trimethylammonium)-ethyl]-methanethiosulfonate) replaces the hydrogen atom with

the $-S(O_2)-CH_2-CH_2-N^+(CH_3)_3$. The negatively charged MTSES (2-sulfonatoethyl)methanethiosulfonate) was also effective at blocking the current evoked by ATP and it was suggested that T336 was unlikely to lie within the membrane electrical field (Rassendren *et al.*, 1997a). Also, additional experiments indicated that perhaps T336 lined the ion permeation pathway rather than blocking at the ATP-binding site.

For this reason, mutant T336 was chosen to determine the role of the contribution of individual subunits to channel formation. Two methods were used; co-injection and concatenated cDNAs. Wild-type and mutant $rP2X_2$ receptors were co-injected at varying ratios. ATP-evoked current was no different to currents observed from oocytes expressing wild-type and mutant channels alone but the percentage inhibition of the ATP-evoked current by MTSET increased as the proportion of mutant T336C $P2X_2$ receptor injected increased. This suggested that the inhibition of current through individual channels occurs in proportion to the number of cysteine-containing residues that they express.

Each possible wild-type/mutant $P2X_2$ receptor construct for dimers, trimers, tetramers and hexamers were expressed and tested. Then for all constructs an average amplitude of the current evoked by a saturating concentration of ATP before application of MTSET was calculated. Amplitudes obtained from single subunits were highest, followed by those with trimeric constructs. Stoop concluded that trimeric complexes of identical subunits are vital to the structure of P2X receptor channels but the results from the hexameric constructs suggest that it is not the case that P2X channels exist as hexamers. Nicke *et al* (2003) published results that further agreed a trimeric arrangement of P2X receptors (Nicke *et al.*, 2003).

3. Ding *et al* (1999) examined $P2X_2$ receptors at the single channel level and concluded that there are at least three binding sites in the $P2X_2$ channel (Ding & Sachs, 1999). Since it is a homomer, this would imply that three or more subunits are required to form the channel. This is consistent with a trimeric conformation.
4. Chemical cross-linking combined with direct imaging of $P2X_2$ and $P2X_6$ receptors by atomic force microscopy (AFM) has been used more recently to investigate the composition of the P2X receptor channel (Barrera *et al.*, 2005). AFM allows epitope tagged P2X receptos to be visualized and the measurement of the molecular volume of the receptors and the geometry of the receptors to be calculated. The results of this

study demonstrated the trimeric nature of the $P2X_2$ receptor channel but revealed that the $P2X_6$ receptor subunit is not capable of forming stable oligomers.

Having considered the membrane topology of the P2X receptor family and presented the evidence for the homo or heterotrimeric formation of the P2X receptor channel it is clear that these properties exclude P2X receptors from the other two ligand-gated ion channel familes, cys-loop and glutamate (Figure 1.11).

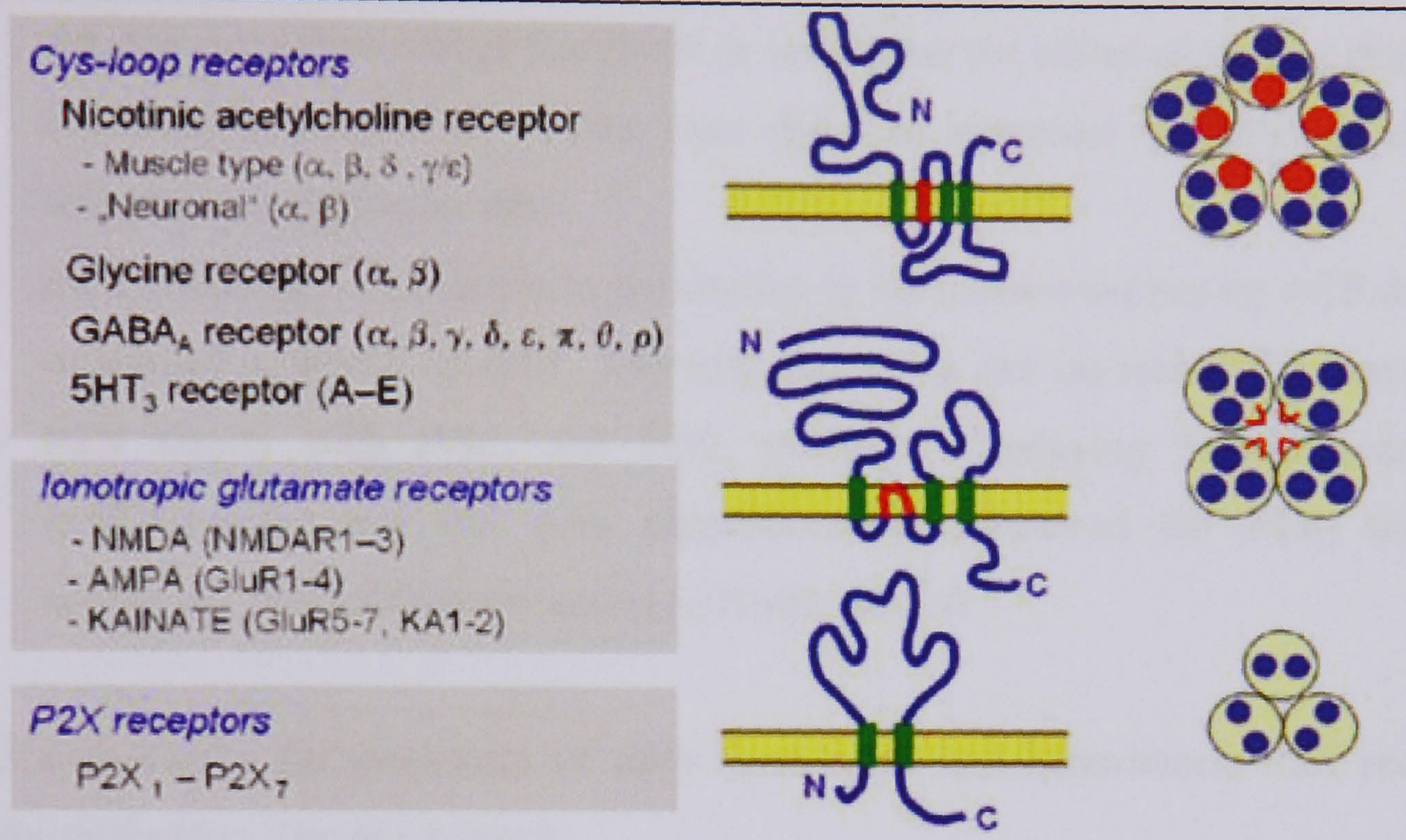


Figure 1.11 Families of ligand-gated ion channels. Due to a lack of evolutionary relationship and distinct topologies the three superfamilies of ligand-gated ion channel (Cys-loop, Glutamate, P2X) must be treated separately. Picture taken from *Assembly and Subtype Differentiation of Ligand-Gated Ion Channels; Nicke, A.*

The receptors of the cys-loop superfamily (nicotinic receptors, 5-HT_3 receptors, $GABA_A$ and $GABA_C$ receptors, glycine receptors, and some glutamate, histamine and serotonin activated anionic channels) consist of five homologous subunits, each with four transmembrane segments with extracellular N and C termini. In contrast, the glutamate activated cationic channels (NMDA receptors, AMPA receptors and Kainate receptors) are made of four homologous subunits, each with three transmembrane segments with the N-terminus lying outside the cell and the C-terminus inside. It is clear that as the P2X receptor family emerged their topology (trimeric channel formed by the homo or heteromeric association of subunits, each with two transmembrane segments and intracellular N and C termini) restricted them to neither the cys-loop or the glutamate family and a third superfamily of ligand-gated ion channels was born.

1.5 Pharmacological Properties of the P2X Receptor Family

P2X receptor channels form as homo or hetero- trimeric channels and this subunit assembly distinguishes them from the other ligand-gated ion channel families. This section examines how the pharmacological properties of the individual members of the P2X receptor family distinguish them from each other with a focus on the homomeric $P2X_1$ receptor. The characteristics of the individual subunits have been described in terms of

i. *Agonists* - which drugs bind to the receptor and activate it

ii. *Antagonists* - those drugs that block or counteract the effect of another drug

iii. *Permeation properties* - those ions that can permeate and the selectivity of receptors for particular ions

iv. *Desensitization* - this refers to the decline in the current elicited by ATP during the continued presence of ATP. The time course of this decline differs amongst the P2X family with $P2X_1$ and $P2X_3$ receptors displaying fast densensitization (milliseconds) and the slow densensitization observed for $P2X_2$ and $P2X_4$ receptors (100-1000 times slower) ((North, 2002)).

Table 1.2 summarises the properties of each homomeric and heteromeric P2X receptor in terms of the properties described above.

1.5.1 Homomeric $P2X_1$ receptor

Distribution

In 1994 Valera *et al* isolated a cDNA encoding the $P2X_1$ receptor by direct expression in *Xenopus* oocytes, starting with a cDNA library made from rat vas deferens. The protein, 399 amino acids long, was thought to be mostly extracellular and contain two transmembrane domains plus a pore forming motif which resembled that of potassium channels (Valera *et al.*, 1994). Methods of detection including, Northern blot analysis and in situ hybridization revealed that the $P2X_1$ receptor was localized to the vas deferens, urinary bladder and the smooth muscle layers of the small arteries and arterioles (Valera *et al.*, 1994). Collo *et al* mapped the distribution of P2X receptor RNAs in the brain and other tissues (Collo *et al.*, 1996). The $P2X_1$ receptor was observed clearly within cells of the cervical spinal cord, dorsal root ganglia, trigeminal ganglia and celiac ganglia (sympathetic ganglia). It was also the only RNA visualized in the smooth muscle of the bladder (Collo *et al.*, 1996). The distribution of the $P2X_1$ receptor was further described by Vulchanova *et al* 1996 using immunocytochemistry with receptor-specific antibodies. The receptor was shown to have a

predominant localization in smooth and cardiac muscle. The only neurons showing $P2X_1$ receptors were primary afferents and immunoreactivity was observed within nerve fibres and terminals in the superficial dorsal horn of the spinal cord (Vulchanova *et al.*, 1996).

Agonists

ATP and αβ-methylene ATP (αβmeATP) are equally potent agonists at the $P2X_1$ receptor each having an EC_{50} value close to 1μM. 2',3'-*O*-(benzoyl-4-benzoyl)-ATP (BzATP) is also an effective agonist. αβmeATP mimics the agonist action of ATP which distinguishes $P2X_1$ and $P2X_3$ receptors from the other homomeric forms. Suggestions to the location of the agonist binding site have been made based mainly on mutagenesis studies but these will be discussed in a later section.

Antagonists

Little is known of the location of the antagonist binding site although $P2X_1$ receptors are blocked by suramin and PPADS. Newer, more selective antagonists exist such as MRS2220 (cyclic pyridoxine-α4,5-mono-phosphate-6-azo-phenyl-2',5'-disulfonate) which blocks at ~10μM but has no effect on currents evoked at $P2X_2$ or $P2X_4$ receptors. TNP-ATP (2',3'-*O*-(2,4,6-trinitrophenyl)-ATP) is also an antagonist at the $P2X_1$ receptor but only the $P2X_3$ homomers and the $P2X_2/P2X_3$ heteromers are similarly sensitive.

Permeation Properties

The homomeric $P2X_1$ receptor is a cation-selective receptor that shows little selectivity for sodium over potassium. It has low permeability to larger organic cations such as Tris or N-methyl-D-glucamine when tested with brief agonist applications at least. It has a relatively high permeability to calcium and extracellular calcium has little or no inhibitory effect on $P2X_1$ receptor currents in contrast to the $P2X_2$ receptor.

Receptor Name	Agonists	Antagonists	Permeation Properties	Desensitization Characteristics	Tissue	Other Notes
Homomeric Receptors						
$P2X_1$	ATP, αβ-meATP, BzATP effective partial agonist	Suramin, PPADS, TNP-ATP (1000fold more effective blocking ATP at $P2X_1$ than at $P2X_{2,4,7}$)	Cation selective, rel high permeability to Ca^{2+}	Fast	Expressed on blood cells, neurons and variety of smooth muscle cells	
$P2X_2$	ATP	Suramin, PPADS, reactive blue 2	Increase concentration Ca^{2+} decreases amplitude	Slow	Widely expressed in the nervous system, and found in the intestine, bladder and vas deferens	Unique due to potentiation by low pH
$P2X_3$	ATP, αβ-meATP, Ap5A	Suramin, PPADS, TNP-ATP	Rat $P2X_3$ receptors are cation-selective channels.	Fast & concentration dependant	DRG cells	Contribute to pain sensation, detection of bladder distension and gut peristalsis
$P2X_4$	ATP	Effect of ATP blocked by high concentrations of Suramin and PPADS	When application of ATP is of short duration, operate as cation-selective channels;Ca2+ permeability is quite high. Application continued for several seconds, channel permeable to larger cations, NMDG	Slow, intermediate between $P2X_1$ and $P2X_2$	Distributed throughout the central and peripheral nervous system, strongly expressed in several glandular tissues	Phenotypically resemble homomeric $P2X_2$ receptors.
$P2X_5$	ATP, small amplitude	Suramin, PPADS	Relatively high permeability to chloride ions	Currents show little desensitization during agonist applications of several seconds	Detected in brain, spinal cord, heart and adrenal gland.	Functional properties resemble $P2X_2$
$P2X_6$	ATP, αβ-meATP	Suramin, PPADS	not determined	Slow	Widely expressed in CNS	
$P2X_7$	More sensitive to BzATP (10-30 times more) than ATP	Suramin and TNP-ATP are weak antagonists	In low concentration, both magnesium and zinc ions increase amplitude	In HEK293 cells, inward current evoked by ATP or BzATP show no desensitisation	Expressed in the brain, spinal cord, lung, spleen	
Heteromeric Receptors						
$P2X_{1/2}$	ATP, αβ-meATP	not determined	not determined	Fast and slow	SCG (Calvert et al)	Currents potentiated in both acidic and alkaline conditions
$P2X_{1/5}$	ATP, αβ-meATP	Suramin, PPADS, TNP-ATP(sensitivity intermediate to homomeric $P2X_1$ and homomeric $P2X_5$	Less permeable to Ca^{2+} than homomeric $P2X_1$, NMDG permeability similar to that of $P2X_2$ or $P2X_4$ receptors	Fast and slow	Possible spinal cord	Currents inhibited by an increase or decrease of the extracellular pH
$P2X_{2/3}$	ATP, αβ-meATP	TNP-ATP, PPADS, suramin	Calcium permeability similar close to $P2X_3$ subunit	Relatively slow	Sensory neurons	Potentiated by low pH
$P2X_{2/6}$	ATP	Inhibition by suramin is biphasic	Zinc ions of concentration 1-10μM increase amplitude	Fast and slow		
$P2X_{4/6}$	αβ-meATP, ATP	Suramin, PPADS, reactive blue 2	Zinc ions of concentration 10μM increase amplitude	Slow		

Table 1.2 Pharmacological Properties of homo- and hetero-meric P2X receptors

Desensitization

$P2X_1$ receptors undergo fast desensitization (milliseconds). Recovery from desensitization is very slow and second and subsequent applications of ATP do not elicit as large currents as the first application and they must be made at long intervals (>15mins) for reproducible responses to be observed. A study with $P2X_1$ and $P2X_2$ receptors and chimeric constructs aimed to determine the domains involved in desensitization (Werner *et al.*, 1996). Two domains were identified as involved in desensitization at the $P2X_1$ receptor. Each domain is 34 amino acids long consisting of the transmembrane segment and the contiguous residues (~14) on its intracellular side. These results suggest that closure of the channel during the continued presence of agonist required a conformational change involving both transmembrane segments.

For a more detailed discussion of all the receptor subtypes please refer to (North, 2002).

1.6 Possible Therapeutic Benefits of the P2X Receptor Family

The motivation to study this family of receptors lies in the wide potential of drug targets they represent and therefore the great therapeutic potential of purinergic compounds in the treatment of a wide range of diseases. This section will examine some of the different areas of possible therapeutic benefit of the P2X receptor family with a focus on the $P2X_1$ receptor.

Nervous System

P2X receptors are highly expressed in the nervous system. Agonists and antagonists are being explored as potential therapeutic agents for a number of neurological conditions. $P2X_2$, $P2X_4$ and $P2X_6$ receptors are expressed in the prepiriform cortex and microinjection of ATP analogues into the prepiriform cortex induces generalized motor seizures (Burnstock, 2002). It has been suggested that a P2X receptor antagonist may have potential as a neuroleptic agent. $P2X_3$ receptors are localized on sensory neurons. ATP may contribute to the pain associated with angina, migraine and cancer pain to name a few (Burnstock, 1996), and so a $P2X_3$ receptor antagonist may represent a novel target for an analgesic agent. In addition nerve damage that can lead to normally painless tactile stimuli becoming painful (allodynia) may be reversed by $P2X_4$ receptor anti-sense treatment (Tsuda *et al.*, 2003).

Gastroenterology

Puringeric signalling plays a role in the different activities of the gut. It has been suggested that during moderate distension low-threshold intrinsic enteric sensory fibres may be activated via $P2X_3$ receptors by ATP released from mucosal epithelial cells, leading to reflex propulsion of material down the gut (Burnstock, 2001). $P2X_3$ receptor expression is increased in human inflammatory bowel disease which may indicate a new therapeutic target in the treatment of dysmotility and pain (Yiangou *et al.*, 2001).

Urogenital System

P2X receptors are present in the kidney and ATP has been used to protect the kidneys from renal ischaemic-reperfusion injury and is being investigated for the treatment of chronic renal failure and transplantation-induced erythrocytosis (Burnstock, 2002). Purinergic signalling appears to play a role in afferent sensation from the bladder and ATP is released from urothelial cells when the bladder is distended. Indications from sensory-nerve recordings suggest that $P2X_3$ receptors are involved in mediating the nerve responses to bladder distension. This may be a possible treatment of detrusor instability (Burnstock, 2002).

1.6.1 $P2X_1$ receptor

Pharmacological profiles have been a very useful tool to characterize the phenotypes of specific subunits in native tissues. However, the development of $P2X_1$ knockout mice has allowed the functional roles of the receptor subtype to be assessed in great detail. This was a major advance in the study of $P2X_1$ receptors as the lack of potent, subtype-selective P2X receptor antagonists made studying the physiological role of $P2X_1$ receptors difficult. Since the generation of $P2X_1$ receptor deficient (-/-) mice $P2X_1$ receptors have been implicated in a wide range of functions.

Male Fertility

P2X receptors mediate the contractile response of the vas deferens to sympathetic nerve stimulation which propels sperm into the ejaculate (Burnstock, 1997). Mulryan *et al* (2000), using male $P2X_1$ KO mice, found that male fertility was reduced by ~90% (Mulryan *et al.*, 2000). However, reduced fertility was due to a reduction of sperm in the ejaculate rather than dysfunctional sperm. In addition contraction of the vas deferens to symapathetic nerve stimulation is reduced by up to 60% and response to P2X receptor agonists are abolished in the deficient mice. Therefore, $P2X_1$ receptors are essential for the normal functioning of the

male reproductive system. The development of selective $P2X_1$ receptor antagonists may provide an effective non-hormonal male conctraceptive pill and agents that potentiate the actions of ATP at $P2X_1$ receptors may be useful in the treatment of male infertility (Mulryan *et al.*, 2000).

Urinary System

$P2X_1$ receptor deficient mice have also been used to investigate the contribution of the $P2X_1$ receptor to the mouse bladder smooth muscle P2X receptor phenotype (Vial & Evans, 2000). Firstly the distributions of the P2X subtypes in the bladder were studied using subtype-selective antibodies. The $P2X_1$ receptor was found to be expressed in the smooth muscle layer of the bladder along with $P2X_2$ and $P2X_4$ receptors but only the $P2X_1$ receptor appeared to be localized to the extracellular membrane. To determine the role of the $P2X_1$ receptor in the smooth muscle of the bladder, responses to exogenously applied agonists (ATP, αβ-me ATP, L-βγ-meATP) and nerve stimulation were compared in normal and $P2X_1$ receptor deficient mice. The agonists failed to evoke contractions in the $P2X_1$ deficient -/- mice. However, this was not due to an effect on contractile function as contraction to carbachol were unaffected in $P2X_1$ deficient mice. Nerve stimulation evoked contractions in normal bladder smooth muscle with P2X and muscarinic acetylcholine (mACh) receptor mediated components but in $P2X_1$ receptor deficient mice, contractions were mediated by the mACh alone. Also, bladder size and function were no different in +/+ and -/- mice implying that the $P2X_1$ receptor deficiency had no effect on the development or appearance of the bladder. Homomeric $P2X_1$ receptors must be responsible for the bladder smooth muscle P2X receptor phenotype and normal and $P2X_1$ receptor deficient mice could be models of human bladder function.

Cardiovascular

ATP can produce vasoconstriction as it is released from endothelial, blood cells or because of local tissue damage through stimulation of P2X and P2Y receptors. $P2X_1$ receptors have been shown to underlie the artery smooth muscle P2X receptor phenotype and contribute ~50% to sympathetic nerogenic vasoconstriction (Vial & Evans, 2002). Vial concludes that $P2X_1$ receptors may provide a novel drug target for the treatment of cardiovascular disorders including stroke and heart disease.

Platelet Aggregation

$P2X_1$ receptors are also expressed by platelets and megakaryocytes (MKs). MKs are platelet progenitor cells, whose main function is to maintain the normal blood platelet count

(Kaushansky, 1999). Purine nucleotides are released and activate platelet P2 receptors ($P2X_1$, $P2Y_1$, $P2Y_{12}$) at the site or a vascular injury and this results in platelet aggregation and blood clotting. Vial *et al* (2002) showed that ATP evokes a biphasic inward current consisting of initial $P2X_1$ receptor and slower delayed P2Y receptor mediated components and that there is a functional interaction between the $P2X_1$ response and the P2Y response (Vial *et al.*, 2002). A role for $P2X_1$ receptors in thrombosis has been suggested. In addition regulation of $P2X_1$ receptors has a marked effect on the thromboembolism (Roberts *et al.*, 2006). $P2X_1$ receptor deficiency has been shown to provide protection (Hechler *et al.*, 2003) and overexpression of $P2X_1$ receptors promoting mortality (Oury *et al.*, 2003). $P2X_1$ receptors are likely to play a role in the responsiveness to thrombotic stimuli.

Nervous System

Calvert *et al* (2003) characterized the P2X expression within sympathetic postganglionic neurons from the superior cervical ganglia (Calvert & Evans, 2004). Using two techniques, whole-cell patch-clamp and calcium imaging, a heterogeneous population of P2X receptors in ~70% of neurons was revealed. The dominant phenotype was $P2X_2$-like but interestingly 10-15% of neurons were α,β-meATP sensitive, a property not exhibited by $P2X_2$ receptors and this response was reduced to 2% of all neurons in $P2X_1$ receptor-deficient mice thus providing evidence that $P2X_1$ receptors may contribute to the properties of heteromeric P2X receptors in neurons. Due to expression of $P2X_1$ receptors in a range of neurons Calvert also suggests that the regulation of the properties of P2X receptors by the $P2X_1$ subunit is more widespread.

Renal System

It has been established that afferent arterioles express P2X and P2Y receptors. Although full characterization of the P2 receptor composition has not been completed initial evidence suggests the presence of $P2X_1$ and $P2Y_2$ receptors on the preglomerular vascular smooth muscle (Inscho *et al.*, 2003). A critical role for $P2X_1$ receptors in mediating pressure-dependent autoregulatory adjustments in afferent arteriolar diameter has been suggested using selective $P2X_1$ receptor blockade and by showing that $P2X_1$ KO mice exhibit impaired autoregulatory capability (Inscho *et al.*, 2003, 2004).

1.7 Aims: Structural insight to the $P2X_1$ receptor using a combination of bioinformatics and mutagenesis techniques

The lack of homology with other ligand-gated ion channels (North, 1996) and membrane topology distinguished the P2X receptor family as a third class of ligand-gated ion channels. However, the lack of crystal structure for the receptor family has meant that structural models have relied upon experimental techniques to identify residues and regions of importance. This is not the case in the other two superfamilies of ligand-gated ion channels; Cys-loop and glutamate. Examples of receptors from both families have been crystallized and may provide clues to the structure/composition of the P2X receptor channel. These will be discussed in the following sections.

1.7.1 Ligand-binding Domain of Nicotinic Receptors

Nicotinic acetylcholine (nAChR) receptors belong to the pentameric ligand-gated ion channel family, the Cys-loop receptors. The ligand binding site is located at the interface between two subunits (Arias, 2000; Corringer *et al.*, 2000) and the principal part is always formed by the α-subunit residues contributing to the so-called loops, A, B and C whereas the neighbouring subunit residues contributing to loops D, E and F form the complimentory part of the binding pocket (Brejc *et al.*, 2001). The acetylcholine-binding protein (AChBP) aligns with the N-terminal domains of pentameric ligand-gated ion channels and lacks the transmembrane and intracellular domains present in the superfamily (Smit *et al.*, 2001). AChBP is most closely related to the α-subunits of the nAChRs and nearly all residues (including those involved in ligand binding) conserved within the nAChR family are present in AChBP. In addition AChBP bind known nAChR agonists and competitive agonists such as acetylcholine (Smit *et al.*, 2001) and makes AChBP an ideal candidate for modeling the N-terminal domain of an α-subunit of nAChRs.

The crystal structure of the AChBP (Figure 1.12a) proves it is a homologue of the pentameric ligand-gated ion channel superfamily and confirms the predicted location of the binding site at interface of two subunits. This study set a precedent for the idea that residues from two subunits could form a binding site and perhaps this may be true of the P2X receptor channel too. The relative positioning of the principal and complementary part of the ligand-binding site was also established. However, the kinetics of the channel remains unclear due to differences between the AChBP and nAChR. This structure will be extremely relevant for drug design studies that target the Cys-loop family of receptors.

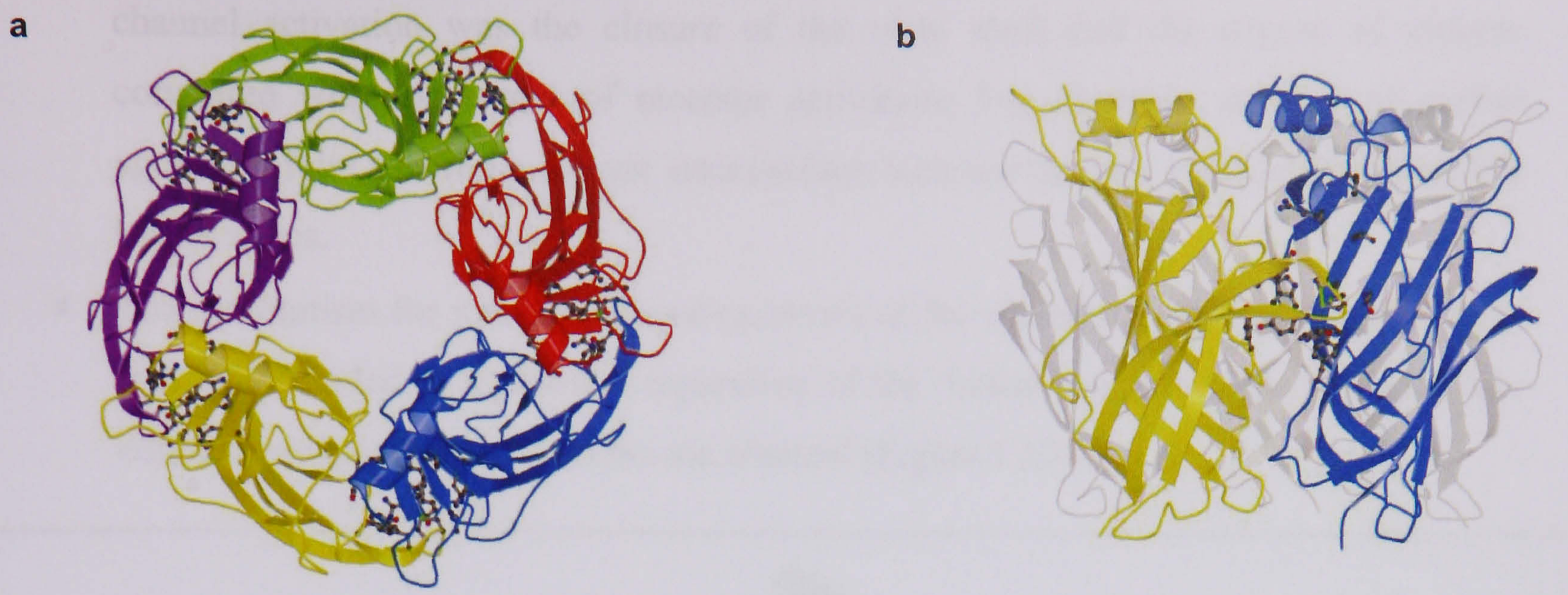


Figure 1.12 Models of the AChBP. a, The Pentameric Structure of AChBP. Each subunit is represented by a different colour. The principal and complementary ligand binding sites are shown by ball and stick representation. **b, Ligand-binding site AChBP.** The ligand binding site is represented by ball and stick. Figure taken from (Brejc *et al.*, 2001)

1.7.2 Crystallization of Glutamate Receptors

Crystal structures can also reveal important insights not only to the location of ligand-binding site and subunit assembly but also to the mechanisms of activation defined by particular structural elements. Glutamate receptors are composed of four subunits and the boundaries of the agonist-binding site have been defined to two regions of the receptor, namely S1 and S2. An S1S2 construct (mini-receptor that only includes the agonist binding core) was produced in large enough quantities in a functionally active state and well-ordered crystals in the presense of partial agonist kainite (Chen *et al.*, 1998) were produced which enabled structural analysis by X-ray diffraction (Gouaux, 2004).

The crystal structure of the S1S2-kainate complex showed that the receptor fragment had a 'clam-shell' like shape and that agonist was bound in the cleft between each shell (Armstrong *et al.*, 1998). It clearly showed the agonist binding pocket and defined key residues mediating agonist-specific interactions but in order to develop structures in key states (for example agonist bound) it was necessary to develop a new construct which involved shortening the amino terminus of S1 and carboxy terminus of S2 as well as shortening the linker between S1 and S2 (Armstrong & Gouaux, 2000). From this new construct, basic conclusions were made:

- The GluR2 S1S1 'clamshell' was most open in the apo state
- By comparing the 'clamshell' conformation in the apo and glutamate/AMPA-bound states it was concluded that the fundamental conformational change involved in

channel activation was the closure of the clam shell and the degree of closure correlated with the extent of receptor activation. For example, the use of partial agonist kainite showed a closure intermediate between the apo and glutamate/AMPA bound states.

- The mechanism for ion channel gating involved the closure of each subunit (agonist-induced) coupled to a physical separation of the 'linker' regions of the protein. This conformational change opens the ion channel (Figure 1.13).

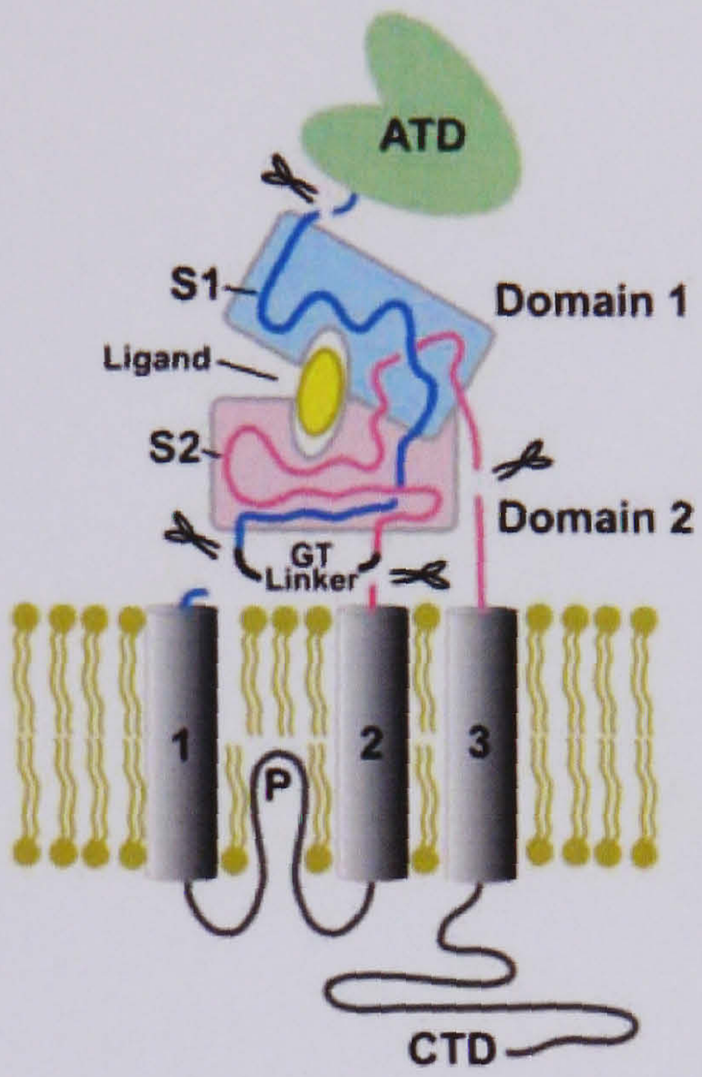


Figure 1.13 Schematic diagram of a single GluR2 receptor polypeptide. The amino terminus which starts at the amino-terminal domain (ATD) is located extracellularly. So is the ligand-binding core, composed of discontinuous polypeptide segments S1 and S2. The ion channel is formed by the membrane-embedded domains 1, P, 2 and 3 while the carboxy terminal domain (CTD) is located inside the cell. The GluR2 S1S2 constructs were generated by deleting the ATD, coupling the end of S1 to the beginning of S2 via a Gly-Thr linker and deleting the final transmembrane segment by ending the polypeptide near the end of S2 (Gouaux, 2004). The clam-shell structure is clearly identified by the pink and blue shaded regions. Figure taken from (Gouaux, 2004).

It is clear from the models available for other members of ligand-gated ion channel superfamilies that structural information can give a clear insight into receptor function. This is crucial for new drug development in the treatment of disease. Can a similar approach be used with the P2X receptor family, the third class of ligand-gated ion channels?

1.7.3 Thesis Aims

The aim of this thesis is to use information from previous mutagenesis studies and bioinformatics tools to build and refine existing structural models of the $P2X_1$ receptor and produce new homology models based on sequence analysis. In addition, mutagenesis will be used as a tool to validate published and new structural models of the $P2X_1$ receptor. A better understanding of the structure of the receptor will provide better insight into the functional roles of the P2X receptor family and aid rational drug design and development.

Chapter Two

2.1 Wet Methods

2.1.1 Site-directed mutagenesis

Cloning of the Human $P2X_1$ receptor

The pcDNA3.1 (Invitrogen) plasmid was used. It had been modified to contain a poly (A) tail adjacent to the 3' untranslated region of the cloned $P2X_1$ cDNA isloted from human bladder. The poly (A) tail increases the stability of the resulting mRNA by protecting it from exonucleases (Ennion *et al.*, 2000).

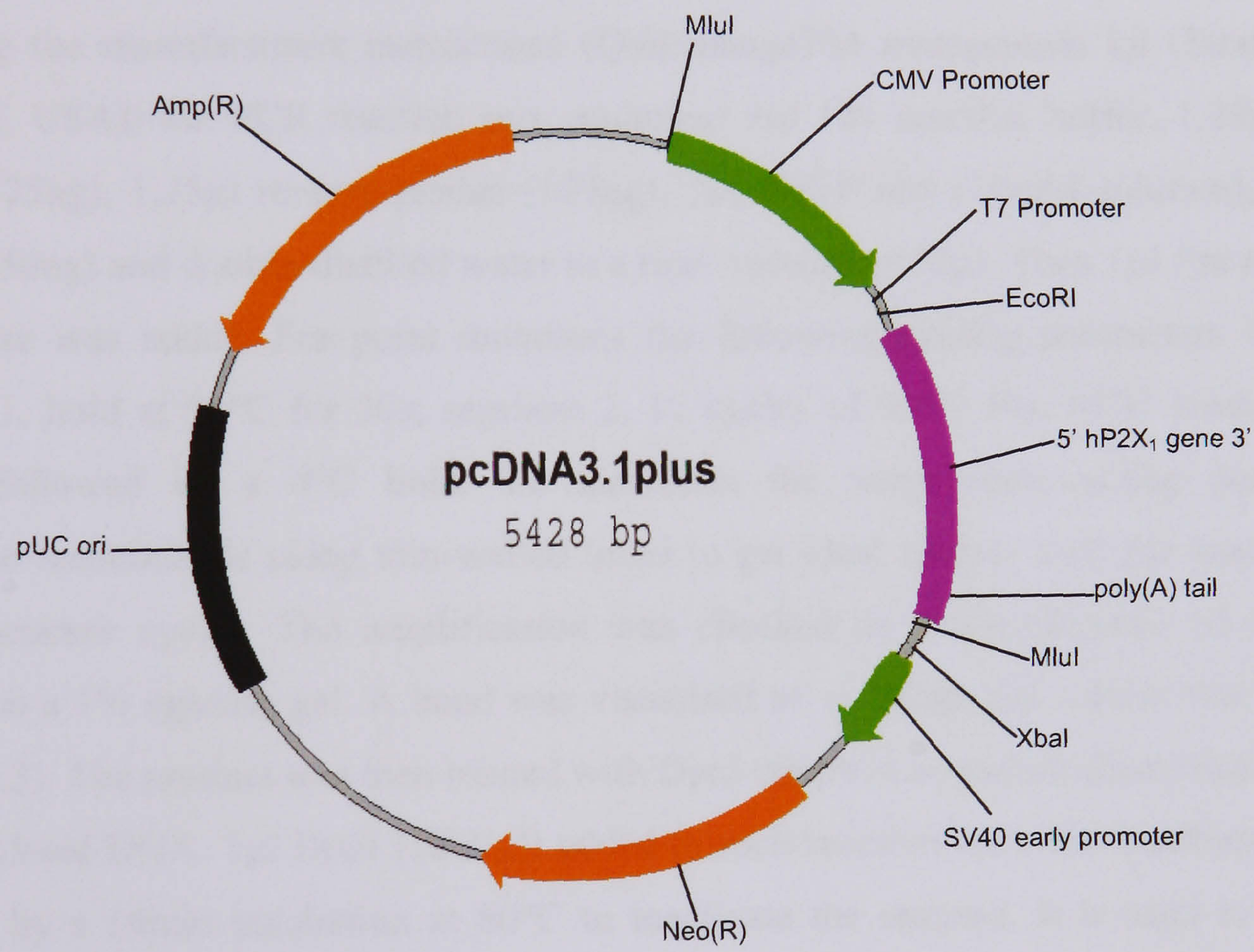


Figure 2.1 pcDNA3.1 plasmid. The $hP2X_1$ receptor gene (pink) was cloned into the pcDNA3.1 plasmid from human bladder. The plasmid was modified to contain a poly(A) tail adjacent to the 3' untranslated region of the cloned $P2X_1$ cDNA. CMV promoter and early promoter SV40 are shown in green. Ampicillin (Amp(R)) and neomycin resistance (Neo(R)) genes are shown in orange and pUC ori in black (for selection and maintenance in *E.* coli). Mlu I, EcoRI and Xbal restriction enzyme sites are labeled and so is the T7 promoter.

Oligonucleotide primers (forward and reverse) were designed to introduce specific point mutations in the $hP2X_1$ receptor (Figure 2.2).

Primers	$P2X_1$ receptor mutant K190C
$P2X_1$ WT	GGCCGAGAACTTCACTCTTTTCATCAAGAACAGCATCAGC
Forward	GGCCGAGAACTTCACTCTTTTCATCTGCAACAGCATCAGC
Reverse	CCGGCTCTTGAAGTGAGAAAAGTAGACGTTGTCGTAGTCG

Figure 2.2 Primer Design. The WT $P2X_1$ receptor DNA sequence is shown for the region containing the residue lysine-190. Sequences for the forward and reverse primers designed for $P2X_1$ receptor mutant K190C are shown where residues in red represent the codon that was altered to mutate amino acid lysine (AAG in WT sequence) to cysteine (TGC in forward primer).

Following the manufacturers instructions (QuikchangeTM mutagenesis kit (Stratagene, La Jolla, CA, USA)) the PCR reaction mix contained 5µl 10x reaction buffer, 1.25µl forward primer (125ng), 1.25µl reverse primer (125ng), 1µl dNTP mix (10mM solution), 1µl $P2X_1$ plasmid (50ng) and double-distilled water to a final volume of 50µl. Then 1µl *Pfu turbo* DNA polymerase was added. For point mutations the following cycling parameters were used; segment 1, hold at 95°C for 30s; segment 2, 12 cycles of 95°C 30s, 55°C 1min and 68°C 15mins followed by a 4°C hold. To maximize the temperature-cycling performance, Stratagene recommends using thin-walled tubes to get ideal contact with the heat blocks of the temperature cycler. The amplification was checked by electrophoresis of 5µl of the product on a 1% agarose gel. A band was visualised at ~1200bp, the correct size for $hP2X_1$ (Figure 2.3). The product was then treated with DpnI which is an endonuclease that is specific for methylated DNA. 1µl DpnI (10U/µl) added to each reaction for a 1hr incubation at 37°C followed by a 10min incubation at 80°C to inactivate the enzyme. It is used to digest the parental DNA template and to select for mutation-containing synthesized DNA. The parent DNA is from a bacterial source and as a result it is methylated so it is recognized by the DpnI whereas the mutated DNA is not methylated as it is synthesized *in vitro*. The vector DNA containing the desired mutations was then transformed into XL1-Blue supercompetent cells. For each reaction, 50µl of the XL1-Blue supercompetent cells were added to pre-chilled Falcon tubes. Then 1µl of the DpnI-treated DNA was transferred into the Falcon tubes containing the XL1-Blue supercompetent and incubated on ice for 30mins. Following incubation, the reactions were heat shocked for 45s at 42°C and then placed on ice for 2mins. Then 0.5ml of S.O.C medium (Invitrogen) was added to the transformation reactions and incubated at 37°C for 1hr with shaking. 50µl of each transformation reaction is then plated

out on agar plates containing ampicillin and carbenocillin (50mg/ml) (reduces the likelihood of satellite colonies) and incubated at 37°C for >16hr.

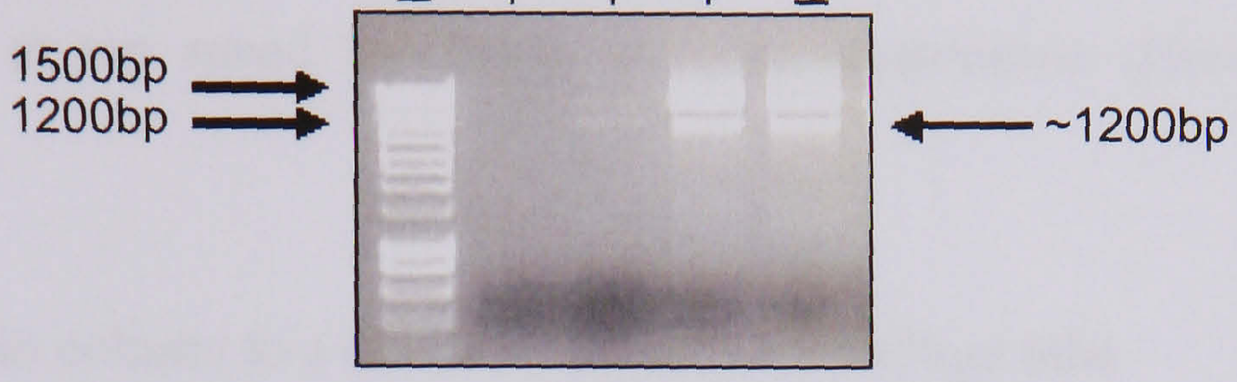


Figure 2.3 PCR Fragment K190C. The PCR product was checked by electrophoresis of 5µl of the product on a 1% agarose gel. Lane one contains the hyperladder for quantification. Lanes 2 and 3 are negative controls (reactions with primers missing in lane 2 and plasmid missing in lane 3). Lane 4 contains a positive control, mutant $P2X_1$ receptor, G147A. A band corresponding to the $P2X_1$ receptor was visualized at ~1200bp.

Individual colonies are picked and grown overnight in 10ml LB broth containing 20µl ampicillin at a temperature of 37 °C. The plasmid DNA was extracted using a Promega mini-prep kit using the centrifugation protocol:

Production of cleared lysate

- Pellet 10ml of overnight culture by centrifugation (Jouan CR412, 3000rpm, 5mins)
- Discard supernatant and completely resuspend in 250µl cell resuspension solution
- Add 250µl cell lysis solution and mix by inverting 4 times
- Add 10µl alkaline protease solution invert 4 times to mix
- Incubate at room temperature for 5mins
- Add 350µl neutralization solution, invert 4 times to mix
- Centrifuge at top speed (Heraeus centrifuge pico, 13000rpm) 10mins room temperature

Binding of Plasmid DNA

- Insert spin column into collection tube
- Decant cleared lysate into spin column
- Centrifuge at top speed for 1min (Heraeus centrifuge pico, 13000rpm)
- Discard through flow and reinsert spin column into collection tube

Washing

- Add 750µl wash solution (95% ethanol added according to manufacturer's instructions)
- Centrifuge at top speed for 1min at room temperature (Heraeus centrifuge pico, 13000rpm)
- Discard through flow and reinsert spin column into collection tube
- Repeat with 250µl wash solution
- Centrifuge at top speed for 2mins at room temperature (Heraeus centrifuge pico, 13000rpm)

Elution

- Transfer spin column to a sterile 1.5ml microcentrifuge tube
- Add 100µl nuclease free water to the spin column
- Centrifuge at top speed 1min room temperature (Heraeus centrifuge pico, 13000rpm)
- Discard column and store DNA at -20°C

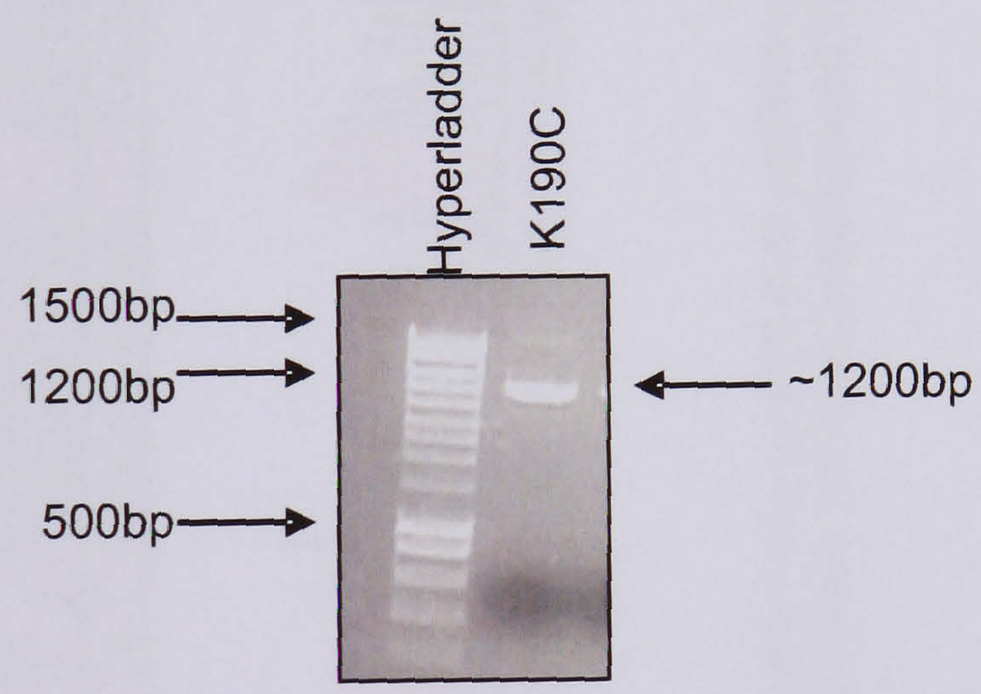


Figure 2.4 Sample of plasmid K190C extracted using Promega miniprep kit. Hyperladder shown in lane one. Mutant $P2X_1$ receptor protein (K190C) visualized at ~1200bp.

Mutations were verified by full DNA sequencing on both strands (Automated ABI sequencing service, Leicester University) (Figure 2.5) using $hP2X_1$ sequencing primers seqA, seqB, seqC, seqE, seqF and seqG (Figure 2.6).

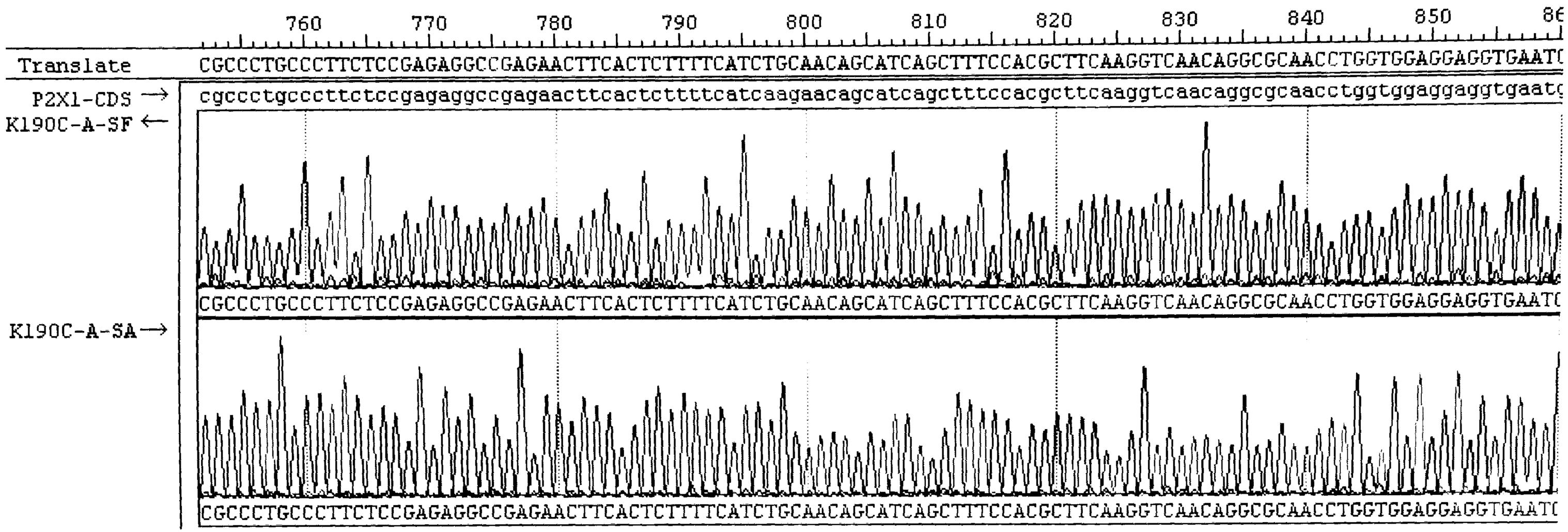


Figure 2.5 Sequencing data K190C. Two sequencing reads for mutant K190C using forward (SA) and backward (SF) primers. The point at which the mutation is sequenced is highlighted on the $P2X_1$ receptor coding sequence (P2X1-CDS) in red. A lysine residue (codon AAG) has been successfully mutated to a cysteine residue (codon TGC).

```
         1 gcctccagct gacctctggc tcctgtcctc tggctccacc tgcaccgccc tgctcttcct
        61 aaggggccag gaagccccca gaagctctac catcgacgtg ggtggtggca cccggctcac
       121 cctgagagca gagggcgtgc aggggctca gttctgagcc cagccggccc acatggcac
       181 ggcggttcca ggaggagctg gccgccttcc tcttcgagta tgacaccccc cgcatggtgc
       241 tggtgcgtaa taagaaggtg ggcgttatct tccgactgat ccagctggtg gtcctggtct
       301 acgtcatcgg gtgggtgttt ctctatgaga agggctacca gacctcgagc ggcctcatca
       361 gcagtgtctc tgtgaaactc aagggcctgg ccgtgaccca gctccctggc ctcggccccc
seqA   421 aggtctggga tgtggctgac tacgtcttcc cagcccaggg ggacaactcc ttcgtggtca
       481 tgaccaattt catcgtgacc ccgaagcaga ctcaaggcta ctgcgcagag cacccagaag
       541 ggggcatatg caaggaagac agtggctgta cccctgggaa ggccaagagg aaggcccaag
       601 gcatccgcac gggcaagtgt gtggccttca acgacactgt gaagacgtgt gagatctttg  seqG
       661 gctggtgccc cgtggaggtg gatgacgaca tcccgcgccc tgcccttctc cgagaggccg
seqB   721 agaacttcac tcttttcatc aagaacag   tcagctttcc acgcttcaag gtcaacaggc
       781 gcaacctggt ggaggaggtg aatgctgccc acatgaagac ctgcctcttt cacaagaccc
       841 tgcacccccct gtgcccagtc ttccagcttg gctacgtggt gcaagagtca ggccagaact
       901 tcagcaccct ggctgagaag ggtggagtgg ttggcatcac catcgactgg cactgtgacc
       961 tggactggca cgtacggcac tgcagaccca tctatgagtt ccatgggctg tacgaagaga  seqF
seqC  1021 aaaatctctc cccaggcttc aacttcaggt ttgccaggca ctttgtggag aacgggacca
      1081 actaccgtca cctcttcaag gtgtttggga ttcgctttga catcctggtg gacggcaagg
      1141 ccgggaagtt tgacatcatc cctacaatga ccaccatcgg ctctggaatt ggcatctttg
      1201 gggtggccac agttctctgt gacctgctgc tgcttcacat cctgcctaag aggcactact
      1261 acaagcagaa gaagttcaaa tacgctgagg acatggggcc aggggcggct gagcgtgacc
      1321 tcgcagctac cagctccacc ctgggcctgc aggagaacat gaggacatcc tgatgctcgg  seqE
```

Figure 2.6 Sequencing primers for hP2X$_1$ receptor sequence. Forward primers, seqA, seqB and seqC are shown in pink. Reverse primers, seqE, seqF and seqG, are shown in blue. Start codon (atg) and stop codon (tga) are shown in red boxes.

2.1.2 Expression in *Xenopus laevis* Oocytes

A *Xenopus laevis* expression system was used to study mutant receptors. Plasmid DNA must be linearized with a restriction enzyme to prevent circular transcripts being produced. Mutant and wild-type plasmids were digested with M*lu*I (10U/µl) (5µl 10x reaction buffer, 4µl plasmid DNA, 4µl M*lu*I (10U/µl) made up to 50µl total volume with double-distilled water). This restriction enzyme cuts to release a 2531 base pair fragment containing the vector T7 RNA polymerase site, the $P2X_1$ cDNA, and the vector poly(A) tail. To confirm that cleavage is complete, the linearized template DNA was checked on a gel (Figure 2.7).

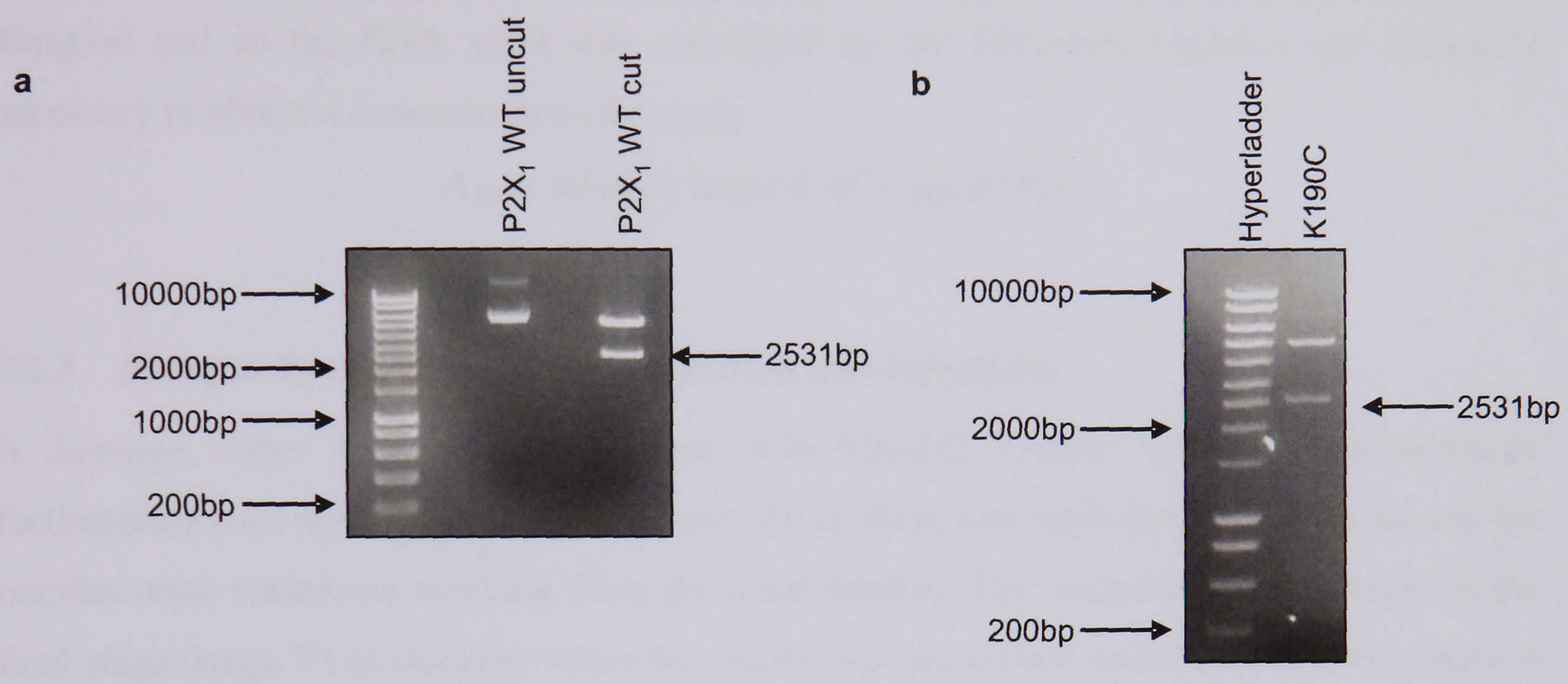


Figure 2.7 Mlu I digest. Bioline Hyperladder I used. **a,** WT $P2X_1$ receptor plasmid and WT $P2X_1$ receptor plasmid cut with restriction enzyme Mlu I. **b**, Mutant $P2X_1$ receptor plasmid K190C, cut with restriction enzyme Mlu I.

It is possible that DNA from some miniprep procedures may be contaminated with residual RNAses. By treating the template DNA with proteinase K (20mg/ml) and 10% SDS for 1hour at 50°C followed by a phenol chloroform extraction (using an equal volume) and ethanol precipitation, any RNase should be removed. Sense strand cRNA was made from this fragment using a T7 mMessage mMachine^{TM} kit (Ambion). All reagents were used as in the kit (Ambion) and kept on ice. A single 20µl reaction consisted of 2µl 10x Reaction Buffer (salts, buffer, dithiothreitol and other ingredients), 10µl 2x NTP/CAP (a neutralized buffered solution containing ATP 15mM, CTP 15mM, UTP 15mM, GTP 2mM and cap analog 12mM), 6µl plasmid and 2µl enzyme mix (buffered 50% glycerol containing RNA polymerase, SUPERase and other components). The reaction was incubated for 2hours at 37°C. DNase treatment was used to remove any template DNA (1µl DNase (2U/µl) I per reaction incubated at 37°C for 15mins). A Lithium Chloride precipitation technique was used to purify the RNA. It will effectively remove unincorporated nucleotides and most proteins.

The reaction was stopped and the RNA precipitated by adding 30µl nuclease free water and 30µl LiCl Precipitation solution (7.5M lithium chloride, 50mM EDTA). It was mixed thoroughly and chilled for 30mins at -20°C and then centrifuged (benchtop centrifuge) at 4°C for 15mins at 13000 rpm to pellet the RNA. The supernatant was carefully removed and the pellet washed once with ~1ml 70% ethanol and re-centrifuged (benchtop centrifuge, 4°C), 2mins) to maximise the removal of unincorporated nucleotides. The 70% ethanol was removed and the RNA re-suspended in nuclease free water and stored at -80°C. Quantification was by spectrophotometry. By reading the A_{260} of a diluted aliquot of the reaction the yield was determined. For single-stranded RNA, 1 A_{260} unit corresponds to 40µg/ml and so the RNA yield was calculated by the following equation and diluted if necessary to obtain a concentration of 1µg/µl.

$$A_{260} \text{ x dilution factor x } 40 = \mu g/ml \text{ RNA}$$

2.1.3 *Xenopus* Preparation, Oocyte Isolation and Injections

A *Xenopus laevis* frog was anesthetised with MS-222 Tricane (ethyl 3-aminobenzoate methanesulfonate salt, Sigma) and sacrificed. An incision was made just above the leg and the oocytes were teased out working from the outer lobules. The oocytes were harvested in the final stage (stage V) of maturity when the oocyte was pigmented such that one hemisphere is dark brown (animal hemisphere) and the other hemisphere (vegetal hemisphere) shows the yellow colour of the egg yolk (Figure 2.8). Clumps of approximately 10-20 oocytes were teased off and placed in a 50ml tube containing OR2 (NaCl 82.5mM, KCl 2.5mM, Na_2HPO_4 1mM, MgCl 1mM, HEPES 5mM pH to 7.5 with NaOH) until ~10-15mls of oocytes had been collected. The oocytes were washed with OR2 until the solution was clear, pouring off excess solution and any lone oocytes each time. Collagenease (1mg/ml) (from Clostridium histolyticum type 1A, Sigma) made up in OR2 was added to the oocytes with a 50ml syringe and 0.45µM acrodisc and placed on a rocker at 12RPM for 30mins. This was to digest the follicular cell layers and connective tissue (Figure 2.9). Following incubation in collagenase, the oocytes were washed with Barth's solution (NaCl 88mM, KCl 1mM, $CaCl_2$ 0.41mM, $Ca(NO_3)_2$ 0.33mM, $MgSO_4$ 1mM, $NaHCO_3$ 2.4mM, HEPES 10mM pH to 7.5 with NaOH) 4 times, placed in a petri dish and manually defolliculated in Barth's solution. Oocytes were stored in an incubator at ~17°C and the media changed daily.

Manually defolliculated stage V *Xenopus laevis* oocytes were injected with 50nl (50ng) of cRNA using an Inject+Matic microinjector and maintained in ND96 buffer (96mM NaCl,

2mM KCl, 1.8mM $CaCl_2$, 1mM $MgCl_2$, 5mM sodium pyruvate, 5mM HEPES, pH7.6) at 18°C. Recordings were made from oocytes 3-7 days post-injection.

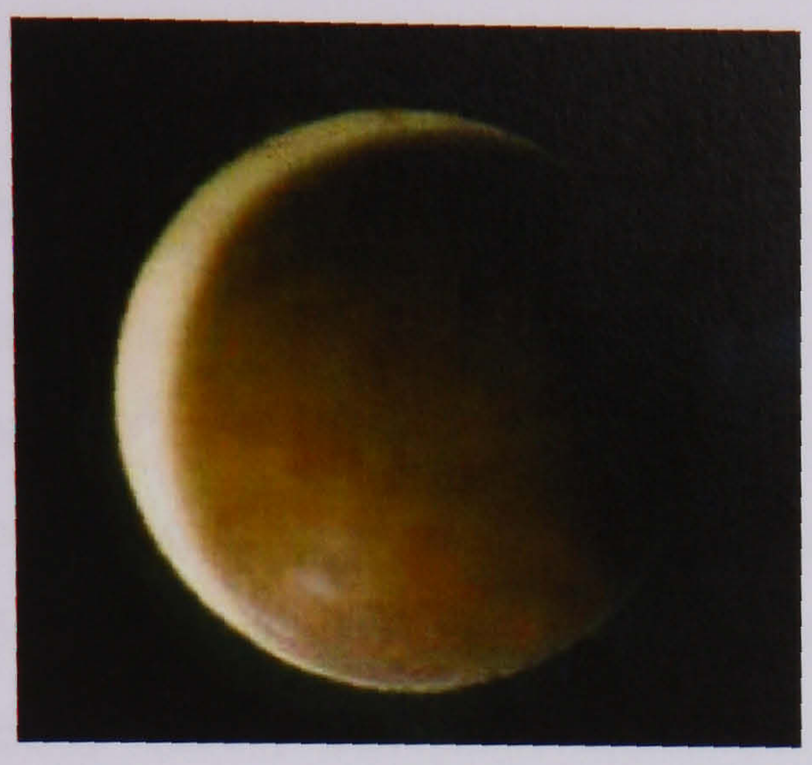


Figure 2.8 A single oocyte with diameter of 1.0-1.2mm. The dark brown, animal hemisphere and yellow coloured vegetal hemispheres are clear.

Figure 2.9 Oocytes wrapped in follicular cell layers that contain blood vessels. The oocytes were manually defolliculated to remove this layer.

2.1.4 Electrophysiological Recordings

Capillary glass (1.5mm width TW150F-4 World Precision Intruments Inc) was pulled on a programmable electrode puller and the electrodes were backfilled with 3M KCl. Two electrode voltage clamp recordings were made from oocytes using a GeneClamp 500B amplifier (Axon instruments) with a Digidata 1322 analog-to-digital converter and pClamp 8.2 acquisition software (Molecular Devices, Menlo Park, CA) as descrived previously (Ennion *et al.*, 2000) with sampling frequency 2kHz and filtering frequency 10Hz. One intracellular electrode measures the membrane potential (the voltage electrode) and the second (namely the current electrode) passes current to maintain the desired voltage clamp. The external solution in the bath consisted of ND96 with 1.8mM $BaCl_2$ replacing $CaCl_2$ in order to prevent activation of endogenous calcium activated chloride channels. Membrane currents were recorded at a holding potential of –60mV. Applications of agonist were applied via a U-tube perfusion system. Due to the large size of oocytes it is difficult to apply solutions rapidly. For desensitizing responses this results in the true time course of the response being underestimated as a result of the relatively slow solution exchange and activation of responses. Experiments on nondesensitizing $P2X_2$ receptors expressed in oocytes indicate that 10-90% solution exchange takes ~300-500ms and means the rise-times of $P2X_1$ receptor currents (~100ms for WT) only give an indication of the speed of activation. The slower decay time ~1s for WT $P2X_1$ receptor currents is less likely to be affected by solution

exchange and therefore can be used to discriminate mutants with a change in time course. Preliminary studies showed 5 min between applications was sufficient for reproducible responses to be recorded.

2.1.5 Data Acquistion and Analysis

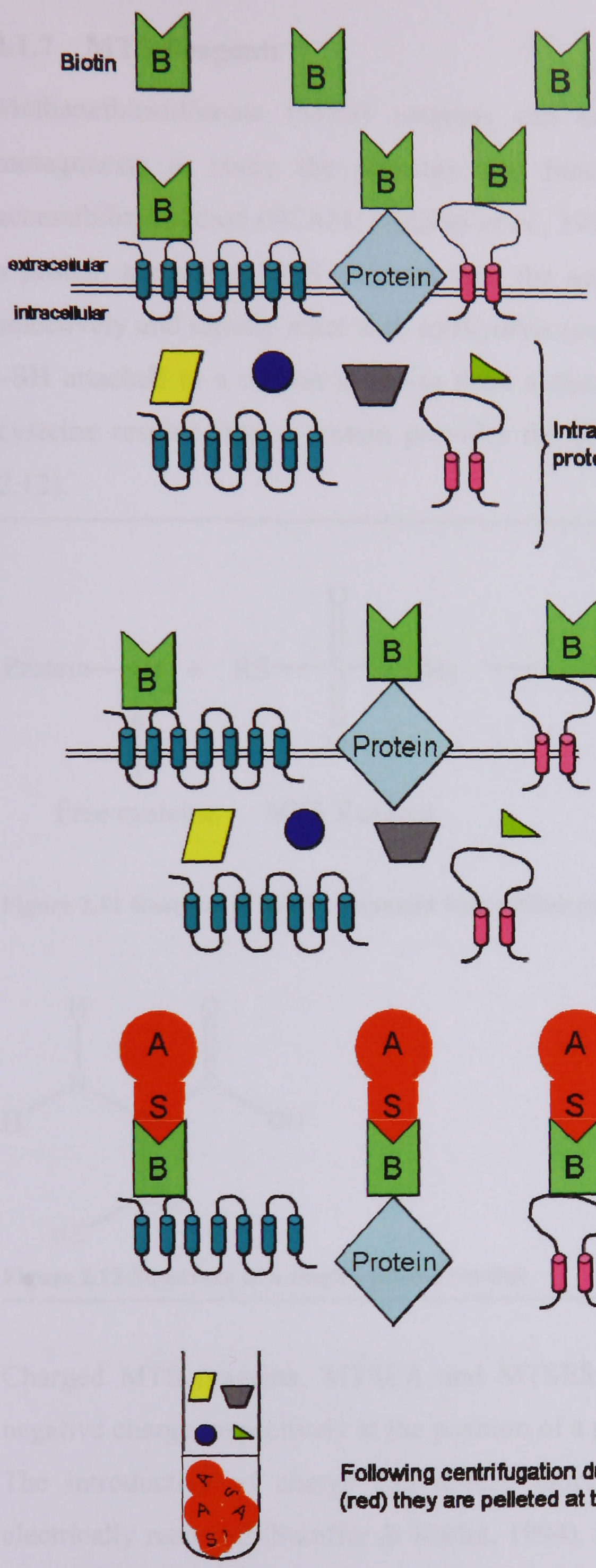
Individual concentration-response curves were fitted with the Hill equation: $Y = [(X)^H \times M]/(X^H + (EC_{50})^H]$, where Y is the response, X is the agonist concentration, H is the Hill coefficient, M is the maximum response and EC_{50} is the concentration of agonist evoking 50% of the maximum response. pEC_{50} is the –log of the EC_{50} value. For the calculation of the EC_{50} values, individual concentration-response curves were generated for each experiment and statistical analysis was performed on the pEC_{50} data generated. In the figures, concentration-response curves are fitted to the mean normalized data. Significant differences between WT and mutant current amplitudes, time courses, agonist potentcy and MTS reagent modification were calculated either by Student's *t* test or one-way ANOVA followed by Dunnett's test for comparisons of individual mutants against control using the SPSS 12.0 for Windows package (SPSS, Chicago, 1L).

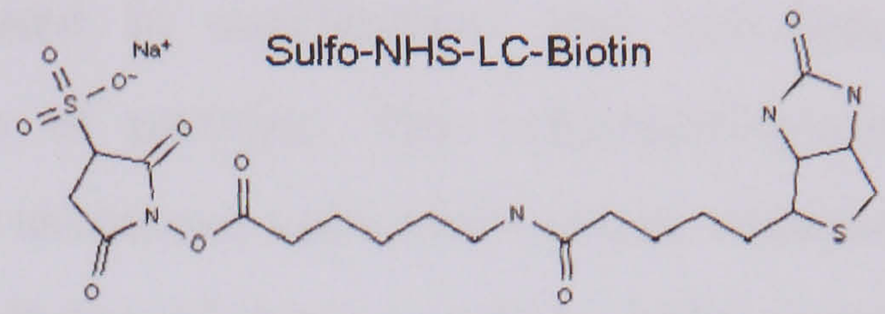
2.1.6 Western Blot Analysis

Western blot analysis was used to estimate levels of wild-type and mutant expression of receptor protein. Oocytes previously injected with wild-type or mutant receptor cRNA (50ng) were homogenized in buffer H (100mM NaCl, 20mM Tris-Cl, pH 7.4, 1% Triton X-100, and 10µl/ml protease inhibitor cocktail) at 20µl/oocyte. After centrifugation at 13000rpm (4°C) for 2 min (benchtop centrifuge), a 10µl aliquot of the supernatant was mixed with an equal volume of gel sample buffer (0.18M Tris, 5.7% SDS, 29% glycerol, 0.003% β-mercaptoethanol), heated for 5 min 95°C and separated on a 10% SDS-PAGE gel, 100V. The gel was transferred to nitrocellulose membrane (0.2µM Schleicher & Schuell) and blocked in TBST (20mM Tris-Cl, pH 7.6, 145mM NaCl, 0.05% Tween 20) plus 5% marvel overnight. Following this, the membrane was incubated with anti-$P2X_1$ antibody (1:500 dilution) (Alamone, Israel) in TBST + 5% marvel for 1.5h, washed 3 times for 5 min in TBST and incubated with anti-rabbit horseradish peroxidase secondary antibody (1:5000 dilution) (Sigma) for 1.5h. After three 5 min washes in TBST the protein bands were visualised using an ECL (Plus) kit (Amersham Pharmacia Biotech) according to manufacturer's instructions.

2.1.6.1 Cell Surface Biotinylation

Cell surface biotinylation was used to label cell surface proteins and determine the level of protein trafficked to the cell surface membrane. Oocytes were incubated in 0.5mg/ml Sulfo-NHS-LC-Biotin in ND96 buffer for 30 min at room temperature. Sulfo-NHS-LC-Biotin is impermeable to the cell membrane and can only biotinylate proteins available at the cell surface. Following incubation, oocytes were washed 5 times in ND96 and homogenised in buffer H (100mM NaCl, 20mM Tris-Cl, pH 7.4, 1% Triton X-100, and 10µl/ml protease inhibitor cocktail) at 20µl/oocyte. After centrifugation at 13000rpm (4°C) for 2 min (benchtop centrifuge), a 140µl aliquot of the supernatant was mixed with buffer H at a ratio of 1:1 and 30µl streptavidin agarose beads (Sigma) and incubated for 3 hours at 4°C on a roller. The samples, containing the agarose beads bound to the biotin labelled surface proteins were centrifuged for 2 mins at top speed (benchtop centrifuge, 13000rpm, 4°C) to pellet the beads. The resulting pellet of streptavidin agarose beads were washed 4 times with buffer H and mixed with 20µl of gel sample buffer (0.18M Tris, 5.7% SDS, 29% glycerol, 0.003% β-mercaptoethanol). They were heated for 5 min at 95°C and separated on a 10% SDS-PAGE gel. The gel was transferred to nitrocellulose and blocked in TBST (20mM Tris-Cl, pH 7.6, 145mM NaCl, 0.05% Tween 20) plus 5% marvel overnight. Following this, the membrane was incubated with anti-$P2X_1$ antibody (1:500 dilution) (Alamone, Israel) in TBST + 5% marvel for 1.5h, washed 3 times for 5 min in TBST and incubated with anti-rabbit horseradish peroxidase secondary antibody (1:5000 dilution) (Sigma) for 1.5h. After three 5-minute washes in TBST the protein bands were visualised using an ECL (Plus) kit (Amersham Pharmacia Biotech) according to manufacturer's instructions. For a schematic of the complete process see Figure 2.10.





Oocytes incubated in 0.5mg/ml Sulfo-NHS-LC-Biotin

Biotin binds to all membrane proteins only

Streptavidin (S) agarose (A) beads (red) bind to biotin

Following centrifugation due to the size of the agarose beads (red) they are pelleted at the bottom.

Figure 2.10 Cell Surface Biotinylation.

2.1.7 MTS Reagents

Methanethiosulfonate (MTS) reagents can be used in combination with site-directed mutagenesis to study the structure and function of proteins. The substituted-cysteine-accessibility method (SCAM; (Akabas *et al.*, 1994)) introduces individual cysteine residues in a protein and using MTS reagents tests the accessibility of those cysteines. MTS reagents selectively and rapidly react with sulfhydryls (compounds which contain the sulfhydryl group –SH attached to a carbon atom) to form a disulfide bond (Figure 2.11). The insertion of a cysteine residue into a protein provides the sulfhydryl group (highlighted in pink, Figure 2.12).

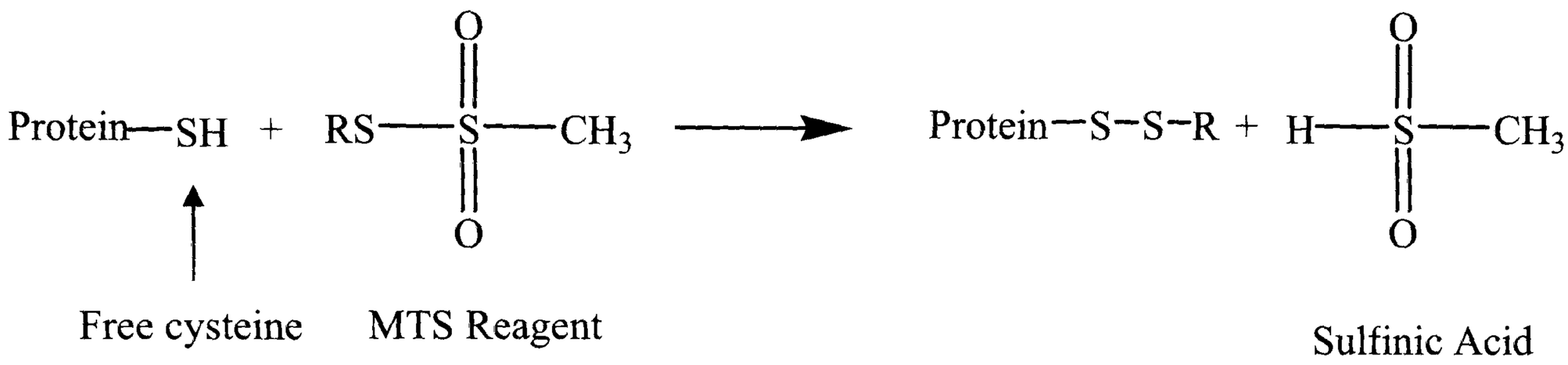


Figure 2.11 Reaction of MTS compound with sulfhdryls.

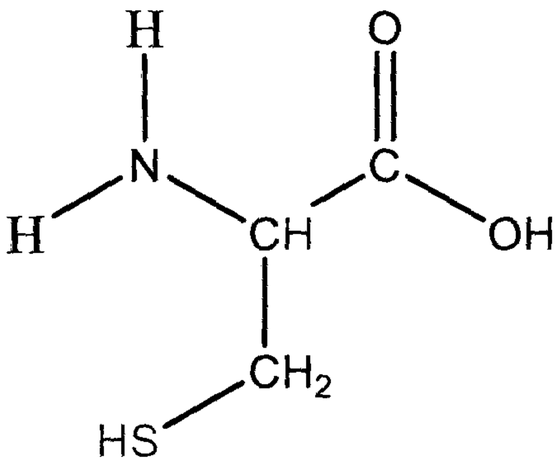
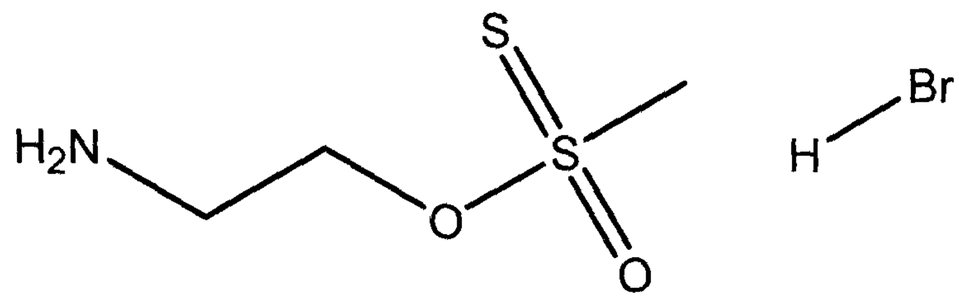


Figure 2.12 Structure of a single cysteine residue.

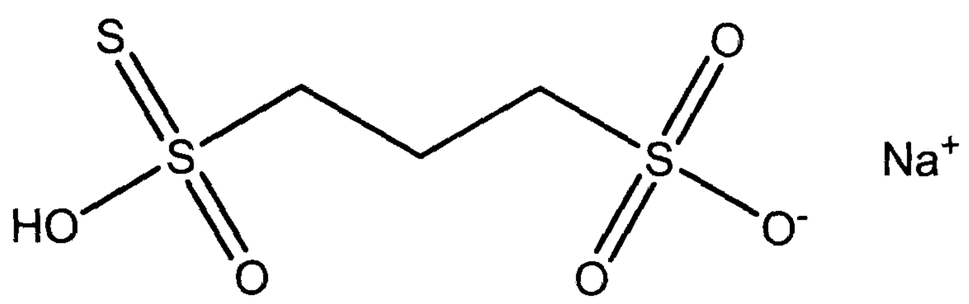
Charged MTS reagents, MTSEA and MTSES (Figure 2.13), can introduce a positive or negative charge respectively at the position of a previously neutrally charged cysteine residue. The introduction of charge can induce functional changes in a channel which can be electrically recorded (Stauffer & Karlin, 1994). By using various MTS reagents that differ in size or charge, it is possible to determine information regarding the physical size of an ion channel or the role of charge at a specific position.

a



2-Aminoethyl Methanethiosulfonate, Hydrobromide **(MTSEA)**

b



Sodium (2-Sulfonatoethyl)methanethiosulfonate **(MTSES)**

Figure 2.13 Structure of a, MTSEA and b, MTSES.

Stock solutions of MTSEA and MTSES (100mM) (Toronto Research Chemicals, Downsview, Ontario, Canada) were made fresh from solids daily and kept in the dark at 4°C. The compounds were dissolved in ND96 buffer (96mM NaCl, 2mM KCl, 1.8mM $CaCl_2$, 1mM $MgCl_2$, 5mM sodium pyruvate, 5mM HEPES, pH 7.6) for use with *Xenopus* oocytes. MTSEA or MTSES (1mM) were bath perfused (5mins) and co-applied with ATP for 3s at EC_{50} concentration, with 5mins between applications. The results are displayed as the mean of the current elicited by ATP+MTSEA/S as a percentage of the current elicited by ATP ± SEM. Data were analyzed by one-way ANOVA (Analysis Of Variance). The advantage of one-way ANOVA is that it compares everything in the group to WT rather than comparing just pairs.

2.1.7.1 MTSEA-Biotin Assay

N-biotinoylaminoethyl methanethiosulfonate (MTSEA-Biotin; Toronto Research Chemicals, Toronto, Canada) is impermeable to the cell membrane and can therefore only biotinylate surface localized proteins in intact cells. MTSEA-Biotin can be used to assess to availability of free cysteine residues (those cysteines that are not part of a disulfide bond and those in an accessible area of the protein tertiary structure) in the extracellular loop of P2X receptors because it will bind sulfhydryls (compounds which contain the sulfhydryl group –SH attached to a carbon atom). Oocytes previously injected with wild-type or mutant $P2X_1$ cRNA were incubated in biotinylation agent, MTSEA-Biotin (0.15 mg/ml) for 30 min in ND96 buffer. Following incubation, oocytes are washed 5 times in ND96 and treated as described for cell surface biotinylation.

2.2 The Model Building Procedure: Methods

The model building procedure is outlined in Figure 2.14. This section aims to describe the methods behind the procedure that starts with an amino acid sequence of a protein of unknown structure and builds a three dimensional model of the protein based on computational techniques. This chapter follows the natural progression through the model building procedure.

2.2.1 Transmembrane Topology Predictions

Transmembrane topology refers to which parts of the protein are inside and which are outside the membrane. Transmembrane topology can be predicted from amino acid sequence using a number of programs. TMpred is a predictive method specifically for transmembrane regions and their orientation. The TMpred program is based on an algorithm, TMbase (Hofmann & Stoffel, 1993). TMbase, derived from SWISS-PROT (Bairoch & Apweiler, 2000), is a database of naturally occurring transmembrane proteins. Predictions are made using a combination of several weight-matrices for scoring. PHDhtm (Rost, 1996) is another example of a program that predicts transmembrane helices. It is related to the PredictProtein secondary structure prediction method (Rost *et al.*, 1994). It is extremely accurate and in addition to predicting the putative transmembrane regions it indicates the orientation of the loop regions with respect to the membrane.

2.2.2 Scanning Amino Acids Sequences against Databases of Known 3D Structures

2.2.2.1 Protein-protein BLAST (blastp)

BLAST, the Basic Local Alignment Search Tool (Altschul & Lipman, 1990) takes a query sequence as input and searches it against a user-specified database for pairs of high scoring local alignments. There are many variations of BLAST allowing for different sequence comparisons to be made. For the purposes of this chapter, standard protein-protein BLAST (blastp) has been used and compares a protein query to a protein database (or subset of). In addition, all BLAST searches have been performed using the BLAST web pages (http://www.ncbi.nlm.nih.gov/BLAST/) rather than stand-alone BLAST.

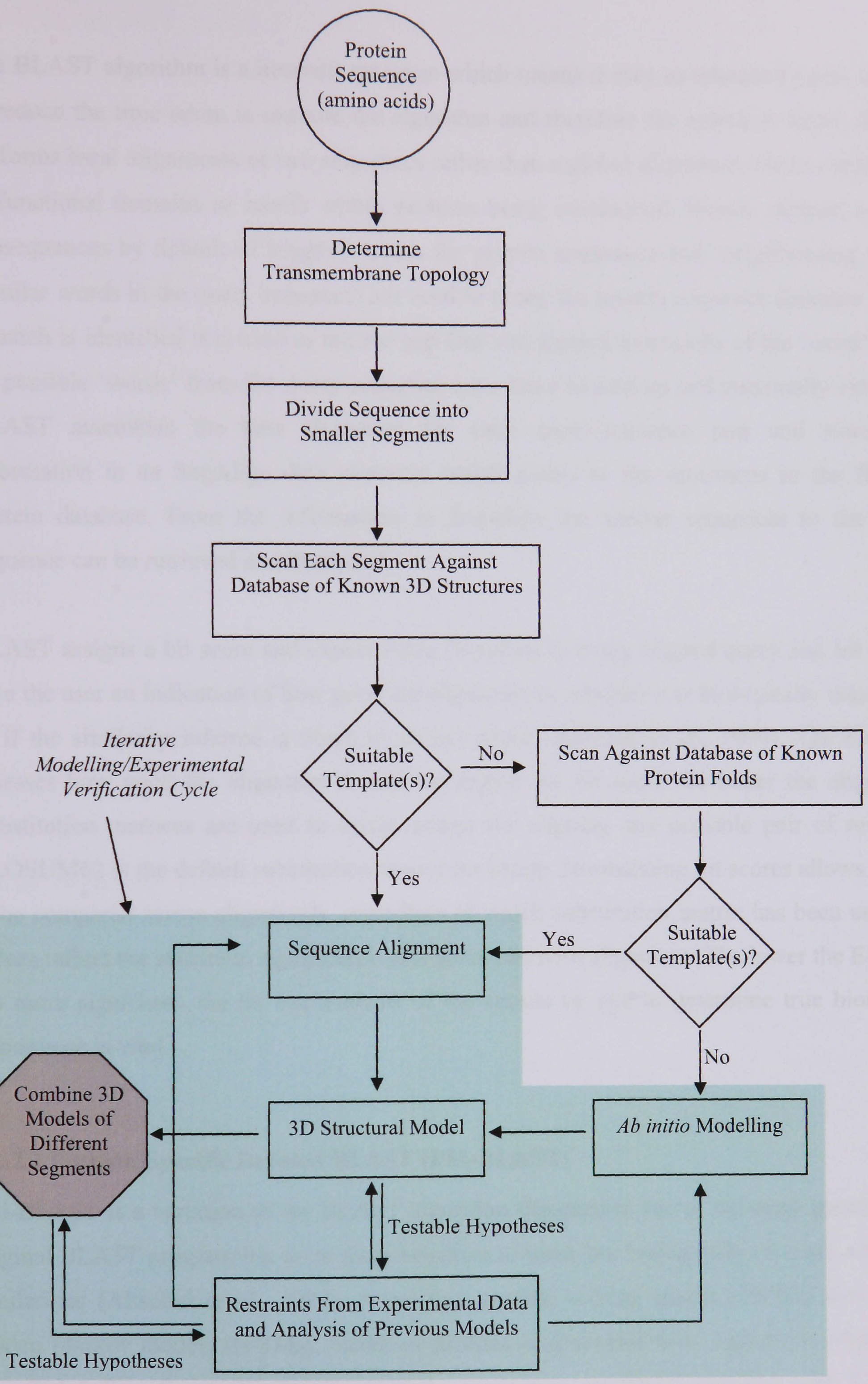


Figure 2.14 Overview of the Model Building Procedure.

The BLAST algorithm is a heuristic program which means it uses an educated guess in order to reduce the time taken to execute the algorithm and therefore the search is faster. BLAST performs local alignments of two sequences rather than a global alignment which could result in functional domains or motifs within proteins being overlooked. Words, defined as short subsequences by default of length 3 letters for protein sequences and 'neighbouring words' (similar words in the query sequence) are used to probe the protein sequence database. When a match is identified it is used to initiate gap-free and gapped extensions of the 'word'. After all possible 'words' from the query sequence have been looked up and maximally extended, BLAST assembles the best alignment for each query-sequence pair and stores this information in an SeqAlign data structure which points to the sequences in the BLAST protein database. From the information in SeqAlign the similar sequences to the query sequence can be retrieved and displayed.

BLAST assigns a bit score and expect value (E-value) to every aligned query and hit pair to give the user an indication of how good the alignment is, whether it is biologically reasonable or if the similarity inferred is down to chance alone (Altschul *et al.*, 1990). The bit score assesses how good the alignment is and the higher the bit score the better the alignment. Substitution matrices are used to assign scores for aligning any possible pair of residues. BLOSUM62 is the default substitution matrix for blastp. Normalising bit scores allows scores to be compared across alignments, regardless of which substitution matrix has been used. E-values reflect the statistical significance of a given pairwise alignment. The lower the E-value, the more significant the hit but analysis of the results by eye to determine true biological importance is vital.

2.2.2.2 Position Specific Iterated BLAST (PSI-BLAST)

PSI-BLAST is a variation of the BLAST algorithm designed to run at the same speed as the original BLAST program but to be more sensitive to weak but biologically relevant sequence similarities (Altschul *et al.*, 1997). A position-specific scoring matrix (PSSM) also called hidden Markov models (HMMs), motifs or profiles is generated from significant alignments in round *i*. Significant alignments are those containing sequences returned with E-values better (lower) than the inclusion threshold (default = 0.005). The PSSM found in round *i* is then used to evaluate the alignment in the next iteration *i+1* (Schneider *et al.*, 1986; Gribskov *et al.*, 1987; Staden, 1988; Tatusov *et al.*, 1994; Bucher *et al.*, 1996). Any new database hits

below the inclusion threshold are included in the generation of the new PSSM. The search converges when no new sequences are added in following iterations. The BLAST algorithm is adapted to take a PSSM as input instead of a query sequence and substitution matrix hence the name PSI-BLAST.

2.2.3 Scanning Amino Acids Sequences against Databases of Known Protein Folds

An alternative strategy to identify a similar protein(s) to a query sequence is to scan the sequence against databases containing protein folds. This method, also known as threading, relies on the principle that proteins with no obvious sequence similarity can show remarkable similarities in their folds. However, it must be noted that fold similarities are not identified as reliably as sequence similarities. Sequences are fitted directly, or 'threaded' onto the backbone of known protein structures (Jones *et al.*, 1992). The belief that there are a limited number of folds in nature adds weight to the reliability of this technique. The sequence is moved position by position through the structure and scored not on amino acid similarity but on physical constraints and the environmental preference of the residues. There are several programs that employ threading methods to identify possible templates for query sequences or model unknown structures of sequences.

2.2.3.1 Threader

The idea of this program is to physically 'thread' a sequence of amino acid side chains onto a backbone structure (fold) and then evaluate the potential 3D structure using a set of pair potentials and a separate salvation potential (Jones *et al.*, 1992). It relies on a similar fold to the native fold of the target protein been found in the library of folds otherwise the whole procedure falls down. It is advisable to chop very long amino acid sequences into single domains as threading cannot be relied upon to locate domain boundaries. Users have the ability to submit PSI-PRED secondary structure predictions as input in addition to an amino acid sequence. Threader reads in the predicted secondary structure and confidence predictions to force the threading alignments generated into agreement with the predicted helices and strands.The results produced by Threader can be complicated to interpret due to the large amount of figures that are generated. An explanation of individual columns of the Threader output is given in Table 2.1 taken from the Threader manual.

Column Number	Description of Result
1	The raw pairwise enery sum for the given threading, units are in kcal/mol
2	This is the energy of the known structure (template) with its own sequence. If the energy of the threaded model (column 1) is less than that of the native structure (column 2) then it is likely this is a false match
3	The Z-score (-(energy – mean)/standard deviation) for the pairwise energies
4	The raw solvation energy sum
5	The Z-score of the salvation energy
6	The weighted sum of pairwise and salvation energy, weighted by standard deviations (Jones *et al.*, 1992)
7	A Z-score of the values in column 6
8	The Z-score of values in column 7 where models with too little matched structure or with other obvious faults are given a default score of -9.99. This is the primary field on which predictions should be based
9	This column gives the sequence similarity score for the model using (using BLOSUM62 score matrix or sequence profile information in the TDB file where available). A high score in this column may indicate common ancestry between the query and template structures
10	This is the number of aligned residues
11	The percentage of the structure which has been aligned
12	The percentage of the sequence which has been aligned
13	The start position for a local domain match, for normal searches this column always contains zero.
14	This gives the Brookhaven PDB code of the structure plus the chain and domain identifier

Table 2.1 Threader Output. The table describes the 14 columns of results that are generated by Threader. Each one has a description of how to interpret the data.

To interpret the filtered combined energy Z-scores (column 8) the THREADER 3 manual supplies the following useful parameters:

$Z > 4.0$	Very significant – probably a correct prediction
$Z > 3.5$	Significant – good chance of being correct
$2.7 < Z < 3.5$	Borderline significant-possibly correct
$2.0 < Z < 2.7$	Poor score –could be right, but needs other confirmation
$Z < 2.0$	Very poor score-probably there are no suitable folds in the library
$Z = -9.99$	Too little of the sequence or template fold have been aligned

These are based on the assumption that the default fold library (over 700 folds) is being used.

2.2.3.2 3D-PSSM

The 3D-PSSM server (Fischer *et al.*, 1999; Kelley *et al.*, 1999) is designed to take a protein sequence, supplied by the user and attempt to predict its 3D structure and probable function. Protein sequences are 'threaded' onto known structures compiled in a library and scored for compatibility. More than one scoring components are used including 1D-PSSMs (sequenced profiles built from relatively close homologues), 3D-PSSMs (more general profiles containing more remote homologues) and matching of secondary structure elements (http://www.sbg.bio.ic.ac.uk/~3dpssm/).

The query sequence is processed by 3D-PSSM in several ways. Firstly it is scanned against an up-to-date sequence library to detect sequence homologues and to generate a sequence profile for the protein of interest. This is done using PSI-BLAST (Altschul *et al.*, 1997). Then PSIPRED (McGuffin *et al.*, 2000) is used to generate a secondary structure prediction. Lastly, the PSI-BLAST results are filtered and keyword information is extracted to provide readable functional information regarding the protein of interest. The query sequence, its profile and secondary structure prediction are scanned against a fold library and the top 20 structural matches are outputted in a table. The table heading are described in the table below (Table 2.2.).

Results Table Heading	Description
Fold	Each entry in the library is a SCOP-defined domain (d[pdbcode][chain][region]) or whole PDB chain (c).
PSSM E-Value	The measure of confidence in the prediction (the closer to 0 the better).
SAWTED E-Value	System used for matching functional keywords between query protein sequence and template structures. The closer the score to 0 the better the match (MacCallum *et al.*, 2000).
E-Value Certainty	A percentage measure of how confident a hit is.

Table 2.2 3D-PSSM Results. The table describes the results generated by 3D-PSSM.

2.2.4 Sequence Alignment

If a suitable template structure is identified then the next step in model building is to produce a good quality sequence alignment between the query and template. The alignment of one or more protein sequences is a comparative method that identifies the similarities and differences between sequences with the aim to infer structure and on the basis of structure, determine function. Two methods of sequence alignment namely pairwise and multiple sequence alignments will be discussed in this next section. For the purposes of this thesis all sequences are protein sequences.

2.2.4.1 Pairwise Alignments

Pairwise alignment refers to the alignment of two sequences to determine whether they show sufficient similarity to conclude whether they are homologous where, two sequences are said to be homologous if they have been derived from a common ancestor. In an alignment, substitutions, insertions and deletions (indels) are an attempt to model the evolutionary steps that occur in sequences during divergence from the common ancestor. Sequences that show an obviously high similarity to each other can be simply aligned using the full length of both sequences (a global alignment) otherwise it is sensible to use a method that will produce a local alignment. A local alignment consists of paired subsequences that may be surrounded by residues that are completely unrelated (Baxevanis & Ouellette, 2001).

2.2.4.2 Multiple Sequence Alignment

Multiple sequence alignments are alignments consisting of more than 2 sequences. Aligning all available sequences even when the similarities of just two sequences are being considered in a given set can improve the accuracy of the alignment between the pair of sequences and is therefore beneficial (Russell & Barton, 1992). Multiple sequence alignments are the starting point for predicting secondary structure and inferring protein function. Analysis of a multiple sequence alignment can give important clues as to which residues are important for protein function, stabilizing secondary and tertiary structure and if information is known about specific residues in one sequence, this information can potentially be transferred to those aligned residues in other sequences. There are many programs that exist for the alignment of multiple protein sequences but simple pre-processing of the sequences and careful analysis of the resulting alignment in a biological context can improve the quality of the alignment. Attempting to align multiple sequences that differ significantly in length will result in a poor quality alignment as most programs are designed to align sequences of a similar length. Therefore it is necessary to edit sequences and only include particular regions of sequences. To increase accuracy, only align sequences that are similar to a query sequence in a database search say with an E-value less than 0.02. Manually inspect and edit the alignment, removing sequences that are obviously disrupting it. Many programs exist for this purpose including Cinema (Parry-Smith *et al.*, 1998), Seaview (Galtier, 1996) and Jalview (Clamp *et al.*, 2004) to name a few. If necessary, re-align the edited sequences.

2.2.4.3 CLUSTALW

CLUSTALW (Thompson *et al.*, 1994) is a free, widely used multiple sequence alignment program that can be accessed through a fast Web service via the EBI server (http:/www.ebi.ac.uk/clustalw). It uses a hierarchical method which is the most accurate, practical method for an automatic multiple alignment. Such methods work by firstly finding a guide tree and then following the guide tree to build the alignment (Baxevanis & Ouellette, 2001). A set of pairwise similarity scores are generated, by comparing all pairs of sequences in the set to be aligned by a pairwise method of sequence comparison. This set of scores is then fed into a cluster analysis or tree calculating program. The tree is calculated such that similar pairs of sequences are placed closer together than those that are less similar. The multiple sequence alignment is then produced by aligning the pair that is closest on the tree, i.e. the most similar pair of sequences and then aligning the next most similar pair and so on until all the sequences have been aligned. However, pairs to be aligned need not be pairs of

sequences; they could be alignments that have been generated earlier in the tree. In this case, gaps that are present in the alignment are preserved. There are many variations of this basic alignment technique (Baxevanis & Ouellette, 2001). CLUSTALW uses a series of different pair-score matrices, biases the location of gaps and allows aligned sequences to be re-aligned to improve the quality of the alignment. It can also read in two pre-existing alignments (two profiles) and align them to each other or align a set of sequences to a profile. CLUSTALW also has phylogenetic tools available. It can take a range of sequence formats as input and has different output formats.

2.2.4.4 Statistical Significance of Alignments

Assigning a score to a given alignment, gives an indication of how 'good' the alignment is but how high a score is necessary to infer homology and how high a score can be expected due to chance alone. There is no mathematical distribution to describe the distribution of scores for global alignments although one method is to compare the alignment score with scores from many alignments constructed from random sequences of similar length to those in the alignment in question (Fitch, 1983).

For local alignments a mathematical model describes the expected distribution of random local alignment scores (Fitch, 1983; Dembo *et al.*, 1984). However, this considers local alignments that are not gapped. Methods have been developed to take into account gaps in local alignments (Waterman & Vingron, 1994; Altschul & Gish, 1996). E-values then give an indication of the statistical significance by relating an observed alignment score to the expected distribution. A simple interpretation of the E-value is the number of alignments with scores at least equal to *S*, an observed alignment score, that would be expected by chance alone (Baxevanis & Ouellette, 2001). General rules of thumb for E-values are:

- $E < 0.02$ = sequences probably homologous
- E in range 0.02 and 1 = homology cannot be ruled out
- $E > 1$ = you would expect this match simply by chance

2.2.4.5 Improving the Quality of an Alignment

A prerequisite to all homology models is a correct alignment with high levels of similarity between the query and template sequences. Manual editing is crucial to obtain a 'good' alignment between query and template sequences especially when sequence identity is less than 40% between the two. In such cases it is the accuracy of the alignment that has most

effect on the quality of the final model produced. To improve the accuracy of the alignment, using structural information from the templates can help. It is possible to generate a sequence alignment based on structure. The functions of proteins are determined by the arrangement of the amino acid side chains in the three-dimensional structure of the protein. If two or more protein 3D structures are compared then residues that are in similar positions in space may indicate residues that are likely to have similar functional roles. Structural alignments most reliably align residues that are functionally important but they are only possible if the 3D structures of all the proteins to be aligned are known. Unfortunately, this is not usually the case so sequence alignment methods have to mimic the structural alignment as close as possible without knowledge of the structure. 3DCOFFEE is a web server for combining sequences and structures into multiple sequence alignments (Poirot *et al.*, 2004).

Using profiles of alignments can also increase accuracy. Structural alignments of templates and multiple sequence alignments of sequences similar to the query sequence (for example other members of a family) can be formed as profiles and then aligned to each other. Such 'profile based alignments' can produce better alignments than pairwise methods as they consider evolutionary and structural information from the templates and evolutionary information from the query sequence (Jaroszewski *et al.*, 2000; Sauder *et al.*, 2000).

Programs such as RASCAL (Thompson *et al.*, 2003) have been developed to improve and correct multiple alignments. It uses a knowledge-based approach and other existing methods to refine alignments by dividing them horizontally and vertically to form a lattice, allowing well-aligned regions and less well aligned regions to be identified. Refinement is then concentrated on the less well aligned regions to improve the overall quality of an alignment.

2.2.5 Homology Modelling: 3D Structural Model

At this point, suitable template structures have been selected and a good quality alignment between query sequence and template structures has been generated. The next stage is to build the model. There are two approaches to model building; fragment-based and single-step.

2.2.5.1 Fragment based approach (e.g. SWISS-MODEL)

This method dissects the known protein structures into structural components; well defined polypeptide backbone, variable loops that connect the well-defined backbone (poorly defined polypeptide backbone) and amino acids sidechains. Then a four-step procedure is followed:

1. Build a well-defined polypeptide backbone-this is a disjointed set of fragments and tends to be secondary structural elements such as α-helices and β-strands.
2. Using loops to connect secondary structural elements, join the fragments.
3. Alter amino acid sidechains, modelling them on their own conformational preferences and on the conformation of the equivalent sidechains in the template structures. Rotamer libraries can be used such as SCWRL (Canutescu *et al.*, 2003) (Sutcliffe *et al.*, 1987).
4. Refine the model. The stereochemistry of the model can be improved by energy minimisation or a molecular dynamics (MD) refinement followed by energy minimisation. The accuracy of the model can also be improved by using more than one template structure.

One example of a program that uses a fragment-based approach to automate sequence-structure comparisons and produce homology models is SWISS-MODEL (Peitsch, 1996). This has several approaches. The First Approach mode (regular) submits an amino acid sequence to be compared with the crystallographic database (ExPdb) and only if a homolog to the query protein is found in the ExPdb is modelling attempted. The First Approach mode also allows user defined templates to be submitted at the same time as the query sequence. Template structures are only selected if there exists at least 25% sequence identity in a region more than 20 amino acids long. If one or more appropriate templates are found in the ExPdb, atomic models are built and energy minimization performed to produce the best model. Models, as atomic co-ordinates in a pdb file, and the structural alignments are then emailed to the user as an attachment. Those results can be re-submitted to SWISS-MODEL using its Optimize mode, which considers other information including biochemical knowledge and adjusts the model accordingly. It is also possible to use the Alignment Interface which allows a user to submit an alignment and specify query and template sequences.

2.2.5.2 Single Step approach (e.g. MODELLER)

In this approach, probability density functions represent individual structural features in the model. If the mainchain conformation is considered and the positioning of hydrogen bonds, then the function will give the probability that a particular hydrogen bond will be formed in the model. Such probabilities are then used as restraints in the building of the model. Unlike a fragment-based approach, a family of models each with different conformations are created that are consistent with the set of distributions produced due to each individual structural

feature being represented by a probability distribution. This approach allows the significance of different interactions and conformations in the final models to be evaluated. MODELLER (Sali & Blundell, 1993) is one such program that employs a single step approach to generate a 3D model of the query sequence containing all mainchain and sidechain non-hydrogen atoms. Users supply an alignment of a query sequence and related known template structures and comparative models are generated by satisfying spatial restraints (Sali & Blundell, 1993; Fiser *et al.*, 2000).

2.2.6 Refine and validate 3D Model using Restraints

Component domains of multi-domain proteins should be combined at this stage. It is important to check for possible errors in the final model and attempt to refine the model by improving the template selection or quality of the alignment between query and template and evaluating whether it agrees with known experimental data for example. Some considerations are listed below:

- Query-template similarity
- Assessment of the model's stereochemistry (bonds, bond angles, dihedral angles, and non-bonded atom-atom distances) and amino acid environment; using programs such as PROCHECK (Laskowski *et al.*, 1993) for main chain and side chain geometry, WHATCHECK (Hooft *et al.*, 1996) for chirality, unsatisfied buried hydrogen bond donors/acceptors, and bad contacts and PROVE (Pontius *et al.*, 1996) for atomic volume.
- Atomic environment. ERRAT (Colovos & Yeates, 1993) is a program for verifying protein structures determined by crystallography and improves models by plotting error values and rejecting regions of the structure that are below confidence levels.
- Energy minimisation to help relieve steric problems
- Consistency with experimental observations, site-directed mutagenesis data, cross-linking data and ligand binding information

The importance of experimentally validating and manually correcting and inspecting the intermediate alignments between query and template sequences is vital to produce a model that is both robust and believable. However, no model should be considered a 'true' representative of the 'real' protein structure without experimental validation and should be considered as another tool in the process of solving structural problems.

2.2.7 Ab Initio Methods

Ab Initio methods provide an alternative technique to the comparative and fold recognition methods previously discussed. They seek to predict the native conformation of the protein from the amino acid sequence alone. *Ab Initio* predictions follow the steps below:

- Define a representation of the polypeptide chain and solvent. This may possibly be a simplified representation.
- Define some kind of energy function that tries to model the interactions found in proteins.
- Search for the simplified chain conformation with the lowest energy

The HMMSTR/Rosetta server employs *ab initio* methods to predict the secondary (helices, beta strands and loops), local (backbone torsion angles phi/psi), supersecondary (context symbols for strands and beta turns) and tertiary structure (co-ordinates) of proteins from sequence (Bystroff & Shao, 2002).

Chapter Three

3.0 Preface

Conserved glycine residues in the extracellular domain of the $hP2X_1$ receptor have been mutated to alanine to determine the effect on channel function. By investigating the role of conserved glycine residues further structural insight into the P2X receptor family was sought. This work in this chapter was published in 2005 (Digby *et al.*, 2005).

3.1 Introduction

The P2X receptor family is structurally distinct from the two other ligand-gated ion channel superfamilies; cys-loop superfamily and the glutamate activated cationic channels (Surprenant *et al.*, 1995). The lack of available crystal structures for the P2X receptor family and the low sequence/structural homology shared with other ATP binding proteins has meant that mutagenesis studies have been relied upon to gain an insight into the molecular basis of the properties of P2X receptors.

The extracellular domain of the $hP2X_1$ receptor has been extensively targeted for mutagenesis studies (Vial *et al.*, 2004). Conserved amino acids are considered to be important across a receptor family as they have remained throughout evolution and are likely to have a functional role. Groups of conserved residues with similar properties can be replaced with alanine or cysteine to identify key residues important for successful agonist binding to the receptor or maintaining the structure of the receptor. Conserved positively charged residues (K68, R292 and K309 $P2X_1$ receptor numbering) close to the pore of the $P2X_1$ receptor are associated with the binding of the negatively charged phosphate tail of ATP (Ennion *et al.*, 2000). Conserved negatively charged residues may play a role by coordinating the magnesium in the complex MgATP or they may interact with the positively polarized amino group in the adenine ring of ATP (Ennion *et al.*, 2001). However, alanine-replacement mutagenesis revealed that individual conserved negatively charged residues (aspartic (D) and glutamic acid (E)) are not essential for ATP recognition by the $hP2X_1$ receptor (Ennion *et al.*, 2001). The contribution of aromatic residues to the recognition of the adenine and ribose functional groups of ATP has been investigated in other ATP-binding proteins which found roles for phenylalanine (Tanner *et al.*, 2003) and tyrosine (Theis *et al.*, 1999). Mutagenesis studies in the $hP2X_1$ receptor identified residues F185 and F291 that have been suggested to coordinate the binding of the adenine ring of ATP (Roberts & Evans, 2004). Conserved uncharged polar

residues, glutamine (Q), asparagine (N) and threonine (T) can also play a part in activating ligand-gated ion channels by hydrogen bonding to the agonist or stabilizing conformational changes within the protein (Roberts & Evans, 2006). When conserved polar residues were replaced with alanine in the $hP2X_1$ receptor, results confirmed the role of two small conserved motifs across the P2X receptor family, $^{185}FT^{186}$ and $^{290}NFR^{292}$ in ATP action at the receptors (Roberts & Evans, 2006).

A distinct model of those conserved residues involved in ATP action at the P2X receptor family has been constructed from different mutagenesis studies concerning some of the seven P2X receptor subtypes. However, there are three amino acids that have the ability to influence the structure and binding/gating movements of the extracellular loop of the $hP2X_1$ receptor. These are cysteine, proline and glycine residues. Focussing on the conserved cysteine and proline residues has led to a clear picture of the structural constraints placed on the protein by these particular amino acids. Cysteine residues contain a reactive sulfhydryl group that allows disulfide bonds to form giving cysteine the ability to stabilize protein structure. There are 10 conserved cysteine residues in the extracellular loop of the P2X family. Single and double cysteine to alanine mutants in the human $P2X_1$ receptor revealed five possible disulfide bond pairings (Ennion & Evans, 2002a) although none of the bonds are thought to be individually essential for channel function. A similar alanine-replacement mutagenesis study of the ten conserved cysteine residues was conducted in rat $P2X_2$ receptors. The same cysteine pairings were predicted in the $rP2X_2$ receptor as for the $hP2X_1$ receptor (Clyne *et al.*, 2002).

Proline residues can influence the secondary structure of proteins by introducing kinks in α-helices due to their unique structure. The side chain of a proline residue is covalently bonded to the main chain nitrogen which results in a ring structure as the main chain amide nitrogen is prohibited from participating in hydrogen bonding. Mutation of individual proline residues in the extracellular domain of the human $P2X_1$ receptor revealed that no single proline residue was essential for normal channel function (Roberts & Evans, 2005).

Information regarding the structure of the P2X receptor is critical to understanding the mechanism of ATP binding to the receptor. The role of conserved glycine residues in the extracellular domain of the $hP2X_1$ receptor have not been investigated but could have important structural consequences due to the unique nature of glycine that allows it to be more flexible than other amino acids. In order to generate homology models of the $hP2X_1$ receptor it is important to consider all three residues, cysteines, prolines and glycines that play a role in

determining the structure of the $hP2X_1$ receptor protein and for this reason the conserved glycine residues of the extracellular loop of the $hP2X_1$ receptor have been mutated to determine their effect on receptor function.

Glycine (Figure 3.1) is the simplest of all amino acids with a single hydrogen atom forming its side chain.

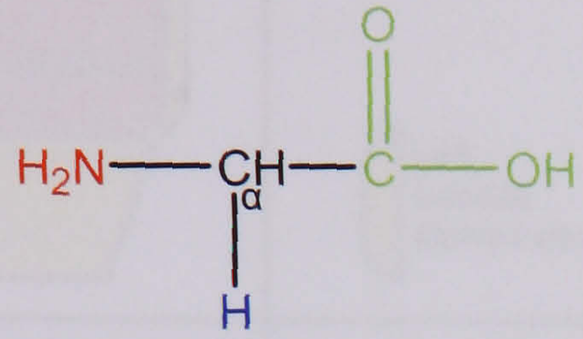


Figure 3.1 Structure of glycine. The general structure of an amino acid molecule contains an amine and carboxyl functional group and an R-group which is dependent on the amino acid. Glycine is unique because it contains an amine group (red) and a carboxylic acid group (green) but its R group consists of a lone hydrogen (blue).

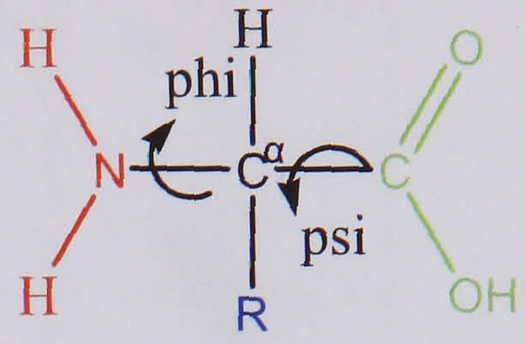


Figure 3.2 Peptide torsion angles. The general structure of an amino acid molecule is shown (amine group is red, carboxyl group is green and R group not specified, blue). The ability of proteins to fold arises from rotation about phi and psi bonds.

The ability of proteins to fold in numerous ways arises from the freedom of rotation about two bonds of each amino acid, namely between the nitrogen and α-carbon (φ phi) and that between the α-carbon and carbonyl carbon atom (ψ psi) (Figure 3.2). Varying the angles about these two bonds renders some protein conformations sterically disallowed for all amino acids except glycine due to its lack of side chain. A Ramachandran plot (Ramachandran *et al.*, 1963) is a 2D plot that visualises the φ and ψ torsion angles of the protein backbone. The φ and ψ angles cluster into several distinct regions in the Ramachandran plot, each corresponding to an element of the protein secondary structure. There are four types of Ramachandran plots depending on the stereochemistry of the amino acid; generic (all 18 amino acids except glycine and proline), glycine, proline and pre-proline (residues preceding a proline (MacArthur & Thornton, 1991)). The generic plot (Figure 3.3a) has become the standard explanation for the observed regions in the Ramachandran plot (Richardson, 1981). Due to glycine's lack of side chain it is less restricted and therefore can occupy a larger area in the Ramachandran plot. The glycine Ramachandran plot has a distinctive distribution (Ho & Brasseur, 2005) (Figure 3.3b).

Therefore, glycine is a unique amino acid because it lacks a side chain and the result of this is that it is less restricted (visualized on a Ramachandran plot) compared to other residues and this property gives glycine its flexible nature.

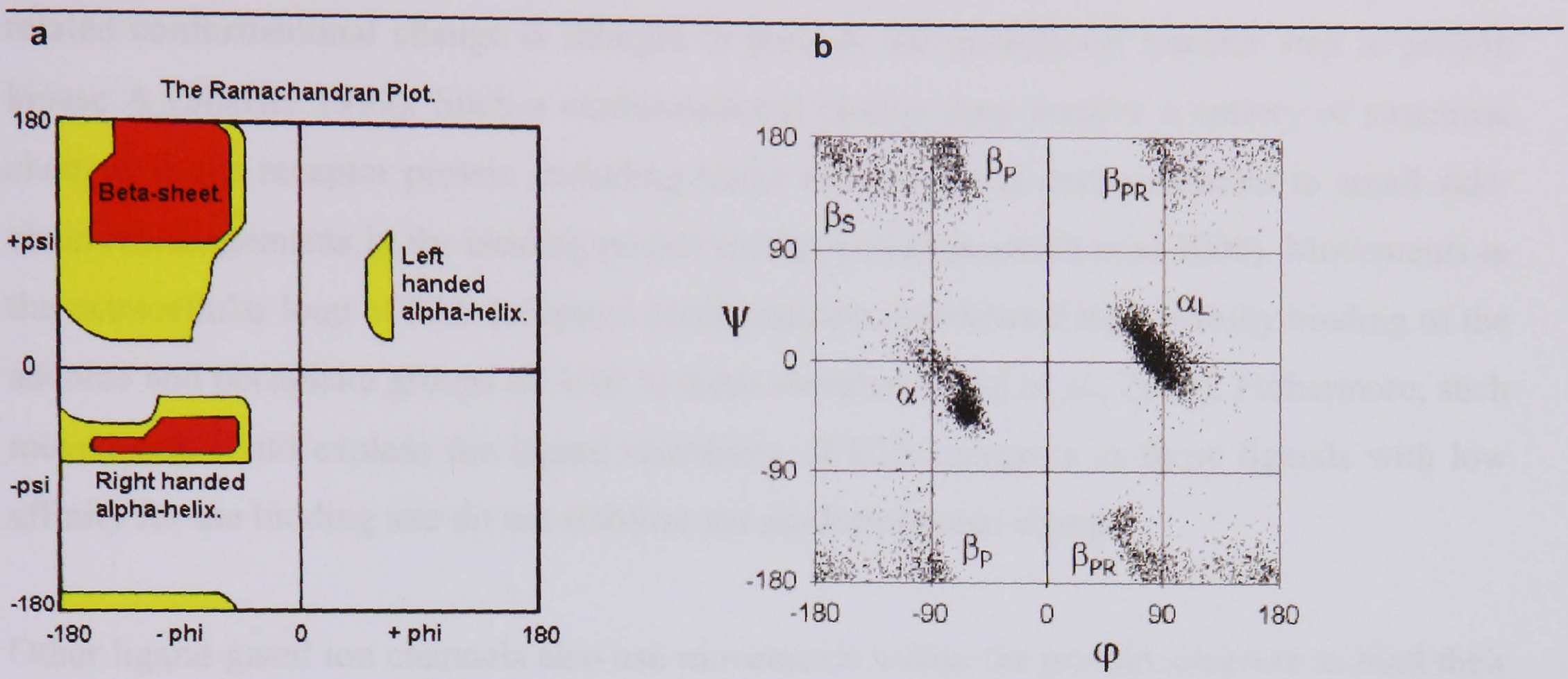


Figure 3.3 Ramachandran Plots. a, generic Ramachandran plot. White areas correspond to conformation where atoms in the polypeptide come closer than the sum of their van der Waals radii. These regions are sterically disallowed for all amino acids except glycine. Red, conformations where there are no steric clashes (alpha-helical and beta-sheet conformations). Yellow areas are the allowed regions if slightly shorter van der Waals radi are used in the calculation. This brings out the region corresponding to the left-handed alpha-helix. Glycine has no side chain and therefore can adopt phi and psi angles in all four quadrants of the Ramachandran plot. Hence it frequently occurs in turn regions of proteins where any other residue would be sterically hindered. **b, Glycine Ramachandran plot**.

Glycine can rotate easily, add flexibility to a protein and as a result, is able to fit into tight spaces other amino acids would not. Conserved glycine residues across a family are strong indicators of loops/secondary structure breakers and their flexible nature would allow certain conformational changes or movement within the protein that may be necessary for the successful binding of an agonist. In addition, glycine residues have been shown to form part of consensus motifs for ATP binding, for example the Walker motif (GxxxxGKxxxxxxxI/V) (J.E. *et al.*, 1982) and a motif conserved in over 95% of known human protein kinases (GxGxxG) (Okkeri *et al.*, 2004). Conserved glycine residues have also been labeled as 'hinge points' allowing the pore of ion channels to open and close. Results have suggested that voltage dependent potassium channels open when the inner helix bends at a conserved glycine hinge (Magidovich & Yifrach, 2004). Crystal structures and mutagenesis studies have supported this role of glycine as a hinge in activation gating in potassium channels (Ding *et*

al., 2005). By considering the conserved glycine residues in the $hP2X_1$ receptor similar roles in channel gating may be discovered.

When ATP binds it can cause a conformational change in the receptor. For example, an ATP-related conformational change is thought to precede the phosphoryl transfer step in protein kinase A (Shaffer 1999). Such a conformational change may involve a variety of structural changes in the receptor protein including hinge movements of entire domains to small side-chain rearrangements in the binding pocket residues (Najmanovich et al 2000). Movements in the extracellular loop of P2X receptors could regulate coordinated high-affinity binding of the adenine and phosphate groups of ATP to these receptors (Vial *et al.*, 2004). Futhermore, such movements could explain the ligand selectivity of P2X receptors as those ligands with low affinity for the binding site do not stabilize the conformational change.

Other ligand-gated ion channels also use movements within the protein structure to bind their agonist, for example glutamate receptors. Glutamate receptors are ligand-gated ion channels that form from four homologous subunits. Each subunit has four discrete regions (Paas, 1998); the extracellular amino terminal domain, the extracellular ligand-binding core (so-called S1 and S2 regions), three transmembrane domains and the intracellular carboxy-terminal domain. Glutamate receptors can be divided into three subgroups based on their responses to selective glutamate analogues: AMPA, kainate and NMDA. The crystal structure of the GluR2 S1S2-kainate complex revealed a 'clam-shell' like structure for the receptor fragments and that agonist was bound in the cleft between each shell (Armstrong *et al.*, 1998). By forming crystals of the GluR2 S1S2 construct in a range of states (for example, pre bound and agonist bound) and with a variety of ligands the role and importance of the 'clam-shell' structure was investigated (Gouaux, 2004). It was found the 'clam-shell' was most open in a pre bound state. It was suggested after studying the conformation of the receptor in the presence of glutamate and AMPA, that the basic conformational change involved in activation of the ion channel was the closure of the clamshell (Gouaux, 2004). The partial agonist kainite revealed that the amount of closure of the S1S2 construct was dependent on the extent of receptor activation. All these results put forward a simple hypothesis suggesting that the binding of agonist and closure of the clam-shell could be coupled to the activation or opening of the ion channel gate (Gouaux, 2004). This result has important consequences on the study of the structure of other ligand gated ion channels and how the structure can relate to the binding mechanisms.

The aim of this chapter was to investigate the role of glycine residues in the structure of the human $P2X_1$ receptor and their part in agonist binding in order to enhance the models already proposed and provide further restraints and validation points in expectation of homology models of the $hP2X_1$ receptor. The extracellular loop of the $P2X_1$ receptor has 16 conserved glycine residues, 11 of which are 100% conserved and the remainder conserved in at least 5 family members (Figure 3.4). All sixteen glycine residues have been mutated to alanine in the $hP2X_1$ receptor to investigate the effects on channel function. Alanine was chosen because it is a similar size to glycine but it does not share the same flexible nature.

```
                                      71                                                                96
hP2X1    -YEKGYQ-TSSGLISSVSVKLKGLAVT-------QLPGLGPQVWDVADYVFPAQGDNSFV
hP2X2    --QKSYQESETGPESSIITKVKGITTS------------EHKVWDVEEYVKPPEGGSVFS
hP2X3    LHEKAYQVRDTAIESSVVTKVKGSGLY------------ANRVMDVSDYVTPPQGTSVFV
hP2X4    ------QETDS-VVSSVTTKVKGVAVTN-------TSKLGFRIWDVADYVIPAQEENSLF
hP2X5    --KKGYQDVDTSLQSAVITKVKGVAFTN-------TSDLGQRIWDVADYVIPAQGENVFF
hP2X6    --KKGYQERDLEPQFSIITKLKGVSVTQ-------IKELGNRLWDVADFVKPPQGENVFF
hP2X7    -SDKLYQ-RKEPVISSVHTKVKGIAEVKEEIVENGVKKLVHSVFDTADYTFPLQG-NSFF

                          115                                  135       143  147
hP2X1    VMTNFIVTPKQTQGYCAEHP-EG-GICKEDSGCTPGKAKRKAQGIRTGKCVAFNDTVK-T
hP2X2    IITRVEATHSQTQGTCPESIRVHNATCLSDADCVAGELDMLGNGLRTGRCVPYYQGPSKT
hP2X3    IITKMIVTENQMQGFCPESE--EKYRCVSDSQCGP--EPLPGGGILTGRCVN-YSSVLRT
hP2X4    VMTNVILTMNQTQGLCPEIPD-ATTVCKSDASCTAGSAGTHSNGVSTGRCVAFNGSVK-T
hP2X5    VVTNLIVTPNQRQNVCAENEGIPDGACSKDSDCHAGEAVTAGNGVKTGRCLRRENLARGT
hP2X6    LVTNFLVTPAQVQGRCPEHPSVPLANCWVDEDCPEGEGGTHSHGVKTGQCVVFNGTHR-T
hP2X7    VMTNFLKTEGQEQRLCPEYP-TRRTLCSSDRGCKKGWMDPQSKGIQTGRCVVHEGNQK-T

hP2X1    CEIFGWCPVEVDDDIPRPALLREAENFTLFIKNSISFPRFKVNRRNLVEEVNAAHMKTCL
hP2X2    CEVFGWCPVEDGASVSQF-LGTMAPNFTILIKNSIHYPKFHFSKGN-IADRTDGYLKRCT
hP2X3    CEIQGWCPTEVDT-VETP-IMMEAENFTIFIKNSIRFPLFNFEKGNLLPNLTARDMKTCR
hP2X4    CEVAAWCPVEDDTHVPQPAFLKAAENFTLLVKNNIWYPKFNFSKRNILPNITTTYLKSCI
hP2X5    CEIFAWCPLETSSRPEEP-FLKEAEDFTIFIKNHIRFPKFNFSN-NVMDVKDRSFLKSCH
hP2X6    CEIWSWCPVESGVVPSRP-LLAQAQNFTLFIKNTVTFSKFNFSKSNALETWDPTYFKHCR
hP2X7    CEVSAWCPIEAVEEAPRPALLNSAENFTVLIKNNIDFPGHNYTTRNILPGLN----ITCT

                              233       240           250  254
hP2X1    FHKTLHPLCPVFQLGYVVQESGQNFSTLAEKGGVVGITIDWHCDLDWHVRHCRPIYEFHG
hP2X2    FHEASDLYCPIFKLGFIVEKAGESFTELAHKGGVIGVIINWDCDLDLPASECNPKYSFRR
hP2X3    FHPDKDPFCPILRVGDVVKFAGQDFAKLARTGGVLGIKIGWVCDLDKAWDQCIPKYSFTR
hP2X4    YDAKTDPFCPIFRLGKIVENAGHSFQDMAVEGGIMGIQVNWDCNLDRAASLCLPRYSFRR
hP2X5    FGPK-NHYCPIFRLGSVIRWAGSDFQDIALEGGVIGINIEWNCDLDKAASECHPHYSFSR
hP2X6    YEPQFSPYCPVFRIGDLVAKAGGTFEDLALLGGSVGIRVHWDCDLDTGDSGCWPHYSFQL
hP2X7    FHKTQNPQCPIFRLGDIFRETGDNFSDVAIQGGIMGIEIYWDCNLDRWFHHCRPKYSFRR

                          288                  301             312          321 324
hP2X1    LYEEKN---LSPGFNFRFARHFVE-NGTNYRHLFKVFGIRFDILVDGKAGKFDIIP----
hP2X2    LD--PKHVPASSGYNFRFAKYYKI-NGTTTRTLIKAYGIRIDVIVHGQAGKFS-------
hP2X3    LDSVSEKSSVSPGYNFRFAKYYKMENGSEYRTLLKAFGIRFDVLVYGNAGKFN-------
hP2X4    LDTRDVEHNVSPGYNFRFAKYYRDLAGNEQRTLIKAYGIRFDIIVFGKAGKFDIIPTMIN
hP2X5    LDN-KLSKSVSSGYNFRFARYYRDAAGVEFRTLMKAYGIR--------------------
hP2X6    QE---------KSYNFRTATHWWEQPGVEARTLLKLYGIRFDILVTGQAGKFG-------
hP2X7    LDDKTTNVSLYPGYNFRYAKYYKE-NNVEKRTLIKVFGIRFDILVFGTGGKFDIIQLV-
```

Figure 3.4 Sequence alignment of extracellular domain of human P2X receptors. Conserved glycine residues across the P2X family are highlighted (**bold red**). Mutated residues of importance for ATP potency are highlighted in green (**bold green**) for the $P2X_1$ receptor. Numbered residues refer to the human $P2X_1$ sequence. SwissProt accession numbers for sequences are P51575, Q9UBl9, P56373, Q99571, Q93086, O15547, Q99572 for $P2X_{1\text{-}7}$ respectively.

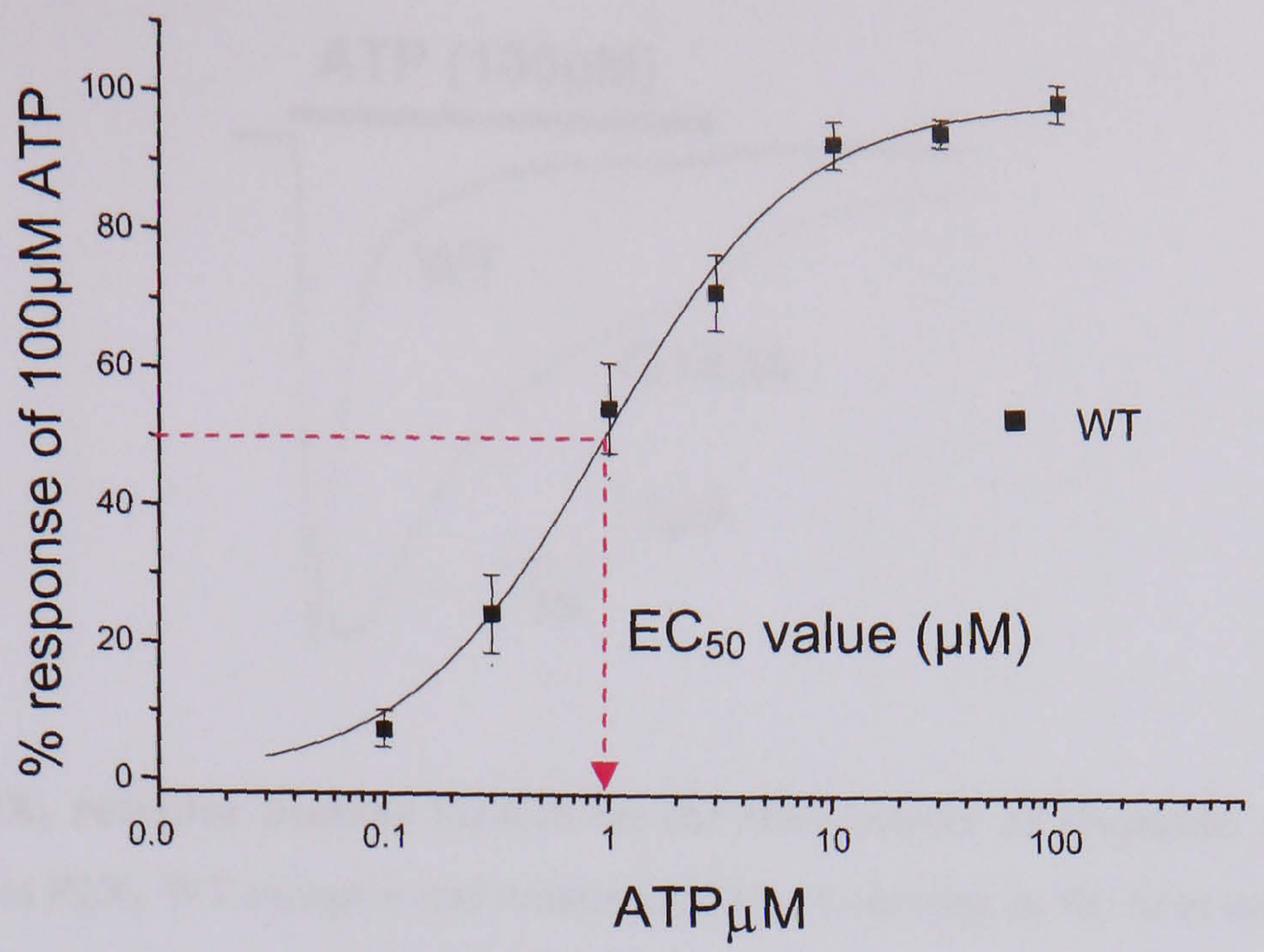
3.2 Results

3.2.1 Effect of alanine replacement of the 16 conserved glycine residues in the extracellular domain of the $P2X_1$ receptor, on the response to ATP

P2X receptors respond to ATP in a concentration dependent manner. Measuring the peak amplitude of the response of WT and mutant $P2X_1$ receptors at different concentrations was used to plot concentration-response curves. The concentration at which the response is 50% of its maximum is called the EC_{50} value (effective concentration) (Figure 3.5a). At wild-type human $P2X_1$ receptors ATP evokes a concentration-dependent desensitizing inward current with an EC_{50} of ~1.4μM (as described previously (Ennion *et al.*, 2000)) (Table 3.1) (Figure 3.8). Other measures we can use to describe the response to a ligand are peak amplitude, rise-time and half-time (Figure 3.5b). Peak data is collected from the first application of a maximal ligand concentration. Rise-time and half-time describe the time-course of response. Rise-time is defined as the time from 10 to 90% of the peak current after a maximal concentration of ATP. Half-time represents the time from peak to 50% decay of the current after application of a maximal concentration of ATP.

For the majority of conserved glycine-to-alanine substitutions (G135A, G147A, G233A, G240A, G251A, G288A, G321A, G324A) little (less then 2-fold) or no effect was seen on ATP potency or time course of the response (Table 3.1) (Figure 3.7). A small increase (~3-fold) in ATP potency with no effect on the time course of the response was seen for mutants G254A and G312A. ATP potency was decreased by 2-3-fold for mutants G115A and G143A. This was accompanied by a slowing in the time course (~3-fold) of the response for mutant G143A (Figure 3.6). The greatest decrease (~6-fold) in ATP potency was recorded for the mutant G71A (Figure 3.8). Various concentrations of ATP ranging from 0.1μM to 100μM failed to evoke responses from three mutants; G96A, G250A and G301A. In order to determine whether there was a large shift in ATP potency for these mutants or whether they were non-responsive to ATP, high concentrations of ATP (up to 10mM) were applied. ATP up to 10mM failed to evoke responses from mutants G96A, G250A and G301A.

a



b

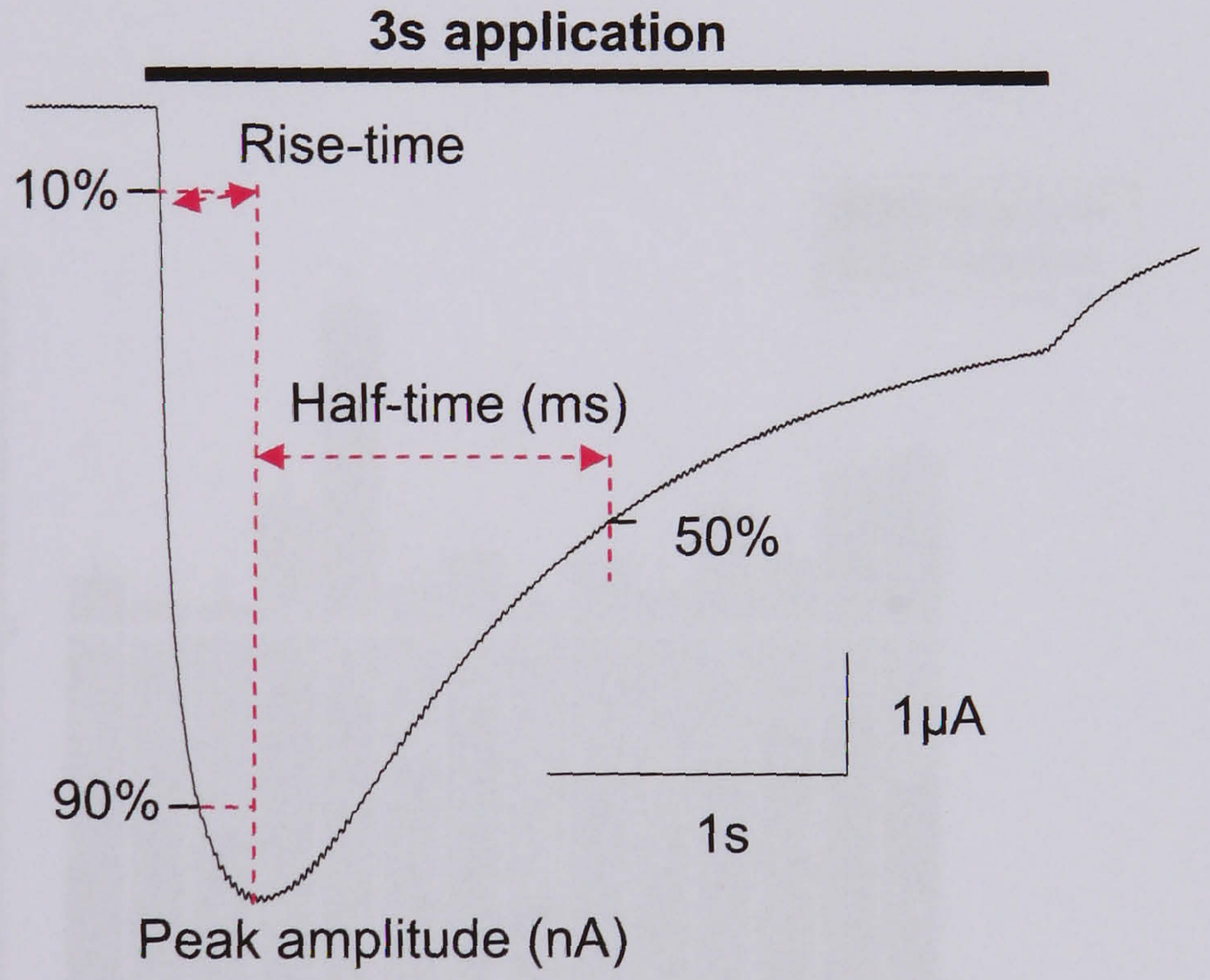


Figure 3.5 Quantification of $P2X_1$ receptor current. a, A typical concentration-response curve of $P2X_1$ WT receptor. The EC_{50} value is shown with a red dashed arrow and is read in μM from the x-axis. **b**, A typical $P2X_1$ response to agonist application (100μM ATP) expressed in *Xenopus* oocytes. Holding potential -60mV. The rise-time and half-time are shown with red dashed arrows, measured in ms.

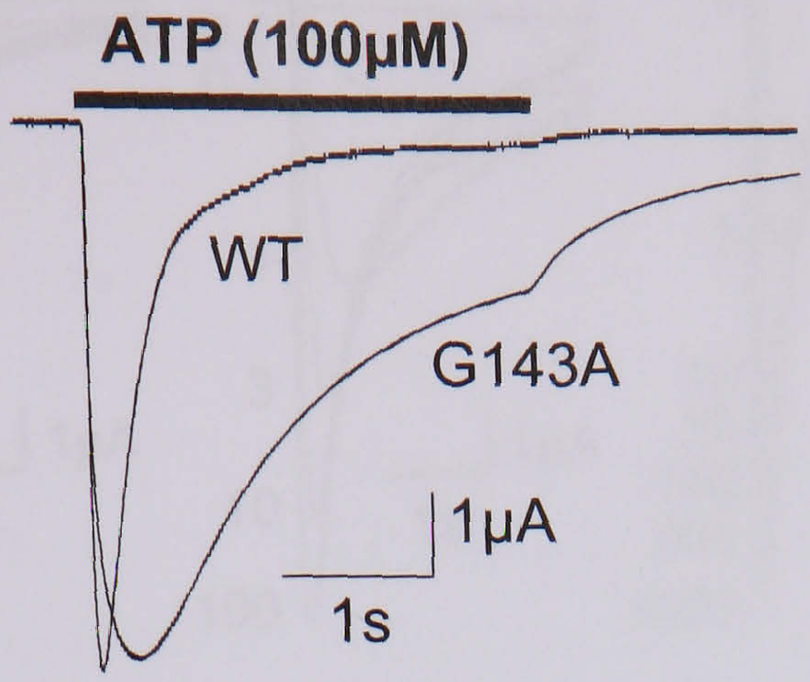


Figure 3.6 Effect of $P2X_1$ receptor mutant G143A on the time course of response. Currents recorded in response to 100μM ATP at $P2X_1$ WT receptor and mutant G143A. A slowing in the time course of response (~3-fold) is recorded for mutant G143A compared with $P2X_1$ WT

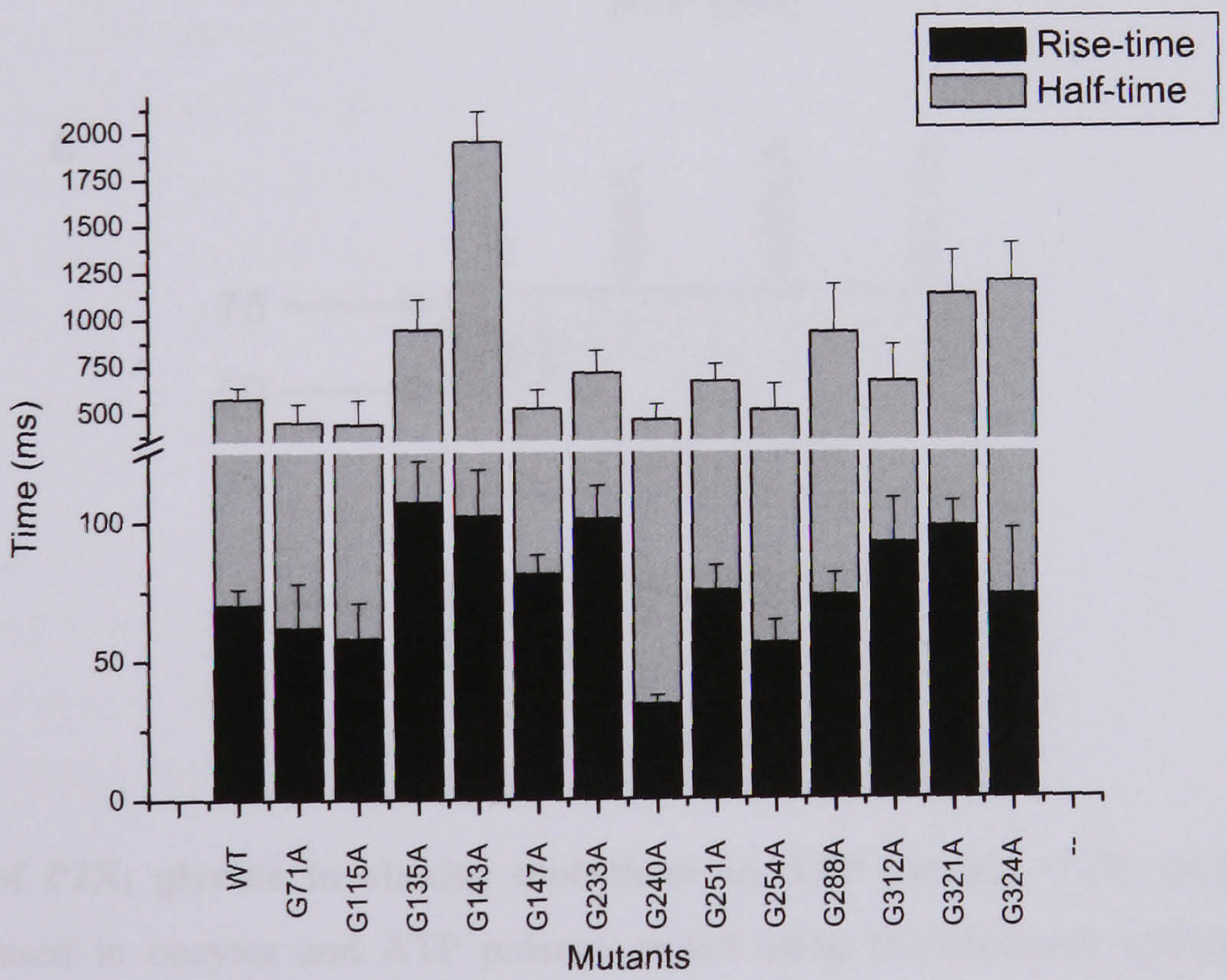


Figure 3.7 Effect of glycine mutations at $P2X_1$ receptors on the time course of response compared to WT $P2X_1$ receptor. Rise-time represents the time from 10 to 90% of peak current and half-time represents the time from peak to 50% decay of the current, both after a maximal stimulating concentration of ATP (n = 4–15). Values shown are representative of the mean ± S.E. of the mean.

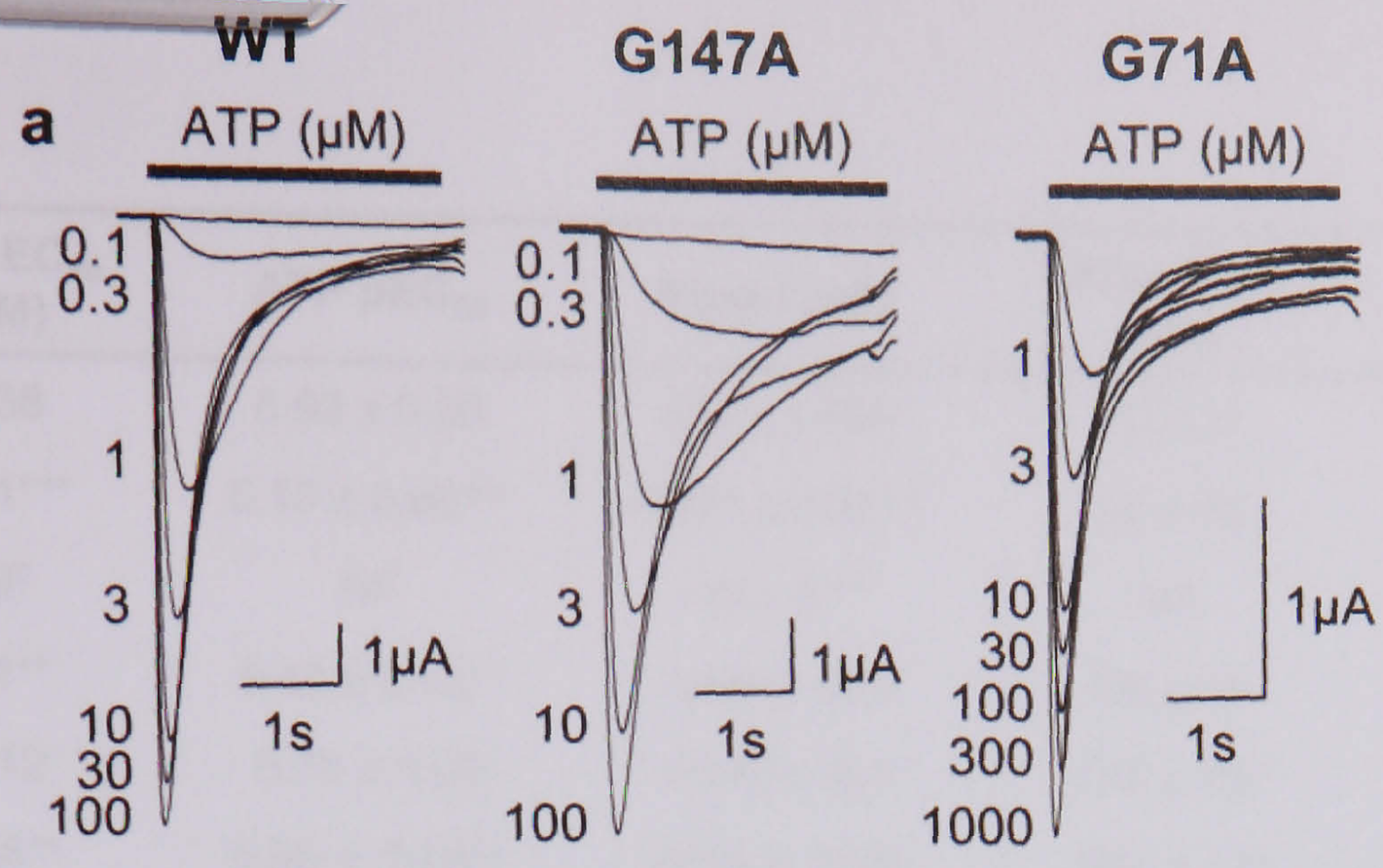


Figure 3.8 Effect of $P2X_1$ glycine to alanine mutations on ATP potency. $P2X_1$ receptor WT or glycine mutants were expressed in oocytes and ATP potency tested using two-electrode voltage clamp. **a**, currents recorded in response to ATP at $P2X_1$ WT and mutants G147A and G71A. **b**, concentration response curves for $P2X_1$ WT and glycine mutants G71A, G96A, G147A, G250A and G301A. Mean currents were normalized to the maximum response (n = 4-9). Holding membrane potential was –60 mV. **c**, mutation of glycine residues 96, 250 and 301 to alanine resulted in non-functional channels. Western blot analysis was carried out using $P2X_1$ antibody on total and cell surface-expressed protein. Cell surface proteins were extracted by sulfo-NHS-LC-Biotin labelling and isolated with streptavidin-agarose beads. Immunoreactivity was observed at ~55 kDa, the correct *Mr* for $P2X_1$ protein and was not seen for non-injected oocytes (data not shown).

Mutant	ATP EC_{50} (μM)	ATP pEC_{50}	Peak I (nA)	Rise-time (ms)	Half-time (ms)
WT	1.38	5.93 ± 0.09	-8571 ± 595	70 ± 6	577 ± 61
G71A	8.31***	5.10 ± 0.06***	-4081 ± 500***	62 ± 16	450 ± 99
G96A	NF	NF	-18 ± 5***	NF	NF
G115A	4.1**	5.42 ± 0.12**	-8480 ± 870	58 ± 13	437 ± 131
G135A	2.12	5.70 ± 0.08	-9546 ± 981	107 ± 15**	945 ± 164*
G143A	2.88**	5.55 ± 0.04**	-9796 ± 1075	102 ± 17*	1953 ± 164***
G147A	1.39	5.95 ± 0.17	-8896 ± 1091	81 ± 7	522 ± 100
G233A	1.44	6.00 ± 0.13	-11484 ± 638**	101 ± 12*	711 ± 121
G240A	0.8	6.10 ± 0.04	-11900 ± 1057**	34 ± 3***	460 ± 83
G250A	NF	NF	-7 ± 4***	NF	NF
G251A	0.85	6.13 ± 0.11	-4975 ± 970**	75 ± 9	665 ± 91
G254A	0.55*	6.27 ± 0.05*	-6973 ± 227	58 ± 8	509 ± 141
G288A	2.07	5.73 ± 0.10	-9444 ± 2407	73 ± 8	928 ± 256*
G301A	NF	NF	-6 ± 1***	NF	NF
G312A	0.53*	6.28 ± 0.04**	-9127 ± 585	92 ± 16	663 ± 195
G321A	2.2	5.90 ± 0.16	-5360 ± 929**	98 ± 9**	1131 ± 230*
G324A	1.47	5.84 ± 0.04	-11854 ± 778**	73 ± 24	1202 ± 201**

Table 3.1 Summary of data for ATP action at $P2X_1$ WT and glycine mutants. EC_{50} data are representative of data collected from three to nine oocytes. pEC_{50} is –log10 of the EC_{50}. Peak data was collected on first application of a maximal ATP concentration (n = 5-17). Rise-time represents the time from 10 to 90% of peak current after an application of a maximal stimulating concentration of ATP (n = 4-15). Half-time represents the time from peak to 50% decay of the current after an application of a maximal stimulating concentration of ATP (n=4-10). Values shown are representative of the mean ± S.E. of the mean.

* $P < 0.05$, different from WT as measured by Student's t test.

** $P < 0.01$, different from WT as measured by Student's t test.

*** $P < 0.001$, different from WT as measured by Student's t test.

NF, non-functional receptor, therefore data not collected

3.2.2 Effect on receptor expression

To determine whether the lack of functional responses at mutants G96A, G250A and G301A resulted from reduced receptor expression, the level of total (western blot using whole cell lysate) and cell surface (western blot to detect cell surface proteins using sulfo-NHS-LC-biotin) $P2X_1$ receptors for wild-type and non-functional mutants (Figure 3.8c) was determined. Wild-type $P2X_1$ receptors were detected in both the total and surface cell lysates as a single band of ~55kDa. This molecular weight corresponds to the fully glycosylated receptor subunit (Valera *et al.*, 1995). All three non-functional mutants (G96A, G250A and G301A) were detected in the total cell lysates, although at reduced levels compared to the wild-type receptor, and were below the level of detection at the cell surface. These results show that alanine cannot substitute for the conserved glycine residues at positions G96, G250 and G301 and produce normal levels of channel expression.

3.2.3 Rescue of mutant receptor functionality

The lack of function and surface expression observed for the three mutants G96A, G250A and G301A suggests that the glycine residue at these positions is important. It may play an important structural role in the receptor enabling it to function normally or it may effect the trafficking of the receptor to the cell surface. For example, the receptor may be recognised as 'faulty' as a result of the mutation and subsequently not trafficked to the cell surface, resulting in a reduction in cell surface expression and a lack of response to ATP is recorded. To test whether the unique structural properties of glycine were essential at these positions, additional point mutations of glycine to either proline or cysteine were generated. Proline was chosen to replace glycine as the lack of a hydrogen bond donor makes proline a classical breaker of α-helical structure, similar to the role glycine can play terminating α-helices. Cysteine substitution has been widely used in scanning mutagenesis studies and is well tolerated (Rassendren *et al.*, 1997a; Egan *et al.*, 1998).

At G96 substitution to proline produced channels that behaved essentially the same as wild-type $P2X_1$ receptors whilst substitution to cysteine produced a small decrease in ATP potency (~4-fold) and a modest ~3-4 fold slowing in time course of the response (Figure 3.9) (Table 3.2). At G301, substitution to proline resulted in a small decrease in ATP potency (~3-fold) and a speeding in the time course of the response. However, when substituted with cysteine the responses were indistinguishable from wild-type $P2X_1$ receptors (Figure 3.10) (Table 3.2).

These results suggest that the unique structural properties of glycine at positions G96 and G301 are not required for normal channel function.

Substitution of G250 to proline or cysteine failed to produce functional channels (Table 3.2). This suggests that in this position, a glycine residue is essential for normal channel function. In an attempt to try and rescue channel function we produced a further series of point mutations at position 250. The substituted residues were chosen to replace G250 based on the classification of their side chains (Figure 3.11a); aspartate (negatively charged), lysine (positively charged), phenylalanine (aromatic) and isoleucine (non-polar). Substitution of glycine in position 250 to each of these residues (G250D, G250K, G250F and G250I) failed to produce functional channels.

Using the BLOSUM62 substitution matrix (Henikoff & Henikoff, 1992) two more polar residues asparagine and serine were also chosen to replace G250. Substitution matrices are used as a scoring system where simple positive and negative scores model an alignment. In such matrices scores for any pair of amino acid can be looked up. Therefore, identical residues are given higher scores than non-identical residues and 'conservative substitutions', those substitutions that are well tolerated in proteins due to the residues sharing physiochemical properties, given an intermediate score. BLOSUM substitution matrices use the BLOCKS database (Henikoff & Henikoff, 1991) which contains local multiple alignments ('blocks') for distantly related sequences to derive substitution scores for amino acids (Henikoff & Henikoff, 1992). BLOSUM obtains frequencies directly from relationships represented in the blocks which represent the most highly conserved regions in proteins, with no regard for evolutionary distance. There is a numbered series of BLOSUM matrices, for example BLOSUM62 where sequences have at least 62% identity are merged into a single sequence. This means that substitution scores are heavily influenced by sequences that are more divergent than this cutoff (Baxevanis & Ouellette, 2001). It was by using the BLOSUM62 substitution matrix that two more mutations, G250N and G250S were generated.

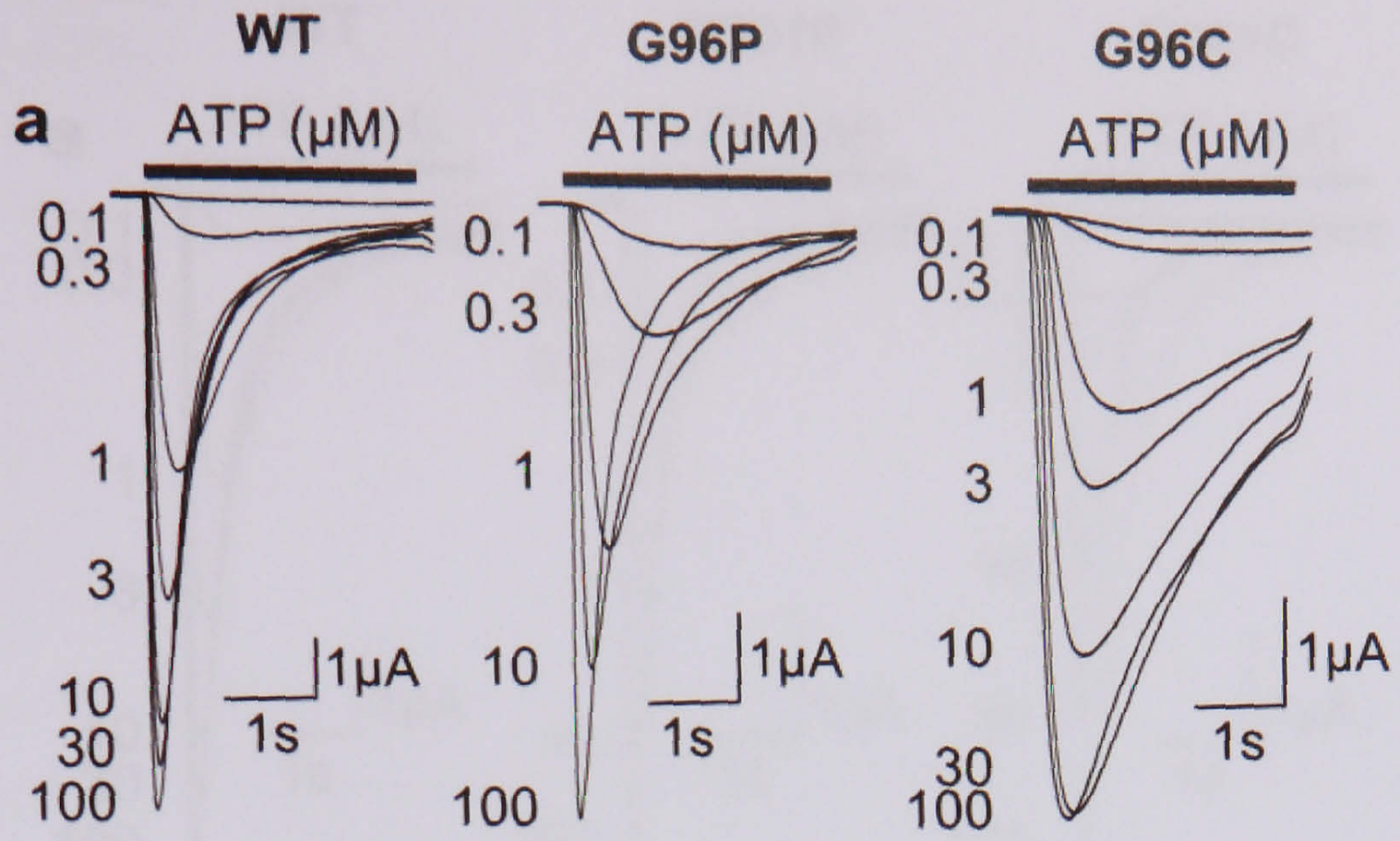


Figure 3.9 Effect of mutating glycine-96 to proline and cysteine on ATP potency. $P2X_1$ WT or mutant receptors were expressed in oocytes and ATP potency tested using a two-electrode voltage clamp. **a**, currents recorded in response to ATP at $P2X_1$ WT and mutants G96P and G96C. **b**, concentration response curves for $P2X_1$ WT and mutants G96A, G96P and G96C. Mean currents were normalized to the maximum response (n = 5-9). Holding membrane potential was –60 mV.

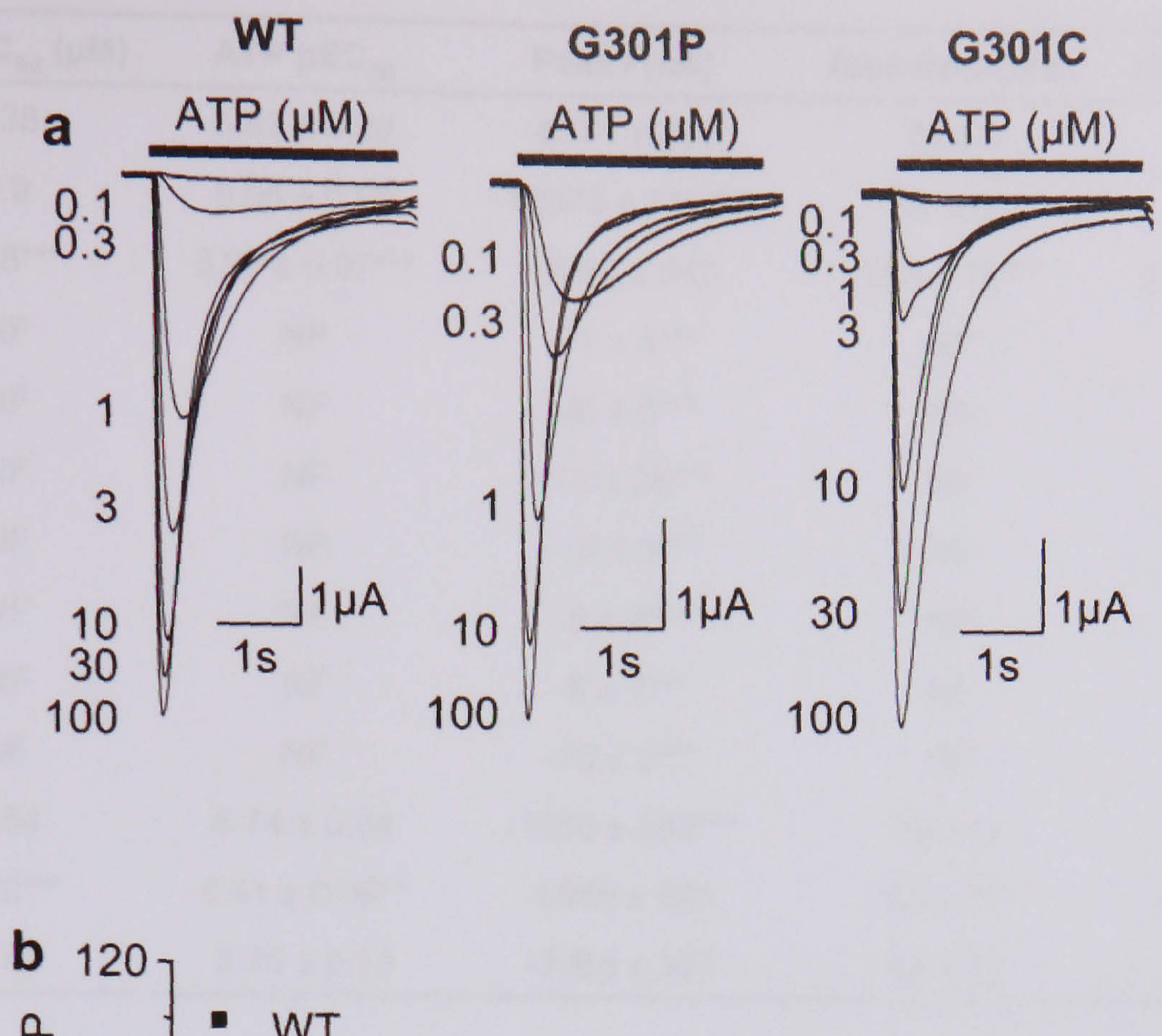


Figure 3.10 Effect of mutating glycine-301 to proline and cysteine on ATP potency. $P2X_1$ WT or mutant receptors were expressed in oocytes and ATP potency tested using a two-electrode voltage clamp. **a**, currents recorded in response to ATP at $P2X_1$ WT and mutants G301P and G301C. **b**, concentration response curves for $P2X_1$ WT and glycine mutants G301A, G301P and G301C. Mean currents were normalized to the maximum response (n = 4-9). Holding membrane potential was –60 mV.

Mutant	ATP EC_{50} (μM)	ATP pEC_{50}	Peak I (nA)	Rise-time (ms)	Half-time (ms)
WT	1.38	5.93 ± 0.09	-8571 ± 595	70 ± 6	577 ± 61
G96P	0.9	6.08 ± 0.08	-12573 ± 1314**	41 ± 5*	484 ± 117
G96C	5.68***	5.27 ± 0.07***	-7084 ± 563	255 ± 10***	2132 ± 164***
G250P	NF	NF	-17 ± 4***	NF	NF
G250C	NF	NF	-38 ± 6***	NF	NF
G250D	NF	NF	- 70 ± 26***	NF	NF
G250F	NF	NF	-10 ± 3***	NF	NF
G250I	NF	NF	-8 ± 2***	NF	NF
G250K	NF	NF	-6 ± 1***	NF	NF
G250N	NF	NF	-16 ± 3***	NF	NF
G250S	1.84	5.74 ± 0.04	-1988 ± 302***	64 ± 4	280 ± 14**
G301P	4.22***	5.41 ± 0.09**	-6986 ± 921	32 ± 3***	178 ± 16***
G301C	1.95	5.76 ± 0.13	-7008 ± 357	61 ± 11	514 ± 62

Table 3.2 Summary of data for ATP action at $P2X_1$ WT and glycine rescue mutants. EC_{50} data are representative of data collected from four to nine oocytes. pEC_{50} is –log10 of the EC_{50}. Peak data was collected on first application of a maximal ATP concentration (n = 5-17). Rise-time represents the time from 10 to 90% peak current after an application of a maximal stimulating concentration of ATP (n = 6-11). Half-time represents the time from peak to 50% decay of the current after an application of a maximal stimulating concentration of ATP (n=6-11). Values shown are representative of the mean ± S.E. of the mean.

* $P < 0.05$, different from WT as measured by Student's t test.

** $P < 0.01$, different from WT as measured by Student's t test.

*** $P < 0.001$, different from WT as measured by Student's t test.

NF, non-functional receptor, therefore data not collected

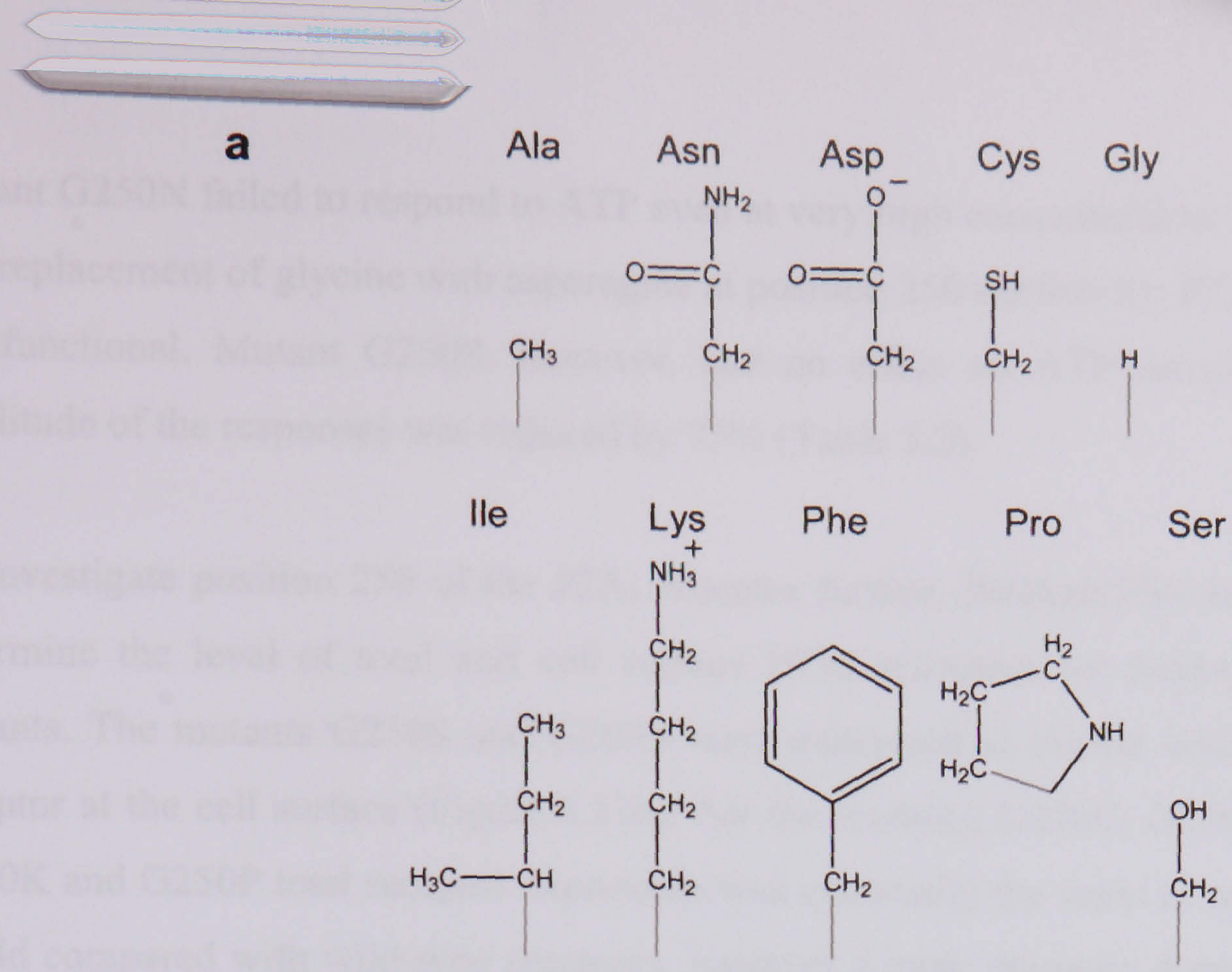


Figure 3.11 Effect of substitution of G250 with a range of amino acids. a, side chains of amino acid substitutions introduced at position 250 of the $P2X_1$ receptor. **b**, Western blot analysis of total and cell surface-expressed $P2X_1$ receptor protein shows that with the exception of G250A the mutant receptors show high levels of total expression. G2250N and G250S were expressed at levels equivalent to WT. At the cell surface the remainder of mutants were expressed at levels ~50% of the WT receptor except for G250P and G250A that were on the limit of detection, indicating that G250P and G250A are inefficiently trafficked. **c**, concentration response curves for $P2X_1$ WT and mutant G250S (n=4).

Mutant G250N failed to respond to ATP even at very high concentrations (10mM) suggesting that replacement of glycine with asparagine at position 250 renders the $P2X_1$ receptor channel non-functional. Mutant G250S, however, had no effect on ATP sensitivity but the peak amplitude of the responses was reduced by 75% (Table 3.2).

To investigate position 250 of the $P2X_1$ receptor further, Western blot analysis was used to determine the level of total and cell surface $P2X_1$ receptors for wild-type and the G250 mutants. The mutants G250S and G250N were expressed at similar levels to the wild-type receptor at the cell surface (Figure 3.11b). For the mutants, G250C, G250D, G250F, G250I, G250K and G250P total receptor expression was essentially the same or reduced by less than 2-fold compared with wild-type receptors, however a more dramatic decrease in cell surface receptor expression was seen (Figure 3.11b). Previous studies have shown that functional responses can be recorded when $P2X_1$ receptor levels are on the limit of detection, considerably lower than those reported in the present study, indicating that the lack of function of these mutants does not result only from decreased surface expression.

3.2.4 Use of partial agonists to investigate mutant G71A

A partial agonist such as, Ap_5A (P(1),P(5)-di(adenosine 5')-pentaphosphate) or BzATP (2',3'-O-(4-benzoyl)-ATP), activates the $P2X_1$ receptor but can only produce a submaximal response unlike a full agonist which can produce a maximal response (the largest response that can be produced by that drug in high concentration). Such compounds have been used to investigate the reason why a decrease in ATP potency is seen for some mutant $P2X_1$ receptors. When an agonist binds to a receptor it triggers the gating process. This is the process whereby an ion channel changes state, for example from a closed state to an open state, i.e. the 'gate' is opened and the ions flow through the channel. The gate can refer to a structural change or property of the channel concerned with the transition between open and closed states. If a change in ATP potency is observed for a mutant it can be due to a change in binding affinity at the receptor and/or a change in the gating process by which the channel opens.

The largest decrease in ATP potency (~10-fold) was recorded for mutant G71A. Partial agonists were used to investigate why a decrease in ATP potency was seen for this mutant in particular. At wild-type human $P2X_1$ receptors, BzATP and Ap_5A (100μM) (are partial agonists (pEC_{50} 6.30 ± 0.04 and 6.46 ± 0.07 respectively) with efficacies of 0.52 ± 0.05 and

0.37 ± 0.06, respectively (n=5) (ATP is a full agonist with an efficacy of 1). Substitution of G71 with alanine resulted in a ~10-fold decrease in ATP potency at the $P2X_1$ receptor (Figure 3.12). It has already been shown that mutant $P2X_1$ receptors that decrease the ATP potency can also reduce the efficacy of the partial agonists BzATP and Ap_5A (Roberts & Evans, 2004). The efficacy of BzATP and Ap_5A (both 100μM) was reduced by ~5-6 fold at mutant G71A compared with wild-type channels (0.11 ± 0.01 and 0.06 ± 0.01 respectively) (n = 6) (Table 3.3). This suggests that G71A may be affecting the binding of agonist to the receptor or an associated conformational change. This would support previous work that suggests residues in this region of the extracellular domain, K68 to G71, are involved in agonist action at the $P2X_1$ receptor.

Mutant	BzATP (ε)	Ap5A (ε)
WT	0.52 ± 0.05	0.37 ± 0.06
G71A	0.11 ± 0.01	0.06 ± 0.01

Table 3.3 Partial agonist efficacies. Relative efficacies of partial agonists BzATP and Ap_5A at wild-type $P2X_1$ and mutant G71A receptors. Relative efficacy (ε) was calculated using the equation: ε = Imax(partial agonist)/Imax(ATP). Values represent mean ± SEM.

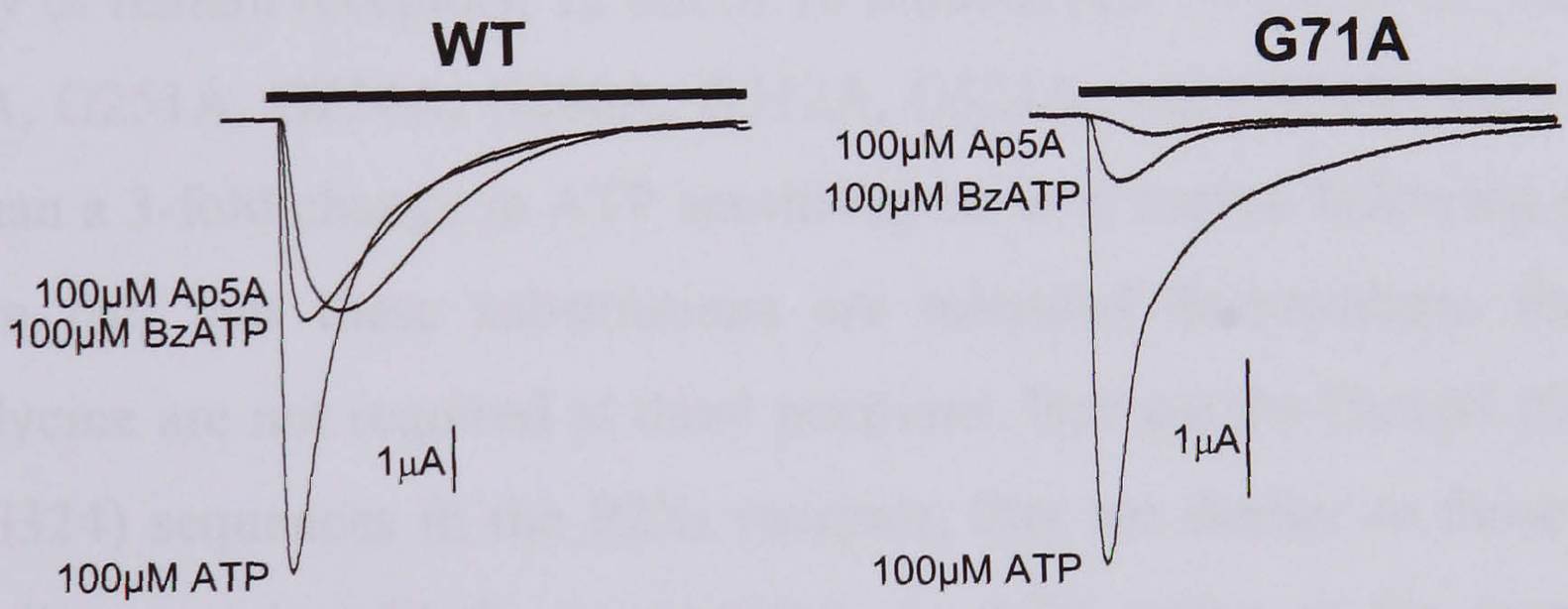


Figure 3.12 Effect of partial agonists at G71A. Typical trace data collected for wild-type $P2X_1$ in response to 100μM ATP, 100μM BzATP and 100μM Ap5A compared with G71A. Black bar represents three second application of agonist.

3.3 Discussion

Conserved glycine residues across a family are strong indicators of turns, secondary structure breakers or regions of flexibility in proteins and are therefore structurally important. Glycine residues are also key components of some ATP binding motifs (Walker) (Okkeri *et al.*, 2004). For these reasons it was important to investigate the role of conserved glycine residues in the $P2X_1$ receptor in order to advance our understanding of the structure of the receptor and give further insight to structural models. The role of 16 conserved glycine residues in the extracellular domain of the human $P2X_1$ receptor were investigated using alanine-replacement mutagenesis which would remove the flexibility glycine would have introduced yet minimize the disruption to the receptor as they are similar in size. This mutagenesis study investigating the role of conserved glycine residues in the extracellular loop of the human $P2X_1$ receptor has been published (Digby *et al.*, 2005). The majority of mutations had little or no effect on channel function. This is a common result for mutagenesis studies of the extracellular domain $hP2X_1$ receptor and it is interesting that of all the conserved mutated residues in the extracellular loop of the $P2X_1$ receptor, ~79% (63/80) of mutations are tolerated. This suggests that those positions where mutations have affected channel properties are functionally important.

For the majority of mutant receptors, 12 out of 16 mutants (G115A, G135A, G143A, G147A, G233A, G240A, G251A, G254A, G288A, G312A, G321A, and G324A) there was either no effect or less than a 3-fold change in ATP sensitivity or time course following point mutation to alanine. The fact that these substitutions are tolerated demonstrates that the unique properties of glycine are not required at these positions. Nor are the GxxxG (G143/G147) or GxxG (G321/G324) sequences in the $P2X_1$ receptor, that are similar to those described for other ATP binding proteins, likely to contribute to ATP action at the receptor. As with previous mutagenesis studies, mutants that have little or no effect on channel function provide important information for refining models of the $P2X_1$ receptor and are critical to determining ATP action at $P2X_1$ receptor.

The largest effect on ATP potency at the $P2X_1$ receptor was recorded for the G71A mutant that had a ~6 fold decrease in potency (EC_{50} 8.31μM compared to EC_{50} 1.38μM wild-type) but no effect on the time course of the response, suggesting it may be involved in agonist action at the receptor. Partial agonists can result in different ligand-binding conformational states that determine the open probability of the channel. Changes in the properties of partial

agonists can therefore help to interpret the effects of agonist binding and subsequent gating of mutant channels (Roberts & Evans, 2004). The efficacy of both partial agonists BzATP and Ap_5A (both 100μM) was reduced by ~5-6fold at the mutant G71A compared with wild-type channels (0.52 ± 0.05 to 0.11 ± 0.01 and 0.37 ± 0.06 to 0.06 ± 0.01 for BzATP and Ap_5A at G71A and wild-type channels respectively, *n*=5). This reduction in efficacy suggests that conformational changes associated with agonist action have been affected. Previous mutagenesis studies have identified the nearby positively charged lysine residues (K68 and K70 $P2X_1$ receptor numbering) to be involved in ATP binding and activation of the P2X receptors (Ennion *et al.*, 2000; Jiang *et al.*, 2000). The conservation of these residues (residues 68-71) throughout the P2X receptor family, suggests that in this region the KxKG motif plays a role in mediating the properties of ATP.

Residues G96 and G301 were sensitive to the nature of the amino acid substitution; alanine mutants produced non-functional channels. Western blot analysis was used to determine whether this was as a result of reduced receptor expression and/or a lack of ability of the channel to open and pass current in response to the binding of ATP. Both mutants (G96A and G301A) were detected in the total cell lysates at a markedly reduced level and were below the level of detection at the cell surface. The $P2X_1$ receptor protein seen in the total cellular lysates corresponds to the fully glycosylated (~55kDa) receptor subunit. The presence of sugars on proteins can be used as markers to track the movement of proteins. Newly made proteins undergo N-linked glycosylation in the ER and then enter the Golgi complex. The presence of fully glycosylated $P2X_1$ receptor protein detected in the total cell lysates indicates that the mutant channels have passed through the ER and Golgi. It is clear that these alanine substitutions at positions 96 and 301 had a marked effect on the normal processing of the $P2X_1$ receptor possibly by inefficient trafficking of the receptor to the cell membrane or by a local misfolding of the protein effecting ATP-binding. Roberts *et al* 2005 saw a similar result in the $P2X_1$ receptor for mutant P272A but was able to 'rescue' the function of the channel by replacing P272 with glycine. Proline has the ability to form *cis* peptide bonds and for this reason is structurally important, acting as a breaker of α-helical structure. It does not have the flexible ability of glycine and due to the nature of its ring structure (described in the introduction to this chapter) it has a unique conformational rigidity compared to other amino acids. Substitution of G96 or G301 with either proline or cysteine produced functional channels that were either identical to wild-type (G96P and G301C) or resulted in a small decrease in ATP potency (G96C and G301P). These results demonstrate that the unique

properties of glycine residues at positions G96C and G301 are not required for normal channel function.

The mutant G250A rendered the $P2X_1$ receptor channel non-functional. Western blot analysis revealed a reduction in total $P2X_1$ receptor protein and levels of the receptor were below the limit of detection at the cell surface. These results suggest that the lack of function as a result of mutant G250A most likely results from a failure in normal processing of the receptor. In contrast to positions G96 and G301, functional channels could not be produced by substitution of G250 with proline or cysteine. A range of amino acids were then selected to substitute for glycine at position 250 based on their amino acid side chain classification (aromatic phenylalanine, negatively charged aspartate, positively charged lysine and non-polar isolecucine). It was also not possible to rescue the function of the channel by substituting G250 with any of these four residues.

A BLOSUM62 substitution scoring matrix was used to determine the likelihood of glycine mutations and to suggest any residues that might be expected to substitute for glycine in position 250. All amino acid substitutions of glycine score less than or equal to zero indicating that all substitutions are non-conservative; substitution of amino acids which do not share similar physiochemical properties (negative scores). The highest scoring residues were asparagine and serine (both scored zero). Asparagine and serine are similar in size to glycine and are both polar residues. Mutants G250N and G250S were generated to test whether the function of the $P2X_1$ receptor channel could be rescued by either of these substitutions. Substitution of G250 with asparagine resulted in a non-functional channel but interestingly mutant G250S was able to rescue the function of the channel. The expression levels of mutant G250 $P2X_1$ receptors were dependent on the amino acid substitution, with alanine and proline substituted receptors below the limit of detection at the cell surface. However, for the majority of substitutions $P2X_1$ receptor protein was detected at the cell surface suggesting that the mutant may affect a local protein conformation effecting ATP-binding. Therefore it was concluded that the processing and trafficking of these mutants is dependent on the amino acid substitution. These results suggest that the unique structural properties of glycine or a small polar residue are required at position 250 for normal receptor function and that this should be incorporated into structural models of the human $P2X_1$ receptor.

A similar P2X receptor mutagenesis study of this region of conserved glycine residues (G240-254) was published for the $P2X_2$ receptor (Nakazawa & Ohno, 1999). Neighboring residues

G247 and G248 (corresponding to G250 and G251 in the $P2X_1$ receptor) were found to be essential for $P2X_2$ receptor/channel function. Mutation of G247 to alanine resulted in a loss of responsiveness to ATP and sensitivity was reduced in the mutant G248A channel. The extensive substitutions generated for the conserved G250 residue of the $P2X_1$ receptor (G247 $P2X_2$ receptor numbering) were not generated for the $P2X_2$ receptor and therefore it is not possible to conclude that a glycine residue only is necessary in this position for normal channel function. Mutants G237A and G251A (G240 and G254 respectively in the $P2X_1$ receptor) produced channels that behaved similarly to wild-type which is consistent with the results found in this study. Nakazawa suggests that the glycine-rich region from G240 to G254 ($P2X_1$ receptor numbering) may contribute to an ATP-binding motif.

Glycine-rich motifs that form part of the ATP-binding site have been identified in protein kinases (Okkeri 2004). The glycine-rich loop with consensus sequence GxGxxG is conserved in over 95% of known human protein kinases and other nucleotide-binding proteins. Mutation of any of the conserved glycines has been shown to result in enzymatic defects (Spitaler 2000). The GGxxG (250-254, $P2X_1$ receptor numbering) shows some similarity to the kinase/nucleotide-binding glycine-rich loop and this may suggest that the lack of function for the majority of G250 mutants results from disruption of this motif. However, mutation of glycine residues adjacent to G250 in the $P2X_1$ receptor has either no effect on ATP potency (G251A) or results in a ~3-fold increase in ATP potency (G254A) suggesting that it is not the conserved GGxxG motif in the P2X receptors but the conserved G250 residue ($P2X_1$ receptor numbering) alone that is important.

In addition to a role in glycine-rich ATP-binding motifs, glycine residues can give unique structural properties to proteins. The flexible nature of glycine arises from its lack of side chain and gives glycine the ability to exist in tight turns, allow conformational change within proteins and conserved glycine residues can be strong indicators of loops or secondary structure breakers. The study of α-helices revealed that each position in this secondary structural element has its own unique characteristics such that natural helix sequences optimize by minimizing the occurrence of residues that are to be avoided at a given position (Sandeep Kumar, 1998). α-helices are stabilized by hydrogen bonds between the NH and CO groups of the main chain; more specifically the CO group of each amino acid forms a hydrogen bond with the NH group of the amino acid which is four residues (1 complete turn of the helix) ahead in the protein sequence. At a helix-end no next turn exists to provide hydrogen bond partners and so the helix is terminated. About one third of all α-helices

terminate with a glycine at the C-terminus (Aurora *et al.*, 1994) and belong to two glycine-based motifs that are prevalent at the C-termini of α-helices; Schellman motif and the α_L motif (Aurora *et al.*, 1994). Prediction of the secondary structure of the $P2X_1$ receptor (McGuffin *et al.*, 2000) suggests that G250 is located at the C-terminal end of an α-helix (residues 243-249, Figure 3.13). The residues adjacent to G250 in the sequence of the $P2X_1$ receptor correspond to the Schellman C-terminal α-helix terminating motif (Schellman, 1980; Aurora *et al.*, 1994). In this motif glycine adopts a left-handed α-helical conformation; a conformation which asparagine (Richardson, 1981) and (to a lesser extent) aspartate (Srinivasan *et al.*, 1994), but not serine can occupy. However, normal sensitivity to ATP was recorded for the G250S mutant. This suggests that rather than conforming to the Schellman motif, either a glycine or a small polar residue (serine) is required in position 250 to allow for normal function of the $P2X_1$ receptor.

3.4 Conclusion

In conclusion this study has identified G250 as a key residue for the function of $P2X_1$ receptors possibly arising because of its predicted location at the C-terminal end of an α-helix. In addition, G71 may be involved with ligand recognition along with other residues in this region, K68 and K70 also thought to have a role in ATP-binding. These results support previous mutagenesis studies and add more detail to the structural model being developed based on mutagenesis studies.

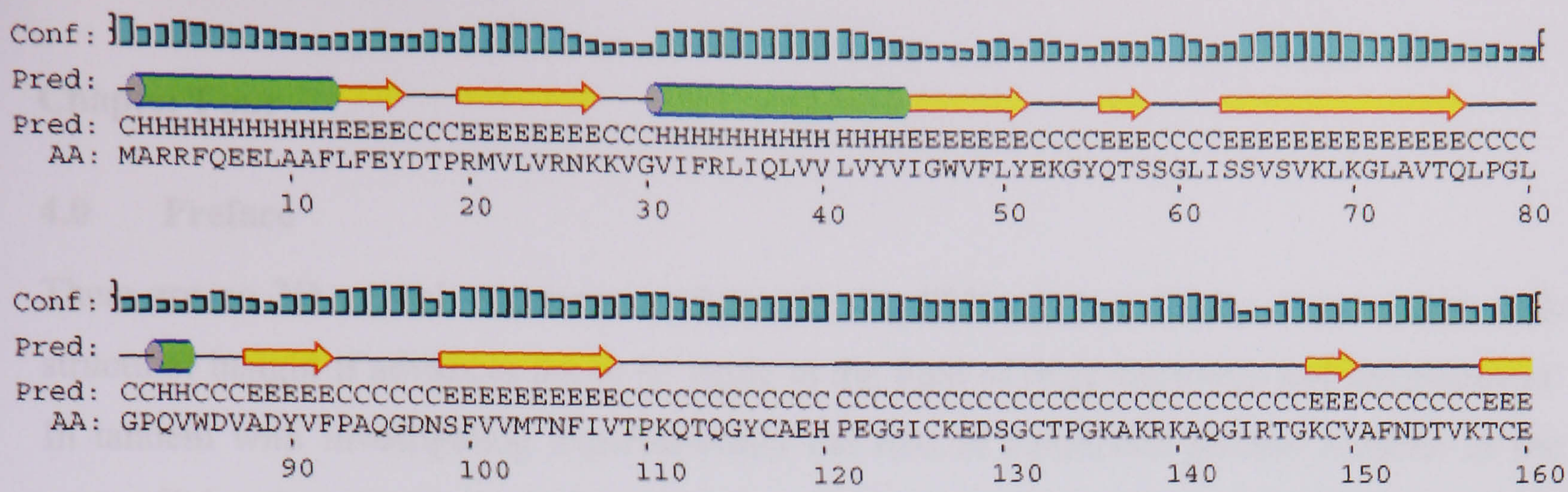


Figure 3.13 PSIPred secondary structure prediction hP2X$_1$ receptor. Alpha helices are depicted by green cylinders, β-strands, yellow arrows. A measure of the confidence of the prediction is shown by blue bars.

Chapter Four

4.0 Preface

There are no 3D crystal structures available for the P2X receptor family. It is vital to seek structural insight if advances are to be made in the field of drug discovery and development. In tandem with investigating experimentally the role of conserved glycine residues in the extracellular domain of the $hP2X_1$ receptor (Chapter 3) further structural information was sought through bioinformatics techniques in this chapter. In particular, understanding the structure of individual domains of the P2X receptor subunit and their contribution to the ligand binding site.

4.1 Introduction

The structure of a protein is directly related to its function so to gain a better understanding of the function of a protein, discovering its structure is crucial. Many structures are available via the Protein Data Bank (PDB) (Berman *et al.*, 2000) but to date, there are no 3D crystal structures available for the P2X receptor family. In this case the amino acid sequence of the P2X receptor is the key to structural insight. However, the folding of amino acid sequences into correct 3D structures is currently unsolved; the so-called 'protein-folding problem'. Therefore, direct sequence analysis and protein structure prediction methods make useful tools in extracting structural information from amino acid sequences. They are effective because although the number of individual proteins is very large there is thought to be a limited set of tertiary structural motifs or folds to which most proteins belong.

In this chapter the $hP2X_1$ receptor sequence will be used to generate 3D structural models based on a model building procedure (Figure 4.1) that can be simplified into 4 main steps which are;

1) Determine the transmembrane topology of the P2X receptor family
2) Divide the amino acid sequence into smaller domains based on step 1
3) Generate 3D structural models of each individual domain
4) Combine the 3D models of the different domains to form a single structure

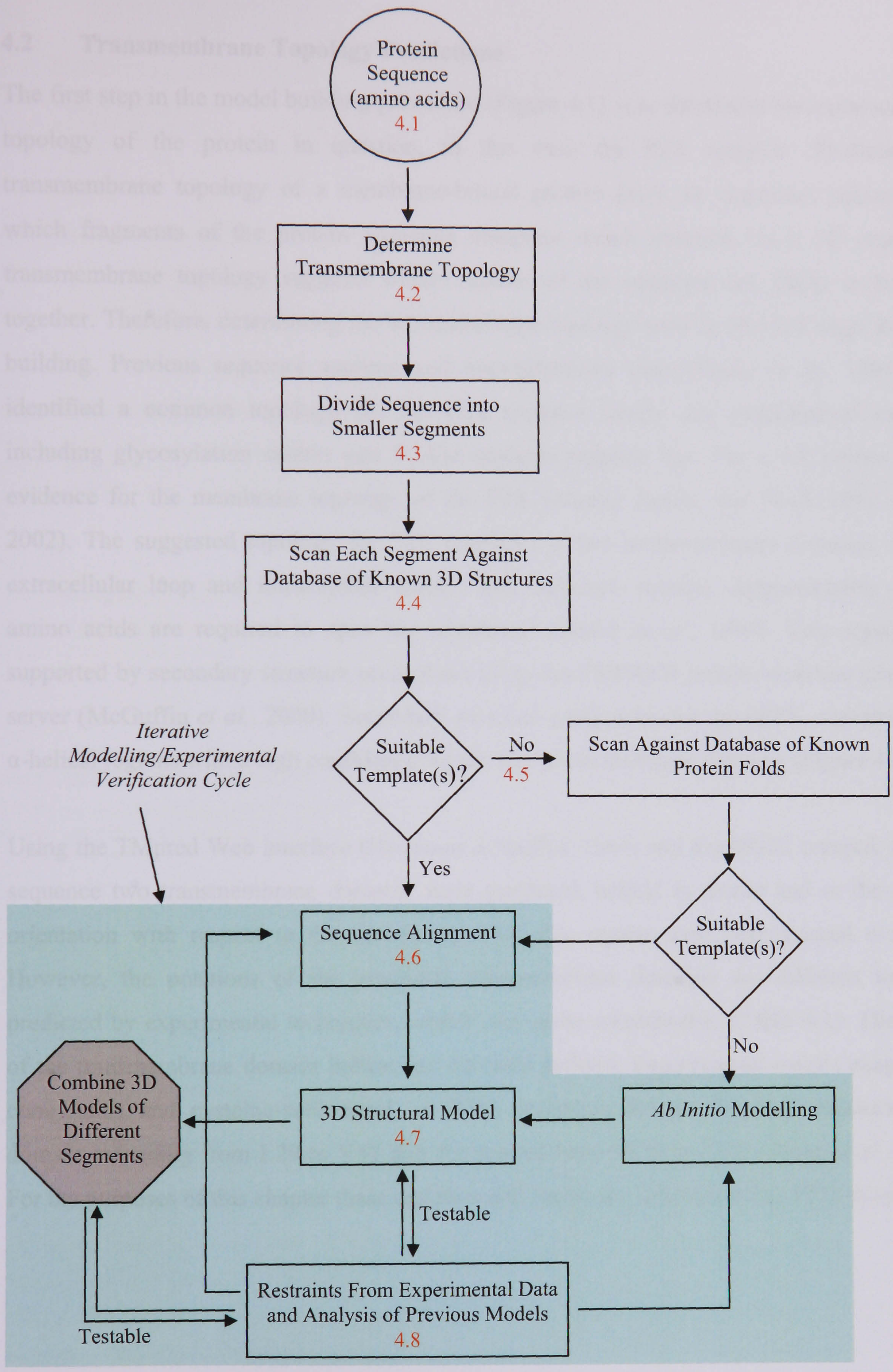


Figure 4.1 Overview of Model Building Procedure. Red numbers refer to relevant chapter headings

4.2 Transmembrane Topology Predictions

The first step in the model building procedure (Figure 4.1) is to determine the transmembrane topology of the protein in question, in this case the P2X receptor. Predicting the transmembrane topology of a membrane-bound protein gives an important indication of which fragments of the protein sequence comprise which domains. In a 3D model the transmembrane topology suggests which regions of the sequence are likely to be close together. Therefore, determining the transmembrane topology must be the first stage in model building. Previous sequence analysis and hydrophobicity plots (Brake *et al.*, 1994) have identified a common topology for the P2X receptor family and experimental evidence including glycosylation studies and SCAM analysis supports this. For a full review of the evidence for the membrane topology of the P2X receptor family, see North 2002 (North, 2002). The suggested topology for P2X receptors is two transmembrane domains, a large extracellular loop and intracellular amino- and carboxyl- termini. Approximately twenty amino acids are required to span the membrane (Brake *et al.*, 1994). This topology is supported by secondary structure predictions using the PSIPRED protein structure prediction server (McGuffin *et al.*, 2000). Secondary structure predictions for the $hP2X_1$ receptor show α-helical regions with a high confidence for the two transmembrane domains (Figure 4.2).

Using the TMpred Web interface (Hofmann & Stoffel, 1993) and the $hP2X_1$ receptor protein sequence two transmembrane domains were predicted, helical in nature and in the correct orientation with respect to the membrane. This also agrees with experimental evidence. However, the positions of the predicted transmembrane domains are different to those predicted by experimental techniques, which vary quite considerably (Table 4.1). The limits of the transmembrane domain helices are not well defined. Experimental results using MTS compounds and cysteine-substituted rat $P2X_2$ receptors defines the first transmembrane domain extending from L29 to V51 and the second from P329 to L353 (Jiang *et al.*, 2001). For the purposes of this chapter these residues will define the structure of the P2X receptor.

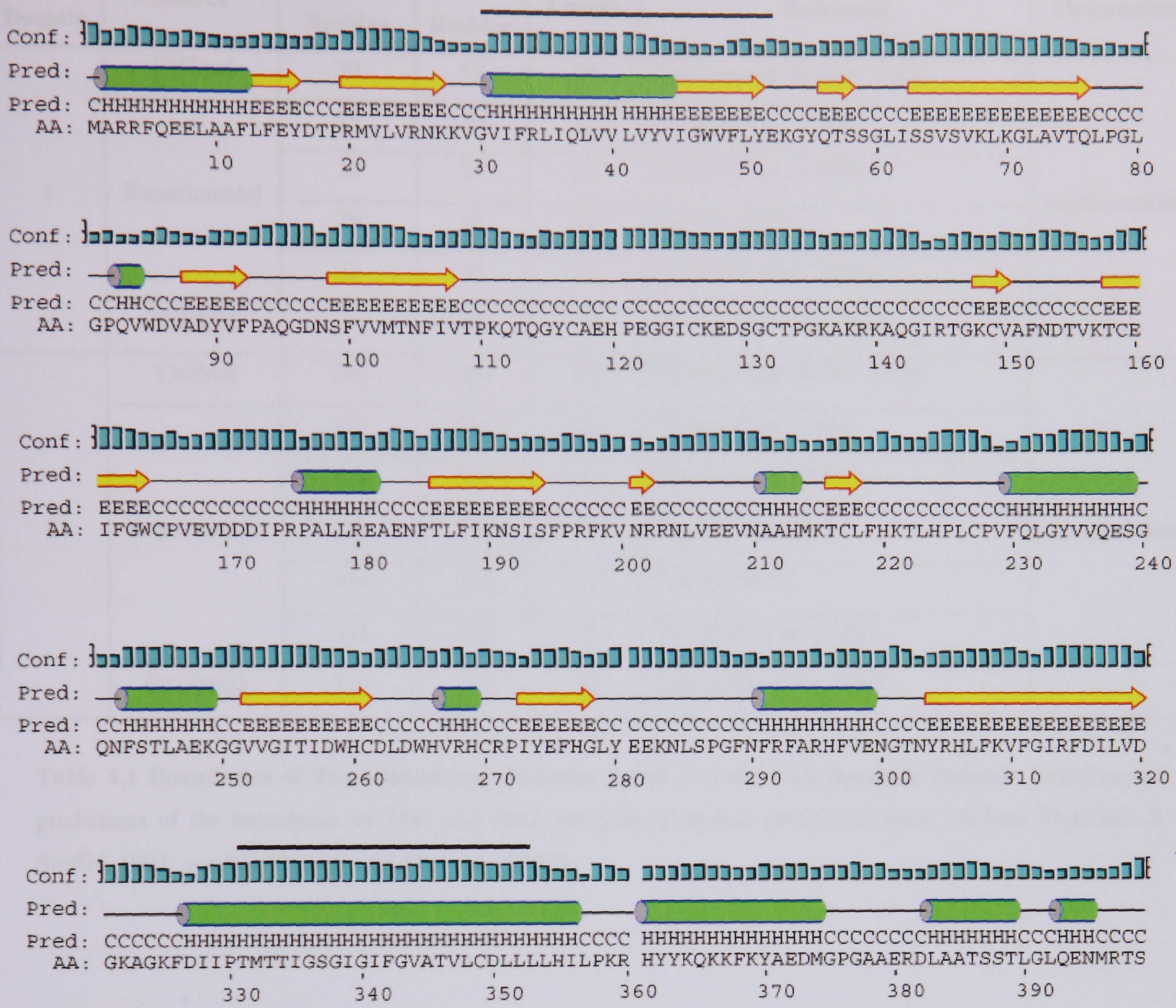


Figure 4.2 PSIPRED secondary structure predictions for the hP2X$_1$ receptor protein sequence. The topology of the P2X receptor family has been predicted using experimental techniques. This includes two transmembrane domains (residues 29-50 and 331-356) thought to be α-helical in nature. Similar results were seen using the PSIPRED secondary structure prediction server. Two α-helical regions were predicted for transmembrane domains, TM1 and TM2, represented by black bars.

TM Domain	Source	From Residue	To Residue	Length	Reference	Orientation
1	TMPred	39	58	20	(Hofmann & Stoffel, 1993)	inside-outside
	Experimental	29	50	22	(Valera *et al.*, 1994)	
		33	54	22	(Torres *et al.*, 1998b)	
		29	50	22	(North, 2002)	
		31	51	21	(Newbolt *et al.*, 1998)	
	PSIPRED	31	44	14	(McGuffin *et al.*, 2000)	
2	TMPred	346	364	19	(Hofmann & Stoffel, 1993)	outside-inside
	Experimental	339	358	20	(Valera *et al.*, 1994)	
		328	353	26	(Torres *et al.*, 1998b)	
		331	356	26	(Egan *et al.*, 1998)	
		334	354	21	(North, 2002)	
		331	353	23	(Newbolt *et al.*, 1998)	
	PSIPRED	327	356	30	(McGuffin *et al.*, 2000)	

Table 4.1 Boundaries of Transmembrane Domains 1 and 2 of the P2X Receptor Subunit. Experimental predictions of the boundaries of TM1 and TM2 are given alongside predictions from TMPred (Hofmann & Stoffel, 1993) and PSIPRED (McGuffin *et al.*, 2000).

4.2.1 Topology Summary

The topology of the P2X receptor subunit has been determined using experimental predictions from published P2X receptor studies and by sequence analysis using bioinformatic techniques to reveal a simple structure consisting of (Figure 4.3):

- Two transmembrane domains (pink)
- Intracellular amino- and carboxyl- termini (blue)
- A large extracellular loop (green)

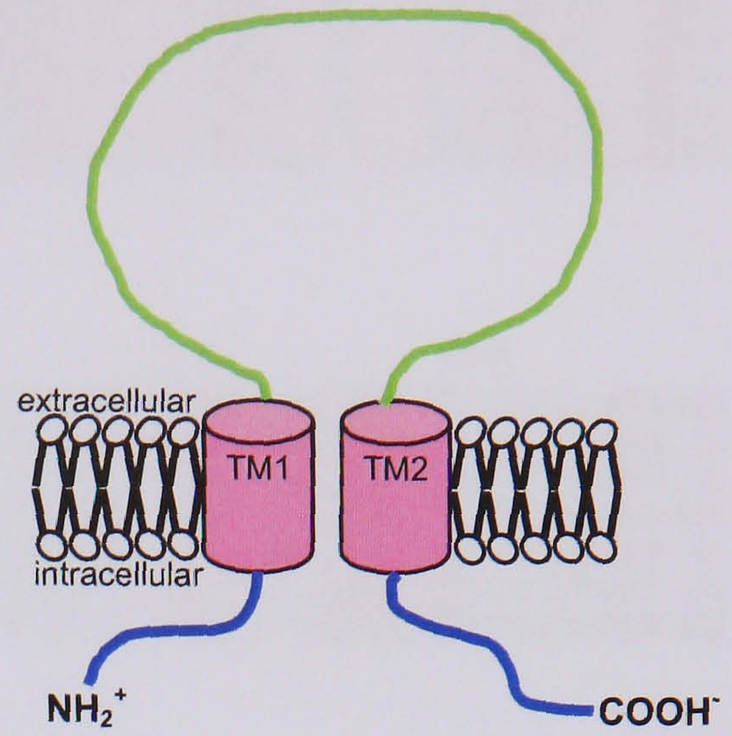


Figure 4.3 Structure of a P2X Receptor. Two transmembrane domains (pink), intracellular amino- and carboxyl- termini (blue) and a large extracellular loop (green).

Using this structural information, smaller segments need to be defined if possible, following the next stage of the model building procedure (Figure 4.1).

4.3 Identification of smaller segments

This chapter will concentrate on modelling the extracellular ligand binding domain of the $hP2X_1$ receptor, in an attempt to locate the ATP binding domain and define its structure. The boundaries of the extracellular domain are loosely defined but the segment is well conserved across the P2X receptor family indicating its functional importance (Figure 4.4). Due to the size of the extracellular domain (~288 amino acids) it is likely that it contains more than one domain. Ten conserved cysteine residues are found in the extracellular segment and they form five disulphide bonds (Ennion & Evans, 2002a). Six of the ten conserved cysteines form a cysteine-rich segment which can be thought of as a fourth segment. The predicted pattern of disulphide bonding in the cysteine-rich segment is similar to that seen in snake toxins and will be modelled separately to the rest of the extracellular segment (see section 4.9). It is important to note that models of the extracellular segment of the $hP2X_1$ receptor can give further insight into the structure and the function of the P2X receptor family and can be used to direct further experimental work. However, such models are no substitute for an experimentally derived 3D structure which we eagerly await for the P2X receptor family.

Figure 4.4 Multiple sequence alignment of the extracellular loop of the hP2X receptor family generated by CLUSTALW (Thompson *et al.*, 1994) and reproduced by ESPript (Gouet *et al.*, 2003). Numbering refers to the first residue belonging to the $hP2X_1$ receptor sequence. A percentage of equivalent residues were selected for the similarity score with a default similarity global score of 0.7. Therefore a percentage of equivalent residues is calculated per column considering physicochemical properties in the list of one-letter amino acids below. All identical residues are boxed in red. A column is framed in blue, with red letters on a white background if more than 70% of its residues are similar according to physicochemical properties **HKR** are polar positive; **DE** are polar negative; **STNQ** are polar neutral; **AVLIM** are non polar aliphatic; **FYW** are non polar aromatic; **PG** and **C**. The black bar indicates the cysteine-rich segment of the extracellular loop.

4.3.1 Summary of Defined Segments

The large extracellular segment (~288 amino acids) has been divided into two segments based on the identification of six fully conserved cysteine residues (Figure 4.5). A second domain has been defined and called the cysteine-rich domain (residues 115 – 165, $P2X_1$ receptor numbering). This domain will be modelled separately to the rest of the extracellular ligand-binding segment (green) (see section 4.9).

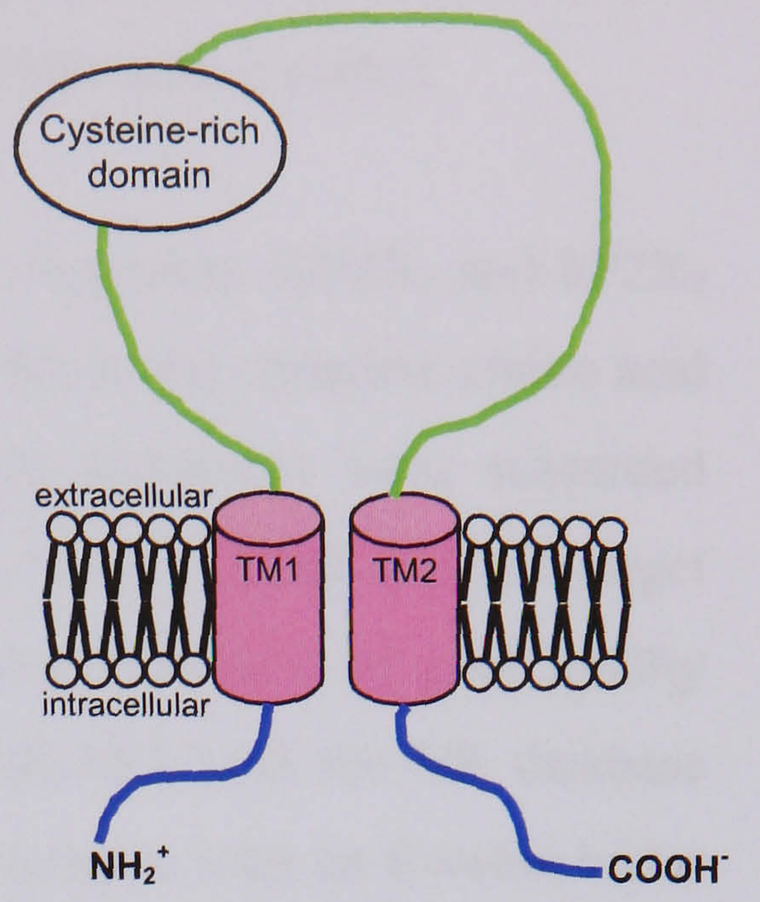


Figure 4.5 Structure of a P2X Receptor. Two transmembrane domains (pink), intracellular amino- and carboxyl- termini (blue) and a large extracellular loop (green) including a cysteine-rich domain.

4.4 Identifying suitable 3D templates

The extracellular domain of the $hP2X_1$ receptor has been identified as a suitable segment for modelling. The next step in the model building procedure is to search for a similar protein(s) of known 3D structure (determined by X-ray crystallography or NMR spectroscopy). This was done using two tools, BLAST (Altschul *et al.*, 1990) and PSI-BLAST (Altschul *et al.*, 1997) which use E-values to reflect the statistical significance of hits (see Chapter 2).

4.4.1 Searching databases of known 3D protein structures

The amino acid sequence of the extracellular segment of the $hP2X_1$ receptor was submitted to the program BLAST (Altschul *et al.*, 1990) which scans a sequence against a database (Protein Data Bank) of protein sequences of known structure and returns the resulting hits. In addition, the $hP2X_7$ receptor sequence was submitted because it is most distantly related to the $hP2X_1$ receptor, the hypothesis being that together they would hit a wider range of sequences and increase the chances of finding suitable templates. However, no significant similarity was found to sequences other than P2X receptor related proteins. The target database for the BLAST program was set to the non-redundant (NR) database and the query sequences re-submitted. Although the NR database may return significant hits to sequences with no known structure it may find similar sequences which can be used as intermediates between the P2X receptor and homologous proteins with solved structures. Again, no significant similarity was found to proteins other than P2X receptor related sequences. In those situations where BLAST has failed to find significant hits it is advisable to use PSI-

BLAST (Altschul *et al.*, 1997), the most sensitive BLAST program which is useful for finding very distantly related proteins. Using the NR database as the target database both query sequences were submitted to PSI-BLAST. No significant hits were reported.

To ensure a distant homolog had not been missed, the query sequences ($hP2X_1$ and $hP2X_7$ receptor sequences) were split into two halves or extended to include the complete amino acid sequence for the $hP2X_1$ and $hP2X_7$ receptor. The six query sequences were submitted individually to the BLAST and PSI-BLAST programs using both the PDB and NR target databases. The majority of the returned hits had high E-values indicative of poor quality alignments and low homology. The seventh iteration of PSI-BLAST with the NR database and the complete $hP2X_7$ receptor amino acid sequence hit a sequence with an E-value better than the threshold; HUNK (Hormonally Upregulated Neu-tumor associated Kinase). HUNK encodes an 80-kDa polypeptide associated with phosphotransferase activity. It is located on the mouse chromosome 16 in a region of conserved synteny with the human chromosome 21q22. It is closely related to the SNF1 family of serine/threonine kinases (Gardner *et al.*, 2000). Serine/threonine kinases catalyse the phosphorylation of serine or threonine residues on target proteins, using ATP as a phosphate donor. The SNF1 subfamily of serine/threonine protein kinases is highly conserved and found in fungi, plants, Drosophila, *C.elegans* and mammals. Members of the SNF1 subfamily appear to be central components of kinase cascades that function as metabolic sensors in eukaryotic cells (Schmidt & McCartney, 2000). The biological significance of kinases and P2X receptors binding ATP was recognised in HUNK making it a possible template for the $hP2X_1$ receptor. However the weakness of this template is clearly illustrated in Figure 4.6. The portion of the $hP2X_7$ receptor sequence that is 'similar' to HUNK is not part of the extracellular domain. This limits the suitability of HUNK as a template for the P2X receptor family. However, it may be useful as an intermediate to find other possible templates. This will be discussed in further detail in section 4.6.

>hP2X$_7$ receptor amino acid sequence

MPACCSCSDVFQYETNKVTRIQSMNYGTIKWFFHVIIFSYVCFALV**SDKLYQRKEPVISSVHTKVKGIAEVK**
EEIVENGVKKLVHSVFDTADYTFPLQGNSFFVMTNFLKTEGQEQRLCPEYPTRRTLCSSDRGCKKGWMDPQS
KGIQTGRCVVHEGNQKTCEVSAWCPIEAVEEAPRPALLNSAENFTVLIKNNIDFPGHNYTTRNILPGLNITC
TFHKTQNPQCPIFRLGDIFRETGDNFSDVAIQGGIMGIEIYWDCNLDRWFHHCRPKYSFRRLDDKTTNVSLY
PGYNFRYAKYYKENNVEKRTLIKVFGIRFDILVFGTGGKFDIIQLVVYIGSTLSYFGLAAVFIDFLIDTYSS
NCCRSHIYPWCKCCQPCVVNEYYYRKKCESIVEPKPTLKYVSFVDESHIRMVNQQLLGRSLQDVKGQEVPRP
AMDFTDLSRLPLALHDTPPIPGQPEEIQLLRKEATPRSRDSPVWCQCGSCLPSQLPESHRCLEELCCRKKPG
ACITTSELFRKLVLSRHVLQFLLLYQEPLLALDVDSTNSRLRHCAYRCYATWRFGSQDMADFAILPSCCRWR
IRKEFPKSEGQYSGFKSPY

>HUNK amino acid sequence

MPAAAGDGLLGEPAAPGGDGGAEDTTRPAAACEGSFLPAWVSGVSRERLRDFQHHKRVGNY**LIGSRKLGEGS**
FAKVREGLHVLTGEKVAIKVIDKKRAKKDTYVTKNLRREGQIQQMIRHPNITQLLDILETENSYYLVMELCP
GGNLMHKIYEKKRLDEAEARRYIRQLISAVEHLHRAGVVHRDLKIENLLLDEDNNIKLIDFGLSNCAGILGY
SDPFSTQCGSPAYAAPELLARKKYGPKIDVWSIGVNMYAMLTGTLPFTVEPFSLRALYQKMVDKAMNPLPTQ
LSTGAVNFLRSLLEPDPVKRPNIQQALANRWLNENYTGKVPCNVTYPNRISLEDLSPSVVLHMTEKLGYKNS
DVINTVLSNRACHILAIYFLLNKKLERYLSGKSDIQDSICYKTQLYQIEKCRATKEPYEASLDTWTRDFEFH
AVQDKKPKEQEKRGDFLHRPFSKKLDKNLPSHKQPSPSLITQLQSTKALLKDRKASKSGFPDKDSFVCRNLF
RKTSDSNCVASSSMEFIPVPPPRTPRIVKKLEPHQPGPGSASILPKEEPLLLDMVRSFESVDREDHIELLSP
SHHYRILSSPVSLARRNSSERTLSQGLLSGSTSPLQTPLHSTLVSFAHEEKNSPPKEEGVCSPPPVPSNGLL
QPLGSPNCVKSRGRFPMMGIGQMLRKRHQSLQPSSERSLDASMSPLQPTAPSSLSFDMADGVKGQ

Figure 4.6 Sequence similarity of the hP2X$_7$ receptor and HUNK. The FASTA sequences for the hP2X$_7$ receptor and HUNK are shown with the region of similarity between the two sequences identified by PSI-BLAST shown in red. The protein kinase domain of HUNK and the extracellular ATP-binding domain of the hP2X$_7$ receptor are highlighted in **bold.**

4.4.2 Structural Database Searching Summary

Using two database searching tools, BLAST (Altschul *et al.*, 1990) and PSI-BLAST (Altschul *et al.*, 1997) and two databases (PDB and NR) a protein has been identified that is significantly similar to the $P2X_7$ receptor sequence; HUNK. However the structure of HUNK is unknown and the sequence similarity to the $hP2X_7$ receptor was in the C-terminal domain (a portion of which is shown in Figure 4.7, the $hP2X_7$ receptor sequence is in red and the HUNK sequence is shown in black). This does not this mean HUNK should be ignored but a different searching technique (searching databases of known protein folds), used in the following section may reveal a more suitable template.

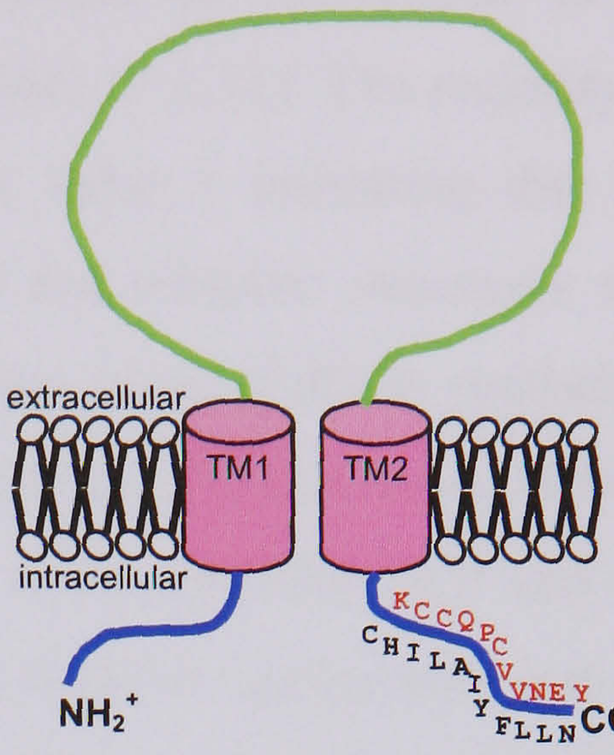


Figure 4.7 Structure of P2X Receptor. Comparison of $hP2X_7$ receptor sequence (red) and HUNK (black) in the C-terminal region.

4.5 Searching databases of known protein folds

The model building procedure (Figure 4.1) suggests that if no suitable templates are found by scanning each segment against a database of known structures then an alternative strategy should be adopted. This strategy is based on the theory that proteins that show little sequence homology can share distinct similarities in their folds. There are many programs that scan sequences against libraries of folds including the 3D-PSSM (Kelley *et al.*, 1999) and UCLA-DOE Structure Prediction Server, TOPITS (Rost, 1995) and Threader (Jones *et al.*, 1992). The query sequences $hP2X_1$ receptor and $hP2X_7$ receptor were submitted to each program and the results will be discussed in the following sections.

4.5.1 3D-PSSM

Human $P2X_1$ and $P2X_7$ receptor amino acid sequences were first submitted for fold recognition to the 3D-PSSM server (Kelley *et al.*, 1999). In addition, the extracellular segments of both receptor sequences were submitted and these segments were divided into two separate sequences and submitted to increase the potential number of hits.

It is advised when analyzing results from 3D-PSSM (Kelley *et al.*, 1999) to take into account all of the results criteria and consider relevant biological information too in order to identify potential templates. 3D-PSSM uses several measures to score a particular protein hit. The E-value is a measure of how statistically significant a hit is and the SAWTED E-value is a

system used to match functional keywords between query protein sequence and template structure. For the $P2X_1$ receptor and the associated query sequences the confidences of the predictions were poor in all cases highlighted by the high E-values (> 2.13). The majority of the SAWTED (MacCallum *et al.*, 2000) E-values scored the value 1 indicating that the number of keywords in common between the query sequence and template structures was very low. A member of the class II aminoacyl-tRNA synthetases, aspartyl-tRNA synthetase (PDB code 1ASZ) scored better on the SAWTED (MacCallum *et al.*, 2000) E-value (0.888) than the other hits. Class II aminoacyl-tRNA synthetases are known to bind ATP and the similarity between the catalytic domain of the class II aminoacyl-tRNA synthetases and the extracellular domain of the P2X receptor family is well documented in the literature and will be discussed in Chapter 5. However it is important to note at this stage that the confidence of the prediction and the sequence identity is poor. The $P2X_7$ receptor and associated query sequences did not hit any suitable templates. The SAWTED (MacCallum *et al.*, 2000) E-values suggested that the results were not closely related to the P2X receptor family.

In summary, although the 3D-PSSM server did not identify any suitable templates for the extracellular segment of the P2X receptor family, the synthetase family was identified and the similarities between the P2X receptor family and the class II aminoacyl-tRNA synthetase will be discussed in more detail in Chapter 5.

4.5.2 UCLA-DOE Fold Server

Human $P2X_1$ and $P2X_7$ receptor sequences or the extracellular segments were submitted to the UCLA server individually. The protein sequences that were hit had little biological relevance to the P2X receptor family and alignments between query and structure sequences were poor.

4.5.3 TOPITS

Human $P2X_1$ and $P2X_7$ receptor sequences or the extracellular segments of both receptors were submitted to TOPITS (Rost, 1995), via the PredictProtein server (Rost, 1996). The structures hit by the $hP2X_1$ receptor and the extracellular domain of the $hP2X_1$ receptor were unreliable (all alignment Z-scores<3) and the low (26-34%) pairwise sequence identity scores confirmed this. The number of residues inserted and the number of insertions were high indicating poor quality alignments of template structures to the $P2X_1$ receptor sequence. The

template structures also had little biological relevance to the P2X receptor. The similarity of the P2X receptor family and the aminoacyl-tRNA synthetases was highlighted by structures isoleucyl-tRNA synthetase and phenylalanyl-tRNA synthetase sharing a degree of similarity to the $P2X_1$ receptor.

The alignments between the $P2X_7$ receptor sequence and the template sequences were more reliable (Z-scores>3) than that seen for the $P2X_1$ receptor sequence. The most interesting hits were the hexokinases (PDB code 1DGK and 1QHA). The high alignment Z-scores suggests they may have been a reliable match. However the differences in sequence length between the hexokinases and the P2X receptor family produced poor global alignments and poor quality resulting models (alignments and models omitted for this reason).

4.5.4 Threader

The amino acid sequences of the extracellular domains of the $hP2X_1$ and $hP2X_7$ receptors were submitted to THREADER 3. In addition THREADER allows the submission of secondary structure predictions to be searched against its database so these were generated by PSIPRED (McGuffin *et al.*, 2000) for the $hP2X_1$ and $hP2X_7$ receptors and submitted to THREADER as additional queries. Threader generates large amounts of data and for this reason full result tables can be found in the Appendix. Recommendations for analyzing the results tables can be found in Chapter 2. In order to simplify the data generated by THREADER the top scoring structures for the $hP2X_1$ receptor queries will be discussed first, followed by the top scoring structures for the $hP2X_7$ receptor queries.

4.5.4.1 $hP2X_1$ receptor results

Generally the Z-scores for the hits with the $hP2X_1$ receptor sequence were poor or borderline significant and needed further confirmation. Of the top scoring structures, several were likely to be false matches as energy scores of the threaded models were lower than that of the native structures. Also, the partial alignments of secondary structural elements were further indicators of false matches between query sequence and template structure. A member of the kinase family, pyruvate phosphate dikinase (PDB code 1KBL) was found but the similarity identified was not with the ATP-binding domain of the structure and the sequence similarity score was poor. Analyzing those hits with a relatively high sequence similarity score revealed a class II aminoacyl-tRNA synthetase (threonyl-tRNA synthetase PDB code 1QF6) but the other hits had little biological significance. The secondary structure of the extracellular

domain of the $P2X_1$ receptor produced a list of hits with poor and borderline Z-scores. The highest scoring Z-scores related to structures with little biological relevance to the P2X receptor family and all had low sequence similarity scores for the models. A structure that scored highly on the sequence similarity field was a cytokine binding region of gp130 (1BQU) but with a lower energy score for the model than the native structure it was likely to be a false match.

4.5.4.2 $hP2X_7$ receptor results

A similar story was seen when the extracellular domain of the $hP2X_7$ receptor was scanned by Threader. Z-scores indicated poor and borderline significant hits with little biological relevance to the P2X receptor family. Two structures, transcription inhibitor (PDB code 1ERJ) and elastase complexed with trifluoroacetyl-l-lysyl-l-prolyl-p-isopropylanilide (PDB code 1ELA) had high sequence similarity scores (57.1 and 37.6 respectively) with 70-80% of the structure aligned but with little biological relevance to the P2X receptor family they were unlikely to be suitable template structures. The highest Z-score of 3.12 was achieved by a winged bean albumin (PDB code 1WBA) by increasing the depth of search variable to 1000 from the default 60. The structure was given a high sequence similarity score, but the length of the sequence was ~100 residues less than the extracellular domain of the $P2X_7$ receptor. A sequence alignment of the $hP2X_7$ receptor and 1WBA using CLUSTALW (Thompson *et al.*, 1994) quickly realizes that the difference in length produces a poor quality alignment and structural elements in the P2X receptor family are not conserved (alignment omitted for this reason). Therefore it is not a suitable template. The secondary structure prediction of the $hP2X_7$ receptor did not identify any novel structures.

4.5.5 Protein Folds Database Searching Summary

In conclusion, using different programs that employ different techniques based on the threading principle has not been successful in identifying template structures for the P2X receptor family. However, further similarities have been found with the class II aminoacyl-tRNA synthetase family and this will discussed in Chapter 5. Due to the large number of protein structures but relatively small number of protein folds it is likely that the P2X receptor family may share a likeness with other protein structures but they have not been discovered in this study.

Searching for suitable templates for the P2X receptor family using databases of protein structures and protein folds has revealed one possible sequence that may be used for the modeling exercise, HUNK. The following sections will use this sequence to generate a model for the extracellular binding domain of the $hP2X_1$ receptor.

4.6 Sequence Alignment

The biological significance of HUNK and the P2X receptor family makes it the best template available for modelling the extracellular domain of the $hP2X_1$ receptor. There is no structure available for HUNK so it must be used as an 'intermediary' to find structural templates. Intermediate sequence search (ISS) methods (as described by (Li *et al.*, 2000)) attempt to recognise distant homologues using transitive sequences. An intermediate sequence, say sequence B, with significant alignment scores to two remote homologs, say sequences A and C, can allow the similarity between sequences A and C to be established.

Following the model building procedure, the human HUNK sequence was divided into smaller domains and the protein kinase domain selected to scan against databases of known 3D structures. A list of significant hits was obtained with a PSI-BLAST (Altschul *et al.*, 1997) search. The sequence was scanned against databases of known protein folds to try and identify the most suitable structures to use as templates. The 3DPSSM (Kelley *et al.*, 1999) server returned four significant hits all belonging to the serine/threonine kinase family; 1A06, 1PHK, 1APM and 1FOT. However, 1FOT is a homolog of 1APM so three suitable template structures were identified for modelling; 1A06, 1PHK and 1APM. All three templates were also identified by the PSI-BLAST search with relatively good scores and E-values.

A correct, high quality alignment between the query sequence and template structures is essential to achieving a reliable model. The extracellular domain of the $hP2X_1$ receptor and template structures 1A06, 1APM, 1PHK were aligned and the alignment was manually adjusted and improved to result in the best possible alignment for the sequence and template structures. The length of the extracellular loop of the $P2X_1$ receptor is approximately 280 residues long. To increase the accuracy of the alignment it is important to use sequences of a similar length. The protein kinase domain of the templates has been used and varies in length from 255 to 269. A profile method of aligning two sets of sequences has been employed. The first profile consists of a sequence alignment of the extracellular domains of the P2X family produced with ClustalW (Thompson *et al.*, 1994). The second is a structural alignment of the protein kinase domains of the three templates produced by the 3DCOFFEE web server (Poirot *et al.*, 2004). The full alignment of query and template sequences is viewed in Jalview (Clamp

et al., 2004) and the sequences coloured in blue indicating the percentage identity of the residues in that column (Figure 4.8). Template structure 1PHK is bound to ATP and this structure was used to model the bound ATP molecule in the 3D model of the extracellular domain hP2X$_1$ receptor. It is important to ensure the residues that interact with ATP in the template structure are not aligned to 'gaps' in the alignment. ATP and metal (manganese II ion) interactions in 1PHK are highlighted in the alignment (Figure 4.8) by red and green circles respectively. It is interesting to note that all the interactions with the ligand and metal are with residues located in the first half of the protein sequence. This will have an effect on the resulting model of the extracellular domain hP2X$_1$ receptor.

4.7 Model Building of the hP2X$_1$ receptor Extracellular Segment

A 3D model of the extracellular domain of the hP2X$_1$ receptor with ATP bound and two magnesium ions was generated using MODELLER (Fiser & Sali, 2003). The ATP molecule and magnesium ions were modelled from the structure 1PHK which has ATP and manganese (II) ions bound (Figure 4.9) and can be found in a similar conformation in the model.

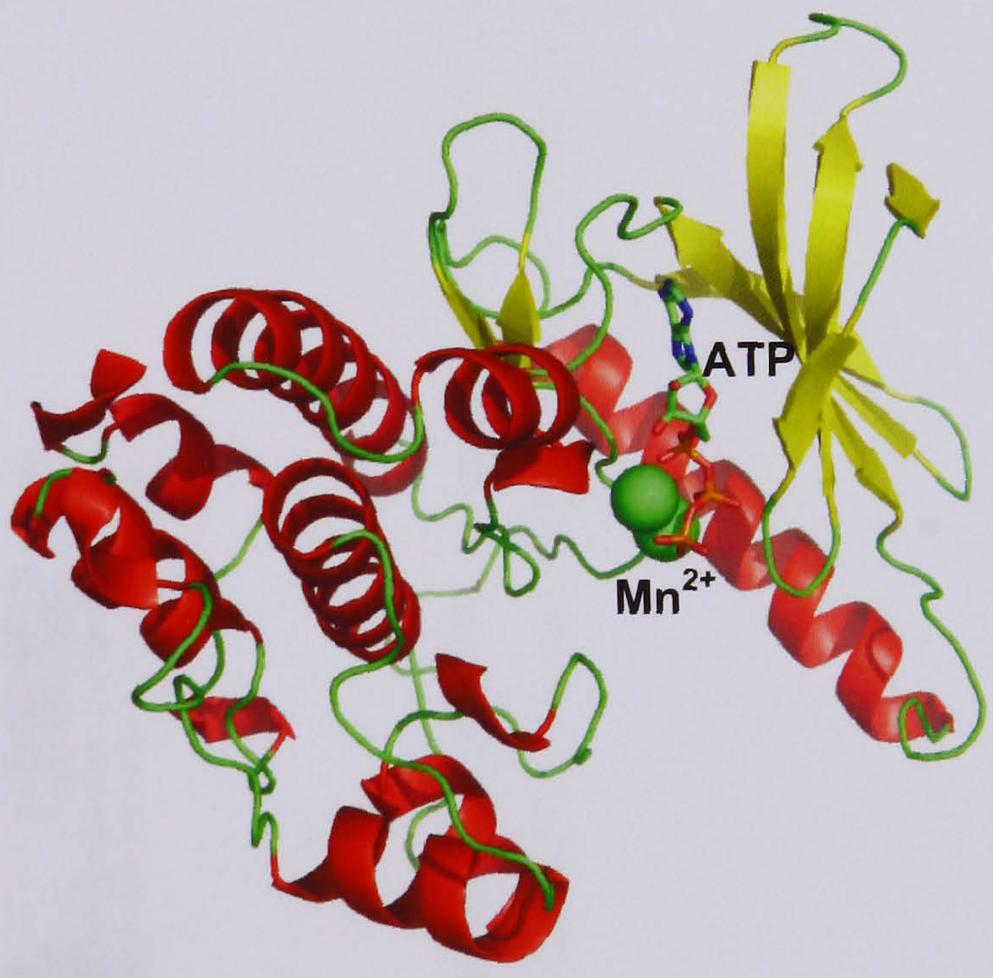


Figure 4.9 X-ray structure of phosphorylase kinase (PDB code 1PHK). Secondary structural elements α-helices are shown in red, β-strands in yellow and coil in green. Manganese (II) ions are shown as green spheres. ATP is coloured according to atom (nitrogen blue, oxygen red and phosphate orange).

Figure 4.8 Profile alignment of extracellular domain of hP2X$_1$ receptor and three structural templates 1PHK, 1A06 and 1APM. Alignment is coloured according to percentage sequence identity. Red circles represent interactions with the ligand ATP for structure 1PHK and green circles represent interactions with the metal (Mn).

Many ATP-binding proteins bind ATP complexed with magnesium cations which help to stabilize the conformation of ATP (Yan *et al.*, 2005). The model also contains 5 disulphide bonds which were restrained in MODELLER (Fiser & Sali, 2003). The disulphides are as predicted for the $hP2X_1$ receptor by experimental evidence (C117-C165, C126-C149, C132-C159, C217-C227 and C261-C270) (Ennion & Evans, 2002a).

The majority of the structure of the 3D model of the extracellular domain $hP2X_1$ receptor is α-helical with one β-sheet consisting of three β-strands in the first 91 residues of the domain (Figure 4.10a). This is consistent with an observation in literature that suggests the ATP-binding domain of the P2X receptor family is built upon a β-sheet structure, similar to that seen in the catalytic domain of the class II aminoacyl-tRNA synthetases (Freist *et al.*, 1998). The secondary structure predictions for the $hP2X_1$ receptor (McGuffin *et al.*, 2000) do not agree with the 3D model. The entire second half of the model (from residue 91 onwards) is α-helical in nature or large loops. The distance between the two transmembrane domains on the extracellular surface is known to be ~8.4 Å apart (Jiang *et al.*, 2001; Spelta *et al.*, 2003). This was determined by mutant P2X channels expressing two cysteine residues at positions 48 and 328 and observing the formation of a disulphide bonds between them. In the 3D model of the $P2X_1$ receptor it is clear that the distance between residues preceding TM2 and those immediately after TM1 is too large to form the channel pore (~50 Å).

4.8 Model Analysis

It is important to rationalise the mutagenesis data available for the ATP binding site $hP2X_1$ receptor with respect to the 3D model. From the alignment of the template structures and the $hP2X_1$ receptor extracellular domain (Figure 4.8) it is clear that the percentage identity between the protein sequences is low and therefore the quality of the resulting 3D model will be compromised. It is also apparent that the interactions between the protein and ligand/metal are all located in the first half of the template structure 1PHK. The effect of this in the 3D model of the $hP2X_1$ receptor is clearly seen in Figure 4.10b where the residues thought to interact with ATP in the $hP2X_1$ receptor, as predicted by mutagenesis data are not in close proximity to the ATP molecule. These residues include the motif $^{290}NFR^{292}$ ($P2X_1$ receptor numbering) and K309, which are thought to play a role in ATP binding at the $P2X_1$ receptor. Due to the size of the extracellular domain of the $P2X_1$ receptor (~288 residues) this may be an issue with other 3D models as mutagenesis predictions involve the entire length of the extracellular domain in ATP action at the $P2X_1$ receptor. One of the limitations of homology

modelling in this case is the ability to model the P2X receptor channel as a trimer of receptor subunits forming a functional channel. The impact of the trimeric structure on the number and location of ATP binding sites cannot be investigated. Positive residues K68 and K70 both of which are thought to bind the negatively charged phosphate groups of the ATP molecule are within 4Å distance of the ATP molecule (Figure 4.11a). The nitrogen of K68 interacts with the oxygen of the α-phosphate of ATP as predicted ($P2X_1$ receptor numbering) (Figure 4.11b). Other interactions with the ATP and magnesium ions in the 3D model are detailed in the Table 4.2 including any further information known about these amino acids. Of those residues that are 100% conserved across the P2X receptor family (K68, D86, N191) all have been mutated in past studies however, only K68 had an effect on ATP potency at the $hP2X_1$ receptor. D191 is close to a region thought to be important in ATP binding at the $hP2X_1$ receptor (region containing motif $^{185}FT^{186}$ $P2X_1$ receptor numbering). S64 is conserved in six out of seven members of the P2X receptor family but there is no mutagenesis data available because past mutagenesis studies have concentrated on those residues that are 100% conserved across the family. Its proximity to K68 and K70, residues which are known to be important in ATP binding at the $P2X_1$ receptor, may implicate this region of the $P2X_1$ receptor in ligand binding. The remainder of the interacting residues listed in Table 4.2 are not well conserved across the P2X receptor family which suggests they are unlikely to have a functional role at the $P2X_1$ receptor.

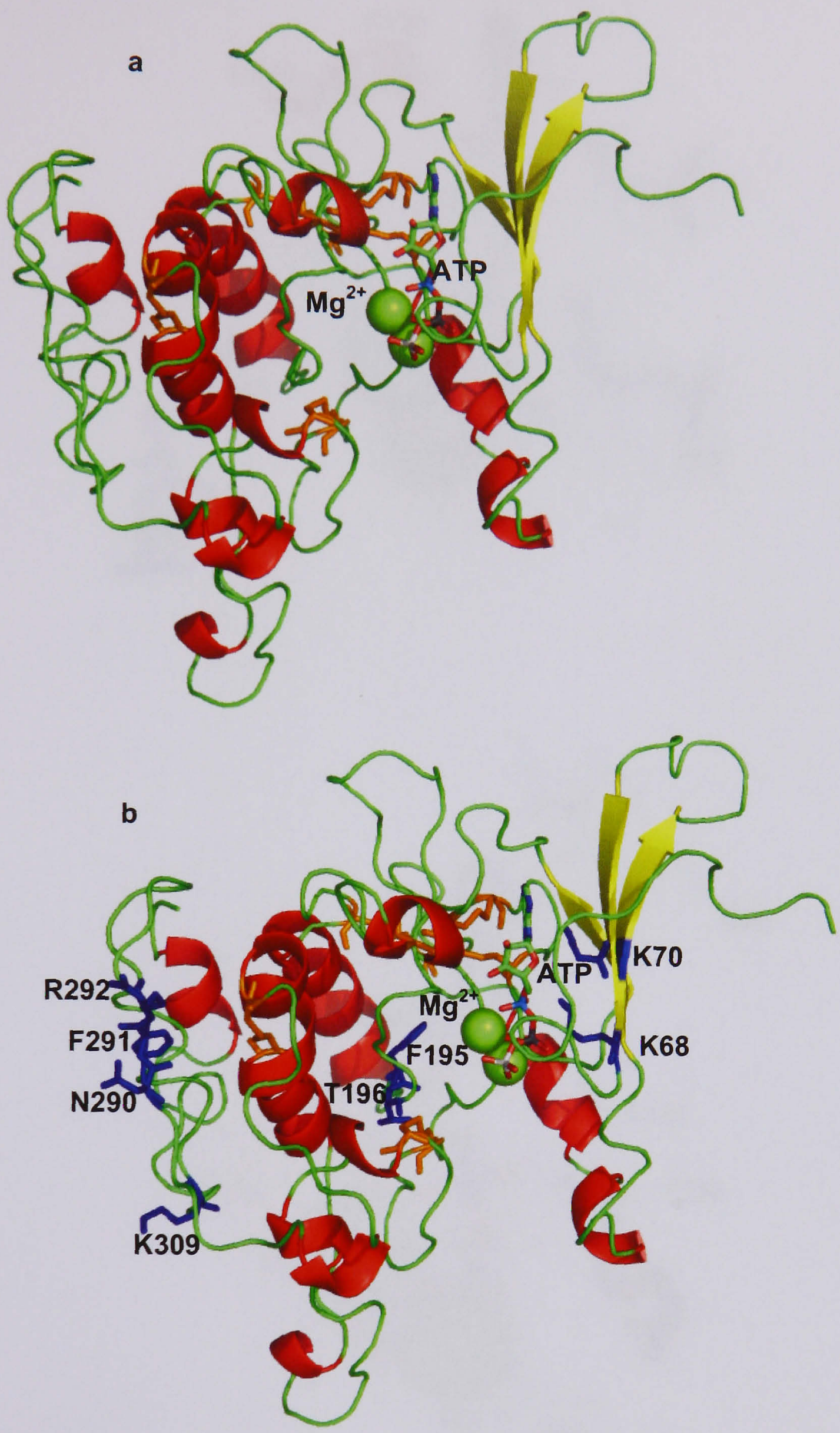


Figure 4.10 Three-dimensional model of extracellular domain hP2X$_1$ receptor with ATP bound. Secondary structure coloured such that α-helix are red, β-strands yellow and coil is green. Cysteine residues are coloured orange. ATP is bound and coloured green by element and magnesium ions are shown as green spheres. **a**, Structure of extracellular loop hP2X$_1$ receptor. **b**, Residues shown in blue are thought to be involved in ATP binding at the hP2X$_1$ receptor by mutagenesis data (P2X$_1$ receptor numbering).

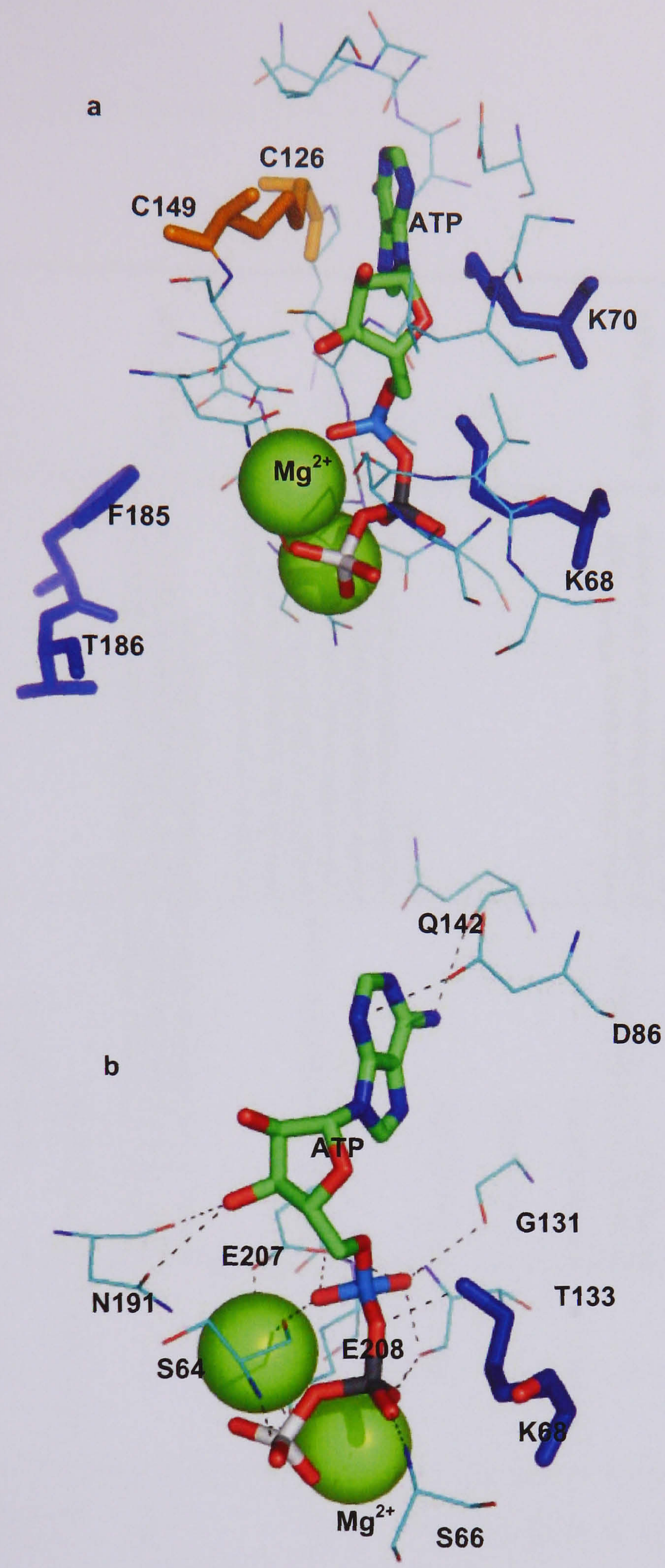


Figure 4.11 Three-dimensional model of the ATP-binding site of hP2X$_1$ receptor (hP2X$_1$ receptor numbering). ATP is bound and coloured green by element and Mg^{2+} are shown as green spheres. **a**, Residues within 4Å of the ATP molecule. Cysteines are coloured orange. Residues shown in blue are thought to be involved in ATP binding at the hP2X$_1$ receptor by mutagenesis data. **b**, Polar contacts of ATP and Mg^{2+} shown by dashed black line. Positively charged K68 is predicted by mutagenesis data to bind the negatively charged phosphate groups of ATP and is coloured blue.

Residue	Interactions		Conservation across P2XR family	Mutagenesis data	Binding Information	References
	Residue atom	ATP atom				
S64	O	O2A	6/7	no data available		
	N	O1G				
S66	N	O1B	non-conserved	no data available		
K68	NZ	O3A	7/7	Important for ATP potency, mutant K68A >1800 fold decrease in ATP potency, mutant K68R non functional	K68 might be involved in coordinating the binding of the negatively charged β and γ phosphates in the tail of ATP	Ennion, 2000
D86	OD2	N3	7/7	mutant D86A, reduced peak current due to expression	Individual conserved negatively charged residues are not essential for ATP recognition by the $hP2X_1R$ and cooordinated binding of the positive charge on magnesium complexed ATP by negatively charged amino acids is not required.	Ennion, 2001
G131	O	O1A	2/7	no data available		
	N	O1A				
T133	O	O1A	2/7	no data available		
	O	O2B				
Q142	O	N6	non-conserved	no data available		
N191	OD1	O3'	7/7	mutant N191A no effect seen on ATP potency	residue located close to $^{185}FT^{186}$ motif thought to be involved in ATP action at the $hP2X_1$ receptor	Roberts, 2004
	O	O3'				
E207	OE1	O1A	1/7	no data available		
	OE1	O2A				
E208	OE2	O3G	non-conserved	no data available		
	OE2	O2G				

Table 4.2 ATP interactions in the 3D model extracellular domain $hP2X_1$ receptor. All numbering refers to $hP2X_1$ receptor.

4.9 Modelling the Cysteine-Rich Domain of the Extracellular Segment of the human $P2X_1$ receptor

P2X receptor subunits consist of two transmembrane domains, intracellular amino and carboxy termini and a large extracellular ligand-binding segment. The seven subunits ($P2X_{1-7}$) vary from 384 ($P2X_4$) to 595 ($P2X_7$) amino acids in length with the bulk of this contained in the extracellular segment. The extracellular segment features 10 cysteine residues that are 100% conserved across the P2X receptor family. The first half of the extracellular domain contains six cysteines (C117, C126, C132, C149, C159, C165; $P2X_1$ receptor numbering) and a further four cysteine residues (C217, C227, C261 and C270; $P2X_1$ receptor numbering) are contained in the latter section of the extracellular domain (Figure 4.12).

These ten conserved cysteines can form five disulphide bonds that act to stabilize the conformation of the receptor (Hansen *et al.*, 1997). The pattern of disulphide bonding in the P2X family has been investigated using cysteine to alanine substitutions. The effects of such substitutions on ATP sensitivity were considered in the rat $P2X_2$ receptor. Based on maximum current EC_{50} values and experiments involving zinc and pH potentiation, the following cysteine pairs were suggested to form disulphide bonds; C113-C164, C214-C224 and C258-C267. It was also suggested that C124, C130, C147 and C158 form two additional disulphide bonds but the specific pairing could not be determined (Clyne *et al.*, 2002). A similar method of mutagenesis was employed in the $hP2X_1$ receptor but an additional approach using MTSEA-biotin was used (Ennion & Evans, 2002a). This compound is known to bind free cysteine residues and coupled with systematic cysteine to alanine substitutions; possible disulphide bonding partners were assigned. Direct biochemical evidence was found for the existence of three disulphide bonds (C126-C149, C132-C159 and C217-C227) and the results suggested the remaining 4 cysteine residues form two more disulphide bonds (C117-C165 and C261-C270) (Figure 4.13).

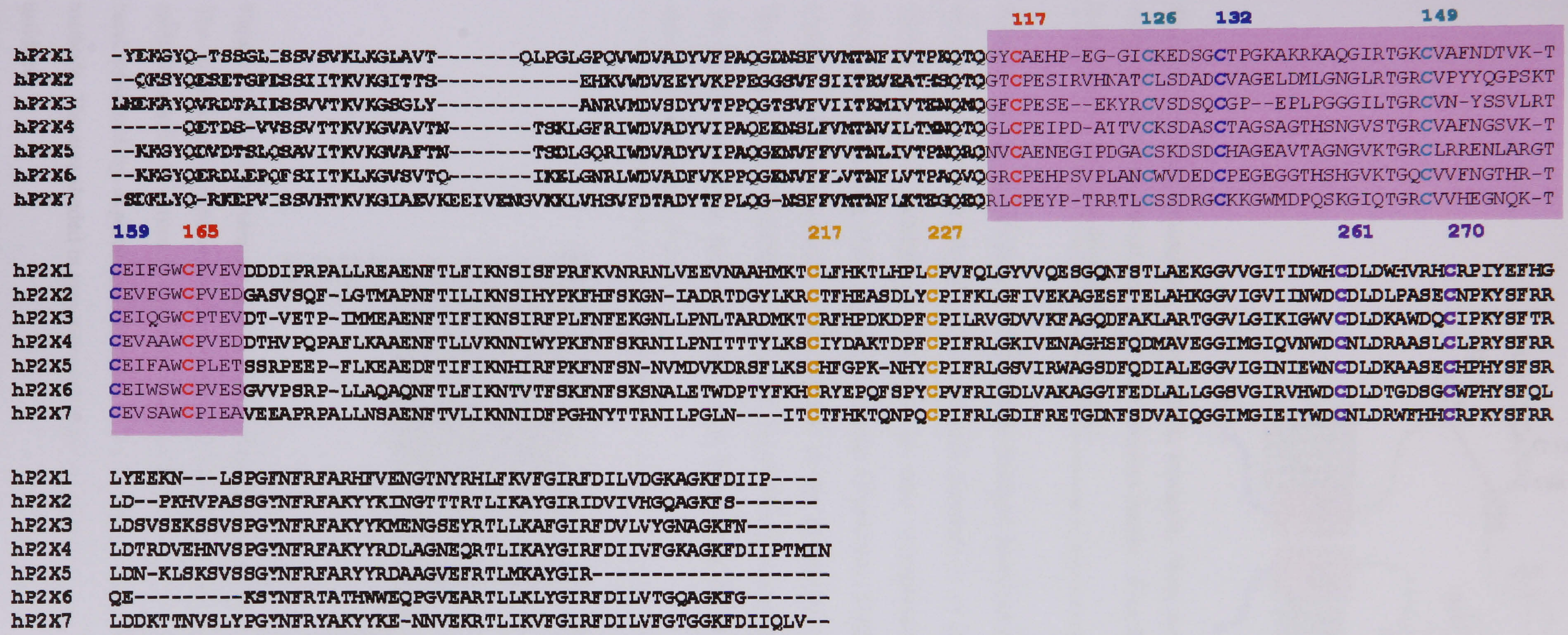


Figure 4.12 Multiple Sequence Alignment Extracellular Domain hP2X Receptor Family. Conserved cysteine residues are coloured and bold. Cysteine residues of the same colour represent disulfide bonds. There five predicted disulphide bonds, C117-C165, C126-C149, C132-C159, C217-C227, C261-C270. Region with pink background represents the cysteine-rich domain. Numbering refers to the $hP2X_1$ receptor.

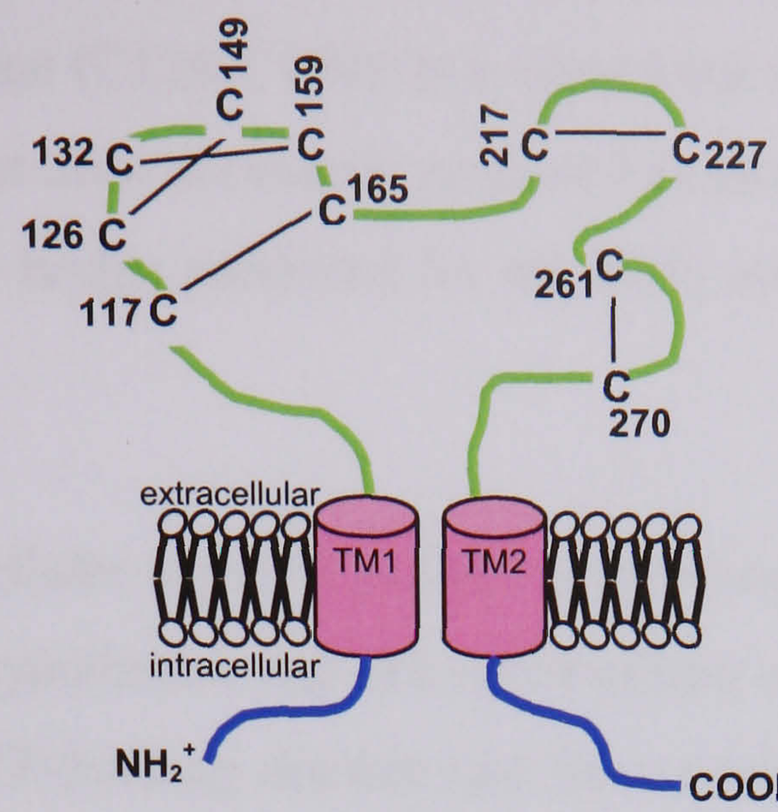


Figure 4.13 Cysteine bonding pattern hP2X$_1$ receptor. There are ten cysteine residues in the hP2X$_1$ receptor which are 100% conserved across the P2X receptor family. Five disulphide bonds are predicted to form in the pattern shown above (Ennion & Evans, 2002a).

Servers for the prediction of disulphide bridges from amino acid sequence can be a useful tool in the study of structural and functional properties of proteins. DISULFIND (Ceroni *et al.*, 2006) predicts disulphide patterns in two computational stages; the first predicts the disulphide bonding state of each cysteine (Frasconi, 2002; Ceroni, 2003) and the second pairs cysteine residues known to participate in the formation of bridges (Vullo & Frasconi, 2004). By default DISULFIND returns the most likely connectivity pattern. The bonding pairs for the hP2X$_1$ receptor are summarised in the matrix below compared to those experimentally determined (Figure 4.14).

Cys	117	126	132	149	159	165	217	227	261	270
117			B			B				
126				BB						
132					B					
149										
159									B	
165							B			
217								B		
227										B
261										B
270										

Figure 4.14 Matrix comparing disulphide bonding predicted by experimental or computational methods. The 10 conserved cysteine residues (**Cys**) in P2X receptors (P2X$_1$ receptor numbering) are compared against each other and a B written to represent a disulphide bond between a pair of cysteines. The black **B** represents a bond predicted by experimental methods, a red **B** represents a disulphide bond predicted by computational methods. The area shaded in grey represents the cystine-rich domain that contains 6 of the 10 conserved cysteine residues.

The experimentally predicted disulphide bonds differ to those predicted by computational methods in all cases except one (C126-C149). It is clear from the literature that there may be some disulphides that have not been accurately predicted as direct biochemical evidence is not available for each disulphide bridge predicted for the $P2X_1$ receptor for example (Ennion & Evans, 2002a).

Due to the size of the extracellular segment and the positioning of the ten conserved cysteine residues it is reasonable to hypothesise that the extracellular segment of the $hP2X_1$ receptor consists of two domains; ATP-binding domain and the cysteine-rich domain (residues 115-169 $hP2X_1$ receptor numbering). The cysteine-rich domain would contain the first six conserved cysteine residues (C117, C126, C132, C149, C159 and C165) and three disulphide bonds. The remainder of the extracellular loop would form the ATP-binding domain. This is consistent with mutagenesis studies which have identified residues that play a role in ATP-action at the $P2X_1$ receptor as they are all contained in the putative ATP-binding domain.

4.9.1 Identifying suitable 3D templates

According to the model building procedure (Figure 4.1) the cysteine-rich segment must be scanned against databases of known 3D structure and known protein folds in order to find suitable templates on which to base a structural model. The results of both searches will be discussed in the next section.

4.9.1.1 Searching databases of known 3D structures

The cysteine-rich domain of the $hP2X_1$ receptor (residues 115-169) was submitted to PSI-BLAST using the PDB as the target database. No significant similarity was found. Using PSI-BLAST with the non redundant database a list of sequences producing significant alignments was generated but the majority of hits were P2X receptor protein related. No relevant hits were found.

4.9.1.2 Searching databases of known protein folds

An alternative method of scanning databases of known protein folds was employed. The 3D-PSSM server generated a list of less than confident hits (E-values of certainty less than 50%). The SAWTED function that matches functional keywords between the query protein sequence and template structures had poor values for the list of hits (this measure is assessed similar to E-values; the closer to zero the score the better the textual match). Alignments

contained numerous gaps and secondary structure elements were poorly aligned. However, of all the hits it was interesting to note that a number of them belonged to a toxin family.

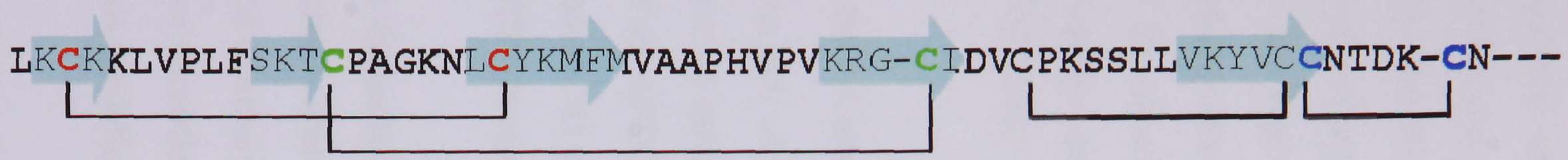
The cysteine-rich segment was then submitted to THREADER. Of the top ranked hits, all had Z-scores less than 2.7 indicating that they were poor matches. The pattern of disulphide bonding in the top ranked hits was analyzed and compared to the predicted pattern of bonding in the $hP2X_1$ receptor. Although the specific pattern of bonding was not replicated, the positions of the cysteine were conserved in many of the structures. Due to the high percentage of structure/sequence aligned (91.7%), 1TGX was chosen as the best fitting template for the cysteine-rich domain. In addition it is a member of the cardiotoxin family which had been identified by 3D-PSSM.

4.9.2 Sequence Alignment

A homology model should be based on at least two template structures so the amino acid sequence of 1TGX was submitted to the BLAST (Altschul *et al.*, 1990) program and the PDB database scanned for short, nearly exact matches. Two cytotoxins were selected from the list of hits based on their high bit scores indicating good alignments and low E-values indicating significant hits; 1CCQ and 1UG4. Importantly both structures shared identical cysteine-bonding patterns to 1TGX and all three structures were of a similar length to the cysteine-rich domain of the $hP2X_1$ receptor.

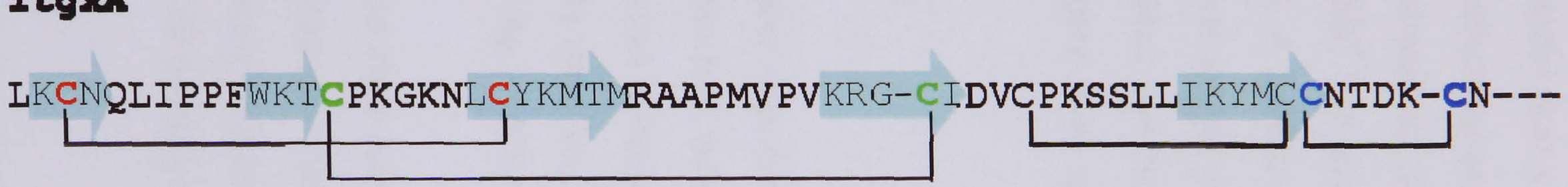
The three templates contain eight cysteine residues that form four disulphide bonds. An alignment of the template sequences to the query sequence was generated using ClustalW (Thompson *et al.*, 1994) (Figure 4.15). The secondary structure of the templates and the predicted secondary structure for the cysteine-rich domain of $hP2X_1$ receptor, as predicted by PSIPRED (McGuffin *et al.*, 2000) were also compared.

1ccqA



1ug4A

1tgxA



cysteine-rich domain hP2X$_1$

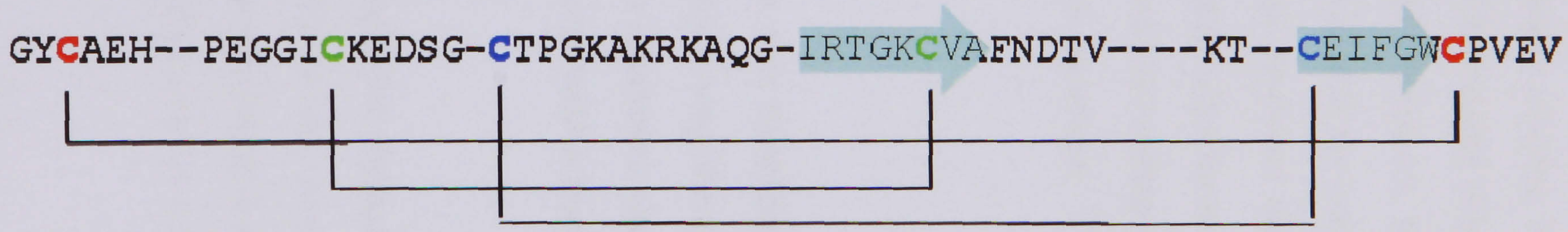


Figure 4.15 Sequence alignment of template sequences and cysteine-rich domain hP2X$_1$ receptor. Brackets indicate the disulphide bonds between pairs of cysteine residues. Pairs of cysteine residues that form disulphide bonds are the same colour. Beta strands are shown as light blue arrows. Those shown for the hP2X$_1$ receptor are predicted by PSIPRED (McGuffin *et al.*, 2000).

4.9.3 Model Building of the hP2X$_1$ receptor Cysteine-Rich Domain

A 3D model of the cysteine-rich domain of the hP2X$_1$ receptor was generated using MODELLER (Fiser & Sali, 2003). The disulphide bonds were specified for three pairs of cysteine residues in the same pattern of bonding as the template structures. The initial model had no secondary structural elements (Figure 4.16). It was clear after overlaying the template structures onto the P2X$_1$ receptor model that a loop (residue starting N153 P2X$_1$ receptor numbering) was responsible for the structural problems of the model (Figure 4.17). To solve this, the N-terminal of the last β-strand of the template structures was shortened. This involved removal of residues ^{49}IKYMC^{53} of structure 1TGX and the corresponding residues in the other two template structures. A further model was produced that contained two β-strands corresponding to the first two β-strands of the template structures and the first β-sheet of the structure (Figure 4.18). A third and fourth β-strand according to the alignment of template structures to query sequence was specified at positions ^{131}GCTPGKA^{137} and ^{144}IRTGKC^{149} (Figure 4.19).

4.10 Conclusions

The aim of this chapter was to combine experimental data with bioinformatics techniques to give an insight into the structure of the P2X receptor family. Using the amino acid sequence of the hP2X$_1$ receptor to probe databases of known protein structures, similar sequences with known 3D structure would be used as templates on which to build a homology model of the hP2X$_1$ receptor but more specifically the ATP-binding site.

In the early stages of this study, it was clear that the extracellular segment of a P2X receptor subunit was likely to contain more than one domain due to its size. A cysteine-rich domain and a putative ATP-binding domain were defined. This was consistent with experimental data that identified residues involved in ATP-binding at the hP2X$_1$ receptor. Each of these domains was modelled separately.

Unfortunately no structures were identified that could be used directly to build a homology model of the ATP-binding domain of the hP2X$_1$ receptor; instead an intermediate sequence was used to identify three template structures. A useful model of the ATP-binding site was generated and tested using mutagenesis data. However, due to the ligand interactions in the template structures, the second half of the ATP-binding domain did not contact the ATP in the model and as a result many of the residues thought to bind ATP at the P2X$_1$ receptor were not

in close proximity. This model further hints at the hypothesis that more than one subunit it likely to bind ATP and that the ATP-binding site may be located at the interface of two subunits. The trimeric structure of the P2X receptor channel has not been taken into account in this model and future studies will need to consider the impact such a formation will have on the ligand-binding site before a more 'believable' model is achieved.

The model of the cysteine-rich domain of the $hP2X_1$ receptor agrees with secondary structural predictions for this region of the $hP2X_1$ receptor. However, the pattern of disulphide bonding in the model does not agree with experimental data. Based on this study the most likely disulphide bonding pattern for the $hP2X_1$ receptor is C126-C149, C132-C159 and C217-C227 for six of the ten conserved cysteines based on the experimental evidence. The bonding pattern for the residues C117, C165, C261 and C270 is not clear as direct biochemical evidence was not found for the existence of the bonds C117-C165 and C261-C270 (Ennion & Evans, 2002a). Therefore, it is possible if one bond is incorrectly predicted, another pairing must also be wrong. It would seem likely that the cysteine-rich domain forms three disulphide bonds which is predicted and the other four cystine residues, two more. Based on the model and experimental evidence this is as much as can be said.

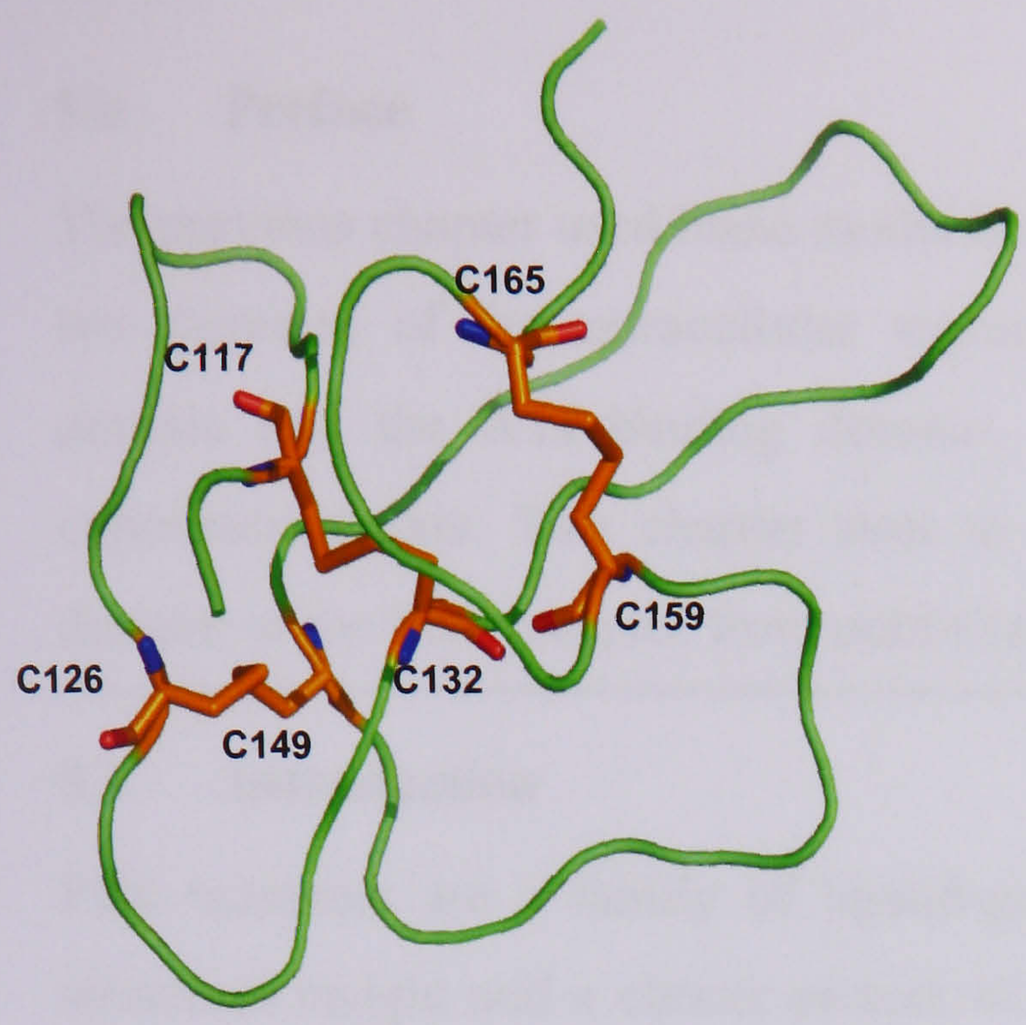


Figure 4.16 Homology model for cysteine-rich domain of hP2X$_1$ receptor. Cysteine residues are coloured orange (hP2X$_1$ receptor numbering). Three disulphide bridges are shown (C117-C132, C126-C149 and C159-C165). Coil is green.

Figure 4.17 Overlay of structures for 1TGX, 1CCQ, 1UG4 and homology model for cysteine-rich domain of hP2X$_1$ receptor. Cysteine residues are coloured orange. Each template structure contains five β-strands (shown as arrows) and 2 β-sheets. The template structures are coloured cyan (11TGX), magenta (1CCQ) and yellow (1UG4). The cysteine-rich domain model is coloured green. The arrow points to a loop region of the cysteine-rich domain of the hP2X$_1$ receptor

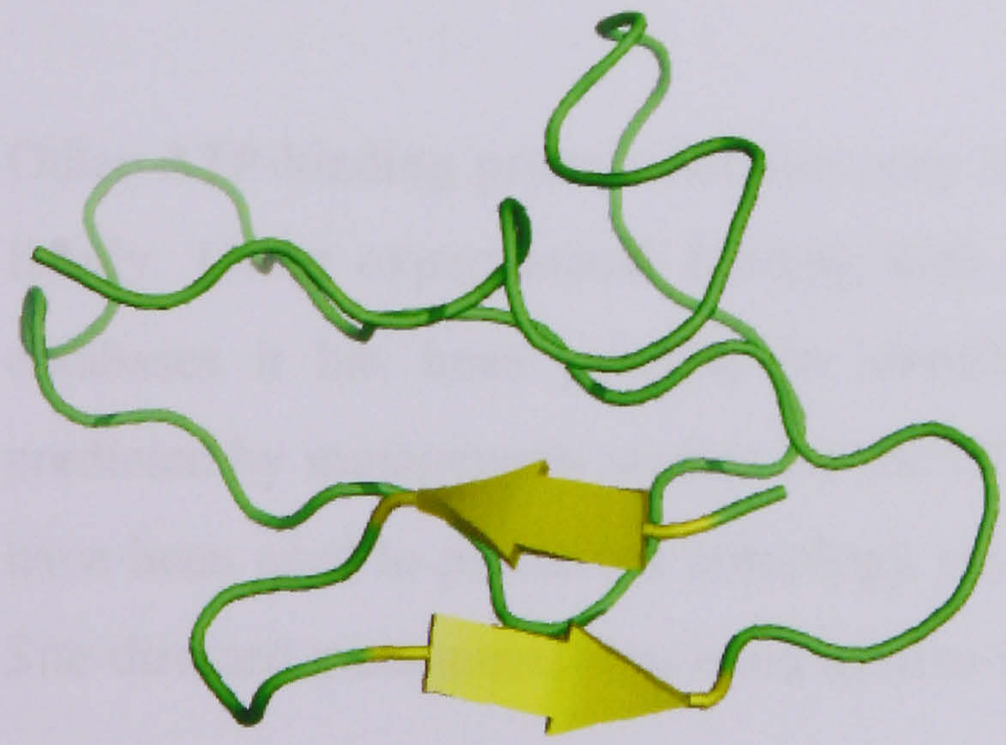


Figure 4.18 Homology model for cysteine-rich domain of hP2X$_1$ receptor. Model coloured by secondary structure, beta-strands are yellow. Two beta strands form the first beta sheet.

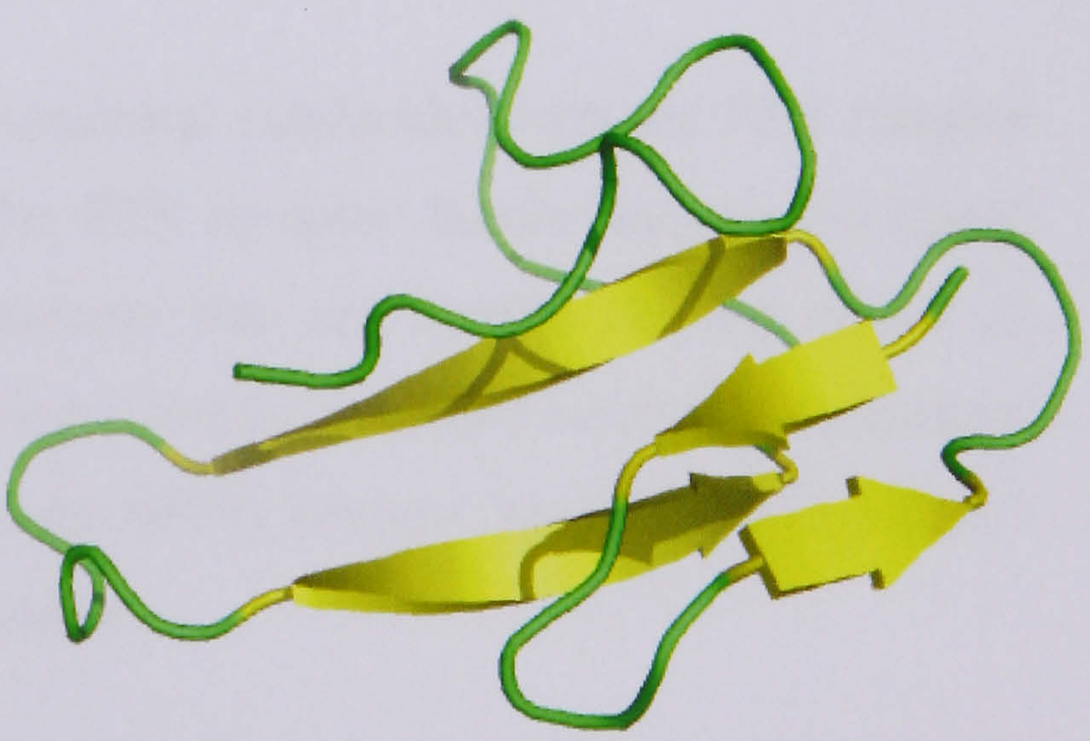


Figure 4.19 Homology model for cysteine-rich domain of hP2X$_1$ receptor. Model coloured by secondary structure, beta-strands are yellow.

Chapter Five

5.0 Preface

The previous chapter used basic model building principles to derive structural information for two domains of the extracellular segment of the P2X receptor family; the cysteine-rich domain and the ATP-binding domain. The models produced were not consistent with experimental data. This chapter aims to derive structural information for the ATP-binding domain of the P2X receptor from published models and approaches.

5.1 Introduction

P2X receptors are a family of ligand-gated ion channels that bind ATP. To gain further structural insight and a clearer picture of the ATP-binding domain an alternative strategy to the model building procedure used in chapter 4 has been adopted. Instead of identifying individual protein sequences as structural templates for the P2X receptor, families of proteins have been examined.

Freist suggested that the extracellular segment of the P2X receptor family contains a portion of sequence (residues 170–330) that exhibits similarities with the catalytic domains of the class II aminoacyl-tRNA synthetase family (Freist *et al.*, 1998). This similarity has been used as the basis to generate a 3D homology model of the ATP-binding domain of the $rP2X_4$ receptor based on three templates belonging to the class II aminoacyl-tRNA synthetases (Yan *et al.*, 2005).

Other ATP-binding protein families may share structural similarities with the P2X receptor family. Using experimental binding data for the P2X receptor family and protein-ligand databases it has been possible to identify structures that use similar binding motifs as predicted by mutagenesis studies for the P2X receptor family. Standard modelling techniques have been used to produce a homology model of the $hP2X_1$ receptor ligand binding domain. Site-directed mutagenesis has been used to test models.

5.2 The P2X receptor family are similar to the class II aminoacyl-tRNA synthetase family

This section will explore the hypothesis made by Freist in 1998 that a portion of the extracellular segment of the P2X receptor family is similar to the catalytic domain of the class II aminoacyl-tRNA synthestases. The models published based on this hypothesis will be

discussed and tested using mutagenesis and lastly an attempt to reproduce such models will be made.

5.2.1 Introduction to Aminoacyl-tRNA Synthetases

Aminoacyl-tRNA synthetases play a crucial part in translation - the process that converts mRNA codon sequences into amino acid sequences - guaranteeing the accuracy of protein synthesis. They covalently bond each tRNA with the proper amino acid, thus ensuring the correct translation from the genetic code of DNA into the amino acid code of proteins. This is a two step process and the energy is obtained by conversion of ATP to AMP and PP_i :

Step 1: amino acid + ATP → aminoacyl-AMP + PP_i
Step 2: aminoacyl-AMP + tRNA → aminoacyl-tRNA + AMP

Most living cells possess twenty different aminoacyl-tRNA synthetases, one for each amino acid. Each synthetase is optimized for function with its own particular amino acid and the set of tRNA molecules related to that specific amino acid. The family of aminoacyl-tRNA synthetases is divided into two groups of ten enzymes each, first identified by sequence analysis (Eriani *et al.*, 1990) and later classified by their tertiary structure. Class I aminoacyl-tRNA synthetases have in common a catalytic domain structure based on the Rossmann fold and are mostly monomeric (Sugiura *et al.*, 2000). They can be further divided into three subclasses a, b and c according to sequence homology. The synthetases specific for arginine, cysteine, glutamic acid, glutamine, isoleucine, leucine, methionine, tyrosine, tryptophan and valine belong to class I (Szymanski *et al.*, 2001).

Class II aminoacyl-tRNA synthetases share a novel six-stranded anti-parallel beta-sheet formation and are mostly dimeric or multimeric. The synthetases specific for alanine, asparagine, aspartic acid, glycine, histidine, lysine, phenylalanine, proline, serine, and threonine belong to class II (Szymanski *et al.*, 2001). Class II synthetases contain three homologous sequence motifs namely 1, 2 and 3. Motifs 2 and 3 are regions of a tertiary fold built around the central six-stranded (β_I, β_{II}, β_{III}, β_{IV}, β_V, β_{VI}) anti-parallel beta sheet and long α helix (Eriani *et al.*, 1995). They form the active site that contains the ATP-binding site.

5.2.2 Structural similarities of the ATP binding site of P2X receptor proteins and class II aminoacyl-tRNA synthetases

It was shown by secondary structure predictions and sequence analysis, that a portion of the extracellular loop of P2X receptor proteins shared a degree of homology with the catalytic domains of class II aminoacyl-tRNA synthetases (Freist *et al.*, 1998). The results were based on a sequence alignment of thirteen characteristic sequence regions of the P2X receptor family and nine class II aminoacyl-tRNA synthetases shown in Figure 5.1. The predicted secondary structure of the P2X receptor family and the secondary structure of the class II synthetases based on crystal structures (arrow for β-sheet, helical structure for α-helix and box for transmembrane domain) are also shown. Figure 5.2 shows the six-stranded anti parallel beta sheet in prolyl-tRNA synthetase chain A, (Berman *et al.*, 2000) accession code 1H4Q.

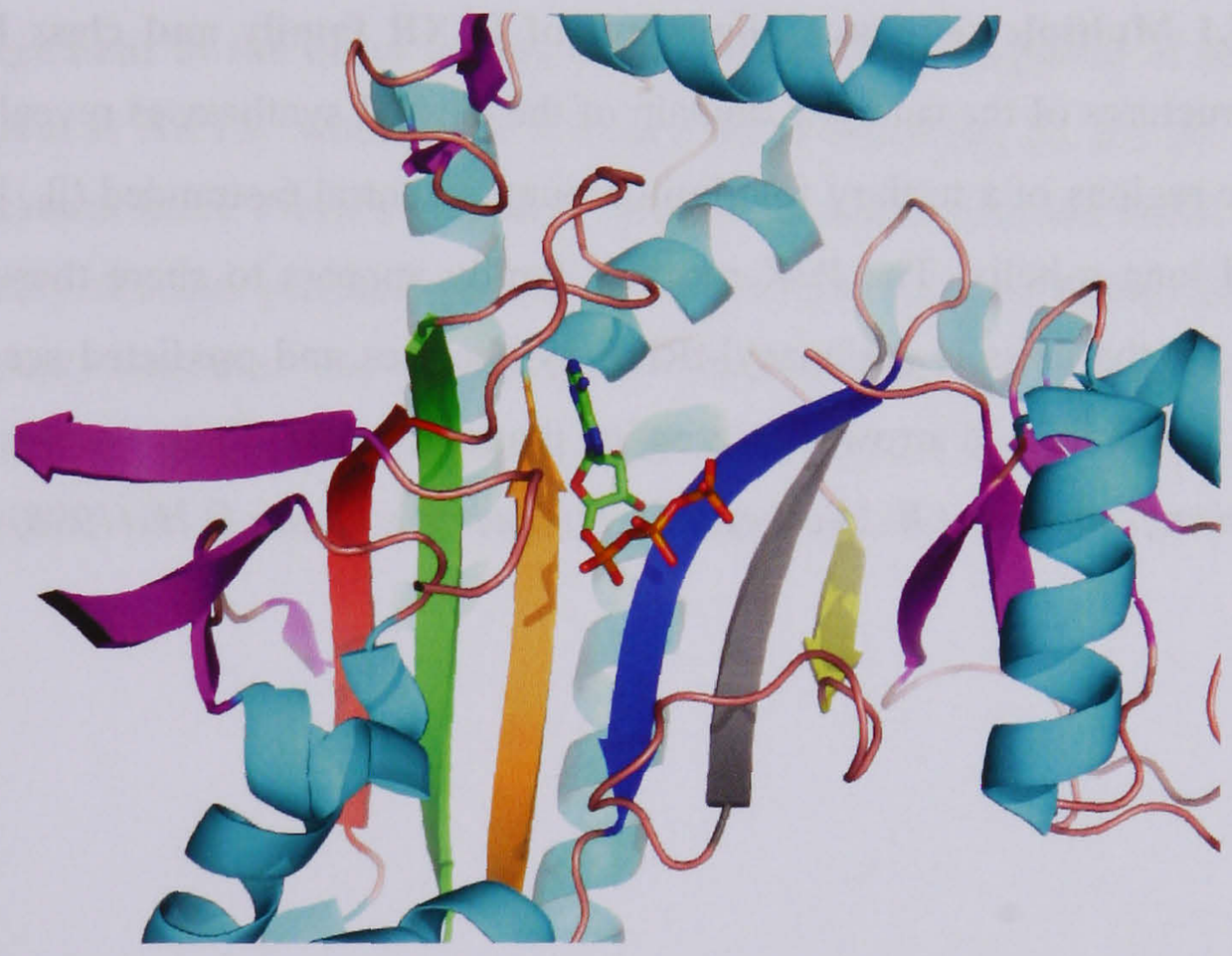


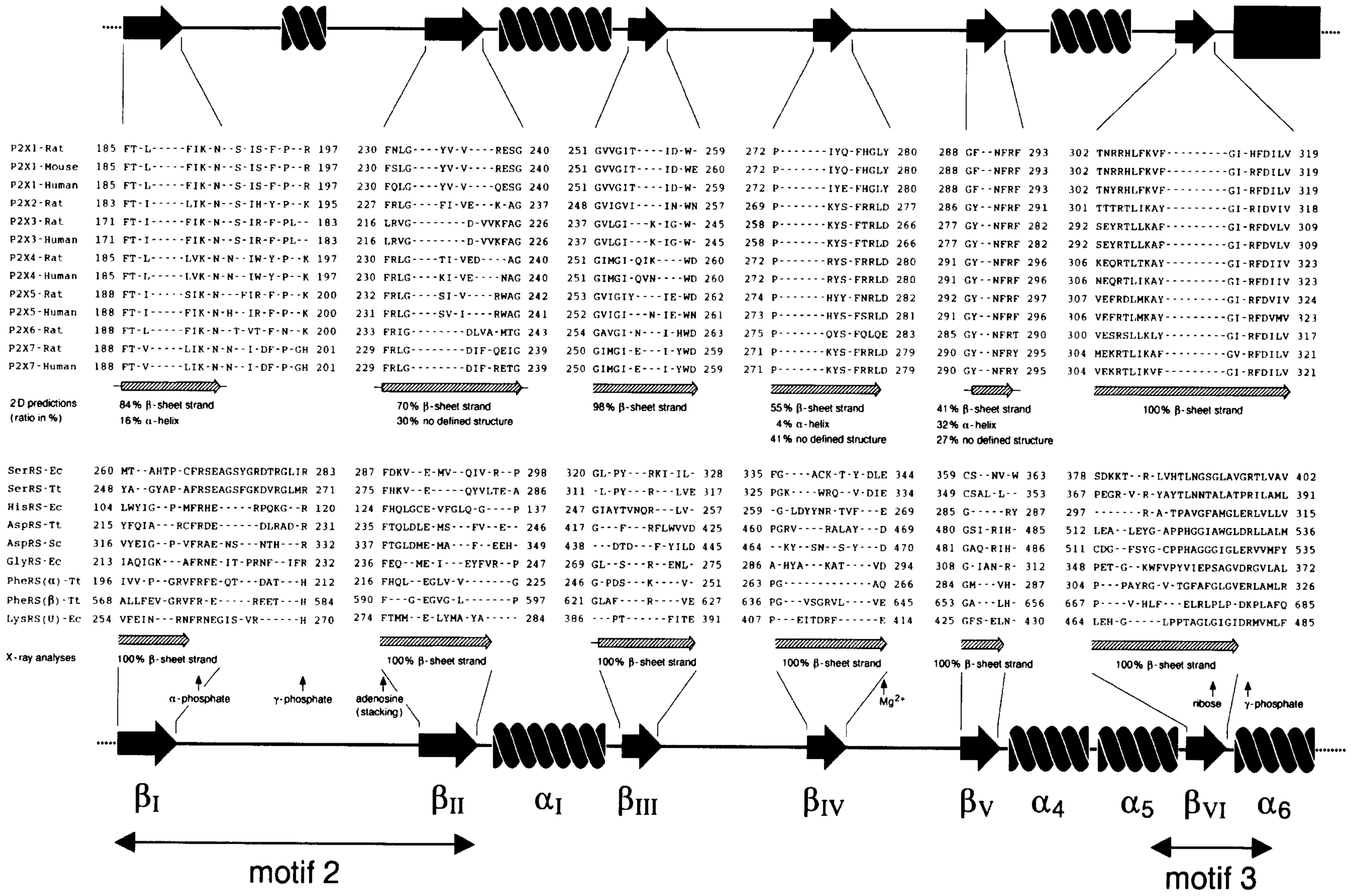
Figure 5.2 Structure of prolyl-tRNA synthetase (1H4Q). α-helices are shown in cyan and β-strands are shown in magenta. The six-stranded anti-parallel β-sheet that forms the ATP binding site is coloured with β_I (red), β_{II} (green), β_{III} (yellow), β_{IV} (grey), β_V (blue), β_{VI} (orange). The bound ATP molecule is coloured in green by element with an orange phosphate tail, nitrogen in blue and oxygen red.

Class II synthetases contain three motifs. β-strands β_I and β_{II} belong to motif 2 and β_{VI} is a region of motif 3. Freist discusses each β-strand (I-VI) region and the conserved residues that exist between the P2X receptor family and the class II synthetases which is summarised below (see also Figure 5.3) (Freist *et al.*, 1998).

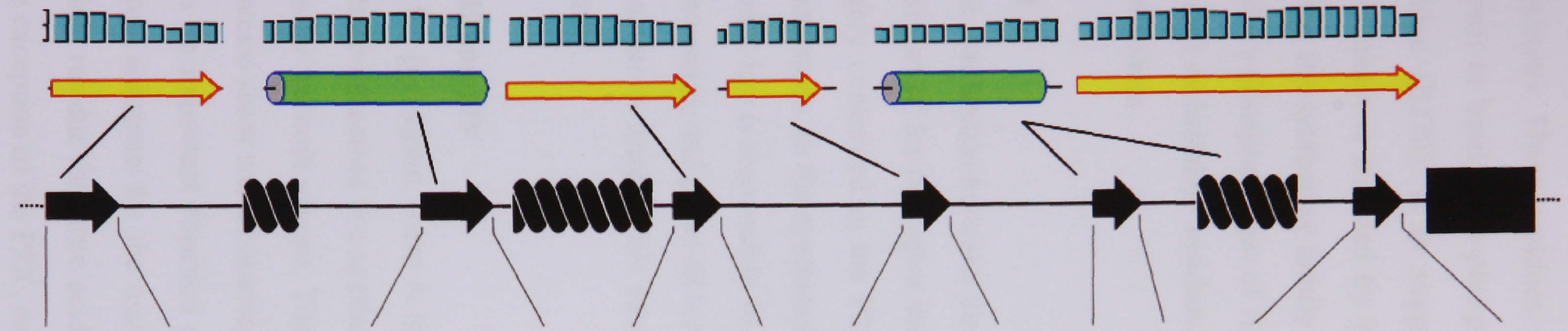
Figure 5.1 Multiple sequence alignment of P2XR family and class II aminoacyl-tRNA synthetases. The crystal structures of the catalytic domain of the class II synthetases reveal common motifs (motif 2 and motif 3) which are regions of a tertiary fold built around a central 6-stranded (β_I, β_{II}, β_{III}, β_{IV}, β_V and β_{VI}) anti-parallel β-sheet and long α-helix. The P2X receptor family appears to share these motifs and structure. The secondary structure for the class II aminoacyl-tRNA synthetases and predicted secondary structure for the P2X receptor family is shown in bold arrows representing β-strands and cylinders representing α-helices.

Figure taken from Friest.W., Verhey.J.F., Stuhmer.W., Gauss.D.H. (1998) FEBS Letters 434 61-65

Figure 5.3 Sequence alignment of the latter region of the extracellular segment of the P2XR family and the catalytic domain of 9 class II AATS. Motif 2 and 3 form a six-stranded (β_I, β_{II}, β_{III}, β_{IV}, β_V, β_{VI}) (represented by black arrows) anti-parallel beta sheet. Freist's secondary structure prediction for the P2XR family is shown in black and the PSIPRED secondary structure prediction is shown for the $P2X_1$ receptor with yellow arrows representing β-strands and green cylinders, α-helices. The confidence of the PSIPRED prediction is shown in blue. Residues highlighted by a green box represent those residues identified by Freist as key residues. *Figure adapted from Freist.W., Verhey.J.F., Stuhmer.W., Gauss.D.H. (1998) FEBS Letters 434 61-65*


P2X1-Rat 185 FT-L-----FIK-N--S-IS-F-P--R 197 230 FNLG----YV-V----RESG 240 251 GVVGIT----ID-W- 259 272 P-------IYQ-FHGLY 280 288 GF--NFRF 293 302 TNRRHLFKVF---------GI-HFDILV 319
P2X1-Mouse 185 FT-L-----FIK-N--S-IS-F-P--R 197 230 FSLG----YV-V----RESG 240 251 GVVGIT----ID-WE 260 272 P-------IYQ-FHGLY 280 288 GF--NFRF 293 302 TNRRHLFKVF---------GI-RFDILV 319
P2X1-Human 185 FT-L-----FIK-N--S-IS-F-P--R 197 230 FQLG----YV-V----QESG 240 251 GVVGIT----ID-W- 259 272 P-------IYE-FHGLY 280 288 GF--NFRF 293 302 TNYRHLFKVF---------GI-RFDILV 319
P2X2-Rat 183 FT-I-----LIK-N--S-IH-Y-P--K 195 227 FRLG----FI-VE---K-AG 237 248 GVIGVI----IN-WN 257 269 P-------KYS-FRRLD 277 286 GY--NFRF 291 301 TTTRTLIKAY---------GI-RIDVIV 318
P2X3-Rat 171 FT-I-----FIK-N--S-IR-F-PL-- 183 216 LRVG--------D-VVKFAG 226 237 GVLGI---K-IG-W- 245 258 P-------KYS-FTRLD 266 277 GY--NFRF 282 292 SEYRTLLKAF---------GI-RFDVLV 309
P2X3-Human 171 FT-I-----FIK-N--S-IR-F-PL-- 183 216 LRVG--------D-VVKFAG 226 237 GVLGI---K-IG-W- 245 258 P-------KYS-FTRLD 266 277 GY--NFRF 282 292 SEYRTLLKAF---------GI-RFDVLV 309
P2X4-Rat 185 FT-L-----LVK-N-N--IW-Y-P--K 197 230 FRLG----TI-VED----AG 240 251 GIMGI-QIK----WD 260 272 P-------RYS-FRRLD 280 291 GY--NFRF 296 306 KEQRTLTKAY---------GI-RFDIIV 323
P2X4-Human 185 FT-L-----LVK-N-N--IW-Y-P--K 197 230 FRLG----KI-VE----NAG 240 251 GIMGI-QVN----WD 260 272 P-------RYS-FRRLD 280 291 GY--NFRF 296 306 NEQRTLIKAY---------GI-RFDIIV 323
P2X5-Rat 188 FT-I-----SIK-N---FIR-F-P--K 200 232 FRLG----SI-V----RWAG 242 253 GVIGIY----IE-WD 262 274 P-------HYY-FNRLD 282 292 GY--NFRF 297 307 VEFRDLMKAY---------GI-RFDVIV 324
P2X5-Human 188 FT-I-----FIK-N-H--IR-F-P--K 200 231 FRLG----SV-I----RWAG 241 252 GVIGI---N-IE-WN 261 273 P-------HYS-FSRLD 281 291 GY--NFRF 296 306 VEFRTLMKAY---------GI-RFDVMV 323
P2X6-Rat 188 FT-L-----FIK-N--T-VT-F-N--K 200 233 FRIG--------DLVA-MTG 243 254 GAVGI-N---I-HWD 263 275 P-------QYS-FQLQE 283 285 GY--NFRT 290 300 VESRSLLKLY---------GI-RFDILV 317
P2X7-Rat 188 FT-V-----LIK-N-N--I-DF-P-GH 201 229 FRLG--------DIF-QEIG 239 250 GIMGI-E---I-YWD 259 271 P-------KYS-FRRLD 279 290 GY--NFRY 295 304 MEKRTLIKAF---------GV-RFDILV 321
P2X7-Human 188 FT-V-----LIK-N-N--I-DF-P-GH 201 229 FRLG--------DIF-RETG 239 250 GIMGI-E---I-YWD 259 271 P-------KYS-FRRLD 279 290 GY--NFRY 295 304 VEKRTLIKVF---------GI-RFDILV 321
2D predictions (ratio in %)
84% β-sheet strand
16% α-helix
70% β-sheet strand
30% no defined structure
98% β-sheet strand
55% β-sheet strand
4% α-helix
41% no defined structure
41% β-sheet strand
32% α-helix
27% no defined structure
100% β-sheet strand
SerRS-Ec 260 MT--AHTP-CFRSEAGSYGRDTRGLIR 283 287 FDKV--E-MV--QIV-R--P 298 320 GL-PY---RKI-IL- 328 335 FG----ACK-T-Y-DLE 344 359 CS--NV-W 363 378 SDKKT--R-LVHTLNGSGLAVGRTLVAV 402
SerRS-Tt 248 YA--GYAP-AFRSEAGSFGKDVRGLMR 271 275 FHKV--E-----QYVLTE-A 286 311 -L-PY---R---LVE 317 325 PGK----WRQ--V-DIE 334 349 CSAL-L-- 353 367 PEGR-V-R-YAYTLNNTALATPRILAML 391
HisRS-Ec 104 LWYIG--P-MFRHE-----RPQKG--R 120 124 FHQLGCE-VFGLQ-G----P 137 247 GIAYTVNQR---LV- 257 259 -G-LDYYNR-TVF---E 269 285 G-----RY 287 297 -------R-A-TPAVGFAMGLERLVLLV 315
AspRS-Tt 215 YFQIA---RCFRDE------DLRAD-R 231 235 FTQLDLE-MS---FV--E-- 246 417 G---F---RFLWVVD 425 460 PGRV----RALAY---D 469 480 GSI-RIH- 485 512 LEA--LEYG-APPHGGIAWGLDRLLALM 536
AspRS-Sc 316 VYEIG--P-VFRAE-NS---NTH---R 332 337 FTGLDME-MA---F--EEH- 349 438 ---DTD---F-YILD 445 464 --KY--SN--S-Y---D 470 481 GAQ-RIH- 486 511 CDG--FSYG-CPPHAGGGIGLERVVMFY 535
GlyRS-Ec 213 IAQIGK---AFRNE-IT-PRNF--IFR 232 236 FEQ--ME-I---EYFVR--P 247 269 GL--S---R--ENL- 275 286 A-HYA---KAT----VD 294 308 G-IAN-R- 312 348 PET-G--KWFVPYVIEPSAGVDRGVLAL 372
PheRS(α)-Tt 196 IVV-P--GRVFRFE-QT---DAT---H 212 216 FHQL--EGLV-V-------G 225 246 G-PDS---K----V- 251 263 PG-------------AQ 266 284 GM---VH- 287 304 P---PAYRG-V-TGFAFGLGVERLAMLR 326
PheRS(β)-Tt 568 ALLFEV-GRVFR-E-----REET---H 584 590 F---G-EGVG-L-------P 597 621 GLAF----R----VE 627 636 PG---VSGRVL----VE 645 653 GA---LH- 656 667 P----V-HLF---ELRLPLP-DKPLAFQ 685
LysRS(U)-Ec 254 VFEIN---RNFRNEGIS-VR------H 270 274 FTMM--E-LYMA-YA----- 284 386 ---PT------FITE 391 407 P---EITDRF------F 414 425 GFS-ELN- 430 464 LEH-G-----LPPTAGLGIGIDRMVMLF 485
X-ray analyses
100% β-sheet strand
α-phosphate
γ-phosphate
100% β-sheet strand
adenosine (stacking)
100% β-sheet strand
100% β-sheet strand
Mg2+
100% β-sheet strand
100% β-sheet strand
ribose
γ-phosphate
βI
βII
αI
βIII
βIV
βV
α4
α5
βVI
α6
motif 2
motif 3

PSIPRED



P2XR Family

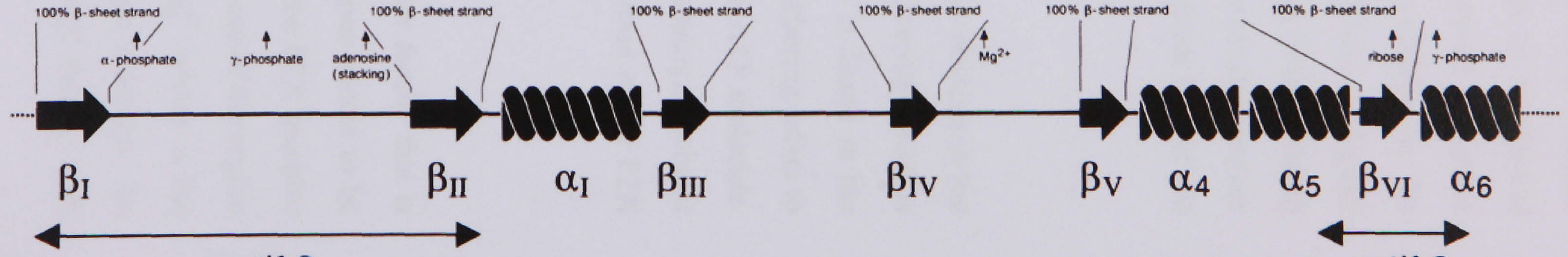
```
P2X1-Rat      185 FT-L-----FIK-N--S-IS-F-P--R 197   230 FNLG----YV-V----RESG 240   251 GVVGIT----ID-W- 259   272 P-------IYQ-FHGLY 280   288 GF--NFRF 293   302 TNRRHLFKVF---------GI-RFDILV 319
P2X1-Mouse    185 FT-L-----FIK-N--S-IS-F-P--R 197   230 FSLG----YV-V----RESG 240   251 GVVGIT----ID-WE 260   272 P-------IYQ-FHGLY 280   288 GF--NFRF 293   302 TNRRHLFKVF---------GI-RFDILV 319
P2X1-Human    185 FT-L-----FIK-N--S-IS-F-P--R 197   230 FQLG----YV-V----QESG 240   251 GVVGIT----ID-W- 259   272 P-------IYE-FHGLY 280   288 GF--NFRF 293   302 TNYRHLFKVF---------GI-RFDILV 319
P2X2-Rat      183 FT-I-----LIK-N--S-IH-Y-P--K 195   227 FRLG----FI-VE---K-AG 237   248 GVIGVI----IN-WN 257   269 P-------KYS-FRRLD 277   286 GY--NFRF 291   301 TTTRTLIKAY---------GI-RIDVIV 318
P2X3-Rat      171 FT-I-----FIK-N--S-IR-F-PL-- 183   216 LRVG--------D-VVKFAG 226   237 GVLGI---K-IG-W- 245   258 P-------KYS-FTRLD 266   277 GY--NFRF 282   292 SEYRTLLKAF---------GI-RFDVLV 309
P2X3-Human    171 FT-I-----FIK-N--S-IR-F-PL-- 183   216 LRVG--------D-VVKFAG 226   237 GVLGI---K-IG-W- 245   258 P-------KYS-FTRLD 266   277 GY--NFRF 282   292 SEYRTLLKAF---------GI-RFDVLV 309
P2X4-Rat      185 FT-L-----LVK-N-N--IW-Y-P--K 197   230 FRLG----TI-VED----AG 240   251 GIMGI-QIK----WD 260   272 P-------RYS-FRRLD 280   291 GY--NFRF 296   306 KEQRTLIKAY---------GI-RFDIIV 323
P2X4-Human    185 FT-L-----LVK-N-N--IW-Y-P--K 197   230 FRLG----KI-VE----NAG 240   251 GIMGI-QVN----WD 260   272 P-------RYS-FRRLD 280   291 GY--NFRF 296   306 NEQRTLIKAY---------GI-RFDIIV 323
P2X5-Rat      188 FT-I-----SIK-N---FIR-F-P--K 200   232 FRLG----SI-V----RWAG 242   253 GVIGIY----IE-WD 262   274 P-------HYY-FNRLD 282   292 GY--NFRF 297   307 VEFRDLMKAY---------GI-RFDVIV 324
P2X5-Human    188 FT-I-----FIK-N-H--IR-F-P--K 200   231 FRLG----SV-I----RWAG 241   252 GVIGI---N-IE-WN 261   273 P-------HYS-FSRLD 281   291 GY--NFRF 296   306 VEFRTLMKAY---------GI-RFDVMV 323
P2X6-Rat      188 FT-L-----FIK-N--T-VT-F-N--K 200   233 FRIG--------DLVA-MTG 243   254 GAVGI-N---I-HWD 263   275 P-------QYS-FQLQE 283   285 GY--NFRT 290   300 VESRSLLKLY---------GI-RFDILV 317
P2X7-Rat      188 FT-V-----LIK-N-N--I-DF-P-GH 201   229 FRLG--------DIF-QEIG 239   250 GIMGI-E---I-YWD 259   271 P-------KYS-FRRLD 279   290 GY--NFRY 295   304 MEKRTLIKAF---------GV-RFDILV 321
P2X7-Human    188 FT-V-----LIK-N-N--I-DF-P-GH 201   229 FRLG--------DIF-RETG 239   250 GIMGI-E---I-YWD 259   271 P-------KYS-FRRLD 279   290 GY--NFRY 295   304 VEKRTLIKVF---------GI-RFDILV 321

2D predictions (ratio in %)
84% β-sheet strand, 16% α-helix
70% β-sheet strand, 30% no defined structure
98% β-sheet strand
55% β-sheet strand, 4% α-helix, 41% no defined structure
41% β-sheet strand, 32% α-helix, 27% no defined structure
100% β-sheet strand
```

Class II AATS

```
SerRS-Ec      260 MT--AHTP-CFRSEAGSYGRDTRGLIR 283   287 FDKV--E-MV--QIV-R--P 298   320 GL-PY---RKI-IL- 328   335 FG----ACK-T-Y-DLE 344   359 CS--NV-W 363   378 SDKKT--R-LVHTLNGSGLAVGRTLVAV 402
SerRS-Tt      248 YA--GYAP-AFRSEAGSFGKDVRGLMR 271   275 FHKV--E-----QYVLTE-A 286   311 -L-PY---R---LVE 317   325 PGK----WRQ--V-DIE 334   349 CSAL-L-- 353   367 PEGR-V-R-YAYTLNNTALATPRILAML 391
HisRS-Ec      104 LWYIG--P-MFRHE-----RPQKG--R 120   124 FHQLGCE-VFGLQ-G----P 137   247 GIAYTVNQR---LV- 257   259 -G-LDYYNR-TVF---E 269   285 G-----RY 287   297 -------R-A-TPAVGFAMGLERLVLLV 315
AspRS-Tt      215 YFQIA---RCFRDE------DLRAD-R 231   235 FTQLDLE-MS---FV--E-- 246   417 G---F---RFLWVVD 425   460 PGRV----RALAY---D 469   480 GSI-RIH- 485   512 LEA--LEYG-APPHGGIAWGLDRLLALM 536
AspRS-Sc      316 VYEIG--P-VFRAE-NS---NTH---R 332   337 FTGLDME-MA---F--EEH- 349   438 ---DTD---F-YILD 445   464 --KY--SN--S-Y---D 470   481 GAQ-RIH- 486   511 CDG--FSYG-CPPHAGGGIGLERVVMFY 535
GlyRS-Ec      213 IAQIGK---AFRNE-IT-PRNF--IFR 232   236 FEQ--ME-I---EYFVR--P 247   269 GL--S---R--ENL- 275   286 A-HYA---KAT----VD 294   308 G-IAN-R- 312   348 PET-G--KWFVPYVIEPSAGVDRGVLAL 372
PheRS(α)-Tt   196 IVV-P--GRVFRFE-QT---DAT---H 212   216 FHQL--EGLV-V-------G 225   246 G-PDS---K----V- 251   263 PG-------------AQ 266   284 GM---VH- 287   304 P---PAYRG-V-TGFAFGLGVERLAMLR 326
PheRS(β)-Tt   568 ALLFEV-GRVFR-E-----REET---H 584   590 F---G-EGVG-L-------P 597   621 GLAF----R----VE 627   636 PG---VSGRVL----VE 645   653 GA---LH- 656   667 P----V-HLF---ELRLPLP-DKPLAFQ 685
LysRS(U)-Ec   254 VFEIN---RNFRNEGIS-VR------H 270   274 FTMM--E-LYMA-YA----- 284   386 ---PT------FITE 391   407 P---EITDRF------E 414   425 GFS-ELN- 430   464 LEH-G-----LPPTAGLGIGIDRMVMLF 485

X-ray analyses
100% β-sheet strand
α-phosphate, γ-phosphate, adenosine (stacking), Mg2+, ribose, γ-phosphate
```



βI

In the β_I region the sequences of the P2X receptor family are quite well conserved and in particular K190 and N191 are 100% conserved (numbering refers to human $P2X_1$ receptor sequence). These residues align with a conserved R268 (numbering refers to SerRS-Ec) known to bind the α-phosphate of ATP across the synthetases and a conserved glutamate residue (E270). It is suggested that the role of the conserved arginine residue in the synthetases is fulfilled by K190 in the P2X receptor family. R283 of SerRS-Ec is aligned across the synthetase family with arginine or histidine residues and participates in the binding of the γ-phosphate part of ATP. In the P2X receptor family it is also aligned with arginine (or lysine) or histidine residues and the function of these residues are inferred from that of the synthetases.

βII

The similarities between the P2X receptor family and the synthetases are more obvious in the first part of the β_{II} region than the second. A ^{230}FRLG^{233} ($hP2X_1$ receptor numbering) motif is highly conserved in the P2X receptor family and a similar motif, FHQLG is found in the synthetases. In the synthetases the conserved phenylalanine residue F287 (numbering refers to SerRS-Ec) is responsible for stacking interactions with the adenine ring of the ATP molecule. Between β_{II} and β_{III} in all tertiary structures of class II aminoacyl-tRNA synthetases an α-helix structure is found. This structure is also predicted for the corresponding position of the P2X family.

βIII and βIV

In the β_{III} region, there is no notable residue conservation in the P2X receptor family that is conserved across the synthetases. However, the general character of this region seems to be similar for both groups. There is a similar scenario in the β_{IV} region, where the P2X receptor proteins show more homology than the synthetases. The acidic residue at the end of the region has an important function complexing the magnesium ion of an $[ATP\text{-}Mg]^{2-}$ which is the actual substrate for the enzyme and probably the agonist at the P2X receptor channels. An acidic residue (aspartic acid) is found in the corresponding position of the P2X family, with the exception of the $P2X_1$ receptor.

βV

Unlike the synthetases the β_V region is well conserved across the P2X receptor family. A glycine residue is conserved across both families (with the exception of seryl-tRNA

synthetase) in the first position of β_V. Also the conserved R292 (numbering refers to $hP2X_1$ receptor sequence) in the P2X receptor family is aligned to positively charged arginine or histidine residues (in the majority of cases) in the synthetases.

βVI

Freist claims that both protein families exhibit relatively conserved sequences and that both families correspond well to each other in the β_{VI} region. A glycine residue is well conserved across both families (G312 in the human $P2X_1$ receptor) and so is an arginine residue (R314 in the human $P2X_1$ receptor). This conserved arginine residue in the aminoacyl-tRNA synthetases interacts with the γ-phosphate part of the ATP molecule. This function could be easily performed by R314 of the P2X receptor family. Also, hydrogen bonds are formed with the ribose part of ATP in the synthetases by a conserved glycine residue (G392 of SerRS-Ec). Freist concludes that if P2X receptor proteins are related to class II aminoacyl-tRNA synthetases then a high sequence homology cannot be expected but more similarities in tertiary and secondary structural elements can be expected.

5.2.3 Testing Freist's hypothesis and the 3D model of the $rP2X_4$ receptor binding site

The model (Freist *et al.*, 1998) for P2X receptor channels, based on class II aminoacyl-tRNA synthetases, was tested by mutagenesis of the rat $P2X_4$ receptor (Yan *et al.*, 2005). Using crystal structures of several class II aminoacyl-tRNA synthetases as templates, a homology model of K180-K326 ectodomain region of rat $P2X_4$ was produced (Yan *et al.*, 2005). The mutagenesis focussed on those residues Freist identified as interacting with the adenine ring (F230) and phosphates (K190, K197, R318) of the ATP molecule and Mg^{2+} (D280) (all numbering refers to the rat $P2X_4$ receptor in Freist's sequence alignment). Mutation of three residues - K190, F230 and D280 - severely affected the function of the rat $P2X_4$ receptor, validating Freist's model. However, experimental data did not support a role for K197 in ATP action and also queried the role for R318 in ATP binding at the receptor.

With experimental evidence to support part of Freist's model, a 3D model of rat $P2X_4$ (K180-K326) receptor was generated using the SWISS-MODEL server (Schwede *et al.*, 2003) with the following crystal structures as templates (Belrhali *et al.*, 1994; Arnez *et al.*, 1997; Arnez *et al.*, 1999): seryl-tRNA synthetase (1SES), histidyl-tRNA synthetase (1KMN) and glycyl-tRNA synthetase (1B76). The model was based on Freist's alignment and adopted a six stranded β-sheet structure and contained an ATP-binding site. The ATP-binding pocket

incorporated three residues (D280, F230, K190) indicated by Freist's model as important for ATP binding and confirmed by mutagenesis to be involved with ATP-binding at the rat $P2X_4$ receptor. In addition, the model implies a role for the β_{IV} strand residue R278 interacting with the γ-phosphate and H286 located at the beginning of the β_V strand interacting with the β-phosphate group of ATP. R318, identified by Freist for its importance in binding to the γ-phosphate of ATP, is incorporated into the model but could be in contact with the deoxyribose ring of ATP instead. The model includes two disulphide bonds, C217-C227 and C261-C270 which were manually restrained. However, both pairs of cysteine residues were in close proximity so no structural changes were made to create the disulphide bonds indicating further that the 180-326 fragment of the rat $P2X_4$ receptor adopts the suggested β-pleated sheet fold (Figure 5.4). The model was experimentally validated by mutagenesis studies.

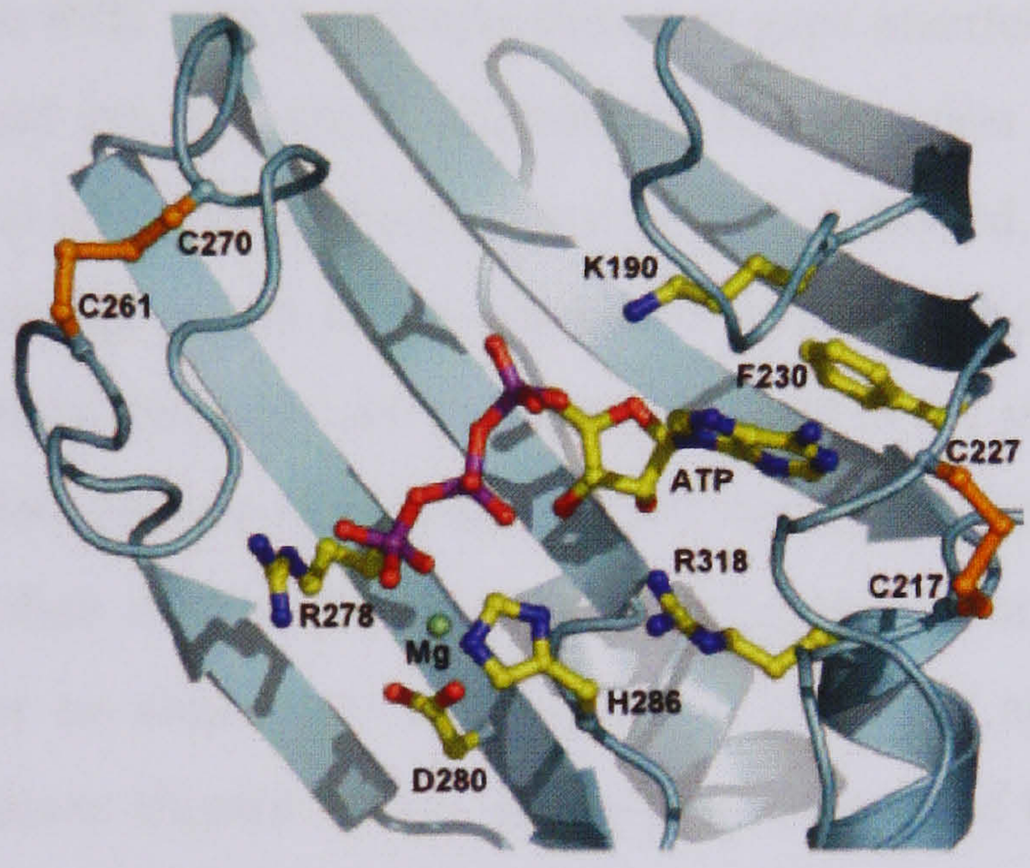


Figure 5.4 ATP-binding site of $rP2X_4$ receptor. 3D model of rat $P2X_4$ (K180-K326) receptor was generated using the SWISS-MODEL server (Yan *et al.*, 2005). The ATP binding site adopts a six-stranded β-sheet structure. ATP is shown in yellow. The model includes two disulphide bonds (shown in orange). Residues that interact with the ATP molecule are shown in yellow (K190, F230, R278, D280, H286 and R318). *Figure taken from Yan.Z.,Liang. et al. (2005) Mol Pharmacol 67 1078-88.*

5.2.4 The suitability of the class II aminoacyl-tRNA synthetase family as templates for the P2X receptor family

The sequence homology of the extracellular loop of the P2X receptor family to the class II aminoacyl-tRNA synthetases is based on sequence alignments and secondary structure predictions (Freist *et al.*, 1998). Friest's alignment considers the catalytic domain of the class II aminoacyl-tRNA synthetases and the second half of the extracellular ligand-binding domain of the P2X receptor family (F185-V319, $P2X_1$ receptor numbering). This removes the

possible role of positively charged K68 ($P2X_1$ receptor numbering) in ATP binding (Ennion *et al.*, 2000) and its corresponding residue K69 in the $P2X_2$ receptor, thought to be important in ATP potency (Jiang *et al.*, 2000) as predicted by experimental studies. Yan *et al.* (Yan *et al.*, 2005) based their model of the $P2X_4$ receptor on Freist's sequence alignment and thus only models a fragment of the $P2X_4$ receptor ectodomain (K180-K326). However it has been shown that insertion of the $P2X_3$ receptor ectodomain in the backbone of the $P2X_{2a}$ receptor and the $P2X_{2b}$ receptor effectively transfers the ligand-binding properties of the $P2X_3$ receptor to chimeric channels (in terms of potency of agonist), indicating the relevance of both ectodomain halves in ligand binding (Zemkova *et al.*, 2004).

The overall quality of the sequence alignment is poor (Freist *et al.*, 1998). Consequently the quality of any homology model based on this alignment will be compromised. Of the six predicted β-strands of the P2X receptor family there are gaps inserted in the middle of the β-strand regions β_I, β_{IV} and β_{VI} in Freist's alignment. The insertion of gaps into secondary structural elements should be avoided. If a β-strand residue is deleted, a corresponding residue must be lost from its pairing strand or an alternative hydrogen bonding network must be created, making it less likely that gaps can be accommodated in the structure. It is more likely to see insertions and deletions in coil regions as their hydrogen bonding leads to the easier insertion of deletion of their residues. Therefore, aligning the secondary structure elements separately could improve an alignment. In this case, a structural alignment of the class II aminoacyl-tRNA synthetases aligned with a sequence alignment of the P2X receptor family may produce a better quality alignment and thus be more appropriate.

The secondary structure of the class II aminoacyl-tRNA synthetases is based on the crystal structures. The secondary structure prediction of the P2X receptor family was made using published algorithms with ~70% accuracy for the individual channel proteins (Chou & Fasman, 1978; Garnier *et al.*, 1978; Kneller *et al.*, 1990; Rost & Sander, 1993; Salamov & Solovyev, 1997). However, using newer secondary structure prediction techniques (McGuffin *et al.*, 2000) and PROF (Rost & Sander, 1993) current studies have identified possible inconsistencies. The second β-strand predicted by Freist for the P2X receptor sequences (region F230-G240 in the human $P2X_1$ receptor sequence) aligned with the β_{II} strand of motif 2 in the synthetases is instead predicted to be an α-helix with relatively high confidence.

The model of the putative ATP-binding domain of the $P2X_4$ receptor proposed by Yan *et al* (Yan *et al.*, 2005) does not incorporate all the experimental data for the $hP2X_1$ receptor.

Firstly, positive residue K68 which is thought to be involved in ATP-binding at the $P2X_1$ receptor (Ennion *et al.*, 2000) is not considered because the model only includes the second half of the extracellular domain, thus excluding the role of conserved residue K68. The first half of the extracellular domain and TM1 of the P2X receptor family have been shown to be involved in signal transduction and gating. Mutant rat $P2X_2$ receptor channels demonstrated that both TM1 and TM2 participate in conformational changes that occur during receptor activation (Li *et al.*, 2004) and further studies on the $P2X_2$ receptor channel suggests that motion at the interface between TM1 and TM2 of neighbouring subunits cause changes in channel permeability (Khakh & Egan, 2005). In addition, tryptophan scanning mutagenesis of TM1 and TM2 in the $rP2X_4$ receptor suggests that both regions are helical in structure and that the outer parts move during receptor activation (Silberberg *et al.*, 2005). Clearly, the first half of the extracellular domain of P2X receptors including TM1 cannot be excluded in structural models of the P2X receptor family. The homology model of the $rP2X_4$ recpetor based on the synthetases (Yan *et al.*, 2005) is restricted due to the alignment on which it is based. Models of the first half of the ectodomain and two TM domains are yet to be produced (Yan *et al.*, 2006).

K190, which is also 100% conserved across the P2X receptor family, is predicted by the Freist model to bind the α-phosphate of the ATP molecule. However, when mutated to alanine in the $P2X_1$ receptor, it showed only a small effect on ATP-potency and the time-course of response and no effect on suramin sensitivity (Ennion *et al.*, 2000) suggesting it does not play a role in ATP-binding. In the $rP2X_4$ receptor model K190 is far from the phosphate group of ATP although experimental results revealed a rightward shift in the sensitivity of this mutant for ATP and a rescue of receptor function by the mutant K190R (Yan *et al.*, 2005). Minor changes to the model accounts for the coordination of the α phosphate of ATP by K190 (Yan *et al.*, 2006).

The acidic residue at the end of the β_{IV} region was suggested by Freist, and supported by the Yan model and experimental studies, to interact with the Mg^{2+}. However, for the $P2X_1$ receptor this residue is not conserved. In the synthetases conserved residue F287 (numbering refers to SerRS-Ec) is responsible for stacking interactions with the adenine ring of the ATP molecule and in Freist's alignment this residue aligns with F230 of the $P2X_1$ receptor. It is suggested that F230, which is conserved across the P2X family except the $P2X_3$ receptor, plays a similar role binding to the adenine ring of ATP in the P2X receptor family. However, for the $P2X_1$ receptor when F230 was mutated to alanine, no change was seen for the ATP

potency or time course of the response. The corresponding residue in the $P2X_2$ receptor, F227, when mutated to hydrophobic residues leucine and isoleucine resulted in the loss of ATP responsiveness (Nakazawa *et al.*, 2004). Interestingly, the corresponding residue to F227 in the $P2X_3$ receptor is a leucine. Five residues, I226, F227, R228, L229 and G230 were mutated in total in the $P2X_2$ receptor and the results obtained show that the residues mutated at odd number positions (i.e. 227 and 229) more dramatically reduced the ATP-responsiveness than at even number positions. Nakazawa *et al* (2004) suggests that if these residues are involved in a β-sheet structure the residues aligning in one side (F227 and L229) may be more influential on the channel function of P2X receptors than those in the other side (I226, R228 and G230).

The lysine at position 313 in the $rP2X_4$ receptor and corresponding lysine residues in the $P2X_1$ (Ennion *et al.*, 2000), $P2X_2$ (Jiang *et al.*, 2000) and $P2X_7$ (Worthington *et al.*, 2002) receptors are well characterised by mutagenesis studies and K313 ($rP2X_4$ receptor numbering) has been identified as critical for ATP-binding (Yan *et al.*, 2005). However, mutant K313R of the $rP2X_4$ receptor (Yan *et al.*, 2006) did not show a significant rescue effect unlike arginine rescue mutants in the $P2X_1$ and $P2X_2$ receptors ((Ennion *et al.*, 2000) and (Jiang *et al.*, 2000) respectively) at the corresponding lysine residues. This suggests that in the $rP2X_4$ receptor this positive residue is not required for ATP-action at the receptor by coordinating the negatively charged phosphate groups of ATP as previously suggested (Yan *et al.*, 2006). The $rP2X_4$ receptor 3D model also does not support a role for K313 in coordination of the α phosphate of ATP due to its close proximity to the 2'-hydroxyl group of the ribose component of ATP. It has been suggested that K313 of the $P2X_4$ receptor may participate in the formation of the intersubunit ATP binding pocket but does not directly contribute to the recognition of the ATP molecule (Yan *et al.*, 2006).

In summary, the similarity of the class II aminoacyl-tRNA synthetases with a fragment of the ectodomain of P2X receptors (Freist *et al.*, 1998) although poor cannot be dismissed. The Yan model (Yan *et al.*, 2005) presents a possible structure for the ATP binding site of the $P2X_4$ receptor validated by mutagenesis experiments. However, this model cannot be directly translated into the ATP binding site of the other members of the P2X family because of the inconsistencies that are presented by the experimental data.

5.3 Mutagenesis assesses whether the model of ATP-binding at the rP2X$_4$ receptor can be transferred to the hP2X$_1$ receptor

The 3D model of the ATP binding site of the rP2X$_4$ receptor based on class II aminoacyl-tRNA synthetases, cannot be directly translated to other members of the P2X receptor family because of the conflicts raised by published mutagenesis data (as previously discussed in section 5.2). This is certainly true for the human P2X$_1$ receptor. However, as a family of receptors it is expected that members of the P2X receptor family should share some similarity in terms of structure and ligand binding site and areas of conserved residues are often functionally important. Therefore models generated for any P2X receptor subunit could provide important information for the other family members and cannot be ignored. A set of five residues (K190, F230, R278, D280 and H286 rP2X$_4$ receptor numbering) have been identified by the rP2X$_4$ receptor 3D model (Yan *et al.*, 2005) to be involved in ATP binding at the receptor. Therefore, the corresponding residues in the human P2X$_1$ receptor may play a role in ATP action at the hP2X$_1$ receptor.

In this study, a multiple sequence alignment for the extracellular domain of the P2X receptor family (from human and rat) has been generated by ClustalW (Thompson *et al.*, 1994) and manually adjusted using Seaview (Galtier, 1996) to ensure a reliable alignment was obtained. Those residues of the hP2X$_1$ receptor that correspond to residues K190, F230, R278, D280 and H286 of the rP2X$_4$ receptor sequence have been mutated to cysteine to investigate the effect on channel properties (Figure 5.5). The five mutant hP2X$_1$ receptor channels generated were K190C, F230C, G278C, Y280C and K283C of which K190C and F230 are 100% conserved across the P2X receptor family. The glycine in position 278 of the hP2X$_1$ receptor is not conserved across the P2X receptor family. The negative charge conserved in all members of the P2X receptor family at position 280 is interestingly not conserved in the P2X$_1$ receptor where there is a tyrosine residue in its place. The positive charge in position 283 of the hP2X$_1$ receptor (position 286 in the rP2X$_4$ receptor) is conserved across the P2X receptor family with the exception of the P2X$_7$ receptor. Cysteine was chosen to replace these five residues (K190, F230, G278, Y280 and K283) in the hP2X$_1$ receptor because it is known to be well tolerated and will allow probing with MTS reagents, which can provide important structural insight.

Figure 5.5 Multiple sequence alignment of residues 159-330 ($hP2X_1$ receptor numbering) of the extracellular domain human and rat P2X receptor family. The rat $P2X_4$ receptor sequence and human $P2X_1$ receptor sequence are highlighted in bold. Residues highlighted in red are those residues identified by the 3D model $rP2X_4$ receptor and the corresponding residues in the P2X receptor family members. Those residues in red in the human $P2X_1$ receptor sequence were mutated to cysteine to determine the effect on channel properties. Numbering refers to the rat $P2X_4$ receptor sequence.

5.3.1 Effect of cysteine replacement of K190, F230, G278, Y280 and K283 in the hP2X$_1$ receptor on the response to ATP

ATP evoked a concentration-dependent desensitizing inward current with an EC_{50} of ~0.95μM at wild-type hP2X$_1$ receptors (Figure 5.6a) (Table 5.1) similar to that previously described (Ennion *et al.*, 2000). For the majority of mutants (F230C, G278C and Y280C) no effect was seen on ATP potency or time course of the response (Figures 5.6c and 5.8). A small yet significant decrease in ATP potency was recorded for mutants K190C and K283C (~6-fold and ~8-fold respectively) (Figure 5.6a, c). This was accompanied by a slowing in the time course of the response for mutant K190C (Figures 5.6b and 5.8). Mean peak amplitudes were not significantly different for any of the mutants compared to WT (Figure 5.7).

5.3.2 Accessibility of cysteine residues to MTSEA-Biotin

An MTSEA-Biotin assay was used to detect accessible free cysteine residues in WT and mutant P2X$_1$ receptors. WT P2X$_1$ receptors showed no biotinylation (Figure 5.9b). This demonstrates that all cysteine residues in the extracellular segment of P2X$_1$ receptors are unavailable either because they are disulfide bonded to another cysteine residue and/or they reside in an area of the molecule that is inaccessible to MTSEA-Biotin. By introducing cysteine residues at a particular position using mutagenesis a 'free cysteine' is created as it is not part of a disulfide bond. Using an MTSEA-Biotin Assay, information regarding the secondary structure of a protein and the accessibility of a region can be assessed. Trimeric channels, dimers and monomers were detected in the total WT and mutant P2X$_1$ receptor lysates at similar levels (Figure 5.9a). A band of ~55kDa represents the monomer and the trimer is triple the molecular weight ~165kDa. The presence of bands other than the monomer suggests that the secondary structure has not been sufficiently broken down by detergent SDS and β-mercaptopethanol. Biotinylation of a free cysteine residue was detected for mutants K190C, F230C, G278C and Y280C (Figure 5.9b). This suggests that the cysteine residue introduced at positions 190, 230, 278, 280 is accessible by MTSEA-Biotin and not buried. The band seen for mutant F230C is detected at a smaller molecular weight than all the other mutants. This could indicate that mutant F230C has in some way disrupted the nearby glycosylation site at position Asn-184 (North, 2002).

5.3.3 Effect of MTS reagents

Methanethiosulfonates (1mM) were bath-perfused and co-applied with ATP for a three second application. Positively charged MTSEA and negatively charged MTSES had no effect on the currents elicited by WT $P2X_1$ receptors and EC_{50} values recorded (Table 5.2) (Figure 5.10). Mutants F230C, G278C, Y280C and K283C were not significantly affected by the application of MTSEA or MTSES (Figure 5.11). MTSEA (1mM) significantly potentiated currents recorded from mutant K190C whereas currents were significantly inhibited with application of MTSES (1mM) (Figure 5.12). MTS reagents are fast acting and an overnight incubation of injected oocytes should ensure the MTS compound has covalently bonded to any available free cysteines. Evidence to suggest that a shorter incubation period is sufficient and that receptor recycling may affect the MTS compound successfully binding is inconclusive. Recordings from K190C following an overnight incubation 1mM MTSEA resulted in a decrease in ATP potency and a leftward shift in concentration dose response curve. Overnight incubation of mutant K190C in 1mM MTSES resulted in an increase in ATP potency and shifted the concentration dose response curve further to the right (Figure 5.13). These results indicate the importance of the positive charge in position 190 to ATP action at the $hP2X_1$ receptor.

Mutant	ATP EC_{50} (μM)	ATP pEC_{50}	Peak (nA)	Rise Time (ms)	Decay Time (ms)
$P2X_1$ WT	0.95	6.02 ± 0.12	8073 ± 1341	94 ± 16	870 ± 115
K190C	5.91	5.23 ± 0.11***	9203 ± 278	135 ± 14	1464 ± 112**
F230C	1.22	5.91 ± 0.05	7695 ± 547	87 ± 13	544 ± 77
G278C	0.72	6.14 ± 0.12	6668 ± 928	130 ± 24	779 ± 125
Y280C	0.74	6.13 ± 0.15	9028 ± 674	90 ± 7	751 ± 69
K283C	7.4	5.13 ± 0.10***	6580 ± 1425	118 ± 20	666 ± 92

Table 5.1 Summary of data for ATP action at $P2X_1$ WT and cysteine mutants

EC_{50} data are representative of data collected from three to seven oocytes. pEC_{50} is –log10 of the EC_{50}. Peak data was collected on the first application of a maximal concentration of ATP (n = 3-7). Rise Time represents the time from 10 to 90% of peak current after an application of a maximal stimulating concentration of ATP (n = 3-7). Decay Time represents the time from 100 to 50% decay of peak current achieved after 3s of a maximal stimulating concentration of ATP (n = 3-6). Values shown are representative of the mean ± S.E. of the mean.

** $P < 0.01$, different from WT as measured by one-way ANOVA.

*** $P < 0.001$, different from WT as measured by one-way ANOVA.

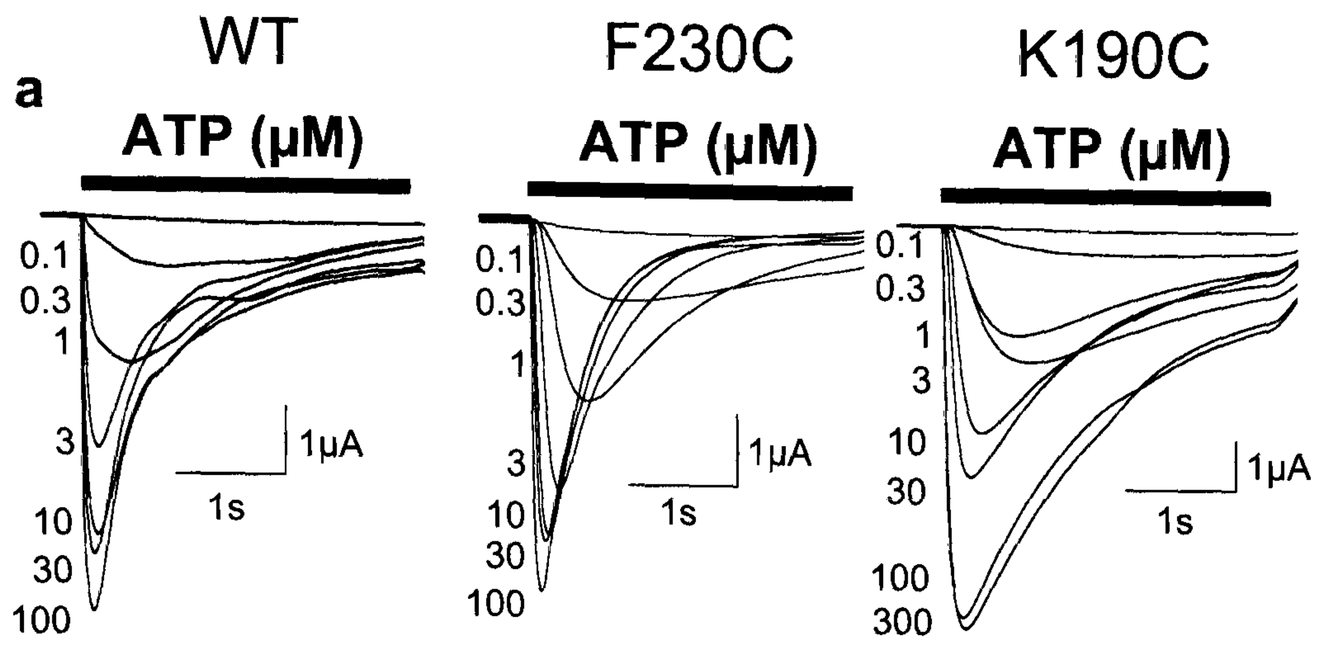


Figure 5.6 Effect of $P2X_1$ cysteine replacement mutants on ATP potency. $P2X_1$ WT or cysteine mutants were expressed in oocytes and ATP potency tested using a two-electrode voltage clamp. **a**, currents recorded in response to ATP at $P2X_1$ WT and mutants F230C and K190C. **b**, currents recorded in response to ATP at $P2X_1$ WT and mutant K190C. A slowing in the time course of the response was recorded for mutant K190C. **c**, concentration response curves for $P2X_1$ WT and mutants K190C, F230C, G278C, Y280C and K283C. Mean currents were normalized to the maximum response (n = 3-7). Holding membrane potential was –60 mV.

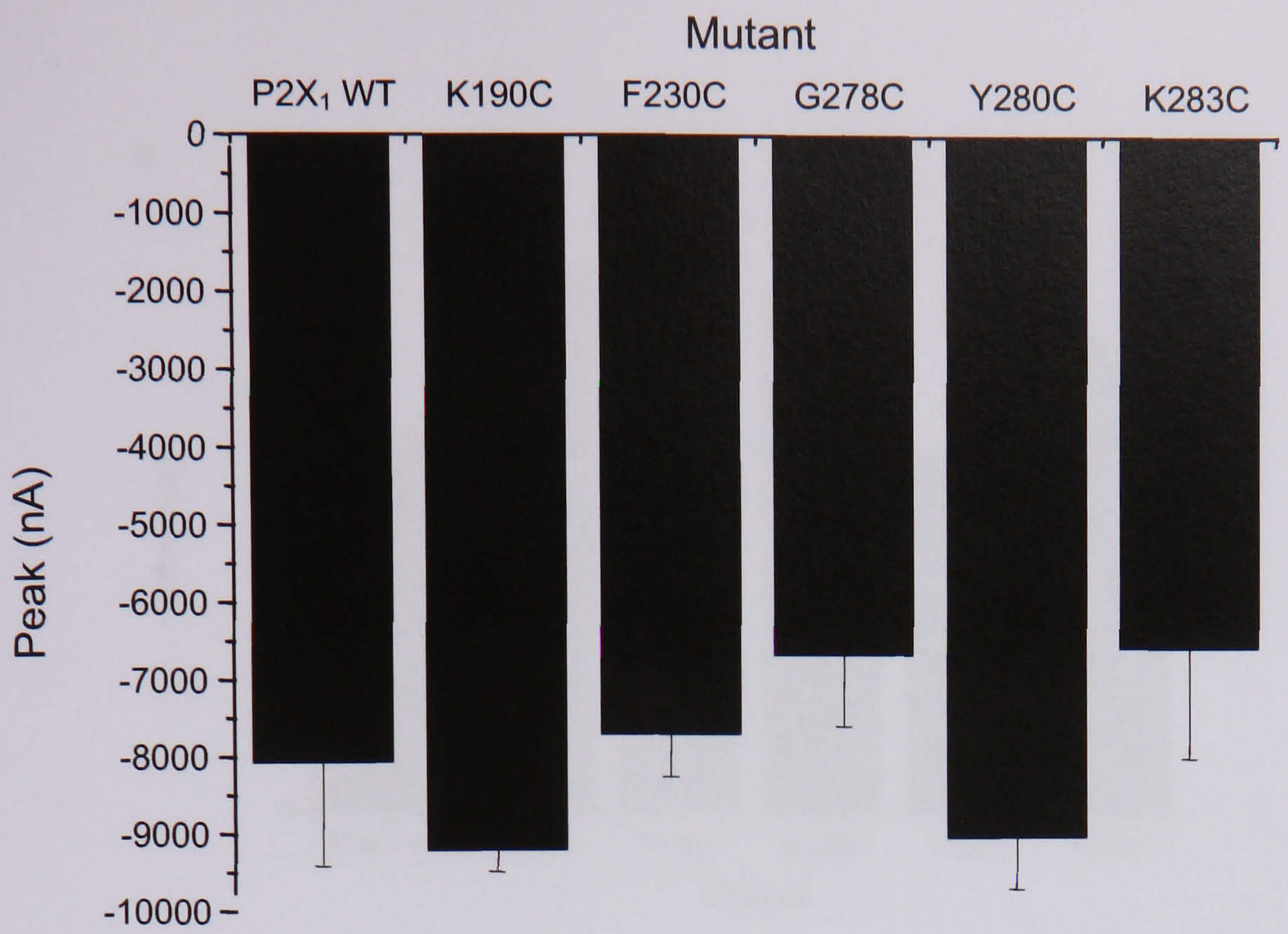


Figure 5.7 Mean peak current of $P2X_1$ WT and cysteine mutants. WT or mutant $P2X_1$ receptor channels were expressed in oocytes and the peak current after an initial maximum concentration of ATP was applied was recorded. The mean peak data is shown (n = 3-7). The mean peak data collected for $P2X_1$ cysteine mutants was not significantly different to $P2X_1$ WT.

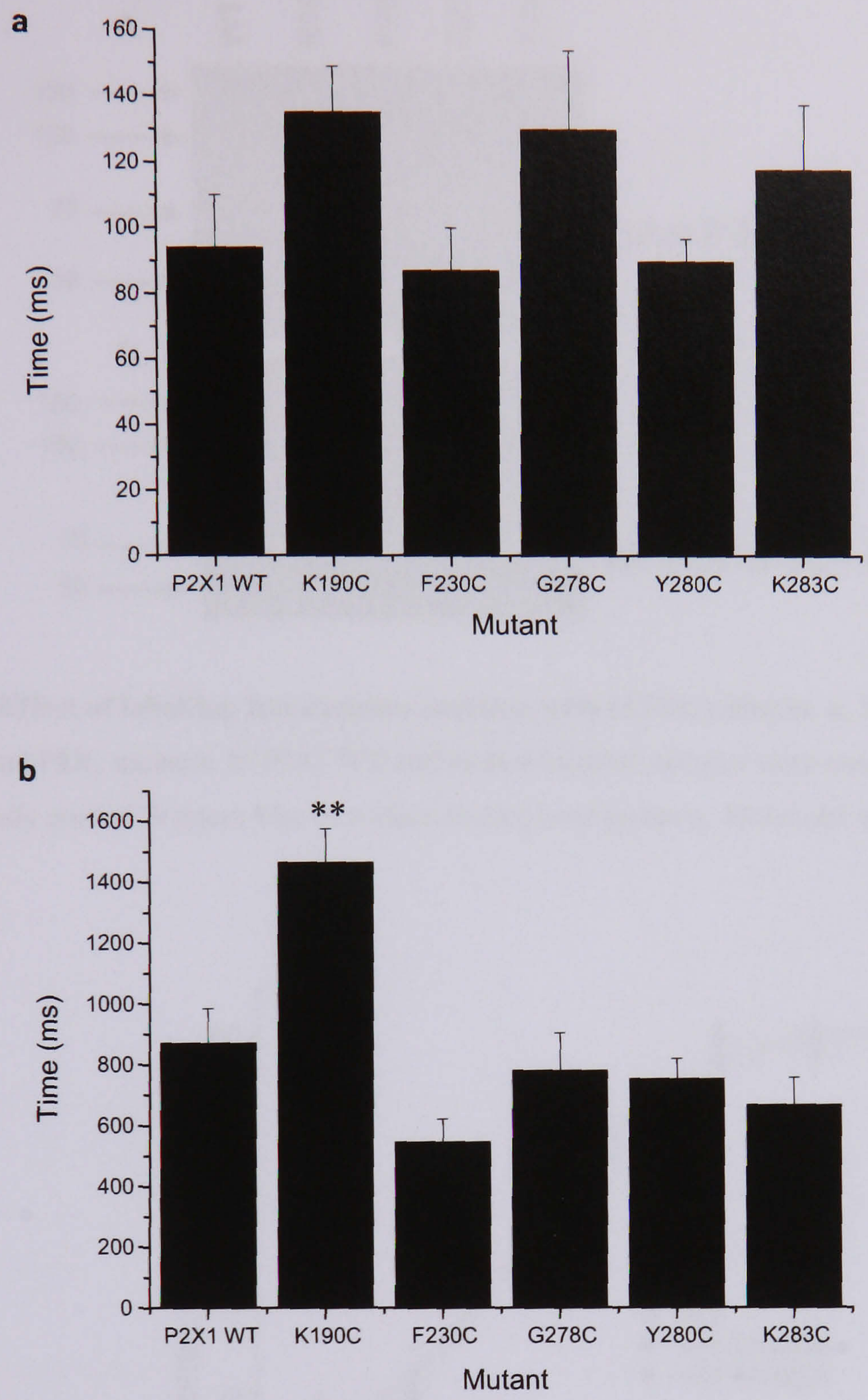


Figure 5.8 Rise Time and Decay Time $P2X_1$ WT and cysteine mutants. WT or mutant $P2X_1$ receptor channels were expressed in oocytes. **a**, the rise-time (time from 10 to 90% of peak current after an initial application of a maximal concentration of ATP) and **b,** decay-time (time from 100 to 50% decay of peak current) was recorded. The mean data is shown (n = 3-7) ± S.E of the mean.

** $P < 0.01$, different from WT as measured by one-way ANOVA.

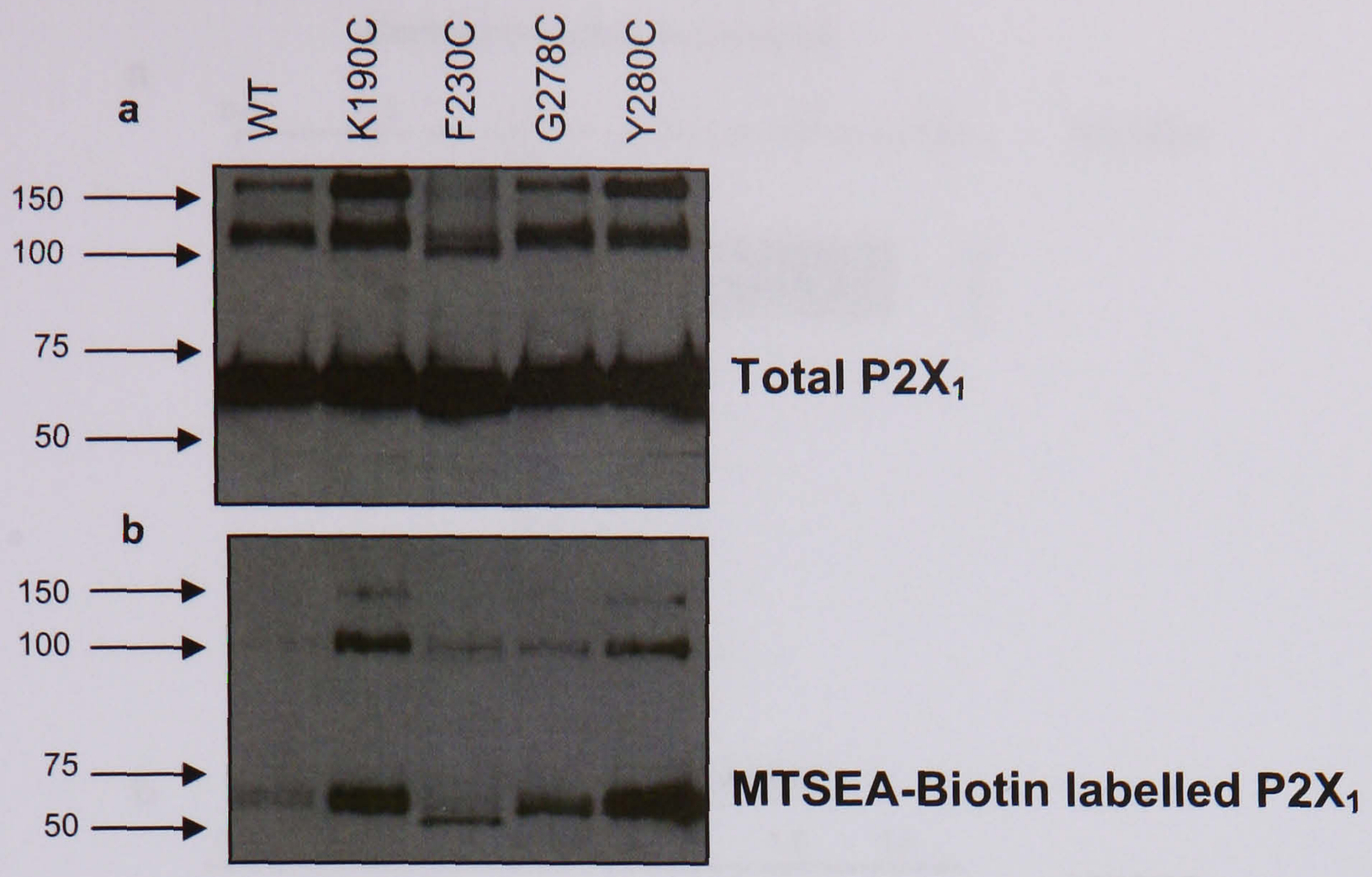


Figure 5.9 Effect of labelling free cysteine residues with MTSEA-Biotin. a, Total protein Western blot of $P2X_1$ WT and $P2X_1$ mutants. **b**, $P2X_1$ WT and mutant injected oocytes were incubated in 1mM MTSEA-Biotin. $P2X_1$ antibody probed Western blot of surface biotinylated proteins. Molecular weight shown in kDa.

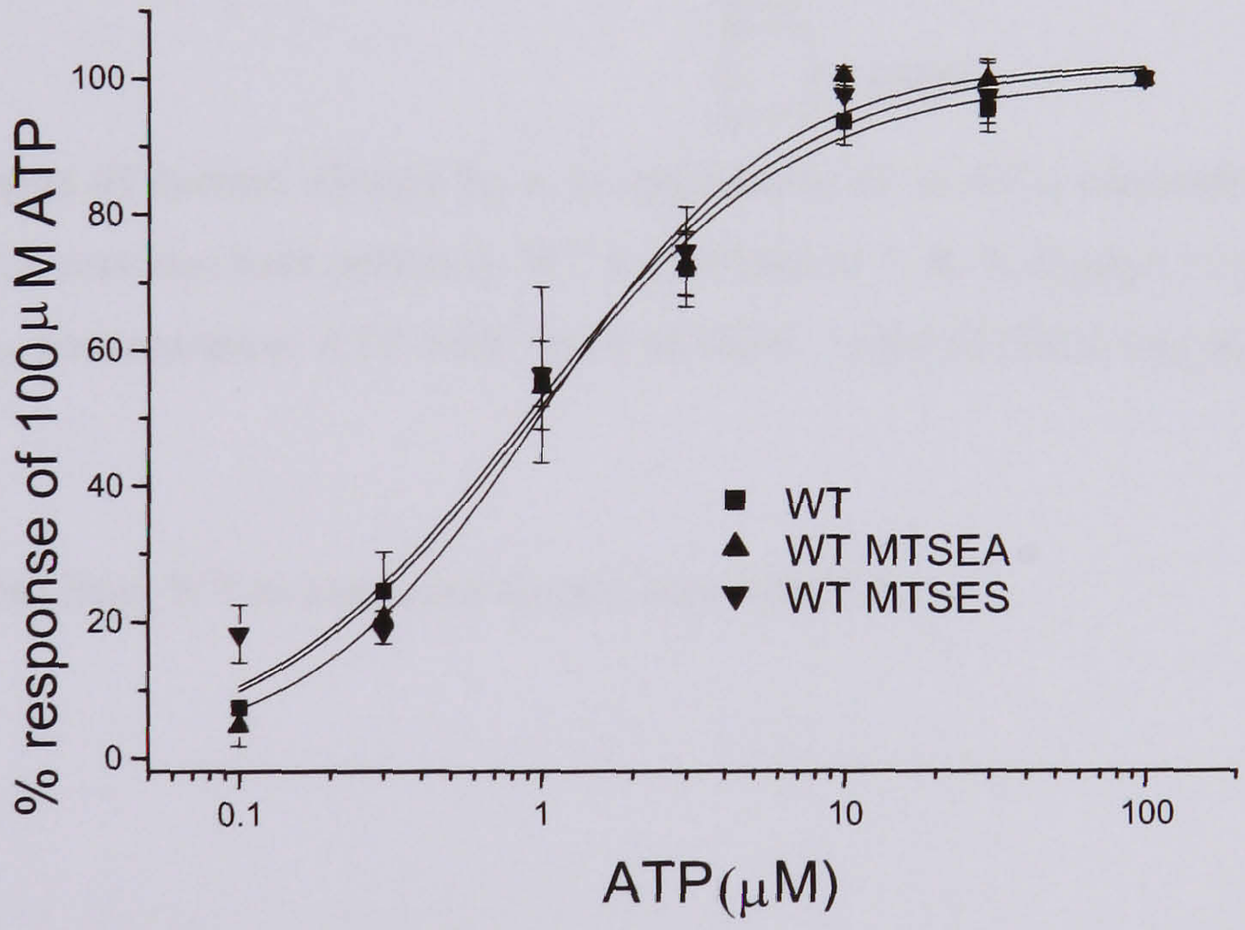


Figure 5.10 Effect of MTSEA and MTSES $P2X_1$ WT on ATP potency. $P2X_1$ WT were expressed in oocytes and incubated overnight in 1mM MTSEA or MTSES. ATP potency was tested using a two-electrode voltage clamp. Concentration response curves for $P2X_1$ WT, $P2X_1$ WT+MTSEA, $P2X_1$ WT+MTSES. Mean currents were normalized to the maximum response (n = 3-4). Holding membrane potential was –60 mV.

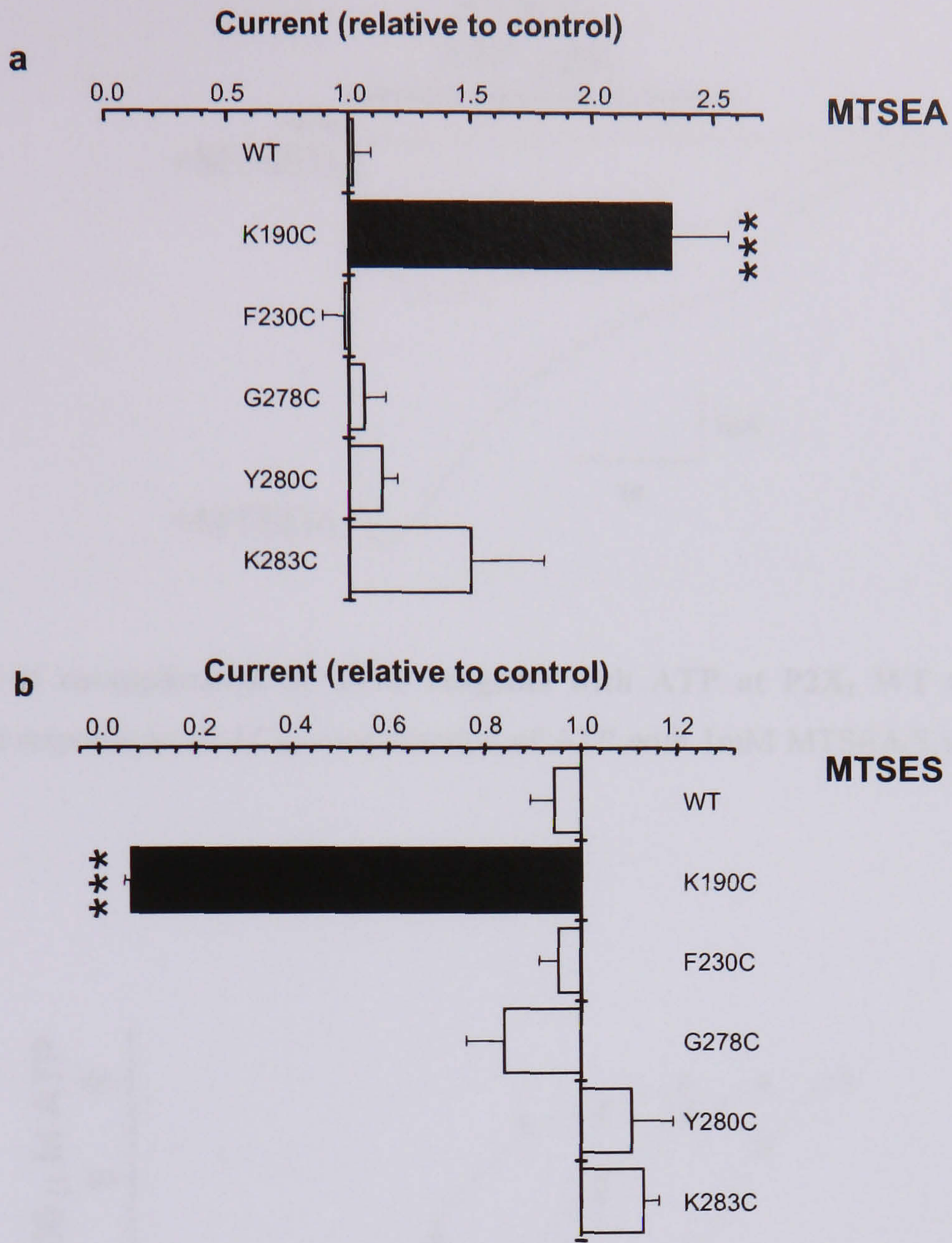


Figure 5.11 a, % change of current elicited by a 3s application of an EC_{50} concentration ATP with 1mM MTSEA. 1mM MTSEA was also bath perfused. WT normalised to 1. **b**, % change of current elicited by a 3s application of an EC_{50} concentration ATP with 1mM MTSES. 1mM MTSES was also bath perfused. WT normalised to 1.

*** $P < 0.001$, different from WT as measured by one-way ANOVA.

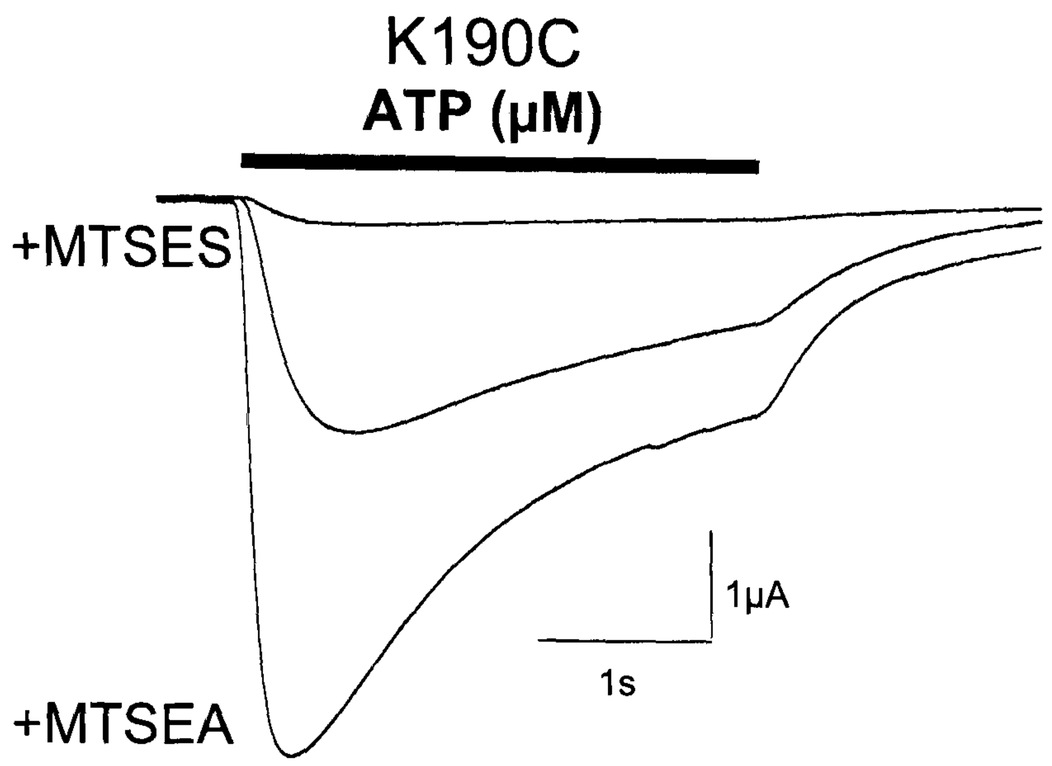


Figure 5.12 Effect of co-application of MTS reagents with ATP at $P2X_1$ WT and mutant receptors. Currents recorded in response to an EC_{50} concentration of ATP with 1mM MTSEA/S at $P2X_1$ mutant receptor, K190C.

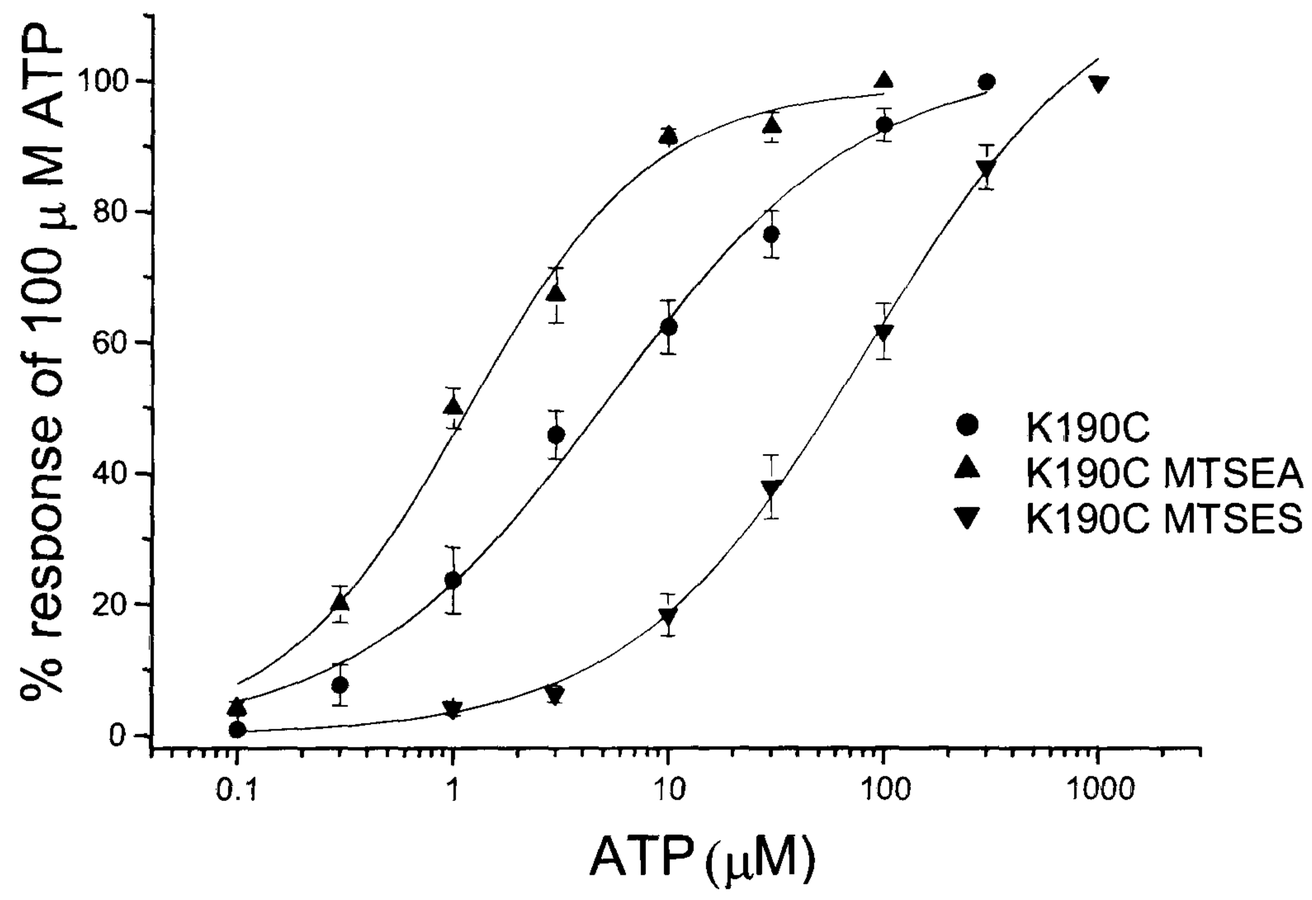


Figure 5.13 Effect of MTSEA and MTSES K190C on ATP potency. Mutant K190C were expressed in oocytes and incubated overnight in 1mM MTSEA or MTSES. ATP potency was tested using a two-electrode voltage clamp. Concentration response curves K190C+MTSEA and K190C+MTSES. Mean currents were normalized to the maximum response (n = 3-4). Holding membrane potential was –60 mV.

Mutant	MTSEA	MTSES
	% relative to control	% relative to control
$P2X_1$ WT	102 ± 7	94 ± 5
K190C	233 ± 24***	7 ± 1***
F230C	98 ± 9	95 ± 4
G278C	106 ± 9	84 ± 8
Y280C	114 ± 6	111 ± 9
K283C	150 ± 30	113 ± 3

Table 5.2 Summary of data for effect of MTS reagents at cysteine mutants relative to $P2X_1$ WT

Data represents % relative to control (n=3-5). Values shown are representative of the mean ± S.E. of the mean.

*** $P < 0.001$, different from WT as measured by one-way ANOVA.

5.3.4 Discussion

The sequence and secondary structure similarity between the catalytic domain of class II aminoacyl-tRNA synthetases and the region K^{180}-K^{326} of $rP2X_4$ receptor (Freist *et al.*, 1998) has been used to produce a 3D homology model of the $rP2X_4$ receptor binding site. The model identified five residues that may play a role in ATP-binding at the $rP2X_4$ receptor and this has been validated using site-directed mutagenesis. It is suggested that D280 coordinates ATP binding via the magnesium ion, F230 coordinates the binding of the adenine ring of ATP and K190, H286 and R278 coordinate the action of the negatively charged α, β and γ phosphate groups respectively (Yan *et al.*, 2005). The aim of this study was to investigate the role of these residues in ATP-binding at the human $P2X_1$ receptor. The corresponding residues in the $hP2X_1$ receptor to K190 (K190), F230 (F230), R278 (G278), D280 (Y280) and H286 (K283) in the $rP2X_4$ receptor were individually replaced by a cysteine residue in the $hP2X_1$ receptor and the effect on the channel properties investigated (bracketed residues are $hP2X_1$ receptor residues).

Alanine replacement mutagenesis of residues thought to be involved in ATP binding at the $rP2X_4$ receptor (K190A, F230A, R278A, D280A and H286A) were characterised in HEK293 cells expressing WT and mutant rat $P2X_4$ receptors. ATP potency was very significantly reduced in mutants K190A, F230A, R278A and D280A with EC_{50} values in excess of 5000μM compared to an EC_{50} 3μM recorded for WT $rP2X_4$ receptors. This decrease in ATP potency was accompanied by a significant decrease in peak current amplitude for mutants K190A, F230A, R278A and D280A. Mutant H286A did not significantly reduce ATP potency compared to WT (EC_{50} 3.7μM compared to EC_{50} 3μM respectively) but a significant decrease in peak current amplitude compared to WT was recorded (Yan *et al.*, 2005).

In contrast, the majority of mutants, F230C, G278C and Y280C had no effect ATP potency at the $P2X_1$ receptor or time course of the response and suggests that mutants F230C, G278C and Y280C do not play a major role in ATP action at the $hP2X_1$ receptor (Table 5.1). Charged MTS reagents had no effect on the current elicited by an EC_{50} concentration of ATP relative to the control. Amino acids phenylalanine, glycine and tyrosine are not charged residues and introducing a charge at this position via the use of charged MTS reagents did not have an affect on the ability of the channel to bind ATP. This confirms that charge is irrelevant in positions 230, 278 and 280 but also that the increase in size/ bulk introduced by MTSEA and MTSES does not affect channel function in these positions.

Mutation of K190 to cysteine resulted in a ~6-fold decrease (EC_{50} 5.91μM compared to EC_{50} 0.95μM WT) in ATP potency at the $P2X_1$ receptor (Figure 5.6). Previous studies have indicated a role for K190 in ATP-binding at the $P2X_1$ receptor. When position K190 was mutated to another positively charged residue, arginine (K190R) a small decrease (~2-fold) in ATP potency was described but when the positive charge was removed at this position (mutant K190A), a decrease in ATP potency (~5-fold) was observed. No effect was seen on the time course of response or suramin sensitivity for mutants K190R or K190A. It was suggested that K190 does not play an essential role in the formation of the ATP binding pocket of P2X receptors but that positive charge on the surface of the protein may attract the negatively charged phosphate groups of ATP toward the binding site of the receptor (Ennion *et al.*, 2000). Mutagenesis studies in the $rP2X_2$ receptor describe a much reduced sensitivity for ATP in mutant K188A (equivalent to K190 in $hP2X_1$ and $rP2X_4$ receptors). This response is 'rescued' and sensitivity to ATP is close to WT receptors, by replacing the positive charge at position 188 with arginine, suggesting it is the positive charge in this position that is important (Jiang *et al.*, 2000). It is clear that the 'impact' on ATP sensitivity by replacement of lysine-190 ($P2X_1$ receptor numbering) by some other residue is dependent on P2X receptor subtype. Lysine-190 is close to the $^{185}FT^{186}$ motif which is 100% conserved across the P2X receptor family and thought to be important in ATP action at the $P2X_1$ receptor (Roberts & Evans, 2004; Roberts & Evans, 2006) and so this region $^{185}FTLFIK^{190}$ ($P2X_1$ receptor sequence and numbering) is likely to be important in ATP-binding based on mutagenesis studies in all the receptor subtypes.

Further confirmation of the role of K190 at the $P2X_1$ receptor is seen by the effects of charged MTS reagents, MTSEA (positively charged) and MTSES (negatively charged). The response of mutant K190C to ATP was inhibited by MTSES and potentiated by MTSEA (Table 5.2) suggesting the importance of charge in this position. Full concentration response curves were recorded from oocytes injected with mutant cRNA encoding K190C and incubated overnight with MTSEA (1mM) or MTSES (1mM) (Figure 5.13). However, due to the fast acting nature of MTS compounds, unpublished work by Roberts *et al* (2006) suggests that an overnight incubation gives the same concentration response curve as a bath perfused application. MTSEA treated K190C increased the potency of ATP so it is not significantly different to WT and MTSES treated K190C significantly decreased ATP potency compared to WT (Figure 5.13). These results suggest that the positive charge at position 190 is important for ATP binding at the $P2X_1$ receptor. Introducing a negative charge at this position further decreases the potency of ATP compared to the neutral charge provided by the cysteine residue

suggesting that a negatively charged residue in position 190 is not preferential for maximal ATP binding.

An MTSEA-Biotin assay determined the accessibility of free cysteines introduced at positions 190, 230, 278 and 280 of the $P2X_1$ receptor. Bands were detected at the cell surface for all mutants (K190C, F230C, G278C and Y280C) suggesting that these positions are not buried deep within the protein but are accessible to the extracellular environment and could potentially contact ATP or other ligands.

Mutant K283C significantly decreased the potency of ATP (~8-fold) at the $P2X_1$ receptor but had no effect on the time course of the response. Previous mutagenesis studies have highlighted the role of the conserved aromatic motif ^{290}NFR^{292} motif in ATP binding at the $P2X_1$ receptor (Roberts & Evans, 2004). The conservation of these residues and the lysine in position 283 of the $P2X_1$ receptor could indicate an important role for this region in ATP-binding at the $P2X_1$ receptor. Charged MTS reagents had no effect on mutant K283C suggesting that charge does not play an important role in ATP binding in position 283 at the $P2X_1$ receptor.

To summarise, mutants F230C, G278C and Y280C do not effect channel function and it is unlikely they play a role in ATP binding at the $P2X_1$ receptor. Mutant K190C agrees with previous mutagenesis data and may contribute to ATP binding at the $P2X_1$ receptor. All four mutants, K190C, F230C, G278C and Y280C are accessible to the extracellular environment and therefore could bind ATP, although this is unlikely for all but K190C due to other experimental results. Mutant K283C highlights a region of the $P2X_1$ receptor that is already considered to be involved in ATP action at the receptor.

5.3.5 Conclusions: Does the model of ATP binding at the rP2X$_4$ receptor reflect ATP binding at the hP2X$_1$ receptor?

The aim of this mutagenesis study was to investigate residues that had previously been identified by a 3D model of the rat $P2X_4$ receptor ATP-binding site and by replacing the equivalent residues in the human $P2X_1$ receptor by cysteine, investigate whether they were involved in ATP-binding at the human $P2X_1$ receptor. Such significant changes to EC_{50} values were not observed for the mutants made in equivalent positions in the $hP2X_1$ receptor (K190C, F230C, G278C, Y280C and K283C) as were reported for the $rP2X_4$ receptor. In fact the majority of mutants in the $hP2X_1$ receptor, F230C, G278C and Y280C had no effect on

channel function. Therefore it is unlikely they play a role in ATP action at the $hP2X_1$ receptor. Mutants K190C and K283C did cause a small decrease in ATP potency at the $P2X_1$ receptor. It is possible K190C may be involved in ATP binding at the $P2X_1$ receptor and it is clear that a positive charge is certainly important for normal channel function at this position. Residue K190 is 100% conserved across the P2X family which indicates that this residue may be functionally important to the family, possibly by coordinating the negatively charged phosphate groups of ATP. The region of $P2X_1$ receptor following K283C has already been implicated in ATP binding (Roberts & Evans, 2004) and the decrease in ATP potency observed for mutant K283C only confirms this.

To conclude it is not possible to use the model of rat $P2X_4$ receptor ATP-binding site directly as a model for the human $P2X_1$ receptor. This is because residues identified as critical for ATP binding at the $rP2X_4$ receptor (F230, G278 and Y280) do not have a similar role at the $hP2X_1$ receptor and should not be implicated in ATP binding in any future models of the $hP2X_1$ ATP-binding site. However, it is clear from this study and the homology model of the $rP2X_4$ receptor that the region containing the aromatic NFR motif is important in ATP binding at P2X receptors. The conserved nature of K190 and extensive mutagenesis work in both $hP2X_1$ and $rP2X_4$ receptor backgrounds suggest a role in ATP action at P2X receptors for this positively charged lysine but the extent of the role may differ between receptor subtypes.

5.4 Aminoacyl-tRNA Synthetases Homology Modelling

5.4.1 Identifying suitable templates

The Model Building Procedure (Chapter 4 Figure 4.1) was used in an attempt to model the extracellular ligand-binding segment of the $hP2X_1$ receptor using the similarity between the catalytic domain of class II aminoacyl-tRNA synthetases and the P2X receptor family (Freist *et al.*, 1998). The templates used all belong to the class II aminocyl-tRNA synthetases and were included in the Freist model (Freist *et al.*, 1998) and used by Yan *et al.* (2005); seryl-tRNA synthetase (Belrhali *et al.*, 1994) (PDB code 1SES), glycyl-tRNA synthetase (Arnez *et al.*, 1999) (PDB code 1B76) and histidyl-tRNA synthetase (Arnez *et al.*, 1997) (PDB code 1KMN).

All three template structures were determined by X-ray crystallography. The resolutions of structures 1SES and 1KMN are 2.50Å and 2.80Å respectively. The higher the resolution (the

lower the number) the more accurate you can expect the structure to be. At these resolutions the main chain will be traced with reasonable confidence and ordered side chains and the planes of peptides will be resolved. Structure 1B76 has a resolution of 3.40Å at which the protein main-chain (usually with some ambiguities) and some side chains will be visible. The R-factors of each structure (0.176 1SES, 0.227 1B76 and 0.249 1KMN) are low which indicates a better agreement between the expected diffraction amplitudes to the observed values. The template structures all have purine based ligands which is similar to the environment the query sequence (putative ATP-binding domain $hP2X_1$ receptor) will be modelled. The ligand-binding site of the templates is contained in a catalytic domain built around a central sheet of six anti-parallel β-sheet strands and a long α-helix (Eriani *et al.*, 1995). It is this domain in each template that has been used to model the query sequence.

5.4.2 Sequence Alignment

The sequence homology between query and template sequences is poor. To improve the quality of the alignment, a structural alignment of the templates has been formed as a profile, created using TCOFFEE (Poirot *et al.*, 2004) and has been aligned to a second profile, a multiple sequence alignment of the putative ATP-binding domains of the P2X receptor family; generated using CLUSTALX (Thompson *et al.*, 1997). 'Profile based alignments' can generate more accurate alignments as both evolutionary and structural information is considered from the templates. The final alignment of the three templates and the putative ATP-binding domain of the $P2X_1$ receptor are shown in Jalview (Clamp *et al.*, 2004) in Figure 5.14.

5.4.3 Generation of a homology model of the putative ATP-binding domain of the $hP2X_1$ receptor using class II aminoacyl-tRNA synthetases

A single-step approach to model building using Modeller (Fiser & Sali, 2003) was used. The 6-stranded anti-parallel β-sheet structure common to the catalytic domain of the class II aminoacyl-tRNA synthetases is clearly identified in the 3D model of the putative ATP-binding domain of the $P2X_1$ receptor (Figure 5.15a). It is this structure that forms the ATP-binding site in the class II synthetases and it has a similar role in the $P2X_1$ receptor model. The distance between those residues preceding the second TM domain and those immediately after TM1 are sufficiently close to form the pore. On closer inspection of the ATP-binding environment it is important to validate the experimental mutagenesis data that is available for the $P2X_1$ receptor. Those residues thought to play a part in ATP-binding at the $P2X_1$ receptor

are highlighted in blue (Figure 5.15b). However, none reside in the ATP-binding environment, within 4Å of the ATP molecule (Figure 5.15c). In addition those residues predicted by the rat $P2X_4$ receptor model based on the same class II synthetase templates as this model (Yan *et al.*, 2005) are not within the ATP-binding environment (Figure 5.15b). However this would largely support mutagenesis data for the corresponding residues in the $P2X_1$ receptor (see section 5.3).

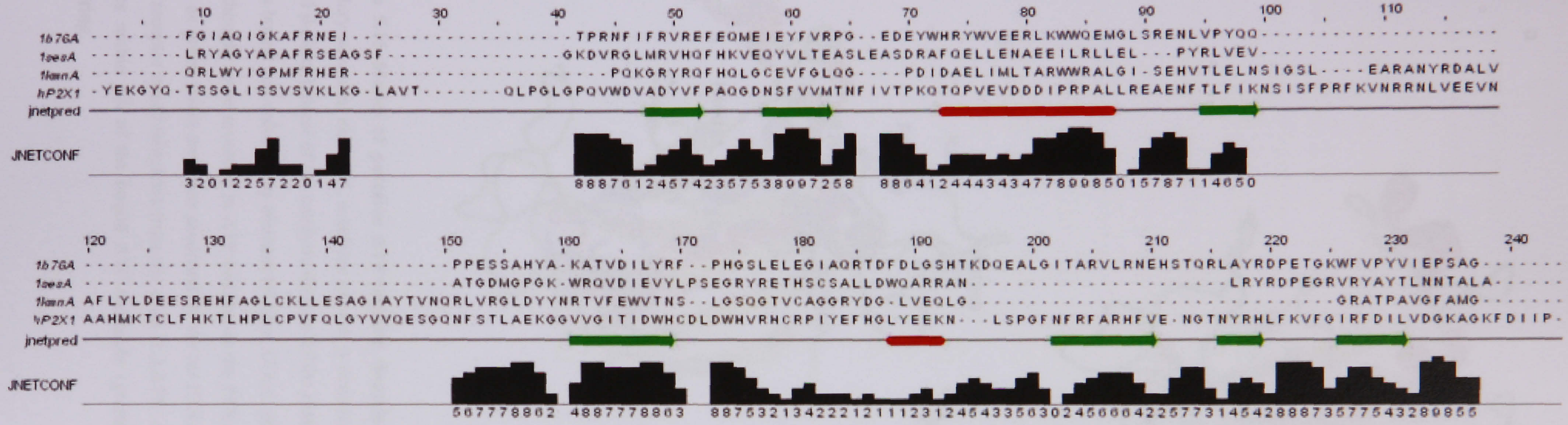


Figure 5.14 Alignment of putative ATP-binding domain hP2X$_1$ receptor and the catalytic domain of three class II aminoacyl-tRNA synthetases templates, 1b76, 1ses and 1kmn. A profile based alignment was generated using CLUSTALX and viewed in Jalview. The secondary structure predictions were made using the JNet method. Helices are marked as red tubes and β-sheets as dark green arrows. The jnetpred is the consensus prediction and the associated confidence of this prediction JNetCONF; with high values denoting a high confidence estimate of the prediction .

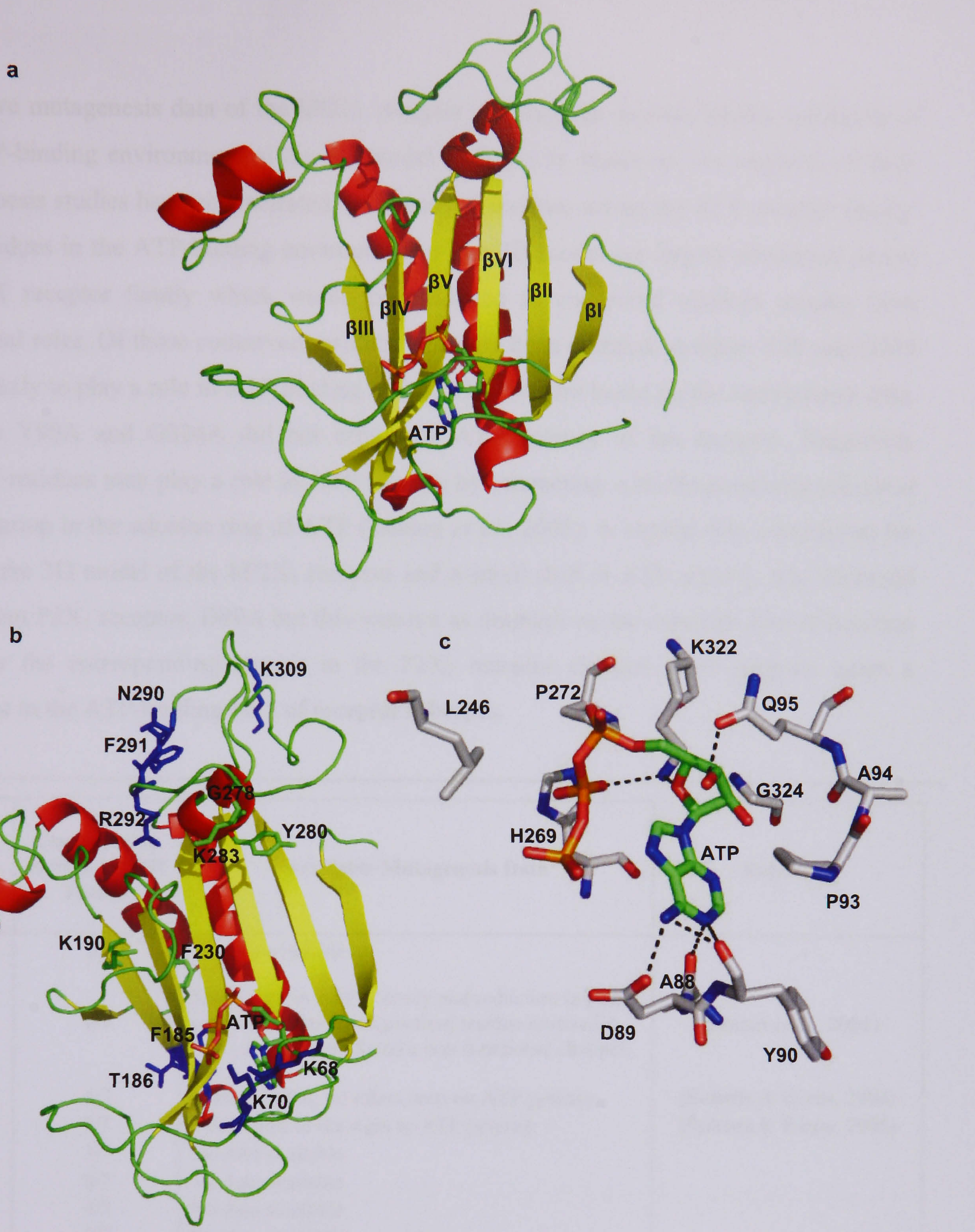


Figure 5.15 Model of putative ATP-binding domain hP2X$_1$ receptor. Model is coloured according the secondary structural elements, α-helices are red, β-strands yellow and random coil, green. ***a***, The 6-stranded anti-parallel β-sheet typical of the catalytic domain of the class II aminoacyl-tRNA synthetases is labelled βI – βVI. ATP is bound and coloured by element (N blue, O red, phosphate orange). ***b***, Residues shown as blue sticks are those thought to be involved in ATP binding at the P2X$_1$ receptor, predicted by mutagenesis studies. Residues shown as green sticks are those predicted by the rat P2X$_4$ receptor model to be involved in ATP binding at the rP2X$_4$ receptor based templates from the class II AATS. ***c***, The ATP binding environment of the P2X$_1$ receptor, residues within 4Å of the bound ATP molecule (green). Dashed lines show polar contacts. P2X$_1$ receptor numbering.

Extensive mutagenesis data of the $hP2X_1$ receptor extracellular domain allows validation of the ATP-binding environment of the 3D model (Table 5.3). However, the majority of these mutagenesis studies have concentrated on conserved residues across the P2X receptor family. The residues in the ATP-binding environment of the 3D model are largely conserved across the P2X receptor family which would be predicted as conserved residues usually have functional roles. Of those conserved residues that have been mutated, residues Y90 and G324 are unlikely to play a role in ATP binding at the $P2X_1$ receptor based on the mutagenesis data. Mutants Y90A and G324A did not effect the ATP potency of the receptor. Negatively charged residues may play a role in ATP binding by interacting with the positively polarized amino group in the adenine ring of ATP (Ennion *et al.*, 2001). A similar role is predicted for D89 in the 3D model of the $hP2X_1$ receptor and a small shift in ATP potency was observed for mutant $P2X_1$ receptor, D89A but this was not as dramatic as the complete loss of function seen for the corresponding mutant in the $P2X_2$ receptor channel. This suggests again a variation in the ATP-binding sites of receptor subtypes.

Residue	Conservation Across hP2XR Family	Available Mutagenesis Data	Reference
A88	5/7	No data available	
D89	6/7	Small shift in ATP potency and reduction in peak current amplitude. Equivalent residue mutated in $P2X_2$ receptor produced a non functional channel.	(Ennion *et al.*, 2001)
Y90	6/7	Mutant Y90A, no effect seen on ATP potency	(Roberts & Evans, 2004)
P93	7/7	4-fold shift to the right in ATP potency	(Roberts & Evans, 2005)
A94	3/7	No data available	
Q95	6/7	No data available	
L246	4/7	No data available	
H269	2/7	No data available	
P272	7/7	Mutant P272A non functional channel	(Roberts & Evans, 2005)
K322	3/7	No data available	
G324	7/7	Mutant G324A no effect seen on ATP potency	(Digby *et al.*, 2005)

Table 5.3. Available mutagenesis data for $P2X_1$ receptor residues.

5.5 An Alternative Method to Identify Suitable Structures for 3D Models of the P2X Receptor Family

Pairwise and multiple protein sequence alignments can give important insight into evolution and protein function. The three-dimensional environments of ligand-binding structures in the PDB (Berman *et al.*, 2000) are often more highly conserved across a functional family than the overall structure. It has been shown that proteins with greater than 20% sequence similarity bind more similar ligands compared to non-homologous proteins (Puvanendrampillai & Mitchell, 2003). However, it is possible for a given ligand to bind to two or more unrelated proteins or to bind to homologous proteins in different binding environments (Shin & Cho, 2005). By analyzing known protein-ligand structures a better understanding of the interaction between protein and ligand can be gained. Adenosine-5'-triphosphate (ATP) (Figure 5.16) and the proteins it binds were investigated based on this hypothesis.

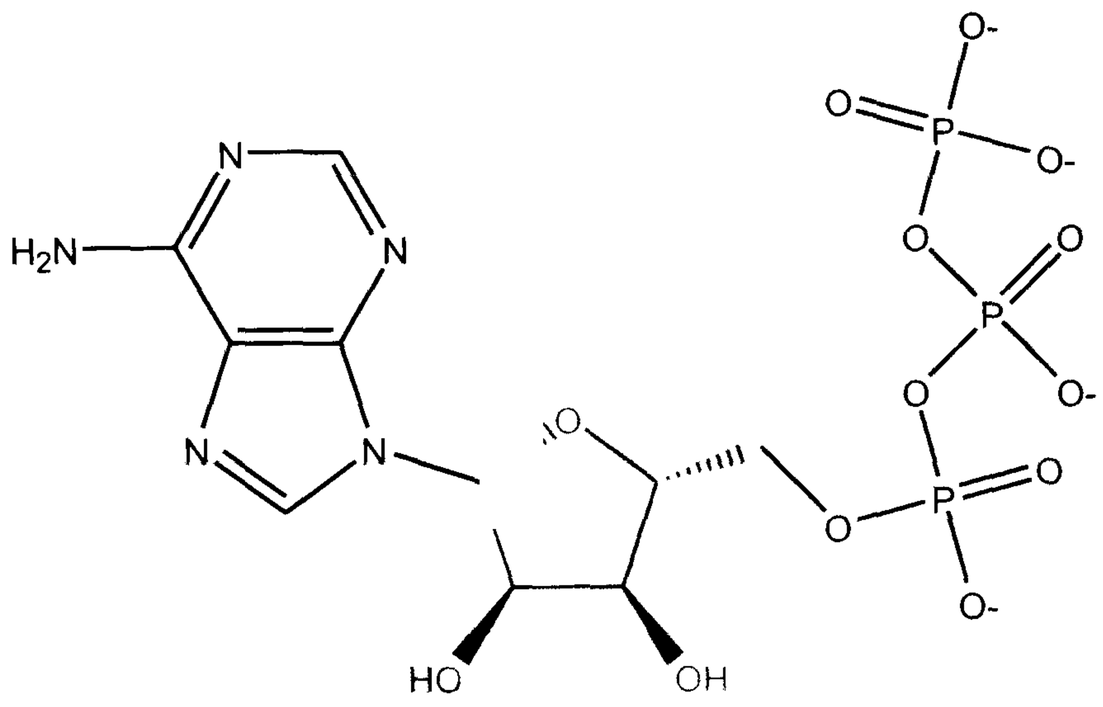


Figure 5.16 ATP molecule. Adenine ring (blue), ribose (green) and phosphate tail (red)

The Model Building Procedure (Chapter 4 Figure 4.1) was adapted to use databases of known structures and mutagenesis data as restraints to search for possible templates. The transmembrane topology has already been discussed and the extracellular segment divided into two domains; a putative ATP-binding domain and a cysteine-rich domain.

5.5.1 Identifying Suitable Templates

The structures of 435 ATP-binding sites are contained in 205 PDB entries (Shin & Cho, 2005). Using MSDsite (Golovin *et al.*, 2005) and data from previous mutagenesis studies of the extracellular domain of the $hP2X_1$ receptor, the 205 PDB entries were filtered according to those residues known to be involved in the binding of ATP to the $P2X_1$ receptor. Currently a model of the $hP2X_1$ receptor binding site has been published using mutagenesis data (Roberts

et al., 2006). Residues that are thought to be involved in ATP action at the receptor have been identified. Conserved positively charged residues K68 and K309 might be involved in coordinating the binding of the negatively charged β and γ phosphates of ATP (Ennion *et al.*, 2000) (Jiang *et al.*, 2000). Aromatic residues F185 and F291 could be associated with binding the adenine ring of ATP with R292 interacting with the α-phosphate of ATP (Roberts & Evans, 2004). It is thought that agonist binding probably induces a conformational change in the extracellular loop. Residues G250 and P272 that allow flexibility in proteins could facilitate such movement (Digby *et al.*, 2005; Roberts & Evans, 2005).

MSDsite allows three-letter amino acid codes to be entered into a ligand environment field which allows the structural database to be queried for ATP-binding environments containing those residues. Using the string of amino acids, `LYS PHE PHE ARG LYS`, the search identified one protein, rat synapsin II (PDB ID 1I7L). The search was widened by omitting one residue constraint; `LYS PHE ARG LYS`. This generated a list of 11 proteins (Table 5.4).

PDB ID	Name	Title
1xdn	Ligase	RNA editing ligase I
1obd	Synthetase	Saicar-synthase complexed with ATP
1nlv	Structural Protein	Dictyostelium discoideum actin complexed with Ca.ATP&human gelsolin segment 1
1nsy	Lyase	nh3-dependent NAD+ synthetase from bacillus subtilis
1pk8	Membrane Protein	Rat synapsin I C-domain complexed Ca.ATP
1wg9 (1xkv)	Transferase	ATP-dependent phosphoenolpyruvate carboxykinase from thermus thermophilus hb8
1atp	Phosphotransferase	cAMP-dependent protein kinase
1i7l	Neuropeptide	Rat synapsin II C-domain with ATP
1tyq	Structural Protein	arp2/3 complex with bound ATP and Ca
1q24	Transferase	Pka double mutant model of pkb in complex with MgATP
1tf7	Circadian Clock Protein	Circadian clock protein KaiC

Table 5.4 Results of ligand environment search, MSDsite. Twelve PDB structures were found but the theoretical model (1pvk) has been omitted.

The ATP-binding environment of each protein was analyzed and the results summarized (Table 5.5). By considering the order the residues appear in the ATP-binding environment within the complete protein sequence a measure of the similarity between the P2X receptor ATP-binding environments can be established. Of the 11 ATP-binding proteins, rat synapsin

II (PDB ID 1I7L) and rat synapsin I (PDB ID 1PK8) have the key ATP-binding residues in the same order as they appear in the $hP2X_1$ receptor sequence. Rat synapsin II also contains a similar motif (^{335}NWK^{337}) to the ^{290}NFR^{292} motif found in the $hP2X_1$ receptor and thought to be involved in ATP action at the receptor. For these reasons they were selected as structural templates for the putative ATP-binding domain of the $hP2X_1$ receptor.

5.5.2 Sequence Alignment

The greater the similarity between the template and query sequences, the higher the more reliable of the resulting model. The putative ATP-binding domain of the $hP2X_1$ receptor and the C-domain of rat synapsin II have low sequence identity and poor sequence similarity, likewise with synapsin I, but an alignment of the three sequences generated by ClustalW and manually edited to ensure secondary structural elements or ATP interactions in synapsin I and II are not aligned to gaps produces an 'average' alignment (Figure 5.17). The correlation between regions that may be involved in ATP binding is reasonable but it is interesting that the first half of the template sequences do not interact with the ligand molecule. This will have an impact on the resulting homology model as it is known that both halves of the extracellular domain of the P2X receptor family are involved in ligand binding (Ennion *et al.*, 2000; Li *et al.*, 2004; Khakh & Egan, 2005; Silberberg *et al.*, 2005).

PDB	1xdn	1obd	1nlv	1nsy	1pk8		1wg9 = 1xkv	1atp	1i7l	1tyq		1q24	1tf7				
Atom	ATP501 (A)	ATP307 (A)	ATP401(A)	ATP2000=ATP3000	ATP800(A)=ATP801=ATP803	ATP802(C)=ATP804=ATP805=ATP806=ATP807	ATP1001(A)	ATP1(E)	ATP800=801(A)	ATP1001(A)	ATP1002(B)	ATP400(A)	ATP1901=ATP2901=ATP4901=ATP6901	ATP1903=ATP2903=ATP4903=ATP6903	ATP3901	ATP3903	ATP
C2		K111					R438		F308(CE1 CD1)	F328(CE1 CZ)							
C4	F209(CE1 CD1)						R438										
C5	F209(CD1)						R438										
C6	F209(CD1)						R438										
C8	V88(O)	F33(CZ)					R438										
N1		H112(N)	F306	Y32a(OH)	I308(N) F307	I308(N) F307	R438	V123(N)	I309(N)								
N3		H75(ND1)			P306(O)	E305(OE2)	T444(OG1)		E306(OE2)					D241a(OD1)			
N6	E86(O)	H110(O) D233(N)			E305(OE2)	P306(O)	F439(O)	E121(O)	P307(O)			E121(O)		S89a(OG)		S89c(OG)	
N7	V88(N)	V232(CB)	K336(NZ)		K269(NZ)			T183(OG1)	K270(NZ)			T183(OG1)		S89a(OG)		S89c(OG)	
N9																	
PG		K19(N)															
O1G	MG	MG	G158(N) V159(N)		N338(ND2) S275(N) H274(ND1) K225(NZ)	K225(NZ) H274(ND1) S275(N)		S53(N) A21i(N) K168(NZ)	R329(NH1)	G173(N) D172(N)	G162(N) D161(N)	K168(NZ)	K457b(NZ)	K224b(NZ) R226b(NH2)	K457d(NZ)	K224d(O1G)	K457f(NZ)
O2G			CA					D184(OD2) MN	N339(ND2)	CA	CA	A21i(N) S53(N) D184(OD1)	R459b(NH2)	T55a(OG1)	R459d(NH2)	T53c(OG1)	R459f(NH2) G291e(N)
O3G	R111(NH2)		S14(N OG)			S275(O3G)	K272(CE)	D184(OD2) MN	S276(OG)	T14(N OG1)		MG	G291a(N)	K224b(NZ)	K457d(NZ)	K224d(NZ)	K457f(NZ)
PB																	
O1B	R309(NH2)		CA K18(NZ)	F129(CE2 CZ)	K225(NZ) G276(N) CA	K225(NZ) G276(N) K279(NZ) CA	T239(N) T236(N OG1) G237(N)	MN K72(NZ)	R329(NH2)		CA	G55(N) F54(N)	K294a(N) T295a(N)	G51a(N) K52a(N)	K294c(N) T295c(N)	K52c(N) G51c(N)	K294e(N) T295e(N)
O2B	MG K87(NZ)	MG	G15(N) M16(N) D157(N)	N133b(ND2 OD1)	G276(N) K279(NZ)	CA K279(NZ) G276(N) K225(NZ)	L233(O) K238(N)	F54(N) G55(N)	K374(NZ) N339(ND2)	G15(N) Y16(N)		K72(NZ) D184(OD2)	G293a(N) G291a(N)	G49a(N)	G293c(N) G291c(N)	G49c(N)	G291e(N)
O3B																	
PA																	
O1A	K87(NZ) K307(NZ)	R21(NE) R21(NH2)	G302(N)	R137(NH2)	K269(NZ)	K269(NZ)	T240(N OG1)	K72(NZ)	K280(CE) K374(NZ)	G324(N)		K72(NZ) N171(OD1)		L54a(N)	L296c(N)	L54c(N)	L296e(N)
O2A	I61(N)	MG	K18(NZ)	K170b(N)	E373(OE1)	E373(OE1)	T239(OG1)	MN D184(OD2)	F386(CE2)	K18(NZ)		D184(OD1)			T295c(OG1)		T295e(OG1)
O3A				K258b(NZ)									T295a(OG1 N)	T53a(OG1)	T295c(N)	T53c(OG1)	T295e(N)
O2	N92(ND2) E159(OE1)	E219(OE2)	R210(CD) K213(NZ) E214(OE2)	D177a(OD2)	D313(OD2) T337(OG1N) W335(O)			E127(OE2) E127(OE2)	D314(OD1)	E229(OE1) K228(NZ)	R214 K217(NZ) E218(OE2)	E127(OE2)	H463b(NE2) R451a(NH2)	H230b(NE2)	H463d(O2) R451c(NH2)	H230d(NE2) G49c(O)	H463f(NE2) K465f(NZ)
O3	R111(NH2 NH1) I59(O)		D157(OD1) K213(NZ) G302(CA)	H257b(NE2)	R328(NH2) K336(CA)	R328(NH2) K336(CA)		E170(O) R18i(NH2)	D314(OD1) R329(NH2)	K228(NZ) D172(OD1) R374	K217(NZ) D161(OD1)	R18i(NH2) E170(O) F327	R451a(NH2) G291a(O)		R451c(NH2) G291c(O)		G291e(O)
O4																	
O5	K87(NZ)																

Table 5.5 Comparison of ATP-binding proteins. The ATP-binding environment of 11 ATP-binding proteins has been compared.

```
                                68
P2X1R          YEKGYQTSSGLISSVSVKLKGLAVTQLPGLGPQVWDVADYVFPAQGDNSFVV--------
1pk8           AARVLLVIDEPHTDWAKYFKGKKIHGEIDIKVEQAEFSDLNLVAHANGGFSVDMEVLRNG
1i7l           KAKVLLVVDEPHTDWAKCFRGKKILGDYDIKVEQAEFSELNLVAHADGTYAVDMQVLRNG
                 :   . .   :. :  ::*  :    .:  :  :.::  : *:.:. : *

                                                                  185
P2X1R          -------MTNFIVTPKQTQPVEVDDD------------IPRPALLREAENFT----LFIK
1pk8           VKVVRSLKPDFVLIRQHAFSMARNGDYRSLVIGLQYAGIPSVNSLHSVYNFCDKPWVFAQ
1i7l           TKVVRSFRPDFVLIRQHAFGMAENEDFRHLVIGMQYAGLPSINSLESIYNFCDKPWVFAQ
                      .:*::  :::  :  : *             :*    *..  **     :* :

P2X1R          NSISFPRFKVNRRNLVEEVNAAHMKTCLFHKTLHPLCPVFQ-------LGYVVQESGQNF
1pk8           MVRLHKKLGTEEFPLIDQTFYPNHKEML---SSTTYPVVVKMGHAHSGMGKVKVDNQHDF
1i7l           MVAIFKTLGGEKFPLIEQTYYPNHREML---TLPTFPVVVKIGHAHSGMGKVKVENHYDF
                  .  :  :.  *:::.  .: :  *   :  .   *.:       :* *  :.  :*

                                                                291
P2X1R          STLAEKGGVVGIT---IDWHCDLDWHVR-HCRPIYEFHGLYEEKNLSPGFNFRFARHFVE
1pk8           QDIA---SVVALTKTYATAEPFIDAKYDVRVQKIGQNYKAYMRTSVSG--NWKTNTGSAM
1i7l           QDIA---SVVALTQTYATAEPFIDAKYDIRVQKIGNNYKAYMRTSISG--NWKTNTGSAM
               . :*   .**.:*     .  :* :   : : * : :  * ...:*   *::    .

                                309
P2X1R          N---GTNYR---------HLFKVFGIRFDILVDGKAGKFDIIPTMT-------------
1pk8           LEQIAMSDRYKLWVDTCSEIFGGLDICAVEALHGKDGRDHIIEVVGSSMPLIGDHQDEDK
1i7l           LEQIAMSDRYKLWVDACSEMFGGLDICAVKAVHGKDGKDYIFEVMDCSMPLIGEHQVEDR
                  . . *         .:*  :.*     :.** *:  *: .:


P2X1R          ---
1pk8           QLI
1i7l           QLI
```

Figure 5.17 Multiple sequence alignment of the putative ATP-binding domain of the $P2X_1$ receptor ($P2X_1R$), rat synapsin II (1i7l) and rat synapsin I (1pk8). ClustalW (Thompson *et al.*, 1994) was used to generate the alignment with manual adjustments. The degree of similarity is illustrated beneath the alignment with consensus symbols such that '*' means the residues in that column are identical in all the sequences in the alignment, ':' means that the conserved substitutions have been observed and '.' means that semi-conserved substitutions are observed. Those residues in red are interactions with the ligand (ATP), those residues in green are residues thought to be important in ATP-binding at the $hP2X_1$ receptor, as predicted by mutagenesis studies. Numbering refers to the $hP2X_1$ receptor.

5.5.3 Building the 3D Model

A single step approach was used to generate a 3D model of the ATP-binding domain of the $hP2X_1$ receptor using the program Modeller (Fiser & Sali, 2003). ATP bound in structure 1I7L was used as a restraint for the $hP2X_1$ receptor model.

The putative ATP-binding domain consists of the majority of the extracellular loop minus the ~60 residues that form the cysteine-rich domain. The model contains 9 β-strands which form 3 β-sheets and an ATP-binding site. Similar to the template structures, the ATP-binding site is formed from one of the β-sheets and the ATP molecule adopts a similar vertical conformation as in the templates (Figure 5.18).

Figure 5.18 3D model of the putative ATP-binding domain $hP2X_1$ receptor. ATP molecule is shown in green and coloured by element, α-helices are red, β-strands are yellow and coil structure is in green.

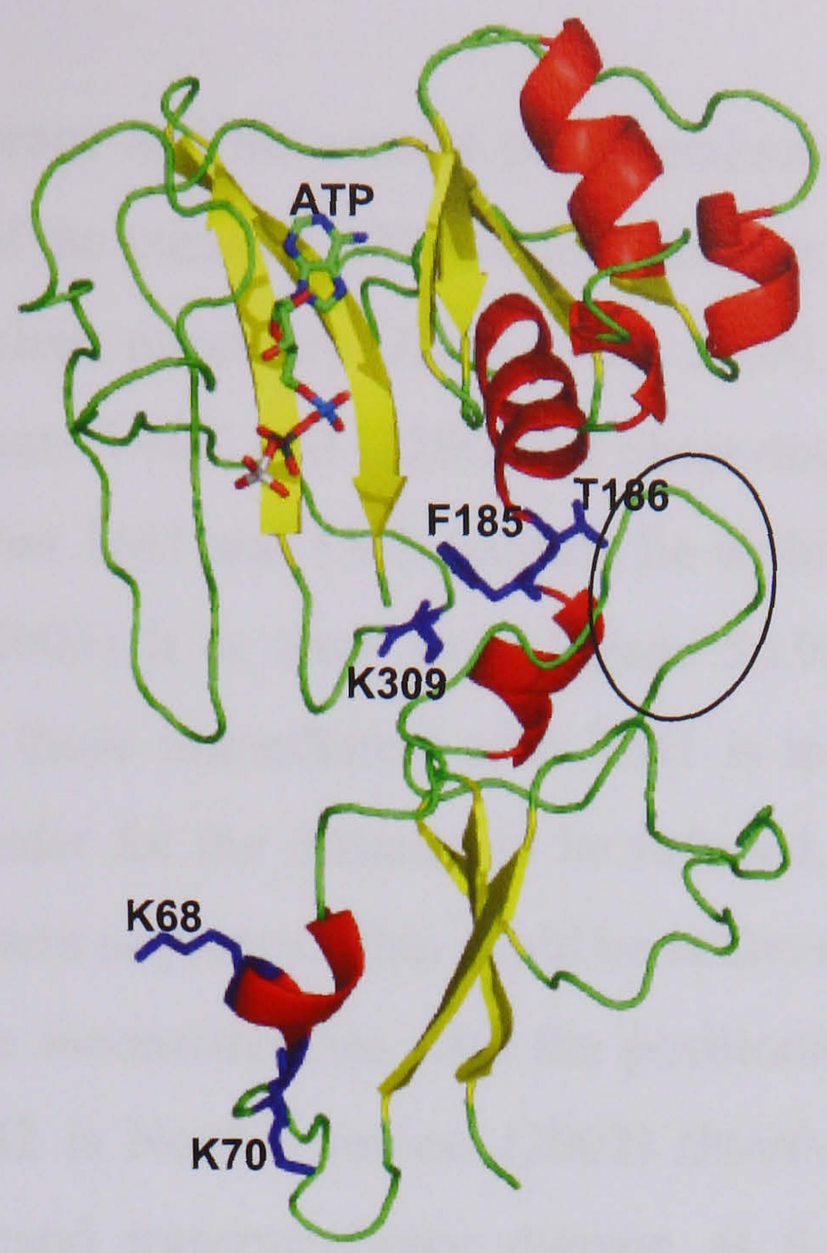
The secondary structure of the 3D model is similar to that predicted for the putative ATP-binding domain of the $hP2X_1$ receptor by PSIPRED (McGuffin *et al.*, 2000). The β-strands in the first section of the model are predicted in this region of the receptor but are not in the predicted positions. There is also an α-helix (residues 15-20) in the model which is not predicted. The $^{185}FT^{186}$ motif of the $hP2X_1$ receptor is predicted to lie between an α-helix and a β-strand although the β-strand is predicted with relatively low confidence. This is the case in the 3D model although both secondary structural elements are α-helices. G250 has been predicted to lie at the end of an α-helix and glycine is known to be a secondary structure

breaker (Digby *et al.*, 2005). In the 3D model of the $hP2X_1$ receptor there is an α-helical region (residues 139-153) which is broken by 4 residues (residues 147-150 inclusive) and is followed by a β-strand (residues 156-160) that terminates on C160 (as predicted for the $hP2X_1$ receptor). G250 lies in the break of the first α-helix but could this helix be successfully rebuilt to include this 'break' region and terminate on G250? The secondary structure breaks down between residues 161 and 211 in the 3D model where in the templates there is an anti-parallel β-sheet that forms the ATP-binding site.

One of the major issues with the 3D model is the fact that the first half of the sequence does not appear to play a role in ATP binding at the receptor. This is a similar case in the template structures. The 3D model of the putative ATP-binding domain $hP2X_1$ receptor consists of two distinct domains separated by a large loop (residues 59-77) (Figure 5.19a).

Mutagenesis data has provided evidence for the roles of K68 and to a lesser extent K70 in ATP binding at the $hP2X_1$ receptor but these cannot be accounted for in the 3D model due to their apparent distance from the bound ATP molecule. The same is true for residues F185, T186 and K309. It is possible to hypothesise that on ATP recognition the receptor goes through a conformational change which would allow the 'closing' of the two domains onto the ATP molecule which would bring these residues in close contact of the ATP molecule. It is also possible that because only one receptor subunit has been modelled and the proposed nature of the P2X receptor channels is trimeric (Stoop *et al.*, 1999) then if three receptors come together to form a functional channel it is possible K68 may be able to interact with the ATP. It is also likely that three molecules of ATP bind (Brake *et al.*, 1994) to a channel perhaps with three binding sites located at the interface between subunits (Vial *et al.*, 2004). This arrangement could then allow for inclusion of residues K68, F185, T186 and K309 into the ATP-binding environment. Therefore, although this 3D model does not fully fit the mutagenesis data for the putative ATP-binding domain of the $hP2X_1$ receptor it is possible to hypothesise about how certain residues and conformations could be allowed to incorporate the mutagenesis and biochemical experimental data.

a



b

Figure 5.19. 3D model of the putative ATP-binding domain hP2X$_1$ receptor. ATP molecule is shown in green, α-helices are red, β-strands are yellow and coil structure is in green. **a**, Residues thought to be involved in ATP binding at the hP2X$_1$ receptor are shown in blue. The large loop separating the two domains of this structure is circled. **b**, the distance between TM1 and TM2 is 30.22Å, highlighted with black circles.

The re-entrance to the membrane and the second transmembrane domain is also a cause for concern with the 3D model of the putative ATP binding domain. At the extracellular surface, the two TM domains are in close proximity (Jiang *et al.*, 2003). This was found by forming channels expressing the mutants V48C and I328C and observing a disulphide bond between the two residues suggesting that TM1 and TM2 residues lie within ~8.4Å of each other (Jiang *et al.*, 2001; Spelta *et al.*, 2003) It is clear from Figure 5.19b that the distance between residues preceding TM2 and those immediately after TM1 is too great to successfully form the pore (Figure 5.19b). In order for the distance to be reduced, additional residues must be added onto the end of the protein sequence. This could be achieved if the predictions for TM2 are incorrect. There are some inconsistencies with the positioning of TM2 in the literature. This thesis reference for TM2 is North's review (2002) (North, 2002) but other published work has positioned the second transmembrane domain as far forward as D327 ($hP2X_1$ receptor numbering) (Vial *et al.*, 2004). In either case I328 must lie in the extracellular domain proven by experimental work with mutant I328C and there must be sufficient residues to span the membrane (~20 residues required). It is not possible for the residues preceding TM2 in the 3D model to be within ~8.4Å of TM1.

The details of the ATP-binding domain are shown in Figure 5.20. F291 possibly stacks over the adenine ring of the ATP molecule and forms part of the $^{290}NFR^{292}$ motif ($hP2X_1$ receptor numbering) which is well characterised by mutagenesis studies as being involved in ATP-binding at the $hP2X_1$ receptor. The residues of this motif are shown in yellow. However, the interaction of R292 with the α-phosphate of ATP as predicted by experimental data is not apparent in this model due to its distance from the phosphate tail. Other interactions with the phosphate tail of ATP include R268 possibly due to its positive charge attracted by the negative charge of the β-phosphate of ATP. In addition Y234 is bound to the α-phosphate. C261 which is 100% conserved across the P2X receptor family interacts with the nitrogen of the adenine ring of the ATP molecule. This cysteine is predicted to form a disulphide bond in the $hP2X_1$ receptor with C270 however this is not shown in this 3D model as the residues are not in close proximity to each other. Previous mutagenesis studies in the $rP2X_2$ receptor show that replacement of this cysteine with alanine (mutant C258A $rP2X_2$ numbering) has no effect on receptor function and had an EC_{50} value similar to the WT $rP2X_2$ receptor (Clyne *et al.*, 2002) suggesting that perhaps this residue does not play a crucial role in ATP binding at the $rP2X_2$ receptor. A294 interacts with the oxygen of the ribose sugar of ATP and is 100% conserved across the P2X receptor family. This residue has been mutated to cysteine (mutant

A294C) but with no effect on ATP potency (unpublished work, Roberts *et al* 2006). Due to the nature of alanine it is unlikely to play a role in ATP-binding and the positioning of this residue in the 3D model is likely to be because of its alignment to the template structures (Figure 5.20).

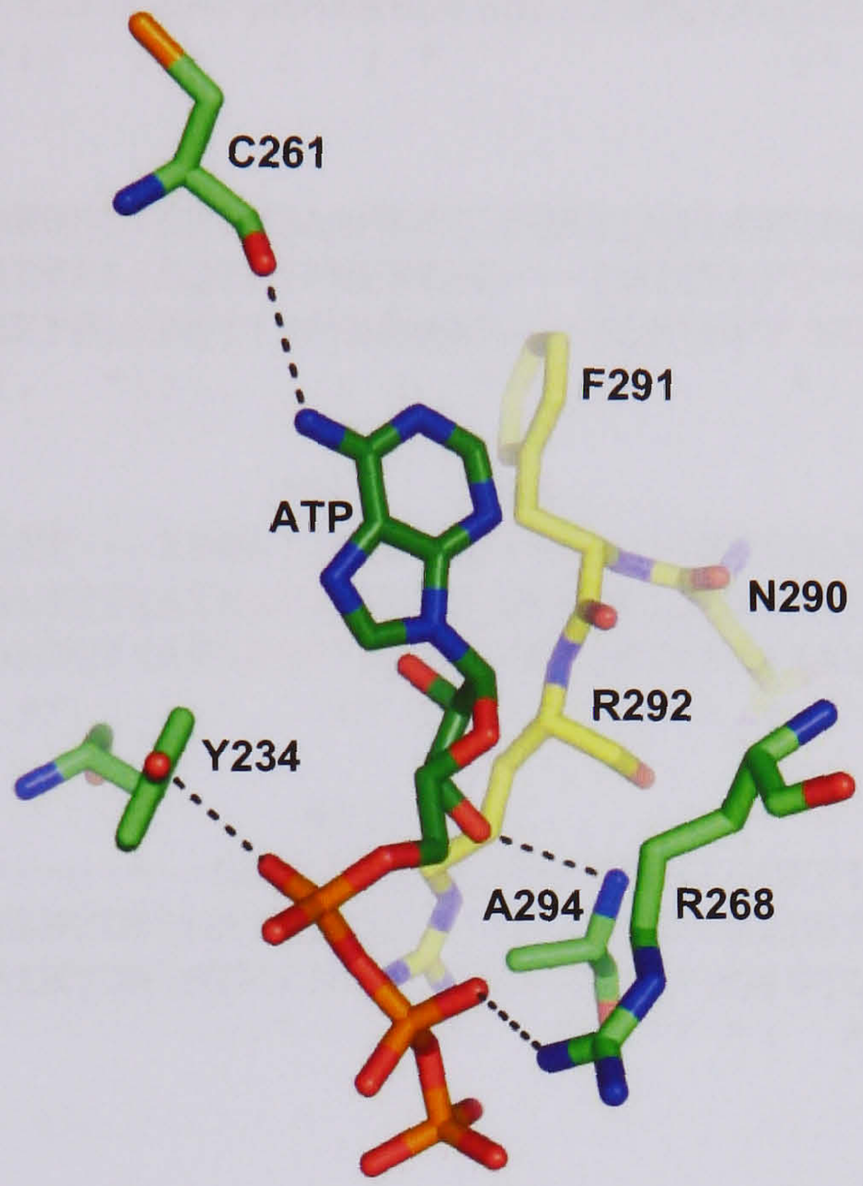


Figure 5.20 The ATP binding site of the 3D model of the putative ATP-binding domain hP2X$_1$ receptor. ATP molecule is shown in green and interactions with the ATP molecule are shown as dashed lines. The ^{290}NFR ^{292} motif is shown in yellow.

5.5.4 Testing the 3D model using mutagenesis

Four residues from the homology model of the putative ATP-binding domain of the hP2X$_1$ receptor that may play a role in ATP-binding at the P2X$_1$ receptor (Figure 5.21) have been identified; C261, A294, Y234 and R268. Of these residues, C261 and A294 are 100% conserved across the P2X receptor family. Mutagenesis data for C261 reported that mutant C261A in the hP2X$_1$ receptor had no effect on ATP potency (Ennion & Evans, 2002a) and a similar result was observed for the rP2X$_2$ receptor (Clyne *et al.*, 2002). Both suggest that it is unlikely C261 plays a role in ATP-binding at P2X receptors. Unpublished mutagenesis data (Roberts *et al* 2006) regarding mutant A294C does not indicate a role for this residue in ATP action at the hP2X$_1$ receptor. Y234 and R268 are both implicated in the 3D model in binding of the negative phosphate groups of ATP. The aromatic residue in position 234 is conserved in the P2X$_2$ receptor but not across the

```
                                68
P2X1R          YEKGYQTSSGLISSVSV L GLAVTQLPGLGPQVWDVADYVFPAQGDNSFVV--------
1pk8           AARVLLVIDEPHTDWAKYFKGKKIHGEIDIKVEQAEFSDLNLVAHANGGFSVDMEVLRNG
1i7l           KAKVLLVVDEPHTDWAKCFRGKKILGDYDIKVEQAEFSELNLVAHADGTYAVDMQVLRNG
                 :   . .   :. :  ::*  :    .:  :  :.::  : *:.:. : *

                                                                    185
P2X1R          -------MTNFIVTPKQTQPVEVDDD------------IPRPALLREAEN T----LFIK
1pk8           VKVVRSLKPDFVLIRQHAFSMARNGDYRSLVIGLQYAGIPSVNSLHSVYNFCDKPWVFAQ
1i7l           TKVVRSFRPDFVLIRQHAFGMAENEDFRHLVIGMQYAGLPSINSLESIYNFCDKPWVFAQ
                       .:*::  :::  :  : *            :*    *..  **     :* :

                                                                   234
P2X1R          NSISFPRFKVNRRNLVEEVNAAHMKTCLFHKTLHPLCPVFQ-------LGYVVQESGQNF
1pk8           MVRLHKKLGTEEFPLIDQTFYPNHKEML---SSTTYPVVVKMGHAHSGMGKVKVDNQHDF
1i7l           MVAIFKTLGGEKFPLIEQTYYPNHREML---TLPTFPVVVKIGHAHSGMGKVKVENHYDF
                    .  :  :.  *:::.  .: :  *   :  .   *.:       :* *  :.  :*

                                261   268                          291 294
P2X1R          STLAEKGGVVGIT---IDWHCDLDWHVR-HCRPIYEFHGLYEEKNLSPGFNFRFARHFVE
1pk8           QDIA---SVVALTKTYATAEPFIDAKYDVRVQKIGQNYKAYMRTSVSG--NWKTNTGSAM
1i7l           QDIA---SVVALTQTYATAEPFIDAKYDIRVQKIGNNYKAYMRTSISG--NWKTNTGSAM
               . :*   .**.:*      .  :* :   : : * : :  * ...:*   *::     .

                               309
P2X1R          N---GTNYR---------HLF VFGIRFDILVDGKAGKFDIIPTMT--------------
1pk8           LEQIAMSDRYKLWVDTCSEIFGGLDICAVEALHGKDGRDHIIEVVGSSMPLIGDHQDEDK
1i7l           LEQIAMSDRYKLWVDACSEMFGGLDICAVKAVHGKDGKDYIFEVMDCSMPLIGEHQVEDR
                   . . *          .:*  :.*     :.** *:  *: .:

P2X1R          ---
1pk8           QLI
1i7l           QLI
```

Figure 5.21 Multiple sequence alignment of the putative ATP-binding domain of the $P2X_1$ receptor (P2X1R), rat synapsin II (1i7l) and rat synapsin I (1pk8). ClustalW (Thompson *et al.*, 1994) was used to generate the alignment with manual adjustments. The degree of similarity is illustrated beneath the alignment with consensus symbols such that '*' means the residues in that column are identical in all the sequences in the alignment, ':' means that the conserved substitutions have been observed and '.' means that semi-conserved substitutions are observed. Those residues in red are interactions with the ligand (ATP) in both the X-ray structures of the templates and the 3D-model of the putative ATP-binding domain of the $hP2X_1$ receptor. Those residues in green are residues thought to be important in ATP-binding at the $hP2X_1$ receptor, as predicted by mutagenesis studies. The residues in blue are the motif NFR thought to be important in ATP action at the $hP2X_1$ receptor and in close proximity to the ATP molecule in the 3D-model of the $hP2X_1$ receptor. Numbering refers to the $hP2X_1$ receptor.

remaining members of the P2X receptor family. A polar residue is conserved in position 268 of the P2X receptor family but the positive charge is only conserved in receptor subunits $P2X_1$ and $P2X_7$. To test the 3D model and the role of these two residues in ATP-binding at the receptor, they were both replaced with cysteine and the mutants characterized. Cysteine was used as it is known to be well-tolerated and allows charged MTS reagents to probe the structure of the mutant receptor. In addition Y234 was mutated to phenylalanine. Phenylalanine is an aromatic residue as is tyrosine but it is not polar. They are similar in size and structure. Any changes observed in channel function will therefore be as a result of the absence of polar tyrosine and not due to change in size for example. If phenylalanine can substitute for tyrosine and no effect is seen in channel function, it may indicate that the properties of a polar aromatic residue are important in this position. R268 was also mutated to methionine due to its similar structure and size but because it is not charged it will remove the positive charge arginine provides in this position. Changes in channel function may be associated with the removal of a positive charge in this position but positively charged MTSEA could mimic the positive charge by binding to the available cysteine residue of mutant receptor R268C to determine the role of charge in this position. Arginine and methionine are similar in size therefore, any effects on ATP potency caused by mutant receptor R268M is likely to be charge- not size- dependent (Figure 5.22).

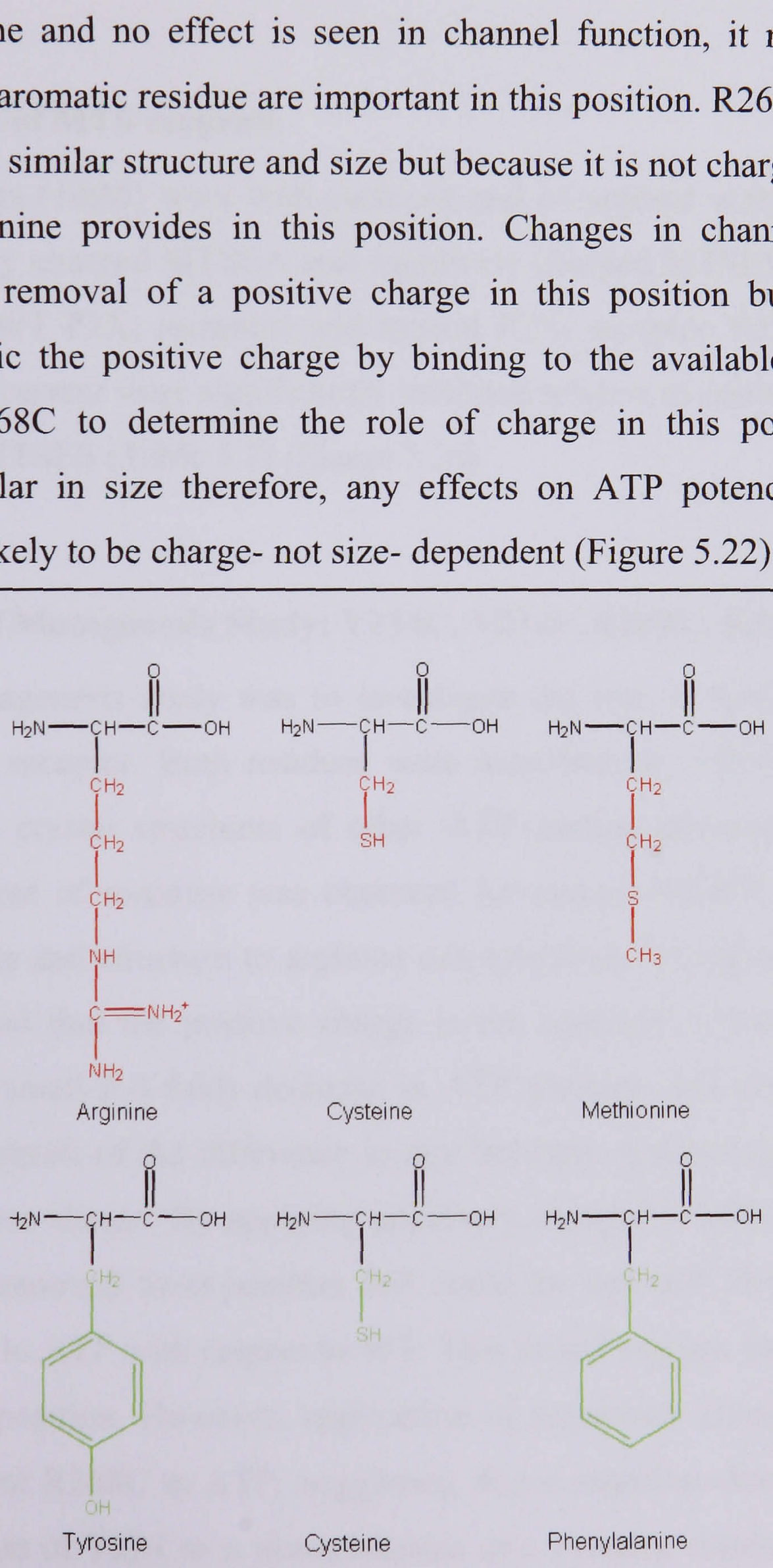


Figure 5.22 Structure of amino acids used in mutagenesis. R groups are highlighted in red and green.

5.5.4.1 Effect of cysteine replacement on the response to ATP

At wild-type $hP2X_1$ receptors ATP evoked a concentration-dependent desensitizing inward current with an EC_{50} of ~0.95μM (Table 5.6) (as previously described (Ennion *et al.*, 2000)). No effect was seen on ATP potency, peak current amplitude or time course of the response for mutant R268M (Figure 5.23) (Figure 5.25). A small yet significant decrease in ATP potency was recorded for mutants R268C (~3-fold) (Figure 5.23), Y234C (~4-fold) and mutant Y234F (~5-fold) (Figure 5.24). No effect on the peak current amplitude and no effect on the time course of the response was recorded for mutants R268C, Y234C and Y234F (Table 5.6) (Figure 5.25).

5.5.4.2 Effect of MTS reagents

Methanethiosulfonates (1mM) were bath-perfused and co-applied with ATP for a 3 second application. Positively charged MTSEA and negatively charged MTSES had no effect on the currents elicited by WT $P2X_1$ receptors and mutant $P2X_1$ receptor Y234C. MTSEA had no effect on R268C but current were significantly inhibited relative to control with application of negatively charged MTSES (Table 5.7) (Figure 5.26).

5.5.4.3 Discussion of Mutagenesis Study; Y234C, Y234F, R268C, R268M

The aim of this mutagenesis study was to investigate the role of R268 and Y234 in ATP-binding at the $P2X_1$ receptor. Both residues were identified by a $P2X_1$ receptor homology model based on the crystal structures of other ATP-binding proteins. No effect on ATP potency or time course of response was observed for mutant R268M. This suggests that a residue of similar size and structure to arginine can substitute for arginine at position 268 in the $P2X_1$ receptor and that the positive charge is not required in this position for normal channel function. A small (~3-fold) decrease in ATP potency was seen for mutant R268C which could be as a result of the difference in size between cysteine and arginine and/or the removal of the positive charge. By applying positively charged MTSEA, the positive charge that mutant R268C removed from position 268 could be replaced. However, no effect was seen on the response to ATP with respect to WT. This would suggest that a positive charge is not essential in this position. However, application of negatively charged MTSES inhibited the response of mutant R268C to ATP, suggesting that a negative charge is not preferred in position 268. Mutation of Y234 to a phenylalanine or a cysteine residue resulted in a small, significant decrease in ATP potency (~5-fold and ~4-fold respectively).

Mutant	ATP EC_{50} (μM)	ATP pEC_{50}	Peak (nA)	Rise Time (ms)	Decay Time (ms)
$P2X_1$ WT	0.95	6.02 ± 0.12	8073 ± 1341	94 ± 16	870 ± 115
R268C	2.7	5.56 ± 0.07*	6751 ± 936	103 ± 31	625 ± 125
R268M	1.67	5.78 ± 0.008	6521 ± 754	125 ± 28	1302 ± 400
Y234C	3.4	5.47 ± 0.07**	6734 ± 797	85 ± 9	1315 ± 21
Y234F	4.92	5.31 ± 0.12***	4684 ± 2076	127 ±11	1532 ± 223

Table 5.6. Summary of data for ATP action at $P2X_1$ WT and mutants receptors

EC_{50} data are representative of data collected from three oocytes. pEC_{50} is –log10 of the EC_{50}. Mean peak data are for 100μM ATP (n = 3-4). Rise-time represents the time from 10 to 90% of peak current after an application of a maximal stimulating concentration of ATP (n = 3-4). Decay time represents the time taken to achieve 50% of the peak current after a maximal stimulating concentration of ATP (n = 3-4). Values shown are representative of the mean ± S.E. of the mean.

* $P < 0.05$, different from WT as measured by one-way ANOVA.

** $P < 0.01$, different from WT as measured by one-way ANOVA.

*** $P < 0.001$, different from WT as measured by one-way ANOVA.

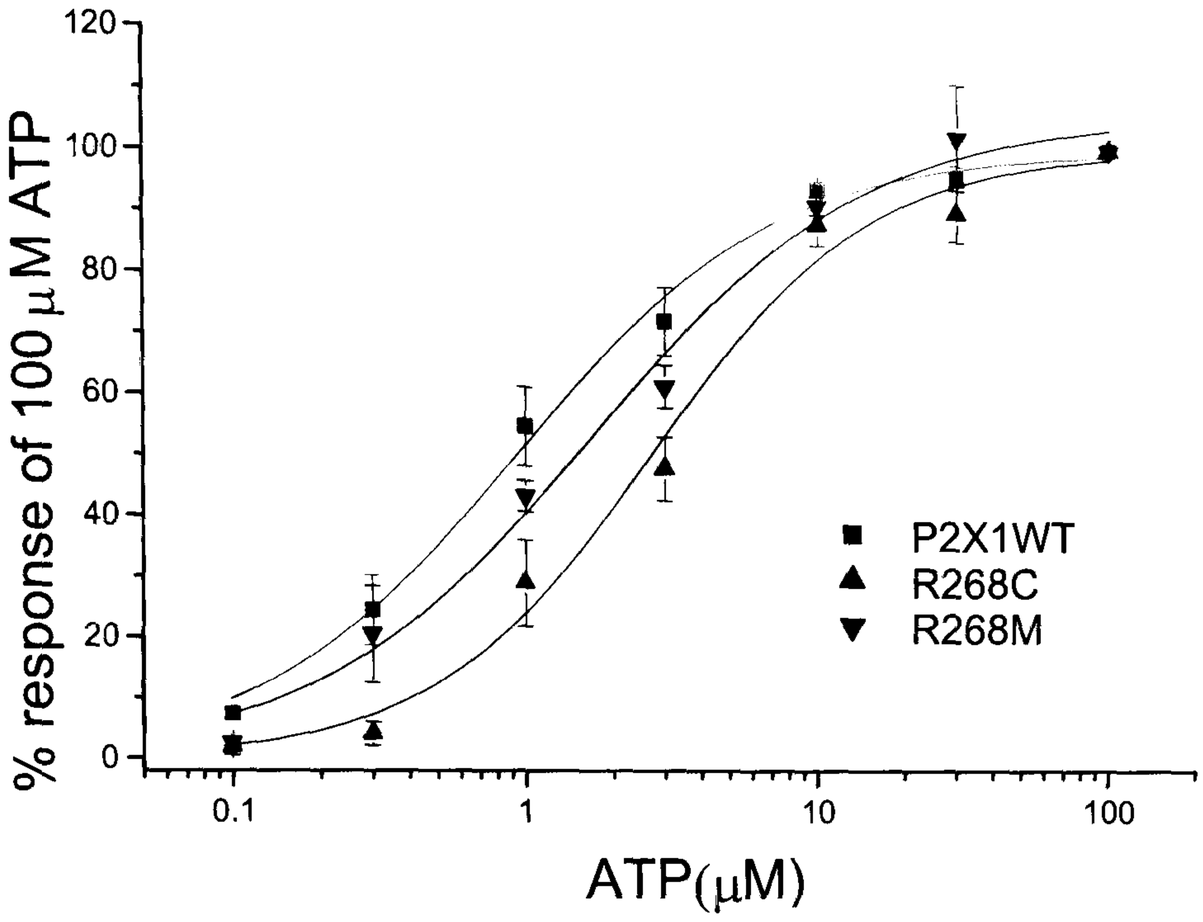


Figure 5.23 Effect of R268 mutants on ATP potency. $P2X_1$ WT or mutant receptors were expressed in oocytes and ATP potency tested using a two-electrode voltage clamp. Concentration response curves for P2X1 WT and mutants R268C and R268M. Mean currents were normalized to the maximum response (n = 4). Holding membrane potential was –60 mV.

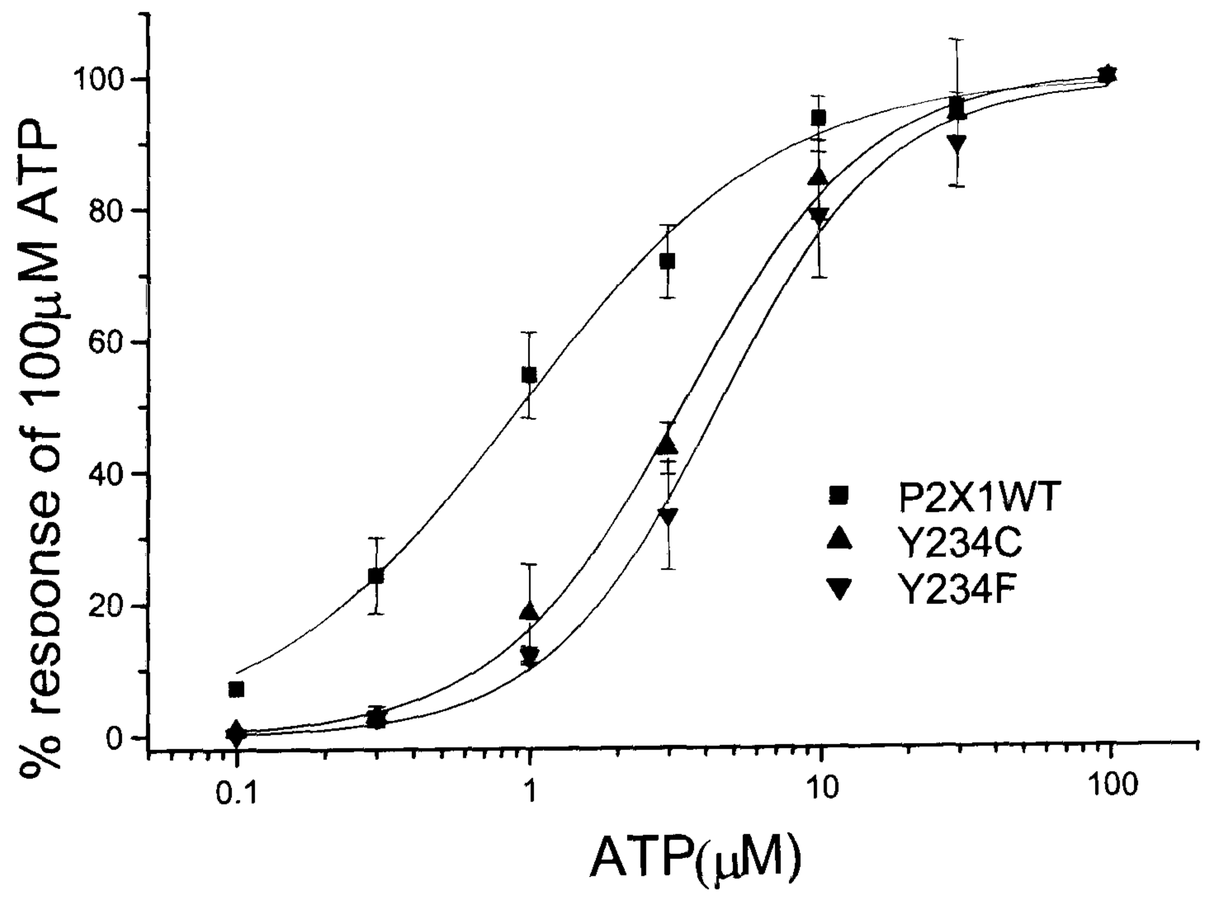


Figure 5.24 Effect of Y234 mutants on ATP potency. $P2X_1$ WT or mutant receptors were expressed in oocytes and ATP potency tested using a two-electrode voltage clamp. Concentration response curves for $P2X_1$ WT and mutants Y234C and Y234F. Mean currents were normalized to the maximum response (n = 3). Holding membrane potential was –60 mV.

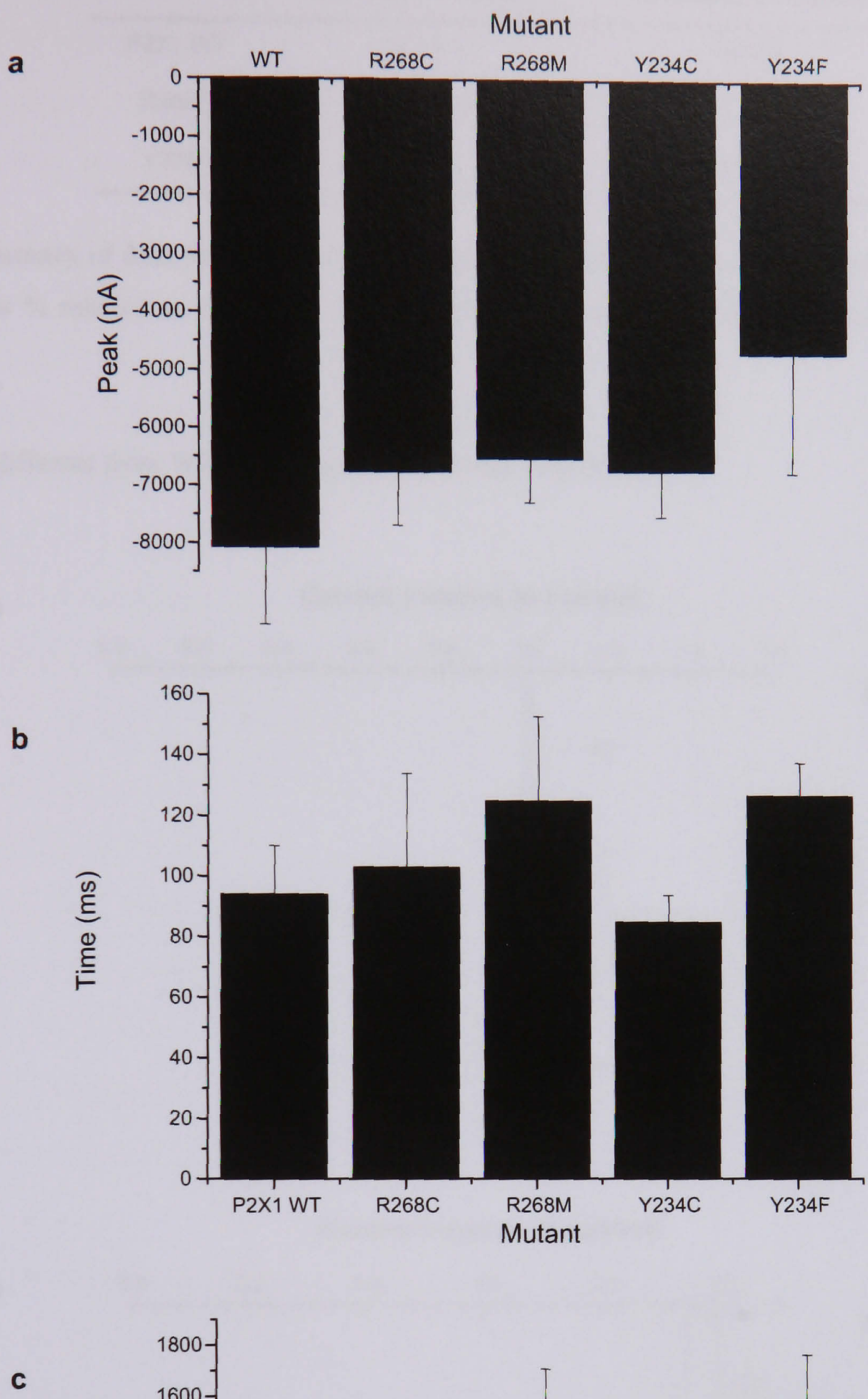


Fig 5.25 a, Mean peak current for $P2X_1$ WT and mutant receptors. **b**, Rise-time (black) and **c**, Decay- time (grey) for $P2X_1$ WT and mutant receptors.

Mutant	MTSEA	MTSES
	% relative to control	% relative to control
$P2X_1$ WT	102 ± 7	94 ± 5
R268C	120 ± 30	78 ± 4*
Y234C	88 ± 2	103 ± 3

Table 5.7. Summary of data for effect of MTS reagents at cysteine mutants relative to $P2X_1$ WT

Data represents % relative to control (n=3-4). Values shown are representative of the mean ± S.E. of the mean.

* $P < 0.05$, different from WT as measured by one-way ANOVA.

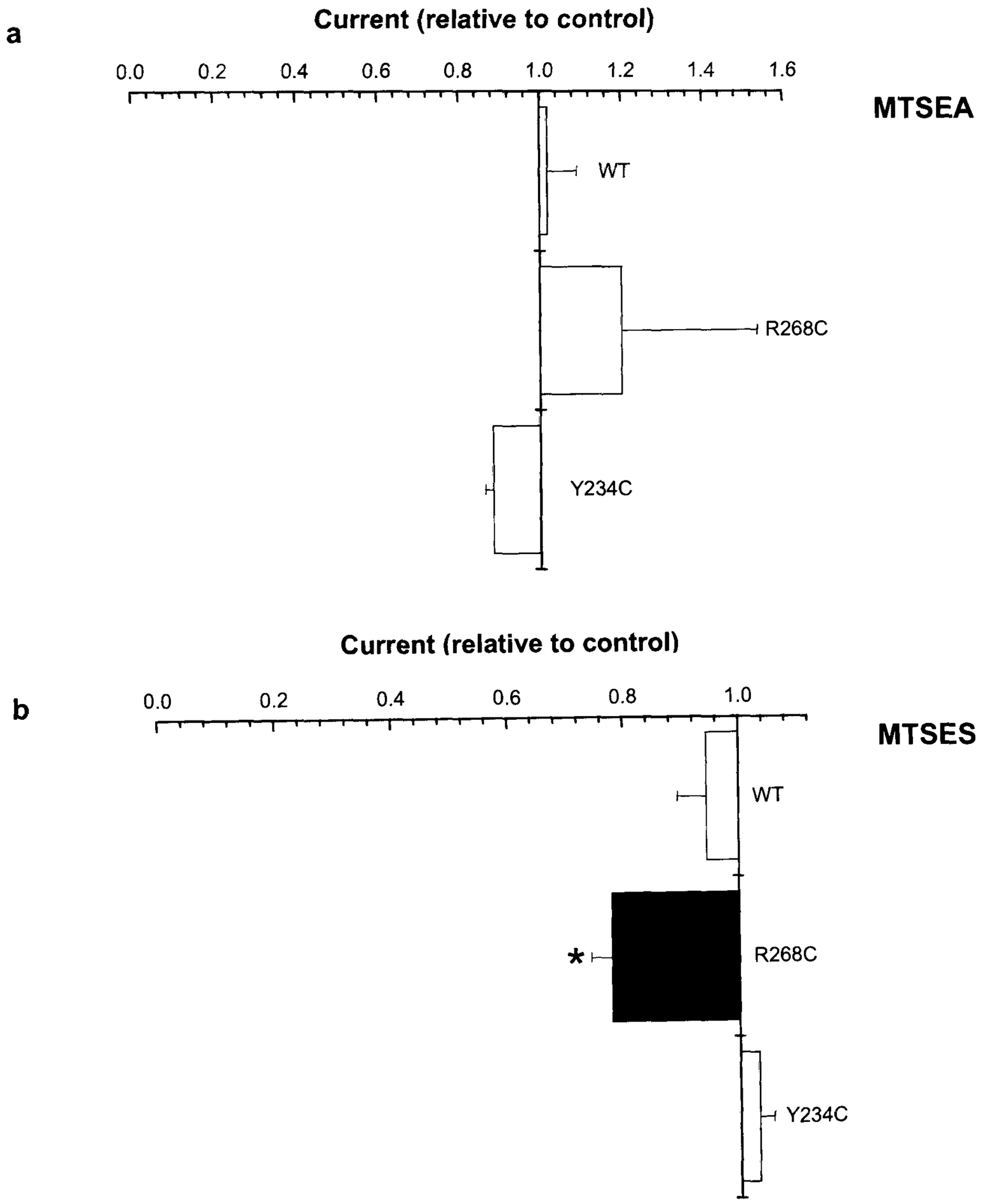


Figure 5.26 a, % change of current elicited by a 3s application of an EC_{50} concentration ATP with 1mM MTSEA. 1mM MTSEA was also bath perfused. WT normalised to 1. **b**, % change of current elicited by a 3s application of an EC_{50} concentration ATP with 1mM MTSES. 1mM MTSES was also bath perfused. WT normalised to 1.

* $P < 0.05$, different from WT as measured by one-way ANOVA.

The time course of response was not affected and charged MTS reagents had no effect on the response recorded for either mutants, Y234F or Y234C. This suggests that charge is not important in this position of the $hP2X_1$ receptor. This is supported across the P2X receptor family where in this position, charge is not conserved. However, it appears that a change in size (caused by mutant Y234C) and the introduction of a non polar residue (Y234F) both decrease the ATP potency observed for both mutant receptors. This suggests that a residue of similar size to tyrosine yet also a polar residue may be important in this position. An interesting mutant to generate would be Y234W which would maintain the polar, aromatic nature of this position.

Residues Y234 and R268 are not conserved across the P2X receptor family. If these residues had an essential role in ATP-binding at P2X receptors you would expect that they would be highly conserved across the receptor family or residues with similar properties would lie in these positions. In position 234, six out of seven family members have a polar residue in this position with the exception of the $P2X_2$ receptor which has phenylalanine. However, the $P2X_1$ and $P2X_2$ receptors are the only family members to have an aromatic residue in position 234. The necessity of a polar residue in position 234 of the $hP2X_1$ receptor has been illustrated with mutant Y234F but also the requirement for a residue of similar size to tyrosine (a decrease in ATP potency was seen for mutant Y234C where cysteine is a polar residue but is much smaller in size to tyrosine) has been observed.

At position 268 a polar residue is found in every P2X receptor family member suggesting that this position is not buried (residues with polar side-chains prefer to reside in an aqueous environment and for this reason are generally exposed on the surface of a protein). However, replacement of R268 with a non-polar residue did not significantly affect ATP potency which suggests that this property is not required in this position for ATP-binding. Replacement of R268 with a polar residue of smaller size did affect the ATP-potency but across the P2X receptor family small residues such as cysteine or serine (seen in four out of seven P2X receptor family members in position 268) are seen in this position.

5.5.5 Conclusions from the Putative ATP-binding Domain $hP2X_1$ Receptor Modelling Exercise and Mutagenesis Study

A 3D model of the putative ATP-binding domain of the $hP2X_1$ receptor based on synapsins I and II has been produced and it fits some of the available mutagenesis data for the $hP2X_1$

receptor. However, the pore has not been successfully created and the involvement of the first half of the domain in ATP-binding is not clear. Two residues, Y234 and R268 were identified from the model as playing a role in ATP binding and they were both mutated to cysteine and individually to phenylalanine and methionine respectively, to investigate this further. Small, yet significant changes in ATP potency of the receptor were observed in both positions when the tyrosine at position 234 and the arginine at position 268 were replaced by other residues. The exact nature of the residue required in both positions is not clear but it is interesting to note that residues in positions 234 and 268 may play a role in ATP action at the $hP2X_1$ receptor. This enhances my 3D model by adding experimental data to the proposed structure but does not suggest that the model is an accurate representation of the $hP2X_1$ receptor.

5.5.6 Conclusions from Chapter 4 and Chapter 5

The results from chapter 4 and chapter 5 are summarized in Figure 5.27. The large extracellular segment of the P2X receptor has been divided into two domains; cysteine-rich domain and ATP-binding domain. Both domains have been modeled separately using different approaches and to varying degrees of success. The model of the cysteine-rich domain of the $hP2X_1$ receptor based on three structural templates 1TGX, 1UG4 and 1CCQ agrees with secondary structural predictions for this region of the $hP2X_1$ receptor although, the pattern of disulphide bonding in the model does not agree with experimental data (Ennion & Evans, 2002a)(see Chapter 4).

Two different approaches for finding structural templates for the ATP-binding domain have been used; searching for individual sequences using databases of known protein structures/folds and using clues from entire protein families that bind ATP. The first approach, used in chapter 4, generated a 3D model of the putative ATP-binding domain of the $hP2X_1$ receptor using a structural intermediate template, HUNK. Inconsistencies with the 3D model and experimental data were found.

In chapter 5, the second approach for finding structural templates identified the family of class II aminoacyl-tRNA synthetases. A model for the $rP2X_4$ receptor has been published based on the class II aminoacyl-tRNA synthetase family (Freist *et al.*, 1998) and validated using mutagenesis (Yan *et al.*, 2005). However, the poor sequence alignment of the P2X receptor family to the synthetase structural templates raised questions about the reliability of the resulting 3D model. Residues thought to be involved in ATP binding at the $rP2X_4$ receptor

were highlighted by this model and to test whether these could be involved in ATP-action at other P2X receptor subtypes, the corresponding residues were mutated to cysteine in the $hP2X_1$ receptor. It was found that it was not possible to transfer the model of ATP-binding in the rat $P2X_4$ receptor to the human $P2X_1$ receptor as residues critical for ATP binding at the $rP2X_4$ receptor (F230, G278 and Y280) do not play a similar role at the $hP2X_1$ receptor. In an attempt to further investigate the similarity of the P2X receptor family to the class II aminoacyl-tRNA synthetases, homology models were generated based on three structural templates all belonging to the synthetase family and the $hP2X_1$ receptor. It was not possible to marry the experimental data with the 3D homology model and for this reason the 3D model was not considered reliable.

Other ATP-binding families were then considered as possible structural templates for the P2X receptor family. Similarities between the P2X receptor family and the synapsin family were used to produce a 3D model. Using cysteine-replacement mutagenesis, residues identified by the model to be involved in ATP-binding were investigated. Residues Y234 and R268 may play a role in ATP-action at the $P2X_1$ receptor but inconsistencies with some of the experimental data in the literature and the 3D model suggest it is not an accurate representation of ATP-binding at the $P2X_1$ receptor.

It is not possible to produce a reliable 3D model of the P2X receptor, despite the publication of such attempts. These chapters have highlighted the inconsistencies with the 3D models generated in this thesis and those investigated from the literature with the mutagenesis information that is available for the P2X receptor subtypes. It is not possible to reconcile the notion of the entire extracellular segment playing a role in agonist binding and the trimeric architecture of the P2X receptor channel. Also, the size and nature of the pore cannot be modelled. The case for the involvement of individual residues in ATP-binding has been further strengthened by the elimination of the role of other residues or the confirmation of what has already been published. For example, it seems extremely likely from this study and the 3D model of the $rP2X_4$ receptor that the region containing the aromatic NFR motif is important in ATP binding at P2X receptors. In addition the conserved nature of K190 and extensive mutagenesis work in both $hP2X_1$ and $rP2X_4$ receptor backgrounds suggest a role in ATP action at P2X receptors for this positively charged lysine but the extent of the role may differ between receptor subtypes.

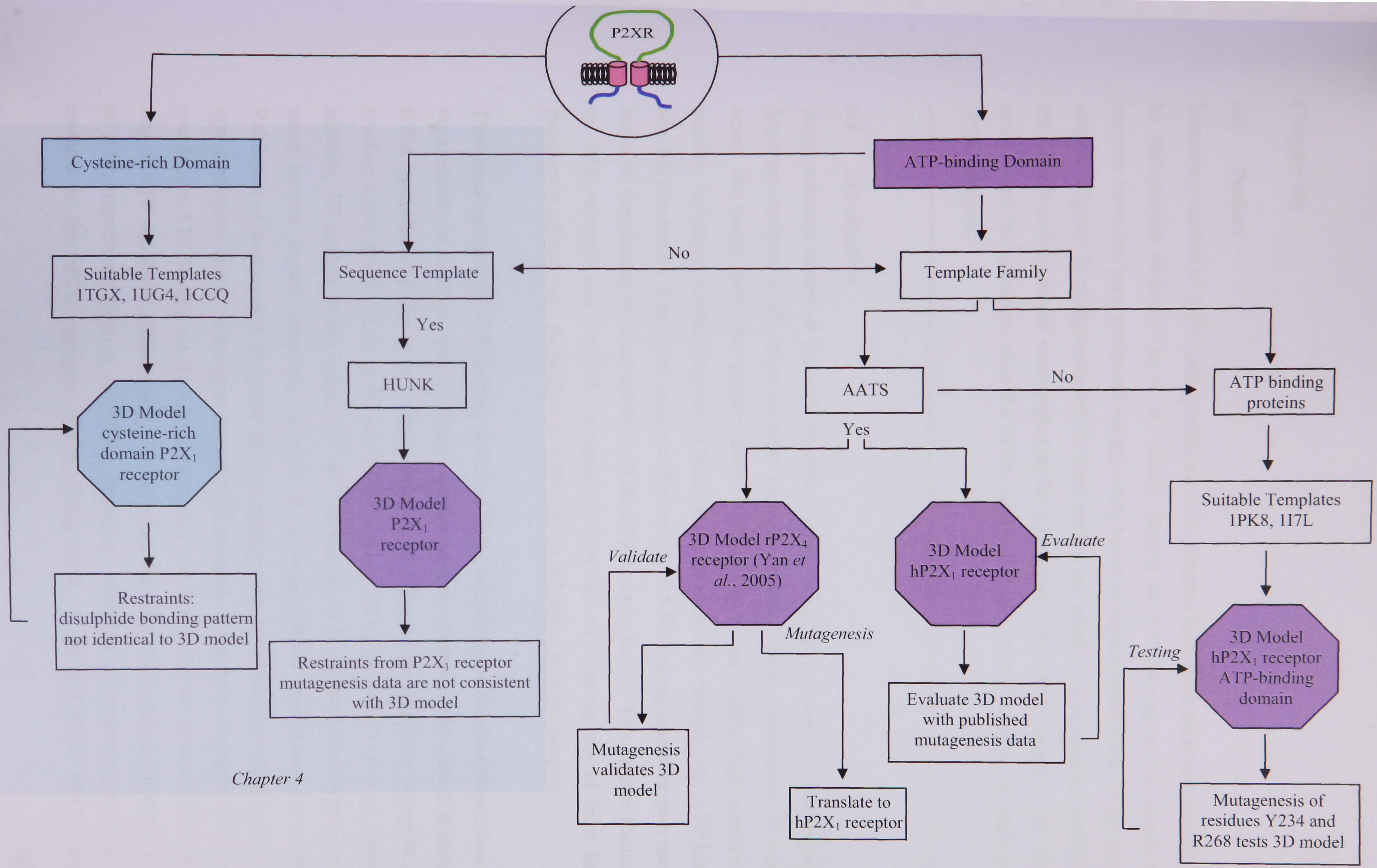


Figure 5.27 Summary of results from Chapter 4 and 5

Chapter Six

6.0 Preface

Homology models of ATP-binding at the rat $P2X_4$ receptor have been generated and validated by mutagenesis studies (Yan *et al.*, 2005). Models of ATP-binding at the human $P2X_1$ receptor have also been generated but are solely based on experimental work. In the previous chapter (Chapter 5) mutagenesis studies were used to test whether those residues identified by the $P2X_4$ homology model were important for ATP action at the $hP2X_1$ receptor. This chapter aims to test the converse; does the $hP2X_1$ receptor model of ATP-binding translate to the $rP2X_4$ receptor.

6.1 Introduction

Seven subunits comprise the P2X receptor family and ATP is known to bind them all. Sequence analysis of all seven receptor subtypes reveals a high conservation of residues across the family (40-55%) (North, 2002) with regions of particularly high conservation possibly indicating residues/regions of functional importance for example, the ATP-binding site. For this reason, it might be expected that models of ATP-binding in one receptor subtype might be transferable to other members of the P2X receptor family. This chapter focuses on the $P2X_1$ receptor and the $P2X_4$ receptor and how the ATP-binding site of these family members is related using the models of ATP-binding published for each receptor.

Theoretical models of ATP-binding at the rat $P2X_4$ receptor have been developed based on the similarity of the P2X receptor family and the catalytic domain of the class II aminoacyl-tRNA synthetases (Freist *et al.*, 1998). These models have been tested by mutagenesis of residues K190, K197, F230, D280, and R318 in the rat $P2X_4$ receptor (Yan *et al.*, 2005). In addition, 3D models of the region K^{180}-K^{326} in the rat $P2X_4$ receptor have been generated using several crystal structures of class II aminoacyl-tRNA synthetases as templates. Validation of such models using site-directed mutagenesis has identified further residues that may participate in ATP binding at the rat $P2X_4$ receptor. It has been suggested that Asp280 coordinates ATP binding via the magnesium ion, F230 coordinates the binding of the adenine ring of ATP and K190, H286, and R278 coordinate the negatively charged phosphate groups of ATP, respectively (Yan *et al.*, 2005). However, when the corresponding residues were mutated in the $hP2X_1$ receptor (K190, F230, G278, Y280 and K283 $hP2X_1$ receptor

numbering) the results suggested that these residues do not play a major role in ATP action at the hP2X$_1$ receptor (see Chapter 5). It is clear that the predictions for the ATP-binding site in the rP2X$_4$ receptor cannot be directly translated to the hP2X$_1$ receptor. The question being asked in this mutagenesis study is can the model of ATP-binding in the hP2X$_1$ receptor be translated to the rP2X$_4$ receptor?

The P2X$_4$ receptor amino acid sequence is most closely related to the other P2X receptor subtypes. This is true whichever species are considered (North, 2002) (Figure 6.1). The P2X$_1$ receptor and P2X$_4$ receptor are both activated by ATP but unusually the rP2X$_4$ receptor is relatively insensitive to classic, non-selective P2X antagonists such as suramin and PPADS (Gever *et al.*, 2006). P2X$_4$ receptors have distinct desensitization characteristics compared to the P2X$_1$ receptor (Table 6.1). Desensitisation at P2X$_4$ receptors is intermediate between currents recorded for P2X$_1$ and P2X$_2$ receptor (North, 2002) (Figure 6.2). Previous mutagenesis studies of the P2X$_4$ receptor have concentrated on transmembrane domains (Silberberg *et al.*, 2005), the C-terminus (Royle *et al.*, 2002; Royle *et al.*, 2005; Fountain & North, 2006; Toulme *et al.*, 2006)and the role of extracellular histidines (Clarke *et al.*, 2000; Coddou *et al.*, 2003; Xiong *et al.*, 2004; Xiong *et al.*, 2005). It is clear that the P2X$_1$ receptor and the P2X$_4$ receptor have different pharmacological profiles but does the sequence homology between the receptor subtypes infer similar ATP-binding sites?

To investigate this, seven residues were mutated to cysteine in the rat P2X$_4$ receptor, K67(68), K69(70), F185(185), T186(186), N293(290), R295(292) and K313(309) which correspond to those residues predicted to be important in ATP action at the human P2X$_1$ receptor (brackets indicate P2X$_1$ receptor numbering). The aim was to test whether residues predicted to be involved in ATP action at the P2X$_1$ receptor play a part in ATP binding at the rP2X$_4$ receptor, thus further testing the rP2X$_4$ homology model (Figure 6.3). Cysteine was chosen to replace the various amino acids because it is a small, non-polar residue and cysteine replacement is known to be well tolerated. Cysteine replacement also allows the use of MTS (methanethiosulfonate) reagents to probe the extracellular domain. Such reagents selectively and rapidly react with sulfhydryls to form a disulfide bond (Figure 6.4). By using various MTS reagents that differ in size or charge, it is possible to determine information regarding the physical size of an ion channel or the role of charge at a specific position.

Properties	$P2X_1$	$P2X_4$
Agonists	ATP, αβ-meATP, BzATP effective partial agonist	ATP
Antagonists	Suramin, PPADS, TNP-ATP (1000 fold more effective blocking ATP at $P2X_1$ than at $P2X_{2,4,7}$)	Effect of ATP blocked by high concentrations of Suramin and PPADS
Permeation Properties	Cation selective, relatively high permeability to Ca^{2+}	When application of ATP is of short duration, operate as cation-selective channels; Ca2+ permeability is quite high. Application continued for several seconds, channel permeable to larger cations, NMDG
Desensitization Characteristics	Fast	Slow, intermediate between $P2X_1$ and $P2X_2$
Tissues	Expressed on blood cells, neurons and variety of smooth muscle cells	Distributed throughout the central and peripheral nervous system, strongly expressed in several glandular tissues

Table 6.1 Pharmacological Properties $P2X_1$ and $P2X_4$ receptor. Agonists defined as drugs that bind to the receptor and activate it, producing a pharmacological response. Antagonists are those drugs that block or counteract the effect of another drug. Permeation properties are those ions that can permeate and the selectivity of receptors for particular ions. Desensitization characteristics describe the decline in current elicited by ATP during the continued presence of ATP. Tissues describe where the receptors are found in the body.

```
                                68
rP2X4   -----QETDSVVSSVTTKAKGVAVTNTSQLGFRIWDVADYVIPAQEENSLFIMTNMIVTV
hP2X4   -----QETDSVVSSVTTKVKGVAVTNTSKLGFRIWDVADYVIPAQEENSLFVMTNVILTM
rP2X1   YEKGYQTSSDLISSVSVKLKGLAVTQLQGLGPQVWDVADYVFPAHGDSSFVVMTNFIVTP
hP2X1   YEKGYQTSSGLISSVSVKLKGLAVTQLPGLGPQVWDVADYVFPAQGDNSFVVMTNFIVTP
             * :..::***:.* **:***:   ** ::*******:**: :.*:.:***.*:*


rP2X4   NQTQSTCPEIPDKTSICNSDADCTPGSVDTHSSGVATGRCVPFNESVKTCEVAAWCPVEN
hP2X4   NQTQGLCPEIPDATTVCKSDASCTAGSAGTHSNGVSTGRCVAFNGSVKTCEVAAWCPVED
rP2X1   QQTQGHCAENPEG-GICQDDSGCTPGKAERKAQGIRTGNCVPFNGTVKTCEIFGWCPVEV
hP2X1   KQTQGYCAEHPEG-GICKEDSGCTPGKAKRKAQGIRTGKCVAFNDTVKTCEIFGWCPVEV
        :***. *.* *:   :*:.*:.**.*.. ::.*: **.**.** :*****: .*****


                       185
rP2X4   DVGVPTPAFLKAAENFTLLVKNNIWYPKFNFSKRNILPNITTSYLKSCIYNAQTDPFCPI
hP2X4   DTHVPQPAFLKAAENFTLLVKNNIWYPKFNFSKRNILPNITTTYLKSCIYDAKTDPFCPI
rP2X1   DDKIPSPALLREAENFTLFIKNSISFPRFKVNRRNLVEEVNGTYMKKCLYHKIQHPLCPV
hP2X1   DDDIPRPALLREAENFTLFIKNSISFPRFKVNRRNLVEEVNAAHMKTCLFHKTLHPLCPV
        *  :* **:*: *******::**.* :*:*:..:**:: ::. :::*.*::.   .*:**:


rP2X4   FRLGTIVEDAGHSFQEMAVEGGIMGIQIKWDCNLDRAASLCLPRYSFRRLDTRDLEHNVS
hP2X4   FRLGKIVENAGHSFQDMAVEGGIMGIQVNWDCNLDRAASLCLPRYSFRRLDTRDVEHNVS
rP2X1   FNLGYVVRESGQDFRSLAEKGGVVGITIDWKCDLDWHVRHCKPIYQFHGLYG---EKNLS
hP2X1   FQLGYVVQESGQNFSTLAEKGGVVGITIDWHCDLDWHVRHCRPIYEFHGLYE---EKNLS
        *.** :*.::*:.*  :* :**::** :.*.*:**  .  * * *.*: *     *:*:*


           290                  309
rP2X4   PGYNFRFAKYYRDLAGKEQRTLTKAYGIRFDIIVFGKAGKFDIIPTMIN----
hP2X4   PGYNFRFAKYYRDLAGNEQRTLIKAYGIRFDIIVFGKAGKFDIIPTMIN----
rP2X1   PGFNFRFARHFVQN-GTNRRHLFKVFGIHFDILVDGKAGKFDIIPTMTTIGSG
hP2X1   PGFNFRFARHFVEN-GTNYRHLFKVFGIRFDILVDGKAGKFDIIPTMTTIGSG
        **:*****:::  :  *.: * * *.:**:***:* ************ .
```

Figure 6.1 Multiple sequence alignment of human and rat $P2X_1$ and $P2X_4$ receptor. ClustalW (Thompson *et al.*, 1994) was used to generate the alignment. The degree of similarity is illustrated beneath the alignment with consensus symbols such that '*' means the residues in that column are identical in all the sequences in the alignment, ':' means that the conserved substitutions have been observed and '.' means that semi-conserved substitutions are observed. Residues in red represent those residues thought to be involved in ATP action at the $hP2X_1$ receptor and the corresponding residues in the $P2X_4$ receptor. Numbering refers to the $hP2X_1$ receptor.

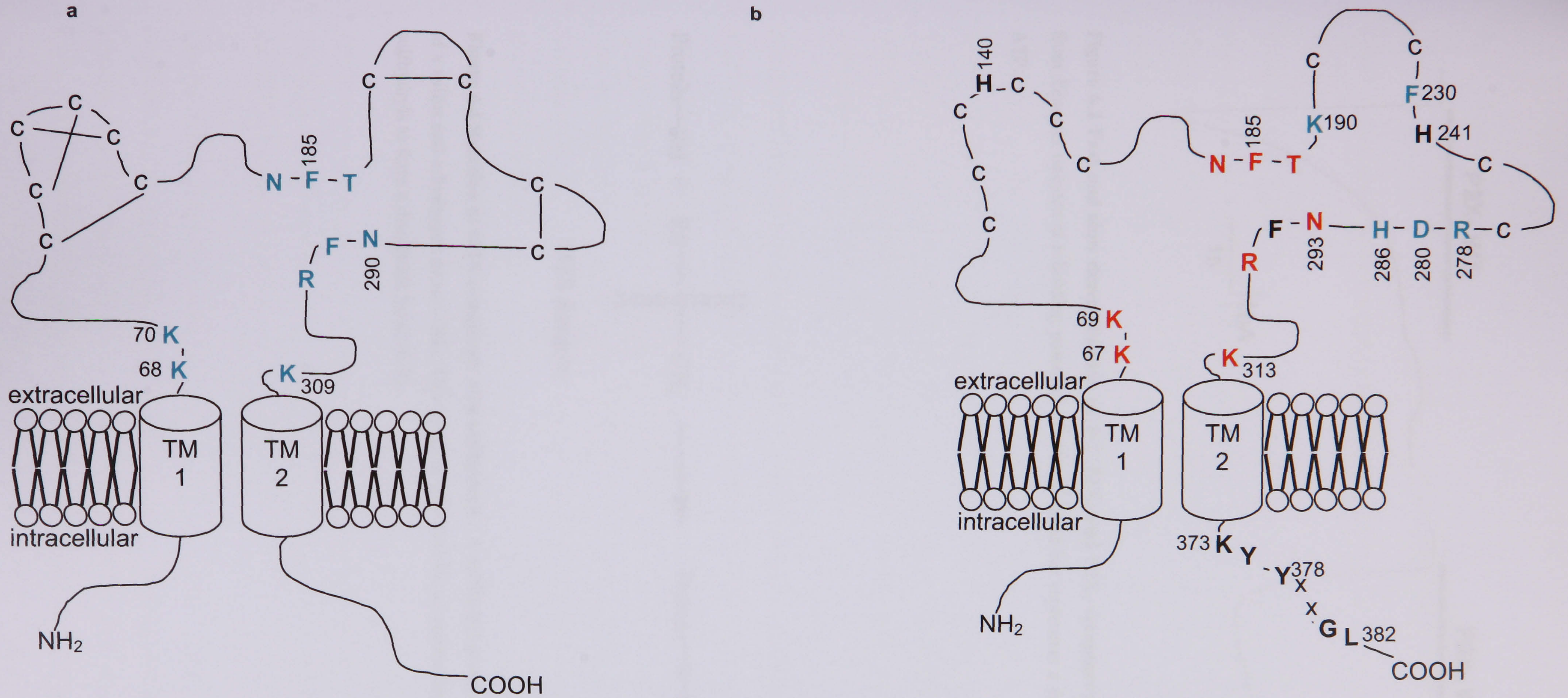


Figure 6.3 Comparison of the $P2X_1$ receptor and the $P2X_4$ receptor. Both subunits share a common topology, two transmembrane domains (TM1, TM2), intracellular amino and carboxy termini and a large ligand-binding extracellular loop. **a,** $hP2X_1$ receptor subunit. Ten conserved cysteine residues are likely to form five disulphide bonds (Ennion & Evans, 2002a). Residues predicted to be involved in binding ATP (Ennion *et al.*, 2000; Roberts & Evans, 2004). **b**, $rP2X_4$ receptor subunit. Ten conserved cysteine residues in the extracellular loop. Residues predicted to be involved in ATP-binding by a homology model based on the catalytic domain of the aminoacyl-tRNA synthestases (Yan *et al.*, 2005). Residues mutated to cysteine that correspond to equivalent positions in the $hP2X_1$ receptor suggested to be involved in ATP-binding.

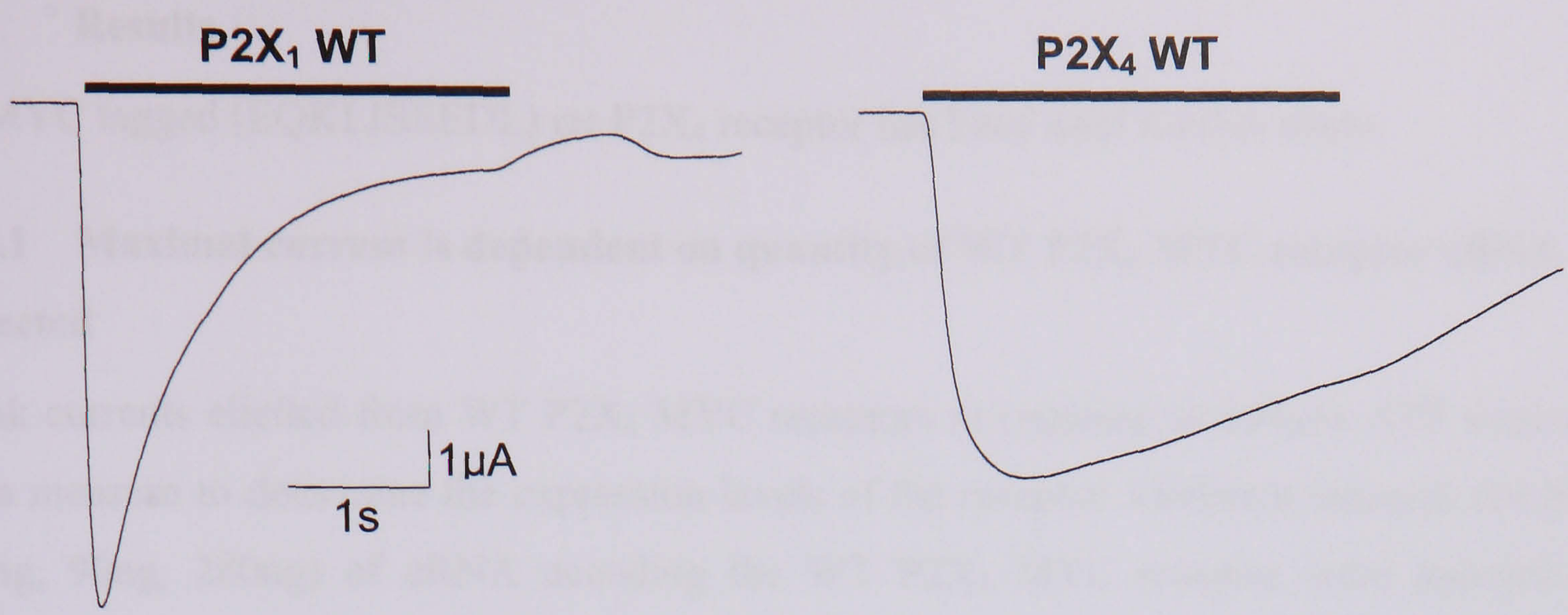


Figure 6.2 Fast and slow desensitization of WT $P2X_1$ and $P2X_4$ receptors respectively. Recordings taken from *Xenopus* oocytes at a holding potential of -60mV. Bold bar represents a three second application 100µM ATP.

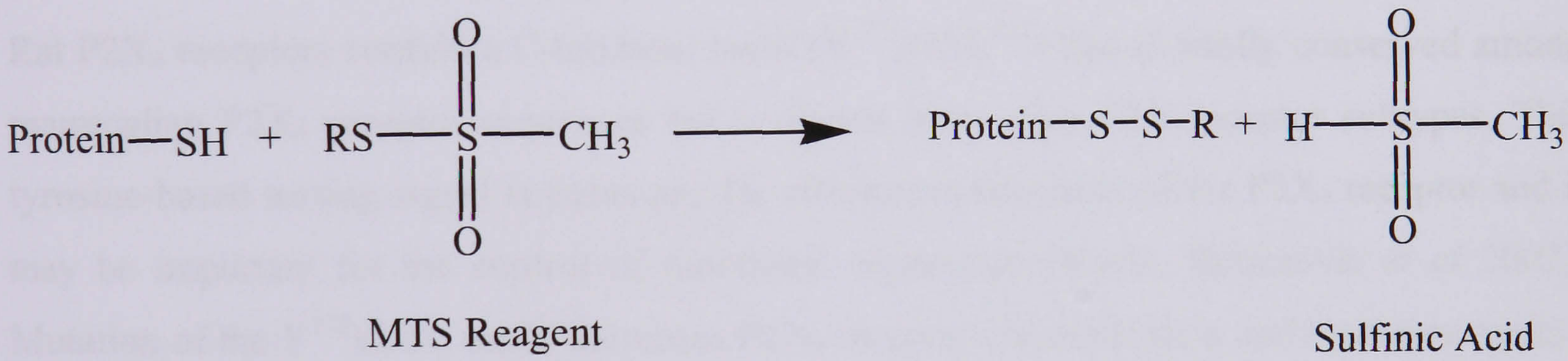


Figure 6.4 Reaction of MTS compound with sulfhydryls. A sulfhydryl group is a functional group composed of a sulfur and a hydrogen atom (-SH). This group is provided by a cysteine residue. MTS reagents react with sulfhydryls to form a disulphide bond (S-S).

6.2 Results

A MYC tagged (EQKLISEEDL) rat $P2X_4$ receptor has been used for this study.

6.2.1 Maximal current is dependent on quantity of WT $P2X_4$ MYC receptor cRNA injected

Peak currents elicited from WT $P2X_4$ MYC receptors in response to 100μM ATP were used as a measure to determine the expression levels of the receptor. Different amounts (16.67ng, 50ng, 90ng, 280ng) of cRNA encoding the WT $P2X_4$ MYC receptor were injected into *Xenopus Laevis* oocytes and currents were recorded, at least 2 days post-injection, in response to 100μM ATP. The mean current amplitudes were plotted against the ng cRNA injected (Figure 6.5). If less than 50ng cRNA was injected, sub maximal amplitudes of current were recorded, greater than 50ng cRNA resulted in the maximal amplitude of current being achieved. Therefore all subsequent injections of WT or mutant $rP2X_4$ receptor were 50ng. Observations of small peak amplitudes of current were seen during these initial experiments with the WT $P2X_4$ MYC receptor but published mutagenesis work suggested this could be improved using a $P2X_4$ receptor mutant, Y378A.

6.2.2 $rP2X_4$ MYC receptor mutant Y378A increases functional responses in oocytes

Rat $P2X_4$ receptors contain a C-terminal motif (Y^{378}xxGL^{382}) that is totally conserved among mammalian $P2X_4$ receptor sequences but is absent from other P2X receptor subtypes. This tyrosine-based sorting signal is necessary for efficient endocytosis of the $P2X_4$ receptor and it may be important for the control of functional expression (Royle, Bobanovic *et al* 2002). Mutation of the Y^{378}xxGL motif decreases $P2X_4$ receptor internalization and increases surface expression (Royle, Bobanovic et al. 2002; Royle, Qureshi *et al* 2005). Mutation of Y378 to alanine in the $rP2X_4$ receptor significantly increased the peak responses compared to the $P2X_4$ WT receptor (Toulme, Soto *et al* 2006). In addition, two residues K373 and Y374 in the C terminus of the $P2X_4$ receptor (located before the endocytosis motif) have been identified by mutagenesis as being key determinants of $P2X_4$ receptor desensitization indicating a role for this region of the receptor in determining the duration of the physiological action of ATP at the $P2X_4$ receptor (Fountain & North, 2006).

Similar results were found by introducing mutant Y378A to the $rP2X_4$ MYC WT receptor (Figure 6.6). The mean peak current amplitude was significantly increased (~70-fold) (n=6-8). For this reason $rP2X_4$ Y378A was used as the WT $P2X_4$ receptor. Any reference to the WT $P2X_4$ receptor from this point will refer to the Y378A $rP2X_4$ MYC receptor.

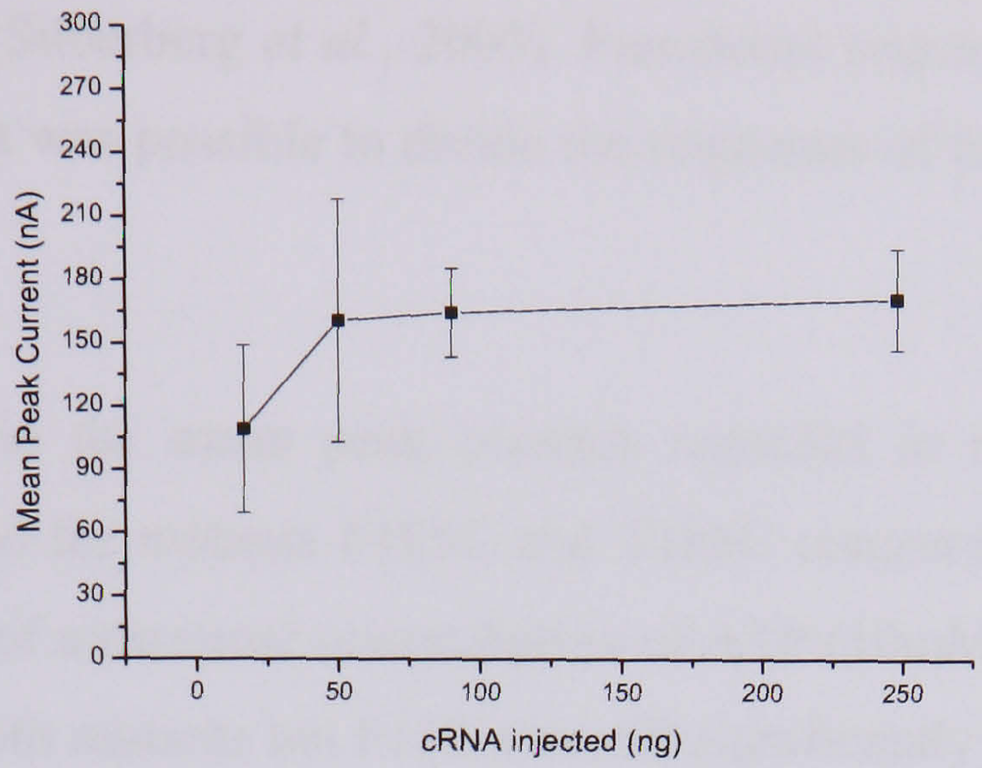


Figure 6.5 Mean peak current amplitude (in response to 100μM ATP). Plot of mean peak current amplitude against ng cRNA injected encoding rP2X$_4$ WT MYC. Currents recorded 2 days after injection in *X. laevis* oocytes. Bars show standard error or the mean (n=5-8).

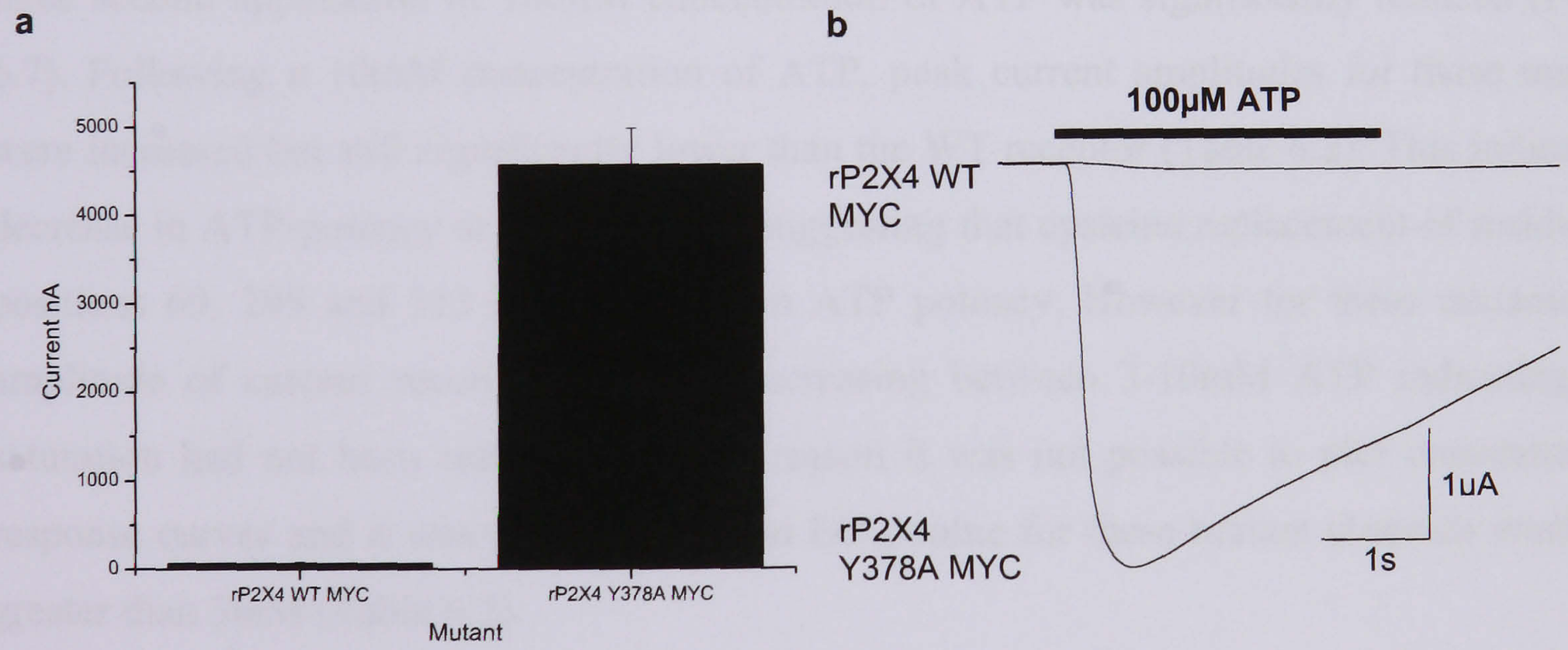


Figure 6.6 Mean current rat P2X$_4$MYC WT and mutant Y378A. a, Mutant Y378A significantly increases the amplitude of current compared to P2X$_4$ WT receptors (***, $p<0.001$) (n=6-8). **b**, currents recorded in response to 100μM ATP at rP2X$_4$ WT MYC receptor and mutant Y378A. Holding membrane potential was -60mV.

6.2.3 Effect of cysteine replacement on the response to ATP

At wild-type (Y378A rP2X$_4$ MYC receptor) rat P2X$_4$ receptors ATP evoked a concentration-dependent desensitizing inward current with an EC$_{50}$ of ~13μM (Table 6.2), similar to that described by Silberberg (Silberberg *et al.*, 2005). Functional responses were recorded for all the mutant channels but it was possible to divide the responses of three groups based on their response to ATP.

A significant reduction in the mean peak currents recorded in response to 100μM ATP (Figure 6.7) was observed for mutants F185C and T186C compared to the WT channel. A three-second application of a maximal concentration of ATP (10mM) increased the amplitude of current recorded for both mutants but F185 was still significantly reduced compared to WT (Figure 6.8). These results suggest a decrease in ATP potency at these mutant channels (~20-fold and ~50-fold for F185C and T186C respectively). Concentration-response curves quantify the effect of mutants F185C and T186C on ATP potency at the P2X$_4$ receptor (Figure 6.9).

The peak current amplitude recorded for mutants K69C, R295C and K313C in response to a three second application of 100μM concentration of ATP was significantly reduced (Figure 6.7). Following a 10mM concentration of ATP, peak current amplitudes for these mutants were increased but still significantly lower than the WT receptor (Table 6.2). This indicates a decrease in ATP-potency at these mutants suggesting that cysteine replacement of residues at positions 69, 295 and 313 has an effect on ATP potency. However for these mutants, the amplitude of current recorded was still increasing between 3-10mM ATP indicating that saturation had not been reached. For this reason it was not possible to plot concentration-response curves and it was predicted that an EC$_{50}$ value for these mutant channels would be greater than 3mM (Table 6.2).

High concentrations of ATP (up to 10mM) evoked a significantly reduced response that was barely detectable compared to WT (4559nA) for mutants K67C and N293C (34nA and 33nA respectively). It was not possible to determine an EC$_{50}$ value, the rise-time or decay-time for these mutants and they were not studied further (Table 6.2). Taken together, these results suggest that the replacement of residues K67, K69, F185, T186, N293, R295 and K313 with cysteine has an effect on ATP action at the rP2X$_4$ receptor which is consistent with mutagenesis data for the hP2X$_1$ receptor.

Mutant	ATP EC_{50} (μM)	ATP pEC_{50}	Peak (nA)	Peak (nA) (max)	Rise Time (ms)	% Peak Current
$P2X_4$ WT	12.86	4.89 ± 0.02	4559 ± 414	4559 ± 414	326 ± 25	75 ± 5
K67C	-	-	13 ± 3***	34 ± 12***	-	-
K69C	>3mM	<4	15 ± 4***	1563 ± 364***	465 ± 88	79 ± 3
F185C	265.1**	3.58 ± 0.07***	88 ± 28***	1818 ± 243***	284 ± 25	64 ± 1
T186C	649.5***	3.19 ± 0.04***	413 ± 117***	3582 ± 656	348 ± 38	74 ± 3
N293C	-	-	6 ± 5***	33 ± 16***	-	-
R295C	>3mM	<4	4 ± 1***	513 ± 101***	569 ± 147	82 ± 4
K313C	>3mM	<4	2 ± 0.49***	487 ± 98***	862 ± 79***	92 ± 3*

Table 6.2 Summary of data for ATP action at $P2X_4$ WT and cysteine mutants

EC_{50} data are representative of data collected from three to four oocytes. pEC_{50} is –log10 of the EC_{50}. Mean peak data are for a 100μM and a 10mM concentration of ATP (n = 3-9) except in the case of WT where the peak amplitude is expressed after a maximal concentration of ATP which was 100μM. Rise Time represents the time from 10 to 90% of peak current after an application of a maximal stimulating concentration of ATP (n = 3-7). % Peak Current represents the % of the peak current left after 3s application ATP (n = 4-7). Values shown are representative of the mean ± S.E. of the mean. Data were analyzed by a Dunnetts one-way ANOVA (Analysis Of Variance) to test the control mean (WT) to all the other means. The significance level was set to 0.05

* $P < 0.05$, different from WT as measured by one-way ANOVA.

** $P < 0.01$, different from WT as measured by one-way ANOVA.

*** $P < 0.001$, different from WT as measured by one-way ANOVA.

- Indicates mutants where the peak amplitudes were significantly smaller than WT and bearly detectable. Therefore, it was not possible to calculate the EC_{50} value, the Rise-time or the % Peak Current.

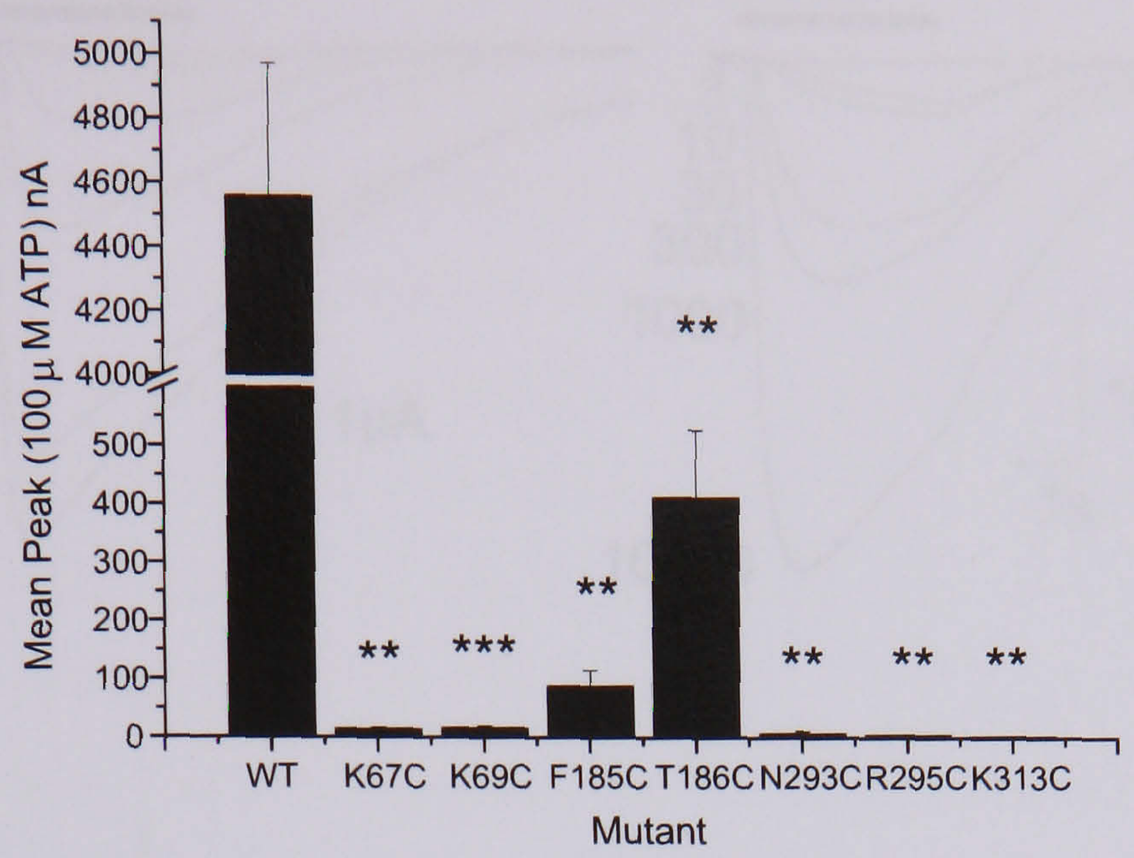


Figure 6.7 Mean peak current (100µM ATP) rat $P2X_4$MYC WT and cysteine replacement mutants. All mutants show a significant decrease in the amplitude of current compared to $P2X_4$ WT receptor (***, $p<0.001$) (n=3-9).

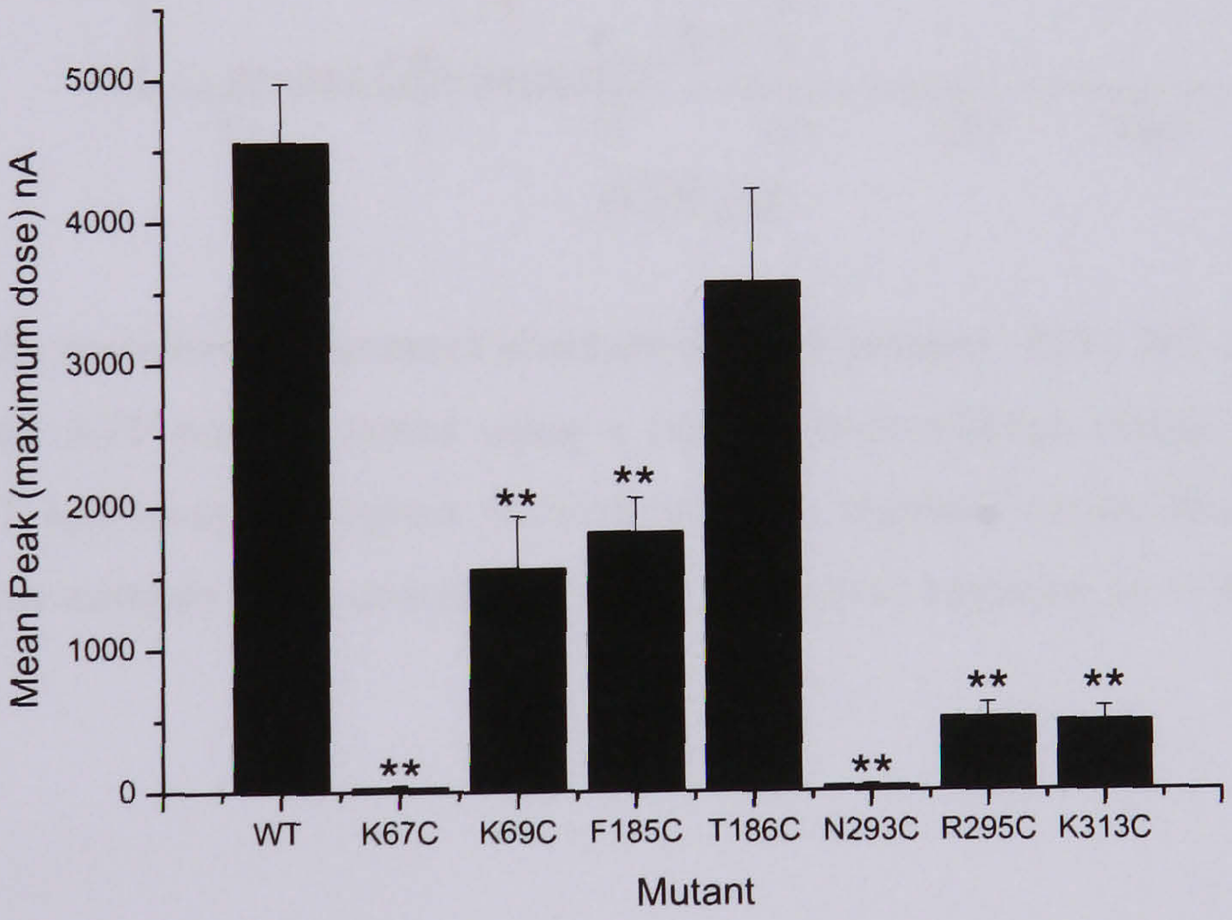


Figure 6.8 Mean peak current (maximal concentration ATP) rat $P2X_4$MYC WT and cysteine replacement mutants. Mutants, K67C, K69C, F185C, N293C, R295C and K313C show a significant decrease in the amplitude of current compared to $P2X_4$ WT receptor (**, $p<0.01$ ***, $p<0.001$) (n=4-9). Mutant T186C is not significantly different to the WT channel.

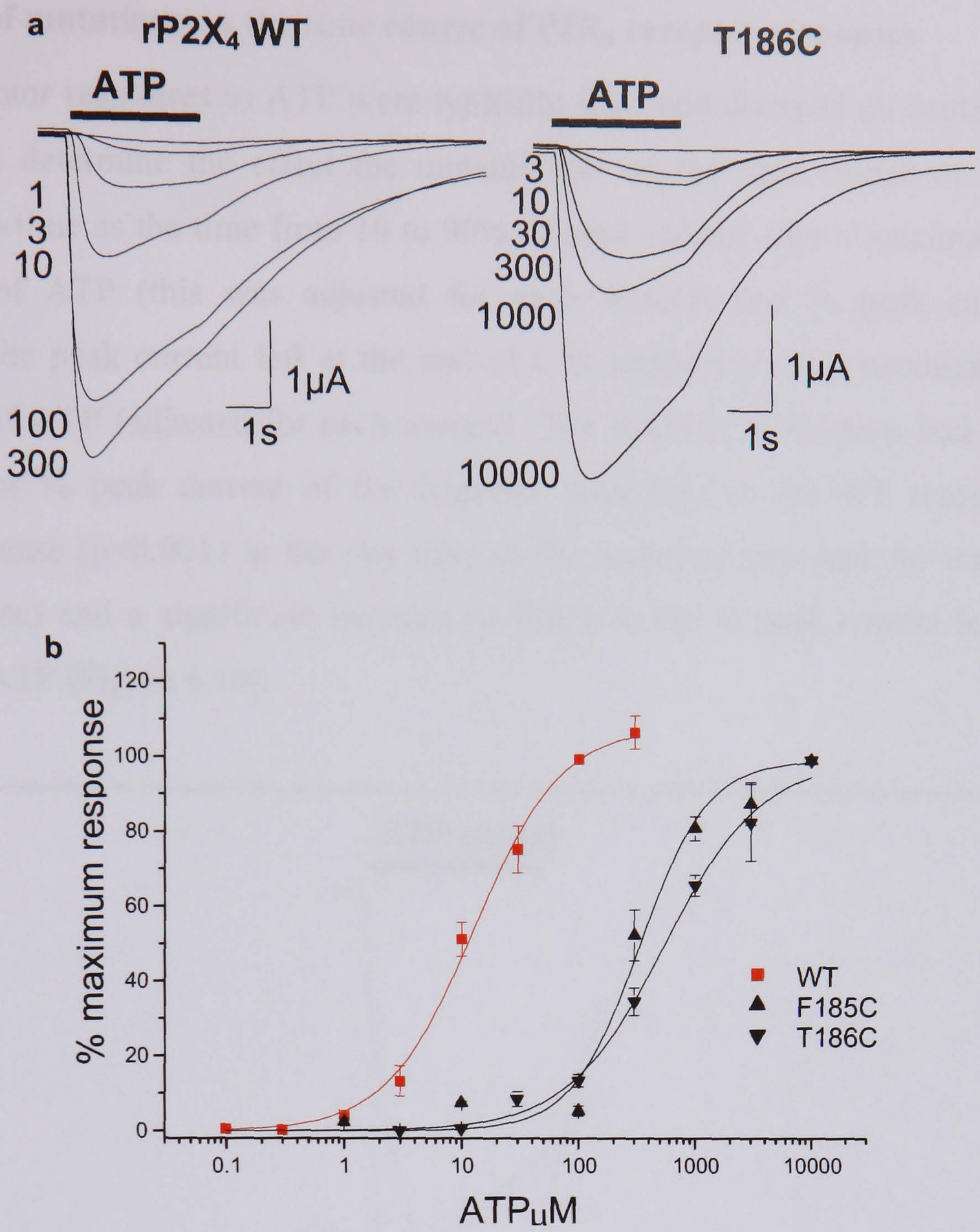


Figure 6.9 Effect of $P2X_4$ cysteine replacement mutants on ATP potency. $P2X_4$ WT or cysteine mutants were expressed in oocytes and ATP potency tested using a two-electrode voltage clamp. **a,** currents recorded in response to ATP at WT and mutant receptors. **b,** Concentration response curves for $P2X_4$ WT and mutants F185C and T186C. Mean currents were normalized to the maximum response (n = 3-4). Holding membrane potential was –60 mV.

6.2.4 Effect of mutations on the time course of $P2X_4$ receptor responses

WT $P2X_4$ receptor responses to ATP were typically slow and decayed on continued agonist application. To determine the effect the mutants had on the time course of response we defined the rise-time as the time from 10 to 90% of peak current after a maximal stimulating concentration of ATP (this was adjusted for each mutant) and % peak current as the percentage of the peak current left at the end of a 3s application of a maximal stimulating concentration of ATP (adjusted for each mutant). The majority of mutants had no effect on the rise time or % peak current of the response compared to the WT receptor. A very significant increase ($p<0.001$) in the rise time of the response was seen for mutant K313C (~3-fold increase) and a significant increase ($p<0.05$) in the % peak current left after a 3s application of ATP (Figure 6.10).

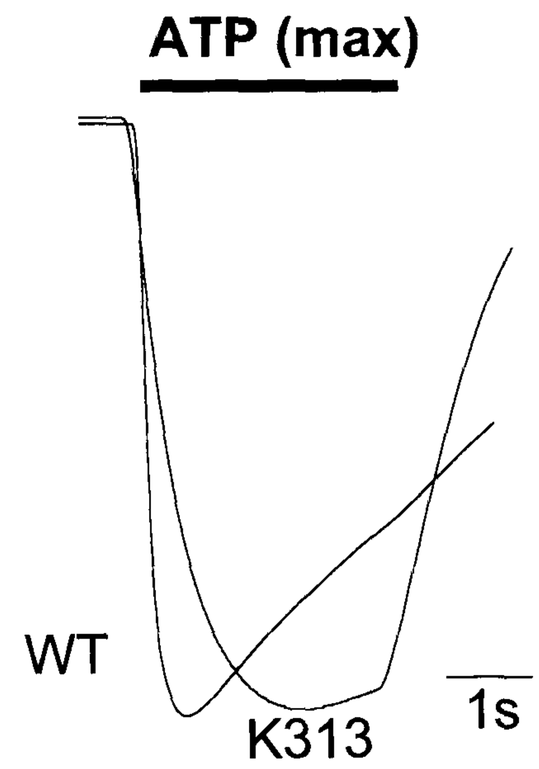


Figure 6.10 Effect of $P2X_4$ mutant K313C on the Time Course of Response. A significant increase in the rise time of response was recorded for mutant K313C (~3-fold) and a significant decrease in the decay time was also seen. Data normalised to peak responses (maximum concentration ATP).

6.2.5 Effect on receptor expression

Significantly reduced peak currents in response to ATP were seen for mutants K67C, K69C, N293C, R295C and K313C compared to WT (Table 6.2). To investigate the level of total $P2X_4$ receptors for WT and mutant expression a $P2X_4$ receptor selective antibody was used. Initial studies to optimise the concentration of anti-$P2X_4$ antibody (Alamone, Israel) suggested that a 1:150 dilution of anti-$P2X_4$ antibody (Alamone) was most effective (Figure 6.11).

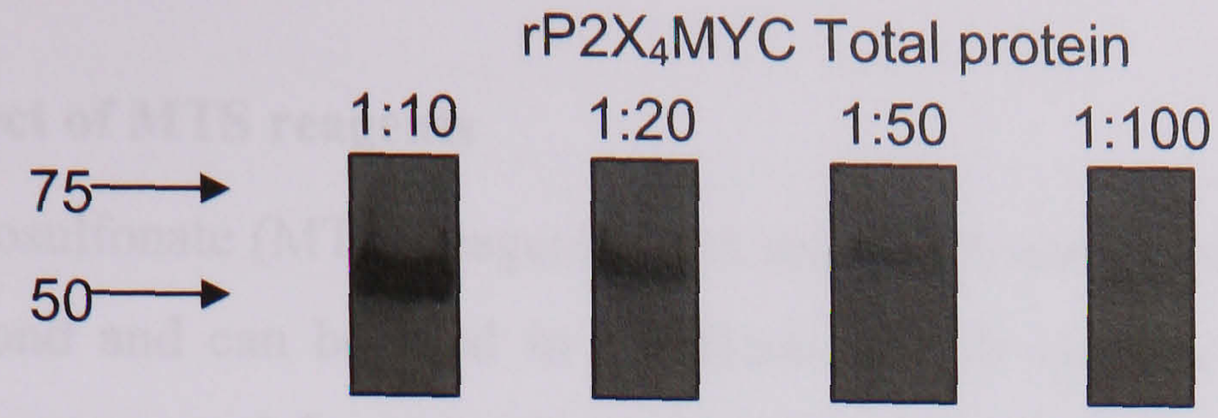


Figure 6.11 Alamone anti-$P2X_4$ antibody used at different concentrations. Western blot analysis using Alamone anti-$P2X_4$ antibody on $rP2X_4$ MYC total cell lysate at different concentrations. Immunoreactivity was observed at the correct *M*r for $P2X_4$ protein.

Sulfo-NHS-LC-Biotin treatment (that labels all cell surface proteins) was used to estimate the expression of $P2X_4$ receptors on the cell-surface. Wild-type $P2X_4$ receptors were detected in both the total and surface cell lysates as a single band of ~65kDa. Mutants K67C, K69C, N293C, R295C and K313C were all detected in the total cell lysates at similar levels to the WT receptor. Surface expression of mutants K67, K69C, N293C, R295C and K313C were seen at similar levels to the WT receptor. This suggests that inefficient receptor trafficking is unlikely to be the cause of the reduced current amplitude recorded for these mutants but it may result in part from changes in the kinetics of channel opening (Figure 6.12)

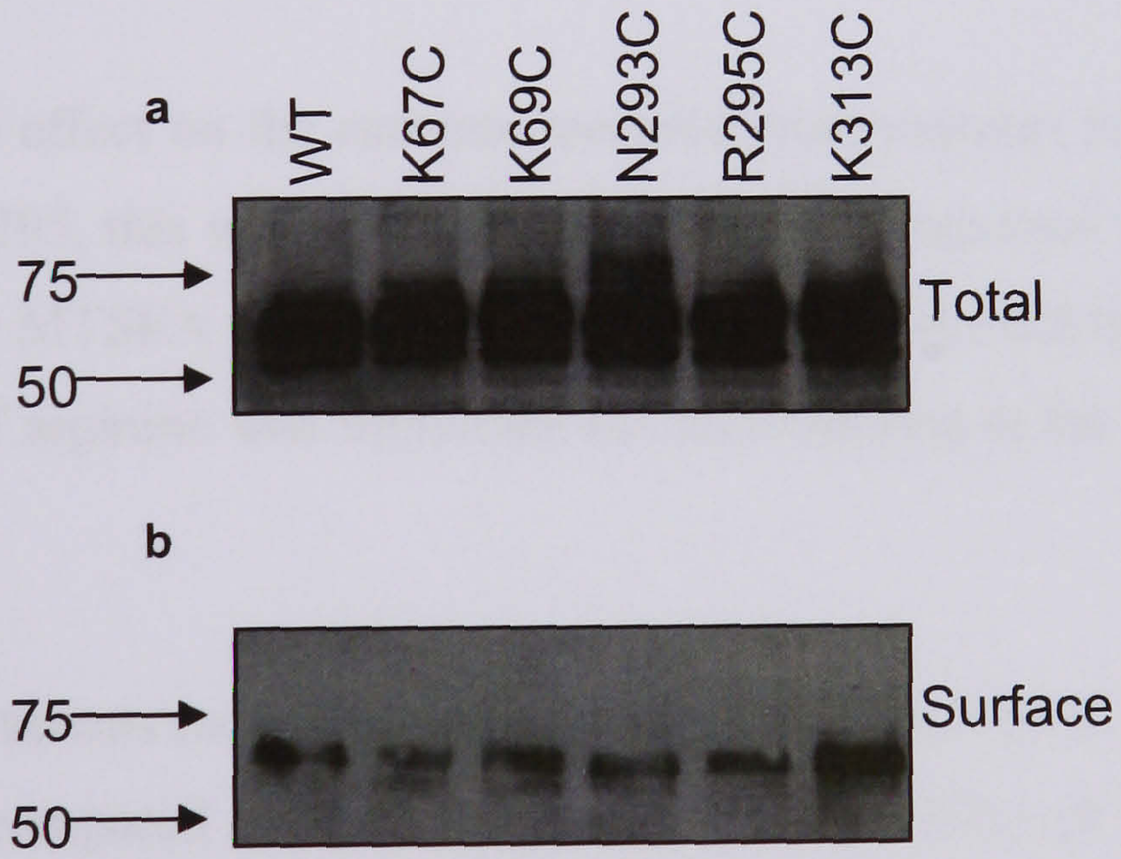


Figure 6.12 Effect of cysteine mutations on cell surface expression. Replacement of residues K67, K69 N293, R295 and K313 with cysteine caused a significant reduction in the mean peak amplitude compared to WT. ***a,b*** to investigate the cause of this, Western blot analysis was carried out using $P2X_4$ antibody on total and cell surface-expressed protein. Cell surface proteins were extracted by sulfo-NHS_LC-Biotin labelling and isolated with streptavidin-agarose beads. Immunoreactivity was observed at the correct *M*r for $P2X_4$ protein and was not seen for non-injected oocytes (data not shown).

6.2.6 Effect of MTS reagents

Methanethiosulfonate (MTS) reagents react selectively and rapidly with sulfhydryls to form a disulfide bond and can be used in combination with cysteine replacement mutagenesis to study the structure and function of proteins. Charged MTS reagents can introduce a positive or negative charge at the position of a previously neutrally charged cysteine residue and this can induce functional changes in a channel that can be electrically recorded (Stauffer & Karlin, 1994). In this study, MTS reagents (1mM) were bath-perfused for 5mins and co-applied with ATP for a 3 second application. Positively charged MTSEA and negatively charged MTSES had no effect on the currents elicited by WT $P2X_4$ receptors. To determine whether MTS reagents had an effect on ATP potency at $P2X_4$ WT receptors, oocytes expressing WT $rP2X_4$ receptors were incubated overnight in MTSEA (1mM) and MTSES (1mM). No effect was observed on the ATP potency of the WT receptors (Figure 6.13) (Roberts & Evans, 2007).

Both K67C and N293C elicited significantly reduced peak currents in response to 10mM ATP compared to the WT $rP2X_4$ MYC receptor. For this reason it was difficult to assess the effects of MTS reagents at these mutant receptors and so the results shown in Table 6.3 and Figure 6.14 are not considered reliable.

MTS reagents had no effect on the currents recorded from mutants F185C or R295C (Table 6.3). In the case of R295, this was interesting as it might be expected that the replacement of the positive charge by MTSEA would potentiate the currents elicited by ATP at this mutant if the positive charge of arginine was important for ATP-binding at the receptor. This was not observed for R295C.

MTSEA potentiated currents recorded from mutants K69C and K313C (increase of 122% and 3453% respectively compared to WT) (Figure 6.14). The effect of the positive charge of MTSEA on the amplitude of current for these mutant $P2X_4$ channels suggests that these residues may play a role in ATP-action at the receptor. However, this change in amplitude of response may be due to a change in ATP potency at the receptor and/or an effect on channel gating and ionic permeation. To find out whether a change in ATP potency was involved, the effect of MTSEA on the concentration-response curves of mutants K69C and K313C was determined. Recordings following an overnight incubation of 1mM MTSEA resulted in a left-shifted concentration response curve for both mutants (Figure 6.15 and Figure 6.16 respectively). This suggests that MTSEA was interfering with ATP-binding at the receptor.

The positive charge of both these lysine residues has been implicated in the $hP2X_1$ receptor in binding of the negatively charged phosphates of ATP (Ennion *et al.*, 2000) and a similar role might be suggested for the $rP2X_4$ receptor based on these results.

Currents recorded from mutant T186C were inhibited in the presence of MTSEA and significantly inhibited in the presence of MTSES (Figure 6.17a). To establish the cause of this result, the effect of MTS reagents on the concentration-response curves of T186C were determined. No or little effect was seen on the EC_{50} value following an overnight incubation of oocytes expressing mutant receptor T186C with both MTSEA (1mM) and MTSES (1mM) respectively (Figure 6.17b). This may suggest that localized charge in this position does not contribute to ATP-binding but the increase in size due to the binding of bulky MTS reagents may affect the functionality of the channel.

Unpublished work carried out after this study showed that mutant F294C ($P2X_4$ receptor numbering) has an EC_{50} ~60μM and is potentiated by both MTSEA and MTSES. Such observations are consistent with a role of F294 in ATP-binding at the $rP2X_4$ receptor which agrees with previous mutagenesis work in the $hP2X_1$ receptor (Roberts & Evans, 2004).

Mutant	MTSEA % relative to control	MTSES % relative to control
$P2X_4$ WT	110 ± 29	90 ± 6
K67C	844 ± 378	89 ± 27
K69C	245 ± 163	86 ± 11
F185C	93 ± 7	95 ± 2
T186C	21 ± 8	4 ± 0.3**
N293C	118 ± 9	74 ± 20
R295C	127 ± 28	61 ± 20
K313C	3908 ± 1642**	81 ± 2

Table 6.3. Summary of data for effect of MTS reagents at cysteine mutants relative to $P2X_4$ WT
Data represents % relative to control (n=3-4). Values shown are representative of the mean ± S.E. of the mean.

** $P < 0.01$, different from WT as measured by one-way ANOVA.

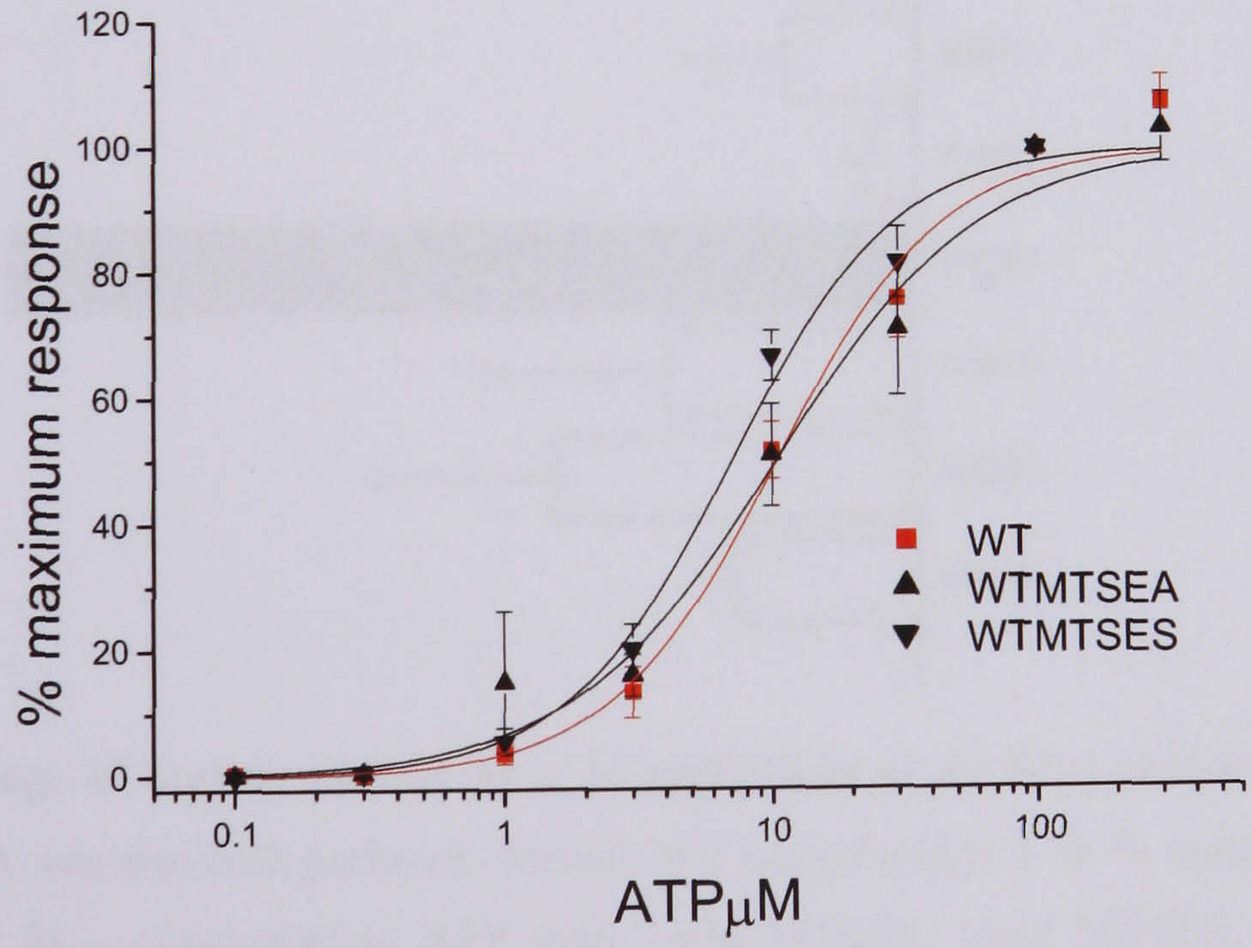
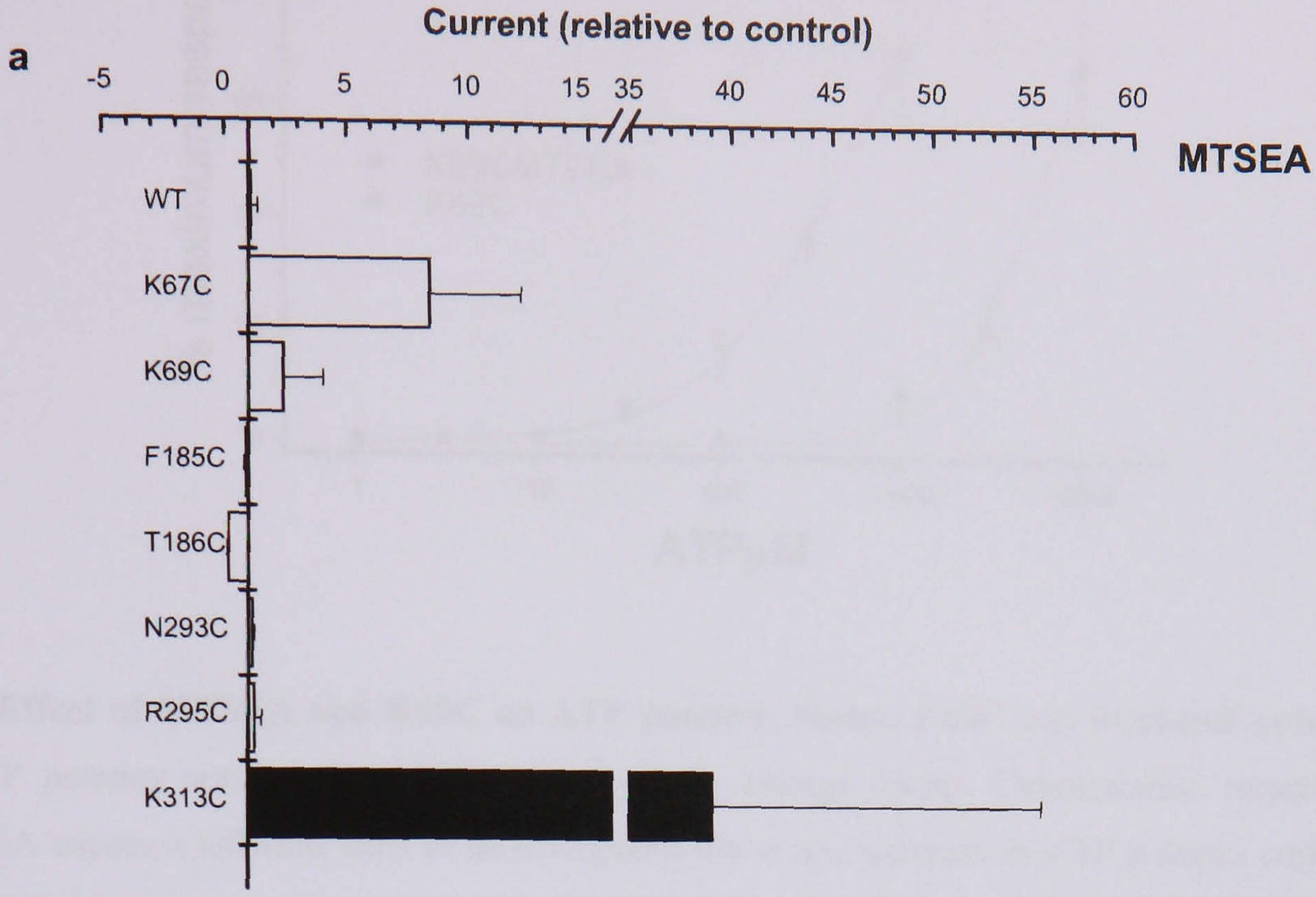


Figure 6.13 Effect of MTSEA and MTSES $rP2X_4$ WT MYC on ATP potency. $rP2X_4$ WT MYC were expressed in oocytes and incubated overnight in 1mM MTSEA or MTSES. ATP potency was tested using a two-electrode voltage clamp. Concentration response curves for $rP2X_4$ WT, $rP2X_4$ WT+MTSEA, $rP2X_4$ WT+MTSES. Mean currents were normalized to the maximum response (n = 3-4). Holding membrane potential was –60 mV.



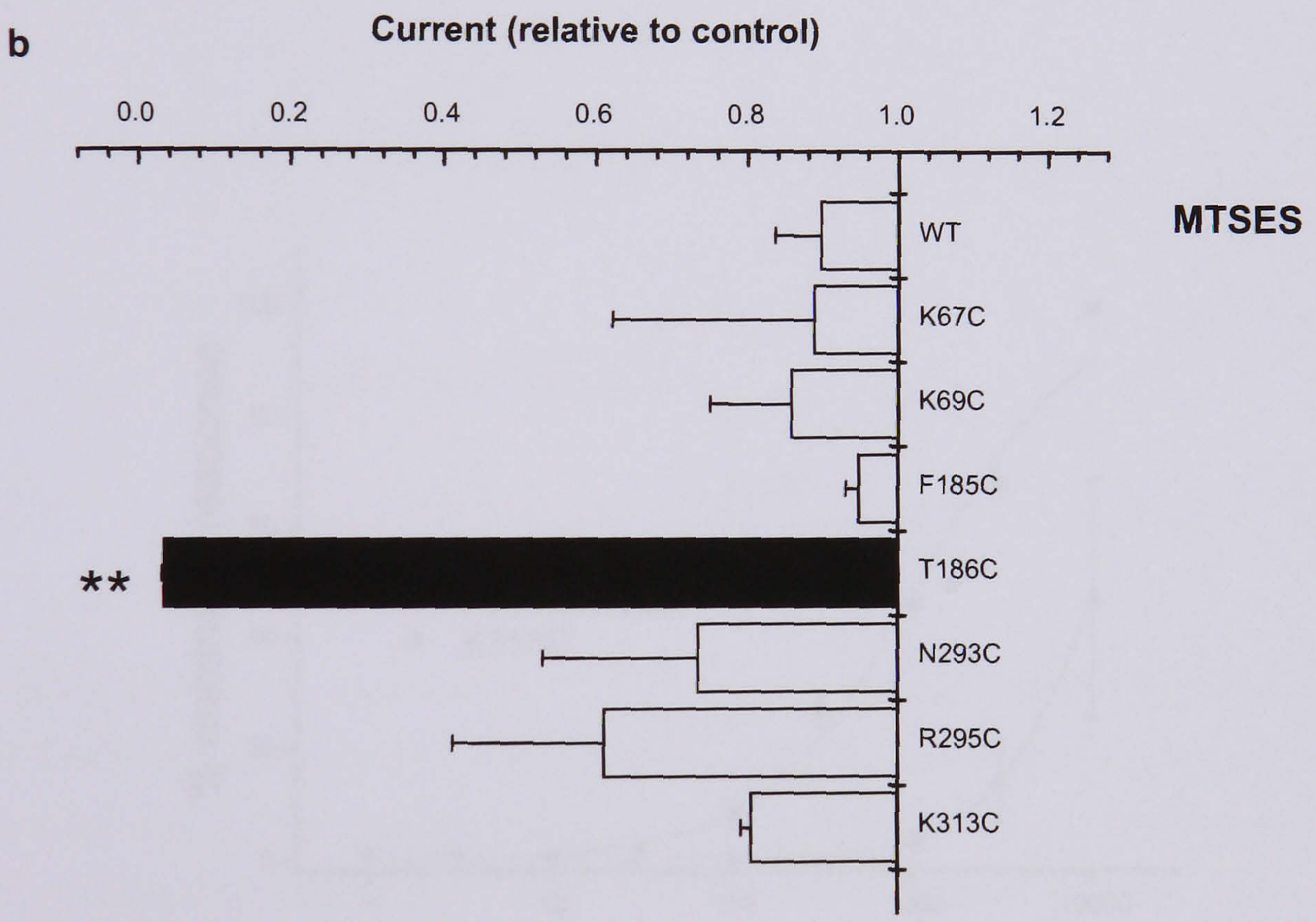


Figure 6.14 a, % change of current elicited by a 3s application of an EC_{50} concentration ATP with 1mM MTSEA. 1mM MTSEA was also bath perfused (5mins). WT normalised to 1. **b**, % change of current elicited by a 3s application of an EC_{50} concentration ATP with 1mM MTSES. 1mM MTSES was also bath perfused (5mins). WT normalised to 1.

** $P < 0.01$, different from WT as measured by one-way ANOVA.

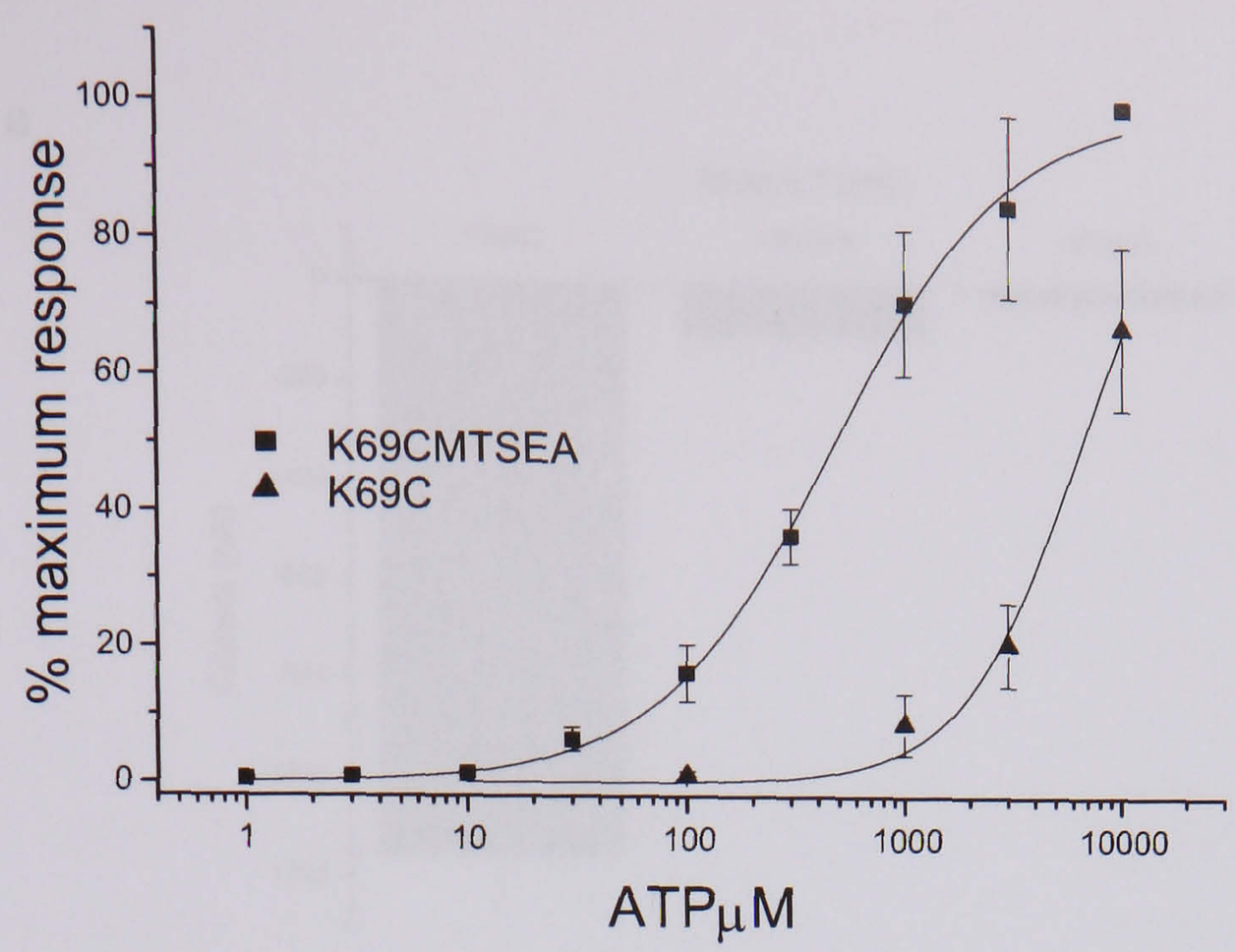


Figure 6.15 Effect of MTSEA and K69C on ATP potency. Mutant K69C was incubated overnight in 1mM MTSEA. ATP potency was tested using a two-electrode voltage clamp. Concentration response curves for K69C. MTSEA causes a leftward shift in dose response curve and increase in ATP potency compared to non-incubated K69C. Mean currents were normalized to the maximum response recorded at 10mM ATP (n = 3-4). Holding membrane potential was –60 mV.

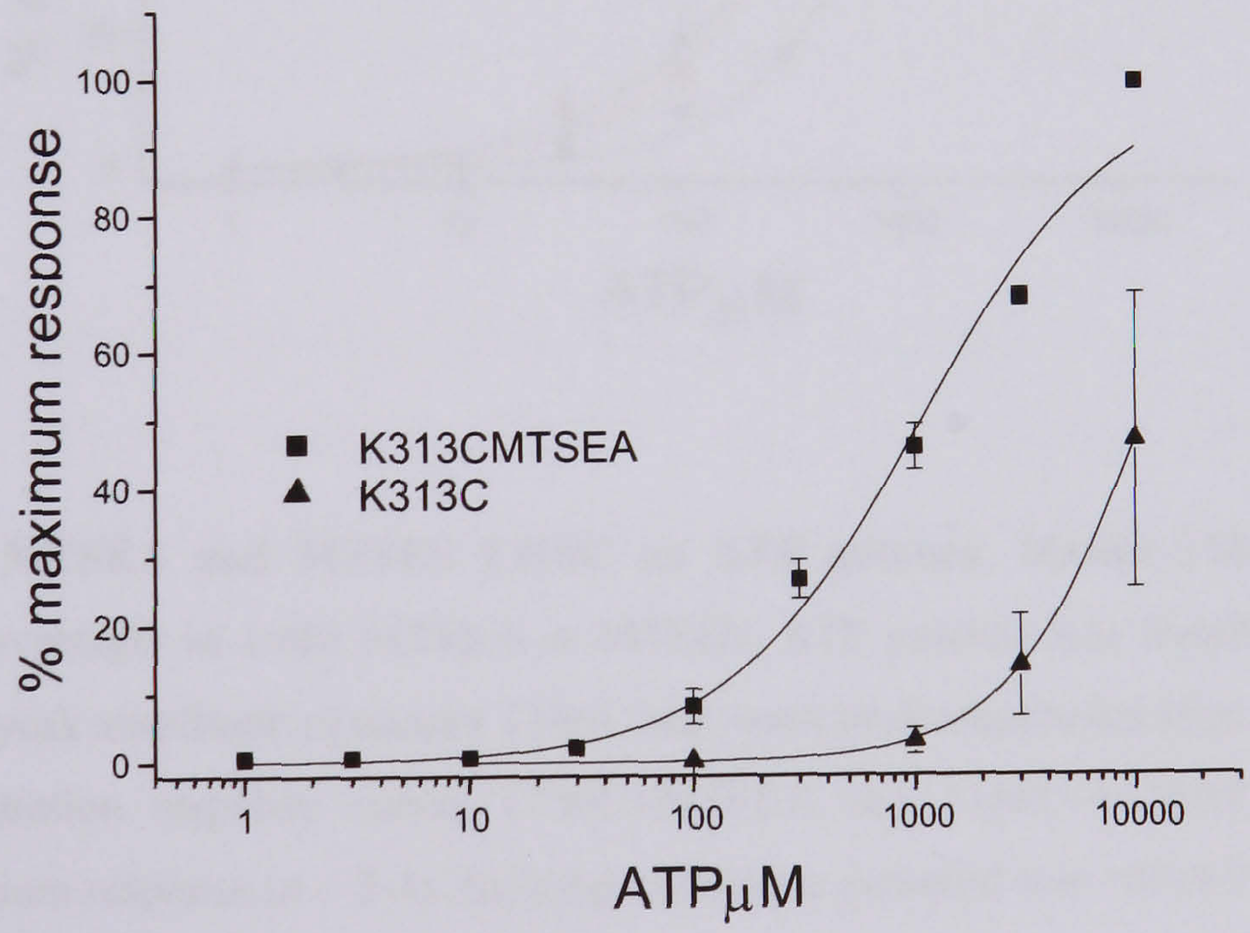


Figure 6.16 Effect of MTSEA and K313C on ATP potency. Mutant K313C was incubated overnight in 1mM MTSEA. ATP potency was tested using a two-electrode voltage clamp. Concentration response curves for K313C. MTSEA causes a leftward shift in dose response curve and increase in ATP potency compared to non-incubated K313C. Mean currents were normalized to the maximum response recorded at 10mM ATP (n = 3). Holding membrane potential was –60 mV.

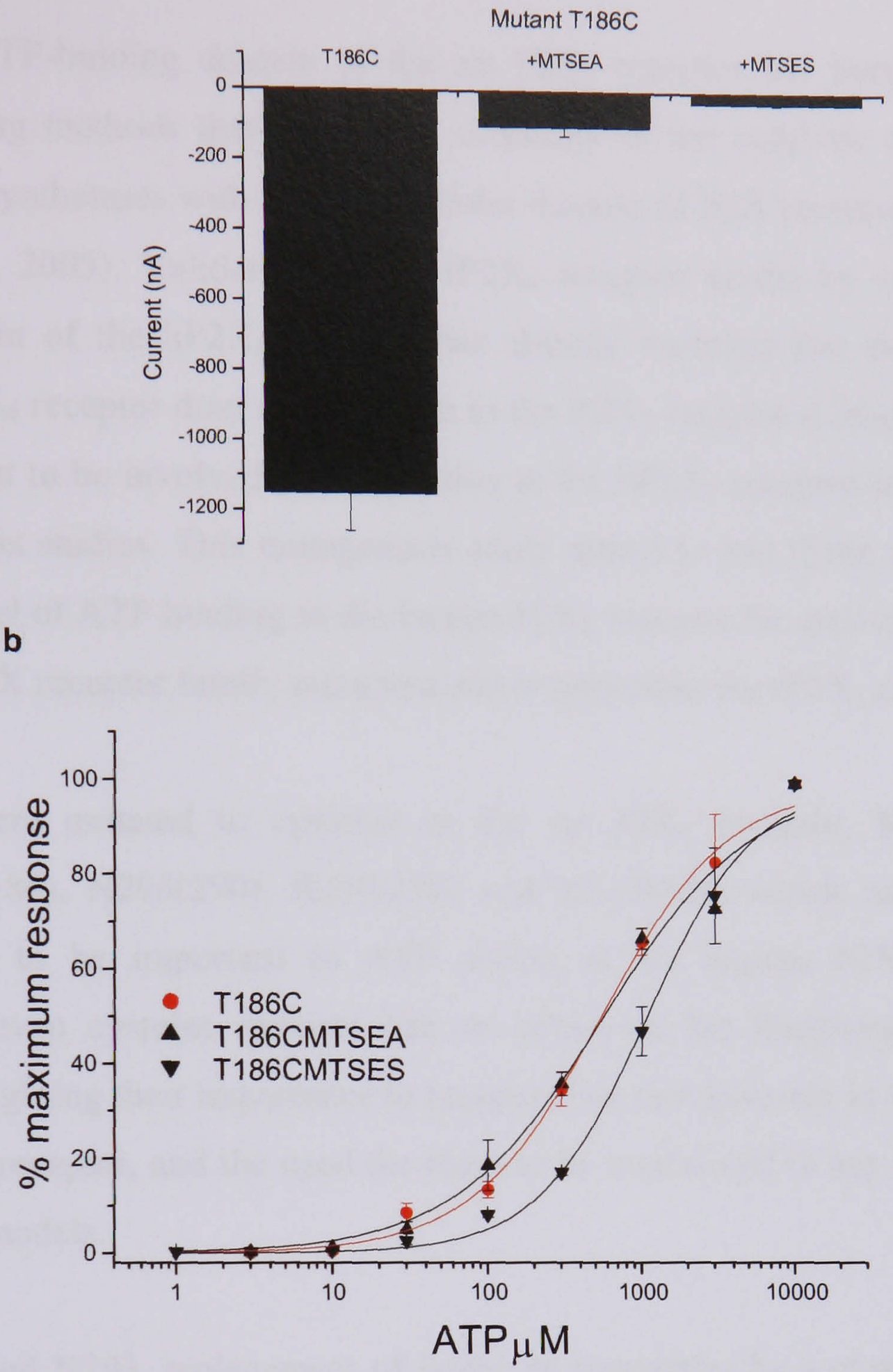


Figure 6.17 Effect of MTSEA and MTSES T186C on ATP potency. Mutant T186C were expressed in oocytes and incubated overnight in 1mM MTSEA or MTSES. ATP potency was tested using a two-electrode voltage clamp. **a**, Mean peak amplitude of mutant T186C and mean peak amplitudes after application of MTSEA or MTSES. **b**, Concentration response curves T186C+MTSEA and T186C+MTSES. Mean currents were normalized to the maximum response (n = 3-4). Holding membrane potential was –60 mV.

6.3 Discussion

A model of the ATP-binding domain of the rat $P2X_4$ receptor has been developed using homology modelling methods that utilise the similarity of the catalytic domain of class II amino-acyl tRNA synthetases with the extracellular domain of P2X receptors to develop a 3D model (Yan *et al.*, 2005). Validation of the $rP2X_4$ receptor model by mutagenesis of the extracellular domain of the $hP2X_1$ receptor has already revealed that the model of ATP-binding at the $rP2X_4$ receptor does not translate to the $P2X_1$ receptor (Chapter 5). A model of the residues thought to be involved in ATP action at the $hP2X_1$ receptor has been developed through mutagenesis studies. This mutagenesis study aimed to test if the reverse holds true; that is can the model of ATP binding in the human $P2X_1$ receptor be used as a model of ATP-binding in other P2X receptor family members and in particular the $rP2X_4$ receptor?

Seven residues were mutated to cysteine in the rat $P2X_4$ receptor, K67(68), K69(70), F185(185), T186(186), N293(290), R295(292) and K313(309) which correspond to those residues predicted to be important in ATP action at the human $P2X_1$ receptor ($P2X_1$ numbering). All seven cysteine mutants had an effect on the functionality of the $rP2X_4$ receptor thus highlighting their importance to receptor function not only in the $hP2X_1$ receptor but also the $rP2X_4$ receptor, and the need for them to be considered in any subsequent $hP2X_1$ or $rP2X_4$ receptor models.

At positions K67 and N293, replacement of lysine or asparagine by a cysteine residue had a significant effect on channel function. Reponses to 10mM ATP (maximal concentration) evoked a significantly reduced response that was barely detectable compared to WT. Western blot analysis was used to determine whether this was as a result of reduced receptor expression and/or a lack of ability of the channel to open and pass current in response to the binding of ATP. Both mutants (K67C and N293C) were detected in the total cell lysates at a similar level to the WT receptor. Cell surface expression for mutant K67C was detectable although reduced compared to the WT receptor. Mutant N293C was expressed at the cell surface at similar levels to the WT receptor. It is likely that the reduced functional response seen for mutant K67C may result from low levels of receptor expression but cannot solely be attributed to this as the expression was not significantly different to WT (observation only). Instead the replacement of lysine in position 67 and asparagine in position 293 by cysteine could be affecting the proteins ability to fold correctly and in turn affecting the binding sites recognition of the ligand. A combination of both of these factors, low surface expression

(seen for K67C) and changes in channel kinetics could explain the significant reduction in the mean peak amplitude compared to the WT receptor recorded for these mutants. Preliminary results of the effect of MTS reagents suggest that MTSEA potentiates currents recorded from mutant K67C and MTSES has no effect on the amplitude of current recorded from K67C. This would be consistent with K67 playing a role in ATP-action at the $P2X_4$ receptor, similar to its role in the $P2X_1$ receptor (Ennion *et al.*, 2000). In addition K67 (K69 $P2X_2$ receptor numbering) has been mutated to alanine in the $P2X_2$ receptor (Jiang *et al.*, 2000). It was found to be critical for the action of ATP. Taken together these results for the $P2X_1$, $P2X_2$ and $P2X_4$ receptor provide strong evidence for the importance of lysine-67 in ATP-action at the P2X receptor family. MTS reagents did not significantly affect currents recorded from mutant N293C in response to an EC_{50} concentration of ATP. In both cases, K67C and N293C, without further experiments this is pure speculation.

F185 and T186 form a motif of the P2X receptor family that is conserved in all seven receptor subtypes. The role of this motif in ATP-action at the $P2X_1$ receptor has been investigated using mutagenesis (Roberts & Evans, 2004) (Roberts & Evans, 2006) and used to refine models of ATP-binding at the $P2X_1$ receptor. However, this is the first time its role in ATP-action at the $P2X_4$ receptor has been investigated. A significant decrease in the amplitude of functional response was seen for mutant F185C. Concentration-response curves revealed a ~20-fold decrease in ATP potency (EC_{50} 265μM compared to 13μM WT) compared to the WT channel. This indicates a role for F185 in ATP-action at the $P2X_4$ receptor. This agrees with experimental data which suggests a similar role for this residue at the $P2X_1$ receptor (Roberts & Evans, 2004). Methanethiosulfonates, MTSEA and MTSES had no effect on the response of F185C to ATP suggesting it is not the introduction of charge at position 185 that modifies ATP potency, but a localized disruption associated with the binding of MTS reagents.

Similarly to mutant F185C, a ~50-fold decrease in ATP potency was observed for mutant T186C (EC_{50} 650μM compared to13μM WT), and a significant decrease in the mean peak amplitude of current was recorded. T186C did not affect the time course of response. To determine the effects of MTS reagents at this mutant receptor, full concentration response curves were recorded from oocytes injected with mutant cRNA encoding T186C and incubated overnight with MTSEA (1mM) or MTSES (1mM) (Figure 6.17). MTSEA treated T186C had no effect on the potency of ATP and MTSES treated T186C decreased ATP

potency. This suggests that a localised negative charge may affect ATP-binding either directly or by affecting protein folding.

Mutant R295C significantly decreased ATP potency at the $P2X_4$ receptor and a significant decrease in mean peak amplitude was recorded. Western blot analysis revealed that the mutant was expressed at similar levels to the WT receptor at the cell surface protein (Figure 6.12) suggesting that inefficient trafficking of the receptor to the cell surface was not the reason for the reduction in mean peak amplitude. Instead a change in channel kinetics may be responsible. These results suggest that R295 is involved in ATP action at the $P2X_4$ receptor. R295C had no effect on the time course of the response and MTS reagents (MTSEA, MTSES) had no effect on the amplitude of response recorded for R295C. This suggests that charge in this position is not responsible for the change in ATP potency but that binding of MTS reagents causes localized disruption to the channel. However it is not clear based on the experimental evidence given here that this is the case. It may be that R295 is inaccessible and therefore MTS reagents are not able to bind. An MTSEA-biotin assay (see Chapter 2 for details) could differentiate water-exposed cysteines and buried cysteines providing further information on the secondary structure and the role of R295 at the $P2X_4$ receptor. Mutation of the equivalent arginine residue in the $hP2X_1$ receptor (R292 $P2X_1$ receptor numbering) to cysteine revealed a decrease in ATP potency of ~20-fold. This is consistent with previous alanine-scanning mutagenesis (Ennion *et al.*, 2000). An inhibition of current by MTSEA and MTSES was also observed for R292C (Roberts & Evans, 2007). The range of results recorded for the effects of charged MTS compounds at this mutant (R292C $P2X_1$ numbering) suggests that it is not charge that modifies ATP potency instead a localized disruption associated with binding of MTS compounds. Therefore it is unlikely that the positive charge of arginine in this position co-ordinates the binding of the negatively charged phosphates of ATP as previously suggested (Ennion *et al.*, 2000; Roberts & Evans, 2007). However for the $P2X_4$ receptor this cannot be confirmed without further work as suggested previously. Future mutagenesis that replaces arginine at position 295 with another positively charged residue (R295K for example) may give more insight into the role of charge in this position.

Mutants K69C and K313C significantly decreased ATP potency. This is consistent with the 1000-fold rightward shift in the sensitivity of $rP2X_4$ receptor mutant K313A (Yan *et al.*, 2006). Western blot analysis revealed both mutants K69C and K313C were present in the total cell lysates at similar levels to WT receptors and were also detected at the cell surface at comparable levels to WT. This demonstrates that the small functional responses to ATP

application observed for both mutants does not result from low levels of receptor expression and may instead be associated with changes in the kinetics of channel opening. Mutant K69C had no effect on the time course of the response. A significant increase in the rise time and percentage decay of the response was recorded for mutant K313C. Such a reduction in ionic flow may affect channel kinetics.

MTSES (negatively charged) had no effect on mutants K69C and K313C but MTSEA potentiated currents from both K69C and K313C (Table 6.3). In both mutants introducing a positive charge at the position of the cysteine residue caused a leftward shift of the concentration response curve, indicating an increase in ATP potency (mean EC_{50} values 874μM and 1301μM respectively) towards the WT receptor. This suggests that the positive charge provided by the lysine residue in position 69 and 313 in the WT $rP2X_4$ receptor is involved in agonist action at the receptor and could bind the negative phosphates of ATP. However, $rP2X_4$ receptor mutant K313R did not show a significant rescue effect (Yan *et al.*, 2006) and would argue against such a role for K313 in the $rP2X_4$ receptor.

Incorporating these residues (K67, K69, F185, T186, N293, R295 and K313) into the published $rP2X_4$ receptor homology model based on class II aminoacyl-tRNA synthetases (Yan *et al.*, 2005) is not possible for residues K68, K70, F185 and T186 because they lie in the first half of the extracellular domain which is not included in the 3D model or considered in the initial alignment of the P2X receptor family and the class II aminoacyl-tRNA synthetases (Freist *et al.*, 1998). K313 has been incorporated into the ATP-binding site of the $rP2X_4$ receptor model using the template seryl-tRNA synthetases (PDB code 1SES). However, the model could not accommodate the role of this residue in coordination of the α-phophate of ATP and the mutant K313R also did not support this role for K313 (Yan *et al.*, 2006). The $^{293}NFR^{295}$ motif suggested to play a critical role in ATP binding in the $hP2X_1$ receptor is not within close proximity of the bound ATP molecule in the $rP2X_4$ receptor model.

6.4 Conclusion

Extensive mutagenesis studies have revealed not only a structural insight into the $hP2X_1$ receptor but an understanding of the residues and regions of the receptor involved in ATP-binding. The aim of this cysteine replacement study was to investigate whether the model of ATP binding in the $hP2X_1$ receptor could be transferred to the $rP2X_4$ receptor. In other words

are the residues thought to be involved in ATP-binding at the $hP2X_1$ receptor, involved in ATP action at the $P2X_4$ receptor? The residues considered (K68, K70, F185, T186, N290, R292 and K309 $P2X_1$ receptor numbering) are 100% conserved across the P2X receptor family which immediately suggests they may play a functional role across the P2X receptor family. This indeed was the case and all seven cysteine mutants had a significant effect on ATP potency at the $P2X_4$ receptor. A family of receptors that share 40-50% identity of residues would be expected to share similarities in ligand binding sites so this is no surprise. But the importance of highlighting these residues in the $rP2X_4$ receptor will have an important effect on future models of the receptor.

Chapter Seven

7.0 General Discussion and Future Directions

The P2X receptor family could provide numerous drug targets for a wide range of diseases including neurological conditions, renal failure and male infertility (see section 1.6), but the problem is the receptor structure is not well defined and the ATP-binding site neither located nor defined. At present there are no crystal structures available for the P2X receptor family and they share little homology with other ATP-binding proteins. The solution to this has been an iterative cycle of mutagenesis studies and modelling techniques to try and build up a picture of the membrane topology of the seven members of the P2X receptor family and the ATP-binding environments. This thesis aimed to combine mutagensis and bioinformatic techniques in tandem to better understand the structure of the P2X receptor family and the ATP-binding site.

Experimental techniques have defined the membrane topology of the receptor family; a large extracellular ligand-binding loop, two transmembrane domains and intracellular amino- and carboxy- termini. Due to the size of the extracellular segment (~288 amino acids long) it would seem reasonable that it would consist of more than one domain. This thesis defined two domains of the extracellular segment; the cysteine-rich domain (residues 115-169 $P2X_1$ receptor numbering) and the putative ATP-binding domain. Previous 3D models have attempted to consider the entire extracellular segment but this has led to poor quality sequence alignments which in turn have generated poor quality models that do not agree with mutagenesis data. One example of this is a 3D model of the mouse $P2X_1$ receptor that was modelled on a similar structural fold with the KcsA protein (Figure 7.1) (Mager *et al.*, 2006). 3D models in this thesis were produced in tandem with mutagenesis work that first considered the role of conserved glycine residues in the extracellular segment of the $P2X_1$ receptor (see Chapter 3). Glycine, due to its lack of sidechain has the ability to introduce flexibility into a protein and may allow for conformational change on agonist binding. A key functional residue, G250, was identified to play a role in ATP action at the $P2X_1$ receptor due to its predicted location at the C-terminal end of an α-helix. Mutation of G250 to a range of amino acids suggested that either a glycine or a small polar residue (serine) is required in position 250 to allow for normal function of the $P2X_1$ receptor. Movement in this region of the receptor may allow for residues K68 and K309, that are thought to be involved in ATP-binding at the $P2X_1$ receptor to move closer together to form the binding site.

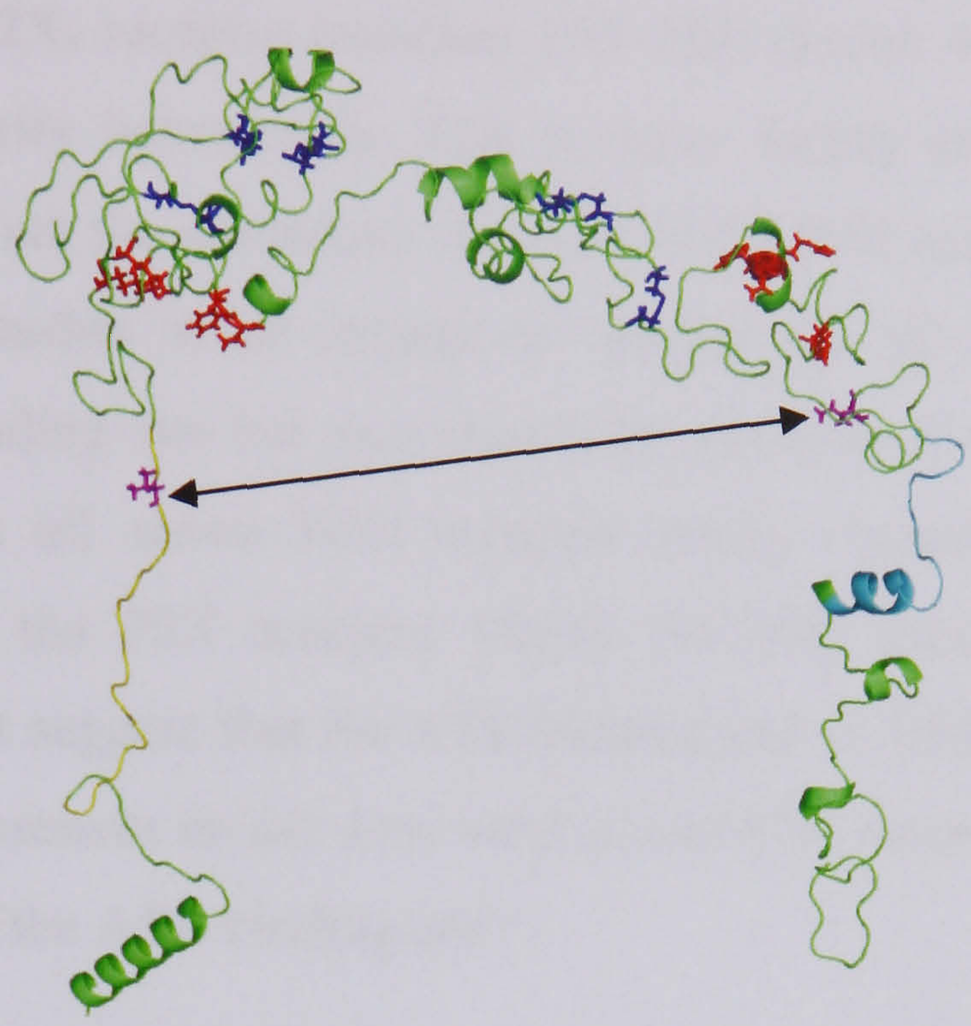


Figure 7.1 The structure of the mouse $P2X_1$ receptor based on a similarity between the $P2X_1$ receptor and the KcsA protein. The five disulphide bonds are shown in blue and residues highlighted in red are those predicted to be important for ATP action at the $P2X_1$ receptor by mutagenesis studies (Ennion *et al.*, 2000; Roberts & Evans, 2004). These residues are not in close proximity to be able to bind ATP unless the other two subunits that form the trimeric structure of the P2X receptor channel contribute. The two residues in magenta (V48 and I329) have been shown to form a disulphide bond (shown by an arrow) when mutated to cysteine (Jiang *et al.*, 2001) and give an impression of the size of the pore. In this model these residues are ~86Å apart which is too far apart to be able to form a disulphide bond, suggesting that this model is not an accurate representation of the $P2X_1$ receptor as it does not agree with the experimental data (Mager *et al.*, 2006).

This further enhanced the P2X receptor model and identified another residue that should be considered in future refinements of P2X receptor structural models.

A portion of the catalytic domain of the aminoacyl-tRNA synthetase family has been likened to a region of the $rP2X_4$ receptor (residues 170–330) (Freist *et al.*, 1998). Homology models based on the similarity between the P2X receptor family and the synthetase family have provided an insight into the secondary structure of the P2X receptor family but have also fed more mutagenesis studies in an attempt to validate the 3D models. These studies further defined the ATP-binding site but also raised the question of whether a consensus model of ATP-binding across all seven P2X receptor family members exists. The high residue conservation across the P2X receptor family (40-50% identical in amino acid sequence (North, 2002)) might suggest that the ATP-binding site in different receptor subtypes may be similar. If such a consensus model does exist across P2X receptor subunits what does this say about the location of the ATP-binding site?

Combining the information from the 3D model of the $rP2X_4$ receptor based on class II aminoacyl-tRNA synthetases and mutagenesis techniques, the idea of a consensus ATP-binding site was tested using the $hP2X_1$ receptor and the $rP2X_4$ receptor. The corresponding residues to those thought to be involved in agonist binding (Yan *et al.*, 2005) at the $rP2X_4$ receptor were mutated to cysteine in the $hP2X_1$ receptor to see the effect on channel function. It was concluded that the model of binding in the $rP2X_4$ receptor could not be directly transferred to the $hP2X_1$ receptor because residues identified as critical for ATP binding at the $rP2X_4$ receptor did not have a similar role at the $hP2X_1$ receptor (see Chapter 5). The converse was tested in Chapter 6 whereby residues predicted to be important in ATP-binding at the $hP2X_1$ receptor were mutated to cysteine in the $rP2X_4$ receptor. All seven cysteine mutants had a significant effect on ATP potency at the $P2X_4$ receptor. Since then, a similar study has been carried out in the $hP2X_2$ receptor in the Evans lab. Initial results suggest that residues predicted to be important for ATP-binding in the $hP2X_1$ receptor are also important in the $hP2X_2$ receptor (unpublished work). These results suggest that there is a consensus ATP-binding site across the P2X receptor subtypes and that this consensus binding site incorporates both halves of the extracellular segment. Future modelling and mutagenesis studies must consider the impact of the predicted trimeric structure of the P2X receptor channel before a more 'believable' model is achieved.

This then raises the question does the ATP-binding site lie intra-subunit or does it lie inter-subunit as with other ligand-gated ion channels? A consensus binding site across subunits would suggest the binding site is inter-subunit which would allow for the entire extracellular segment to be involved in binding perhaps not by a conformational change but by an inter-subunit binding site. There are several studies that have investigated this question and the evidence suggests that residues from two different subunits interact in agonist binding or channel opening (Wilkinson *et al.*, 2006) and that loops from neighbouring P2X subunits form the ATP-binding site in P2X receptors (Marquez-Klaka *et al.*, 2007).

Techniques that attempt to define the structure of the P2X receptor, besides homology modelling and mutagenesis have been used. The $P2X_2$ receptor has been purified and its structure observed using electron microscopy (Mio *et al.*, 2005). Cross-linking experiments confirmed the trimeric structure of the receptor channel and an image of an inverted three-sided pyramid was published. However, until the long-awaited 3D crystal structure of the P2X receptor is published such images and ideas as to the actual structure of the receptor and location of the ligand-binding site cannot be confirmed. In the meantime the collection of mutagenesis data will continue to build a convincing picture of the ATP-binding site and the structure of the channel and will be vital to validate the 3D crystal structure. The arrival of the crystal structure will open up new areas of research as although the structure will be solved, the mechanisms of agonist and antagonist binding and gating for example could then be examined. Also depending on the state the receptor is crystallised, i.e. whether it is at physiological conditions or not, other areas of research will open up including high throughput screening of drugs.

Bibliography

Abbracchio MP, Boeynaems JM, Barnard EA, Boyer JL, Kennedy C, Miras-Portugal MT, King BF, Gachet C, Jacobson KA, Weisman GA & Burnstock G. (2003). Characterization of the UDP-glucose receptor (re-named here the P2Y14 receptor) adds diversity to the P2Y receptor family. *Trends Pharmacol Sci* **24,** 52-55.

Abbracchio MP & Burnstock G. (1994). Purinoceptors: are there families of P2X and P2Y purinoceptors? *Pharmacol Ther* **64,** 445-475.

Abbracchio MP, Burnstock G, Boeynaems J-M, Barnard EA, Boyer JL, Kennedy C, Knight GE, Fumagalli M, Gachet C, Jacobson KA & Weisman GA. (2006). International Union of Pharmacology LVIII: Update on the P2Y G Protein-Coupled Nucleotide Receptors: From Molecular Mechanisms and Pathophysiology to Therapy 10.1124/pr.58.3.3. *Pharmacol Rev* **58,** 281-341.

Akabas MH, Kaufmann C, Archdeacon P & Karlin A. (1994). Identification of acetylcholine receptor channel-lining residues in the entire M2 segment of the alpha subunit. *Neuron* **13,** 919-927.

Altschul SF & Gish W. (1996). Local alignment statistics. *Methods Enzymol* **266,** 460-480.

Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ. (1990). Basic local alignment search tool. *J Mol Biol* **215,** 403-410.

Altschul SF & Lipman DJ. (1990). Protein database searches for multiple alignments. *Proc Natl Acad Sci U S A* **87,** 5509-5513.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W & Lipman DJ. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25,** 3389-3402.

Arias HR. (2000). Localization of agonist and competitive antagonist binding sites on nicotinic acetylcholine receptors. *Neurochem Int* **36,** 595-645.

Armstrong N & Gouaux E. (2000). Mechanisms for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core. *Neuron* **28,** 165-181.

Armstrong N, Sun Y, Chen GQ & Gouaux E. (1998). Structure of a glutamate-receptor ligand-binding core in complex with kainate. *Nature* **395,** 913-917.

Arnez JG, Augustine JG, Moras D & Francklyn CS. (1997). The first step of aminoacylation at the atomic level in histidyl-tRNA synthetase. *Proc Natl Acad Sci U S A* **94,** 7144-7149.

Arnez JG, Dock-Bregeon AC & Moras D. (1999). Glycyl-tRNA synthetase uses a negatively charged pit for specific recognition and activation of glycine. *J Mol Biol* **286,** 1449-1459.

Aurora R, Srinivasan R & Rose GD. (1994). Rules for alpha-helix termination by glycine. *Science* **264,** 1126-1130.

Bairoch A & Apweiler R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res* **28,** 45-48.

Banfi C, Ferrario S, De Vincenti O, Ceruti S, Fumagalli M, Mazzola A, N DA, Volonte C, Fratto P, Vitali E, Burnstock G, Beltrami E, Parolari A, Polvani G, Biglioli P, Tremoli E & Abbracchio MP. (2005). P2 receptors in human heart: upregulation of P2X6 in patients undergoing heart transplantation, interaction with TNFalpha and potential role in myocardial cell death. *J Mol Cell Cardiol* **39,** 929-939.

Barrera NP, Ormond SJ, Henderson RM, Murrell-Lagnado RD & Edwardson JM. (2005). Atomic force microscopy imaging demonstrates that P2X2 receptors are trimers but that P2X6 receptor subunits do not oligomerize. *J Biol Chem* **280,** 10759-10765.

Baxevanis AD & Ouellette BF. (2001). *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*.

Bean BP. (1990). ATP-activated channels in rat and bullfrog sensory neurons: concentration dependence and kinetics. *J Neurosci* **10,** 1-10.

Belrhali H, Yaremchuk A, Tukalo M, Larsen K, Berthet-Colominas C, Leberman R, Beijer B, Sproat B, Als-Nielsen J, Grubel G & et al. (1994). Crystal structures at 2.5 angstrom resolution of seryl-tRNA synthetase complexed with two analogs of seryl adenylate. *Science* **263,** 1432-1436.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN & Bourne PE. (2000). The Protein Data Bank. *Nucleic Acids Res* **28,** 235-242.

Berne RM. (1963). Cardiac nucleotides in hypoxia: possible role in regulation of coronary blood flow. *Am J Physiol* **204,** 317-322.

Bo X, Zhang Y, Nassar M, Burnstock G & Schoepfer R. (1995). A P2X purinoceptor cDNA conferring a novel pharmacological profile. *FEBS Lett* **375,** 129-133.

Bodin P & Burnstock G. (2001). Purinergic signalling: ATP release. *Neurochem Res* **26,** 959-969.

Boue-Grabot E, Archambault V & Seguela P. (2000). A protein kinase C site highly conserved in P2X subunits controls the desensitization kinetics of P2X(2) ATP-gated channels. *J Biol Chem* **275,** 10190-10195.

Brake AJ & Julius D. (1996). Signalling by extracellular nucleotides. *Annu Rev Cell Dev Biol* **12,** 519-541.

Brake AJ, Wagenbach MJ & Julius D. (1994). New structural motif for ligand-gated ion channels defined by an ionotropic ATP receptor. *Nature* **371,** 519-523.

Brejc K, van Dijk WJ, Klaassen RV, Schuurmans M, van Der Oost J, Smit AB & Sixma TK. (2001). Crystal structure of an ACh-binding protein reveals the ligand-binding domain of nicotinic receptors. *Nature* **411,** 269-276.

Bruns RF, Lu GH & Pugsley TA. (1986). Characterization of the A2 adenosine receptor labeled by [3H]NECA in rat striatal membranes. *Mol Pharmacol* **29,** 331-346.

Bucher P, Karplus K, Moeri N & Hofmann K. (1996). A flexible motif search technique based on generalized profiles. *Comput Chem* **20,** 3-23.

Buell G, Collo G & Rassendren F. (1996). P2X receptors: an emerging channel family. *Eur J Neurosci* **8,** 2221-2228.

Burnstock G. (1972). Purinergic nerves. *Pharmacol Rev* **24,** 509-581.

Burnstock G. (1976). Do some nerve cells release more than one transmitter? *Neuroscience* **1,** 239-248.

Burnstock G. (1990). The fifth Heymans memorial lecture-Ghent, February 17, 1990. Co-transmission. *Arch Int Pharmacodyn Ther* **304,** 7-33.

Burnstock G. (1996). A unifying purinergic hypothesis for the initiation of pain. *Lancet* **347,** 1604-1605.

Burnstock G. (1997). The past, present and future of purine nucleotides as signalling molecules. *Neuropharmacology* **36,** 1127-1139.

Burnstock G. (2001). Purine-mediated signalling in pain and visceral perception. *Trends Pharmacol Sci* **22,** 182-188.

Burnstock G. (2002). Potential therapeutic targets in the rapidly expanding field of purinergic signalling. *Clinical Medicine* **2,** 45-53.

Burnstock G & Kennedy C. (1985). Is there a basis for distinguishing two types of P2-purinoceptor? *Gen Pharmacol* **16,** 433-440.

Bystroff C & Shao Y. (2002). Fully automated ab initio protein structure prediction using I-SITES, HMMSTR and ROSETTA. *Bioinformatics* **18 Suppl 1,** S54-61.

Calvert JA & Evans RJ. (2004). Heterogeneity of P2X receptors in sympathetic neurons: contribution of neuronal P2X1 receptors revealed using knockout mice. *Mol Pharmacol* **65,** 139-148.

Canutescu AA, Shelenkov AA & Dunbrack RL, Jr. (2003). A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci* **12,** 2001-2014.

Ceroni A, Passerini A, Vullo A & Frasconi P. (2006). DISULFIND: a disulfide bonding state and cysteine connectivity prediction server. *Nucleic Acids Res* **34,** W177-181.

Ceroni AF, P. Passerini, A. Vullo, A. (2003). Predicting the disulfide bonding state of cysteines with combinations of kernel machines. *Journal of VLSI signal Processing* **35,** 287-295.

Chambers JK, Macdonald LE, Sarau HM, Ames RS, Freeman K, Foley JJ, Zhu Y, McLaughlin MM, Murdock P, McMillan L, Trill J, Swift A, Aiyar N, Taylor P, Vawter L, Naheed S, Szekeres P, Hervieu G, Scott C, Watson JM, Murphy AJ, Duzic

E, Klein C, Bergsma DJ, Wilson S & Livi GP. (2000). A G protein-coupled receptor for UDP-glucose. *J Biol Chem* **275,** 10767-10771.

Chaudry IH. (1982). Does ATP cross the cell plasma membrane. *Yale J Biol Med* **55,** 1-10.

Chen CC, Akopian AN, Sivilotti L, Colquhoun D, Burnstock G & Wood JN. (1995). A P2X purinoceptor expressed by a subset of sensory neurons. *Nature* **377,** 428-431.

Chen GQ, Sun Y, Jin R & Gouaux E. (1998). Probing the ligand binding domain of the GluR2 receptor by proteolysis and deletion mutagenesis defines domain boundaries and yields a crystallizable construct. *Protein Sci* **7,** 2623-2630.

Chou PY & Fasman GD. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. *Adv Enzymol Relat Areas Mol Biol* **47,** 45-148.

Clamp M, Cuff J, Searle SM & Barton GJ. (2004). The Jalview Java alignment editor. *Bioinformatics* **20,** 426-427.

Clarke CE, Benham CD, Bridges A, George AR & Meadows HJ. (2000). Mutation of histidine 286 of the human P2X4 purinoceptor removes extracellular pH sensitivity. *J Physiol* **523 Pt 3,** 697-703.

Clyne JD, Wang LF & Hume RI. (2002). Mutational analysis of the conserved cysteines of the rat P2X2 purinoceptor. *J Neurosci* **22,** 3873-3880.

Coddou C, Morales B, Gonzalez J, Grauso M, Gordillo F, Bull P, Rassendren F & Huidobro-Toro JP. (2003). Histidine 140 plays a key role in the inhibitory modulation of the P2X4 nucleotide receptor by copper but not zinc. *J Biol Chem* **278,** 36777-36785.

Collo G, North RA, Kawashima E, Merlo-Pich E, Neidhart S, Surprenant A & Buell G. (1996). Cloning OF P2X5 and P2X6 receptors and the distribution and properties of an extended family of ATP-gated ion channels. *J Neurosci* **16,** 2495-2507.

Colovos C & Yeates TO. (1993). Verification of protein structures: Patterns of nonbonded atomic interactions. *Protein Sci* **2,** 1511-1519.

Communi D, Gonzalez NS, Detheux M, Brezillon S, Lannoy V, Parmentier M & Boeynaems JM. (2001). Identification of a novel human ADP receptor coupled to G(i). *J Biol Chem* **276,** 41479-41485.

Corringer PJ, Le Novere N & Changeux JP. (2000). Nicotinic receptors at the amino acid level. *Annu Rev Pharmacol Toxicol* **40,** 431-458.

Dembo A, Karlin S & Zeitouni O. (1984). Limit distribution of maximal non-aligned two-sequence segmental score. *Ann Prob* **22,** 2022-2039.

Digby H, Roberts JA, Sutcliffe MJ & Evans RJ. (2005). Contribution of conserved glycine residues to ATP action at human P2X1 receptors: mutagenesis indicates that the unique property of glycine at position 250 is important for channel function. *J Neurochem.*

Ding S, Ingleby L, Ahern CA & Horn R. (2005). Investigating the putative glycine hinge in Shaker potassium channel. *J Gen Physiol* **126,** 213-226.

Ding S & Sachs F. (1999). Single channel properties of P2X2 purinoceptors. *J Gen Physiol* **113,** 695-720.

Edwards FA & Gibb AJ. (1993). ATP--a fast neurotransmitter. *FEBS Lett* **325,** 86-89.

Egan TM, Cox JA & Voigt MM. (2004). Molecular structure of P2X receptors. *Curr Top Med Chem* **4,** 821-829.

Egan TM, Haines WR & Voigt MM. (1998). A domain contributing to the ion channel of ATP-gated P2X2 receptors identified by the substituted cysteine accessibility method. *J Neurosci* **18,** 2350-2359.

Ennion S, Hagan S & Evans RJ. (2000). The Role of Positively Charged Amino Acids in ATP Recognition by Human P2X1 Receptors. *J Biol Chem* **275,** 29361-29367.

Ennion SJ & Evans RJ. (2002a). Conserved Cysteine Residues in the Extracellular Loop of the Human P2X1 Receptor Form Disulfide Bonds and Are Involved in Receptor Trafficking to the Cell Surface. *Mol Pharmacol* **61,** 303-311.

Ennion SJ & Evans RJ. (2002b). P2X1 Receptor Subunit Contribution to Gating Revealed by a Dominant Negative PKC Mutant. *Biochemical and Biophysical Research Communications* **291,** 611-616.

Ennion SJ, Ritson J & Evans RJ. (2001). Conserved Negatively Charged Residues Are Not Required for ATP Action at P2X1 Receptors. *Biochemical and Biophysical Research Communications* **289,** 700-704.

Eriani G, Cavarelli J, Martin F, Ador L, Rees B, Thierry JC, Gangloff J & Moras D. (1995). The class II aminoacyl-tRNA synthetases and their active site: evolutionary conservation of an ATP binding site. *J Mol Evol* **40,** 499-508.

Eriani G, Delarue M, Poch O, Gangloff J & Moras D. (1990). Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* **347,** 203-206.

Fabre JE, Nguyen M, Latour A, Keifer JA, Audoly LP, Coffman TM & Koller BH. (1999). Decreased platelet aggregation, increased bleeding time and resistance to thromboembolism in P2Y1-deficient mice. *Nat Med* **5,** 1199-1202.

Fischer D, Barret C, Bryson K, Elofsson A, Godzik A, Jones D, Karplus KJ, Kelley LA, MacCallum RM, Pawowski K, Rost B, Rychlewski L & Sternberg M. (1999). CAFASP-1: critical assessment of fully automated structure prediction methods. *Proteins* **Suppl 3,** 209-217.

Fiser A, Do RK & Sali A. (2000). Modeling of loops in protein structures. *Protein Sci* **9,** 1753-1773.

Fiser A & Sali A. (2003). Modeller: generation and refinement of homology-based protein structure models. *Methods Enzymol* **374,** 461-491.

Fitch WM. (1983). Random sequences. *J Mol Biol* **163,** 171-176.

Fountain SJ & North RA. (2006). A C-terminal lysine that controls human P2X4 receptor desensitization. *J Biol Chem* **281,** 15044-15049.

Frasconi PP, A. Vullo, A. (2002). A two-stage SVM architecture for predicting the disulfide bonding state of cysteines. 25-34.

Fredholm BB. (1995). Purinoceptors in the nervous system. *Pharmacol Toxicol* **76,** 228-239.

Freist W, Verhey JF, Stuhmer W & Gauss DH. (1998). ATP binding site of P2X channel proteins: structural similarities with class II aminoacyl-tRNA synthetases. *FEBS Lett* **434,** 61-65.

Galtier N, Gouy, M. and Gautier, C. (1996). SeaView and Phylo_win, two graphic tools for sequence alignment and molecular phylogeny. *Comput Applic Biosci,* **12,** 543-548.

Garcia-Guzman M, Soto F, Gomez-Hernandez JM, Lund PE & Stuhmer W. (1997a). Characterization of recombinant human P2X4 receptor reveals pharmacological differences to the rat homologue. *Mol Pharmacol* **51,** 109-118.

Garcia-Guzman M, Stuhmer W & Soto F. (1997b). Molecular characterization and pharmacological properties of the human P2X3 purinoceptor. *Brain Res Mol Brain Res* **47,** 59-66.

Gardner HP, Wertheim GB, Ha SI, Copeland NG, Gilbert DJ, Jenkins NA, Marquis ST & Chodosh LA. (2000). Cloning and characterization of Hunk, a novel mammalian SNF1-related protein kinase. *Genomics* **63,** 46-59.

Garnier J, Osguthorpe DJ & Robson B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J Mol Biol* **120,** 97-120.

Gever JR, Cockayne DA, Dillon MP, Burnstock G & Ford AP. (2006). Pharmacology of P2X channels. *Pflugers Arch* **452,** 513-537.

Glynn IM. (1968). Membrane adenosine triphosphatase and cation transport. *Br Med Bull* **24,** 165-169.

Golovin A, Dimitropoulos D, Oldfield T, Rachedi A & Henrick K. (2005). MSDsite: a database search and retrieval system for the analysis and viewing of bound ligands and active sites. *Proteins* **58,** 190-199.

Gouaux E. (2004). Structure and function of AMPA receptors. *J Physiol* **554,** 249-253.

Gouet P, Robert X & Courcelle E. (2003). ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res* **31,** 3320-3323.

Gribskov M, McLachlan AD & Eisenberg D. (1987). Profile analysis: detection of distantly related proteins. *Proc Natl Acad Sci U S A* **84,** 4355-4358.

Gurin VN, Gurin AV, Melenchuk EV & Spyer KM. (2003). The effects of activation and blockade of central P2X receptors on body temperature. *Neurosci Behav Physiol* **33,** 845-851.

Haines WR, Migita K, Cox JA, Egan TM & Voigt MM. (2001). The first transmembrane domain of the P2X receptor subunit participates in the agonist-induced gating of the channel. *J Biol Chem* **276,** 32793-32798.

Hansen MA, Barden JA, Balcar VJ, Keay KA & Bennett MR. (1997). Structural motif and characteristics of the extracellular domain of P2X receptors. *Biochem Biophys Res Commun* **236,** 670-675.

Hechler B, Cattaneo M & Gachet C. (2005). The P2 receptors in platelet function. *Semin Thromb Hemost* **31,** 150-161.

Hechler B, Lenain N, Marchese P, Vial C, Heim V, Freund M, Cazenave JP, Cattaneo M, Ruggeri ZM, Evans R & Gachet C. (2003). A role of the fast ATP-gated P2X1 cation channel in thrombosis of small arteries in vivo. *J Exp Med* **198,** 661-667.

Hechler B, Vigne P, Leon C, Breittmayer JP, Gachet C & Frelin C. (1998). ATP derivatives are antagonists of the P2Y1 receptor: similarities to the platelet ADP receptor. *Mol Pharmacol* **53,** 727-733.

Henikoff S & Henikoff JG. (1991). Automated assembly of protein blocks for database searching. *Nucleic Acids Res* **19,** 6565-6572.

Henikoff S & Henikoff JG. (1992). Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci U S A* **89,** 10915-10919.

Ho BK & Brasseur R. (2005). The Ramachandran plots of glycine and pre-proline. *BMC Struct Biol* **5,** 14.

Hofmann K & Stoffel W. (1993). TMbase-A database of membrane spanning proteins segments. *Biol Chem* **374,** 166.

Hollopeter G, Jantzen HM, Vincent D, Li G, England L, Ramakrishnan V, Yang RB, Nurden P, Nurden A, Julius D & Conley PB. (2001). Identification of the platelet ADP receptor targeted by antithrombotic drugs. *Nature* **409,** 202-207.

Holton P. (1959). The liberation of adenosine triphosphate on antidromic stimulation of sensory nerves. *J Physiol* **145,** 494-504.

Hooft RW, Vriend G, Sander C & Abola EE. (1996). Errors in protein structures. *Nature* **381,** 272.

http://www.ncbi.nlm.nih.gov/BLAST/.

http://www.sbg.bio.ic.ac.uk/~3dpssm/.

http:/www.ebi.ac.uk/clustalw. ClustalW WWW Service at the European Bioinformatics Institute.

Inscho EW, Cook AK, Imig JD, Vial C & Evans RJ. (2003). Physiological role for P2X1 receptors in renal microvascular autoregulatory behavior. *J Clin Invest* **112,** 1895-1905.

Inscho EW, Cook AK, Imig JD, Vial C & Evans RJ. (2004). Renal autoregulation in P2X1 knockout mice. *Acta Physiol Scand* **181,** 445-453.

J.E. W, M S, M.J R & N.J G. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP requiring enzymes and a common nucleotide binding fold. *Embo J* **1,** 945-951.

Jahr CE & Jessell TM. (1983). ATP excites a subpopulation of rat dorsal horn neurones. *Nature* **304,** 730-733.

Jaroszewski L, Rychlewski L & Godzik A. (2000). Improving the quality of twilight-zone alignments. *Protein Sci* **9,** 1487-1496.

Jiang L-H, Rassendren F, Surprenant A & North RA. (2000). Identification of Amino Acid Residues Contributing to the ATP-binding Site of a Purinergic P2X Receptor. *J Biol Chem* **275,** 34190-34196.

Jiang LH, Kim M, Spelta V, Bo X, Surprenant A & North RA. (2003). Subunit arrangement in P2X receptors. *J Neurosci* **23,** 8903-8910.

Jiang LH, Rassendren F, Spelta V, Surprenant A & North RA. (2001). Amino acid residues involved in gating identified in the first membrane-spanning domain of the rat P2X(2) receptor. *J Biol Chem* **276,** 14902-14908.

Jones DT, Taylor WR & Thornton JM. (1992). A new approach to protein fold recognition. *Nature* **358,** 86-89.

Kaushansky K. (1999). The enigmatic megakaryocyte gradually reveals its secrets. *Bioessays* **21,** 353-360.

Kelley LA, MacCallum RM & Sternberg M. (1999). Recognition of remote protein homolgies using three-dimensional information to generate a position specific scoring matrix in the program 3D-PSSM. ed. Istrail S, Pevzner P & Waterman MS, pp. 218-225. The Association for Computing Machinery.

Khakh BS, Bao XR, Labarca C & Lester HA. (1999). Neuronal P2X transmitter-gated cation channels change their ion selectivity in seconds. *Nat Neurosci* **2,** 322-330.

Khakh BS & Egan TM. (2005). Contribution of transmembrane regions to ATP-gated P2X2 channel permeability dynamics. *J Biol Chem* **280,** 6118-6129.

Khakh BS & North RA. (2006). P2X receptors as cell-surface ATP sensors in health and disease. *Nature* **442,** 527-532.

Kim M, Spelta V, Sim J, North RA & Surprenant A. (2001). Differential assembly of rat purinergic P2X{sub}7 receptor in immune cells of the brain and periphery 10.1074/jbc.M102253200. *J Biol Chem*, M102253200.

Kneller DG, Cohen FE & Langridge R. (1990). Improvements in protein secondary structure prediction by an enhanced neural network. *J Mol Biol* **214,** 171-182.

Koshimizu T, Koshimizu M & Stojilkovic SS. (1999). Contributions of the C-terminal domain to the control of P2X receptor desensitization. *J Biol Chem* **274,** 37651-37657.

Krishtal OA, Marchenko SM & Pidoplichko VI. (1983). Receptor for ATP in the membrane of mammalian sensory neurones. *Neurosci Lett* **35,** 41-45.

Laskowski RA, MacArthur MW, Moss DS & Thornton JM. (1993). PROCHECK:a program to check the stereochemical quality of protein structures. *JAppl Cryst* **26,** 283-291.

Le KT, Babinski K & Seguela P. (1998). Central P2X4 and P2X6 channel subunits coassemble into a novel heteromeric ATP receptor. *J Neurosci* **18,** 7152-7159.

Le KT, Boue-Grabot E, Archambault V & Seguela P. (1999). Functional and biochemical evidence for heteromeric ATP-gated channels composed of P2X1 and P2X5 subunits. *J Biol Chem* **274,** 15415-15419.

Le KT, Paquet M, Nouel D, Babinski K & Seguela P. (1997). Primary structure and expression of a naturally truncated human P2X ATP receptor subunit from brain and immune system. *FEBS Lett* **418,** 195-199.

Leon C, Hechler B, Freund M, Eckly A, Vial C, Ohlmann P, Dierich A, LeMeur M, Cazenave JP & Gachet C. (1999). Defective platelet aggregation and increased resistance to thrombosis in purinergic P2Y(1) receptor-null mice. *J Clin Invest* **104,** 1731-1737.

Leon C, Hechler B, Vial C, Leray C, Cazenave JP & Gachet C. (1997). The P2Y1 receptor is an ADP receptor antagonized by ATP and expressed in platelets and megakaryoblastic cells. *FEBS Lett* **403,** 26-30.

Lesage F, Lauritzen I, Duprat F, Reyes R, Fink M, Heurteaux C & Lazdunski M. (1997). The structure, function and distribution of the mouse TWIK-1 K+ channel. *FEBS Lett* **402,** 28-32.

Lewis C, Neidhart S, Holy C, North RA, Buell G & Surprenant A. (1995). Coexpression of P2X2 and P2X3 receptor subunits can account for ATP-gated currents in sensory neurons. *Nature* **377,** 432-435.

Li W, Pio F, Pawlowski K & Godzik A. (2000). Saturated BLAST: an automated multiple intermediate sequence search used to detect distant homology. *Bioinformatics* **16,** 1105-1110.

Li Z, Migita K, Samways DS, Voigt MM & Egan TM. (2004). Gain and loss of channel function by alanine substitutions in the transmembrane segments of the rat ATP-gated P2X2 receptor. *J Neurosci* **24,** 7378-7386.

Liu GJ, Brockhausen J & Bennett MR. (2003). P2X1 receptor currents after disruption of the PKC site and its surroundings by dominant negative mutations in HEK293 cells. *Autonomic Neuroscience* **108,** 12-16.

Londos C, Cooper DM & Wolff J. (1980). Subclasses of external adenosine receptors. *Proc Natl Acad Sci U S A* **77,** 2551-2554.

Lustig KD, Shiau AK, Brake AJ & Julius D. (1993). Expression cloning of an ATP receptor from mouse neuroblastoma cells. *Proc Natl Acad Sci U S A* **90,** 5113-5117.

Lynch KJ, Touma E, Niforatos W, Kage KL, Burgard EC, van Biesen T, Kowaluk EA & Jarvis MF. (1999). Molecular and functional characterization of human P2X(2) receptors. *Mol Pharmacol* **56,** 1171-1181.

MacArthur MW & Thornton JM. (1991). Influence of proline residues on protein conformation. *J Mol Biol* **218,** 397-412.

MacCallum RM, Kelley LA & Sternberg MJ. (2000). SAWTED: structure assignment with text description--enhanced detection of remote homologues with automated SWISS-PROT annotation comparisons. *Bioinformatics* **16,** 125-129.

Mager PP, Weber A, Sanchez L, Wirkner K & Illes P. (2006). The monomers of the P2X1 receptor model and KcsA protein share a similar structural fold. *Protein Pept Lett* **13,** 77-81.

Magidovich E & Yifrach O. (2004). Conserved gating hinge in ligand- and voltage-dependent K+ channels. *Biochemistry* **43,** 13242-13247.

Marquez-Klaka B, Rettinger J, Bhargava Y, Eisele T & Nicke A. (2007). Identification of an intersubunit cross-link between substituted cysteine residues located in the putative ATP binding site of the P2X1 receptor. *J Neurosci* **27,** 1456-1466.

McGuffin LJ, Bryson K & Jones DT. (2000). The PSIPRED protein structure prediction server. *Bioinformatics* **16,** 404-405.

Meyerhof W, Muller-Brechlin R & Richter D. (1991). Molecular cloning of a novel putative G-protein coupled receptor expressed during rat spermiogenesis. *FEBS Lett* **284,** 155-160.

Migita K, Haines WR, Voigt MM & Egan TM. (2001). Polar residues of the second transmembrane domain influence cation permeability of the ATP-gated P2X(2) receptor. *J Biol Chem* **276,** 30934-30941.

Mio K, Kubo Y, Ogura T, Yamamoto T & Sato C. (2005). Visualization of the trimeric P2X2 receptor with a crown-capped extracellular domain. *Biochem Biophys Res Commun* **337,** 998-1005.

Mulryan K, Gitterman DP, Lewis CJ, Vial C, Leckie BJ, Cobb AL, Brown JE, Conley EC, Buell G, Pritchard CA & Evans RJ. (2000). Reduced vas deferens contraction and male infertility in mice lacking P2X1 receptors. *Nature* **403,** 86-89.

Nakazawa K & Ohno Y. (1999). Neighboring glycine residues are essential for P2X2 receptor/channel function. *Eur J Pharmacol* **370,** R5-6.

Nakazawa K, Ohno Y & Inoue K. (1998). An aspartic acid residue near the second transmembrane segment of ATP receptor/channel regulates agonist sensitivity. *Biochem Biophys Res Commun* **244,** 599-603.

Nakazawa K, Ojima H, Ishii-Nozawa R, Takeuchi K & Ohno Y. (2004). Amino acid substitutions from an indispensable disulfide bond affect P2X2 receptor activation. *Eur J Pharmacol* **483,** 29-35.

Newbolt A, Stoop R, Virginio C, Surprenant A, North RA, Buell G & Rassendren F. (1998). Membrane Topology of an ATP-gated Ion Channel (P2X Receptor). *J Biol Chem* **273,** 15177-15182.

Nicke A, Baumert HG, Rettinger J, Eichele A, Lambrecht G, Mutschler E & Schmalzing G. (1998). P2X1 and P2X3 receptors form stable trimers: a novel structural motif of ligand-gated ion channels. *Embo J* **17,** 3016-3028.

Nicke A, Rettinger J & Schmalzing G. (2003). Monomeric and dimeric byproducts are the principal functional elements of higher order P2X1 concatamers. *Mol Pharmacol* **63,** 243-252.

North RA. (1996). P2X receptors: a third major class of ligand-gated ion channels. *Ciba Found Symp* **198,** 91-105; discussion 105-109.

North RA. (2002). Molecular Physiology of P2X Receptors. *Physiol Rev* **82,** 1013-1067.

O'Connor SE, Dainty IA & Leff P. (1991). Further subclassification of ATP receptors based on agonist studies. *Trends Pharmacol Sci* **12,** 137-141.

Okkeri J, Laakkonen L & Haltia T. (2004). The nucleotide-binding domain of the Zn2+-transporting P-type ATPase from Escherichia coli carries a glycine motif that may be involved in binding of ATP. *Biochem J* **377,** 95-105.

Oury C, Kuijpers MJ, Toth-Zsamboki E, Bonnefoy A, Danloy S, Vreys I, Feijge MA, De Vos R, Vermylen J, Heemskerk JW & Hoylaerts MF. (2003). Overexpression of the platelet P2X1 ion channel in transgenic mice generates a novel prothrombotic phenotype. *Blood* **101,** 3969-3976.

Paas Y. (1998). The macro- and microarchitectures of the ligand-binding domain of glutamate receptors. *Trends Neurosci* **21,** 117-125.

Parry-Smith DJ, Payne AW, Michie AD & Attwood TK. (1998). CINEMA--a novel colour INteractive editor for multiple alignments. *Gene* **221,** GC57-63.

Paukert M, Osteroth R, Geisler HS, Brandle U, Glowatzki E, Ruppersberg JP & Grunder S. (2001). Inflammatory mediators potentiate ATP-gated channels through the P2X(3) subunit. *J Biol Chem* **276,** 21077-21082.

Peitsch MC. (1996). ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling. *Biochem Soc Trans* **24,** 274-279.

Pintor J, Diaz-Rey MA & Miras-Portugal MT. (1993). Ap4A and ADP-beta-S binding to P2 purinoceptors present on rat brain synaptic terminals. *Br J Pharmacol* **108,** 1094-1099.

Pintor J & Miras-Portugal MT. (1995). A novel receptor for diadenosine polyphosphates coupled to calcium increase in rat midbrain synaptosomes. *Br J Pharmacol* **115,** 895-902.

Poirot O, Suhre K, Abergel C, O'Toole E & Notredame C. (2004). 3DCoffee@igs: a web server for combining sequences and structures into a multiple sequence alignment. *Nucleic Acids Res* **32,** W37-40.

Pontius J, Richelle J & Wodak SJ. (1996). Deviations from standard atomic volumes as a quality measure for protein crystal structures. *J Mol Biol* **264,** 121-136.

Puvanendrampillai D & Mitchell JBO. (2003). Protein Ligand Database (PLD): additional understanding of the nature and specificity of protein-ligand complexes. *Bioinformatics* **19,** 1856-1857.

Radford KM, Virginio C, Surprenant A, North RA & Kawashima E. (1997). Baculovirus expression provides direct evidence for heteromeric assembly of P2X2 and P2X3 receptors. *J Neurosci* **17,** 6529-6533.

Ralevic V & Burnstock G. (1998). Receptors for purines and pyrimidines. *Pharmacol Rev* **50,** 413-492.

Ramachandran GN, Ramakrishnan C & Sasisekharan V. (1963). Stereochemistry of polypeptide chain configurations. *J Mol Biol* **7,** 95-99.

Rassendren F, Buell G, Newbolt A, North RA & Surprenant A. (1997a). Identification of amino acid residues contributing to the pore of a P2X receptor. *EMBO J* **16,** 3446-3454.

Rassendren F, Buell GN, Virginio C, Collo G, North RA & Surprenant A. (1997b). The permeabilizing ATP receptor, P2X7. Cloning and expression of a human cDNA. *J Biol Chem* **272,** 5482-5486.

Rettinger J, Aschrafi A & Schmalzing G. (2000). Roles of individual N-glycans for ATP potency and expression of the rat P2X1 receptor. *J Biol Chem* **275,** 33542-33547.

Richardson JS. (1981). The anatomy and taxonomy of protein structure. *Adv Protein Chem* **34,** 167-339.

Roberts JA & Evans RJ. (2004). ATP binding at human P2X1 receptors. Contribution of aromatic and basic amino acids revealed using mutagenesis and partial agonists. *J Biol Chem* **279,** 9043-9055.

Roberts JA & Evans RJ. (2005). Mutagenesis studies of conserved proline residues of human P2X receptors for ATP indicate that proline 272 contributes to channel function. *J Neurochem* **92,** 1256-1264.

Roberts JA & Evans RJ. (2006). Contribution of conserved polar glutamine, asparagine and threonine residues and glycosylation to agonist action at human P2X1 receptors for ATP
doi:10.1111/j.1471-4159.2005.03593.x. *Journal of Neurochemistry* **96,** 843-852.

Roberts JA & Evans RJ. (2007). Cysteine substitution mutants give structural insight and identify ATP binding and activation sites at P2X receptors. *J Neurosci* **27,** 4072-4082.

Roberts JA, Vial C, Digby HR, Agboh KC, Wen H, Atterbury-Thomas A & Evans RJ. (2006). Molecular properties of P2X receptors. *Pflugers Arch.*

Robertson SJ, Ennion SJ, Evans RJ & Edwards FA. (2001). Synaptic P2X receptors. *Curr Opin Neurobiol* **11,** 378-386.

Rost B. (1995). TOPITS: threading one-dimensional predictions into three-dimensional structures. *Proc Int Conf Intell Syst Mol Biol* **3,** 314-321.

Rost B. (1996). PHD: predicting one-dimensional protein structure by profile-based neural networks. *Methods Enzymol* **266,** 525-539.

Rost B & Sander C. (1993). Prediction of protein secondary structure at better than 70% accuracy. *J Mol Biol* **232,** 584-599.

Rost B, Sander C & Schneider R. (1994). PHD--an automatic mail server for protein secondary structure prediction. *Comput Appl Biosci* **10,** 53-60.

Royle SJ, Bobanovic LK & Murrell-Lagnado RD. (2002). Identification of a non-canonical tyrosine-based endocytic motif in an ionotropic receptor. *J Biol Chem* **277,** 35378-35385.

Royle SJ, Qureshi OS, Bobanovic LK, Evans PR, Owen DJ & Murrell-Lagnado RD. (2005). Non-canonical YXXGPhi endocytic motifs: recognition by AP2 and preferential utilization in P2X4 receptors. *J Cell Sci* **118,** 3073-3080.

Russell RB & Barton GJ. (1992). Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels. *Proteins* **14,** 309-323.

Salamov AA & Solovyev VV. (1997). Protein secondary structure prediction using local alignments. *J Mol Biol* **268,** 31-36.

Sali A & Blundell TL. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234,** 779-815.

Sandeep Kumar MB. (1998). Dissecting &agr;-helices: Position-specific analysis of &agr;-helices in globular proteins. *Proteins: Structure, Function, and Genetics* **31,** 460-476.

Sauder JM, Arthur JW & Dunbrack RL, Jr. (2000). Large-scale comparison of protein sequence alignment algorithms with structure alignments. *Proteins* **40,** 6-22.

Schellman C. (1980). The aL conformation at the ends of helices. In: "Protein Folding." Jaenicke, R. (ed.). NewYork: Elsevier/ North-Holland. 53-61.

Schmidt MC & McCartney RR. (2000). beta-subunits of Snf1 kinase are required for kinase function and substrate definition. *Embo J* **19,** 4936-4943.

Schneider TD, Stormo GD, Gold L & Ehrenfeucht A. (1986). Information content of binding sites on nucleotide sequences. *J Mol Biol* **188,** 415-431.

Schwede T, Kopp J, Guex N & Peitsch MC. (2003). SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* **31,** 3381-3385.

Schwiebert EM & Zsembery A. (2003). Extracellular ATP as a signalling molecule for epithelial cells. *Biochim Biophys Acta* **1615,** 7-32.

Seifert R & Schultz G. (1989). Involvement of pyrimidinoceptors in the regulation of cell functions by uridine and by uracil nucleotides. *Trends Pharmacol Sci* **10,** 365-369.

Shin J-M & Cho D-H. (2005). PDB-Ligand: a ligand database based on PDB for the automated and customized classification of ligand-binding structures. *Nucl Acids Res* **33,** D238-241.

Shinozuka K, Bjur RA & Westfall DP. (1988). Characterization of prejunctional purinoceptors on adrenergic nerves of the rat caudal artery. *Naunyn Schmiedebergs Arch Pharmacol* **338,** 221-227.

Silberberg SD, Chang T-H & Swartz KJ. (2005). Secondary Structure and Gating Rearrangements of Transmembrane Segments in Rat P2X4 Receptor Channels 10.1085/jgp.200409221. *J Gen Physiol* **125,** 347-359.

Smit AB, Syed NI, Schaap D, van Minnen J, Klumperman J, Kits KS, Lodder H, van der Schors RC, van Elk R, Sorgedrager B, Brejc K, Sixma TK & Geraerts WP. (2001). A glia-derived acetylcholine-binding protein that modulates synaptic transmission. *Nature* **411,** 261-268.

Solini A, Chiozzi P, Morelli A, Passaro A, Fellin R & Di Virgilio F. (2003). Defective P2Y purinergic receptor function: A possible novel mechanism for impaired glucose transport. *J Cell Physiol* **197,** 435-444.

Soto F, Garcia-Guzman M & Stuhmer W. (1997). Cloned ligand-gated channels activated by extracellular ATP (P2X receptors). *J Membr Biol* **160,** 91-100.

Spelta V, Jiang LH, Bailey RJ, Surprenant A & North RA. (2003). Interaction between cysteines introduced into each transmembrane domain of the rat P2X2 receptor. *Br J Pharmacol* **138,** 131-136.

Srinivasan N, Anuradha VS, Ramakrishnan C, Sowdhamini R & Balaram P. (1994). Conformational characteristics of asparaginyl residues in proteins. *Int J Pept Protein Res* **44,** 112-122.

Staden R. (1988). Methods to define and locate patterns of motifs in sequences. *Comput Appl Biosci* **4,** 53-60.

Stauffer DA & Karlin A. (1994). Electrostatic potential of the acetylcholine binding sites in the nicotinic receptor probed by reactions of binding-site cysteines with charged methanethiosulfonates. *Biochemistry* **33,** 6840-6849.

Stoop R, Thomas S, Rassendren F, Kawashima E, Buell G, Surprenant A & North RA. (1999). Contribution of individual subunits to the multimeric P2X(2) receptor: estimates based on methanethiosulfonate block at T336C. *Mol Pharmacol* **56,** 973-981.

Sugiura I, Nureki O, Ugaji-Yoshikawa Y, Kuwabara S, Shimada A, Tateno M, Lorber B, Giege R, Moras D, Yokoyama S & Konno M. (2000). The 2.0 A crystal structure of Thermus thermophilus methionyl-tRNA synthetase reveals two RNA-binding modules. *Structure Fold Des* **8,** 197-208.

Surprenant A, Buell G & North RA. (1995). P2X receptors bring new structure to ligand-gated ion channels. *Trends in Neurosciences* **18,** 224-229.

Surprenant A, Rassendren F, Kawashima E, North RA & Buell G. (1996). The cytolytic P2Z receptor for extracellular ATP identified as a P2X receptor (P2X7). *Science* **272,** 735-738.

Surprenant A, Schneider DA, Wilson HL, Galligan JJ & North RA. (2000). Functional properties of heteromeric P2X(1/5) receptors expressed in HEK cells and excitatory junction potentials in guinea-pig submucosal arterioles. *J Auton Nerv Syst* **81,** 249-263.

Sutcliffe MJ, Haneef I, Carney D & Blundell TL. (1987). Knowledge based modelling of homologous proteins, Part I: Three-dimensional frameworks derived from the simultaneous superposition of multiple structures. *Protein Eng* **1,** 377-384.

Szymanski M, Deniziak MA & Barciszewski J. (2001). Aminoacyl-tRNA synthetases database. *Nucleic Acids Res* **29,** 288-290.

Tanner NK, Cordin O, Banroques J, Doere M & Linder P. (2003). The Q motif: a newly identified motif in DEAD box helicases may regulate ATP binding and hydrolysis. *Mol Cell* **11,** 127-138.

Tatusov RL, Altschul SF & Koonin EV. (1994). Detection of conserved segments in proteins: iterative scanning of sequence databases with alignment blocks. *Proc Natl Acad Sci U S A* **91,** 12091-12095.

Theis K, Chen PJ, Skorvaga M, Van Houten B & Kisker C. (1999). Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair. *Embo J* **18,** 6899-6907.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F & Higgins DG. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25,** 4876-4882.

Thompson JD, Higgins DG & Gibson TJ. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22,** 4673-4680.

Thompson JD, Thierry JC & Poch O. (2003). RASCAL: rapid scanning and correction of multiple sequence alignments. *Bioinformatics* **19,** 1155-1161.

Torres GE, Egan TM & Voigt MM. (1998a). N-Linked glycosylation is essential for the functional expression of the recombinant P2X2 receptor. *Biochemistry* **37,** 14845-14851.

Torres GE, Egan TM & Voigt MM. (1998b). Topological analysis of the ATP-gated ionotropic [correction of ionotrophic] P2X2 receptor subunit. *FEBS Lett* **425,** 19-23.

Torres GE, Egan TM & Voigt MM. (1999). Identification of a domain involved in ATP-gated ionotropic receptor subunit assembly. *J Biol Chem* **274,** 22359-22365.

Torres GE, Haines WR, Egan TM & Voigt MM. (1998c). Co-expression of P2X1 and P2X5 receptor subunits reveals a novel ATP-gated ion channel. *Mol Pharmacol* **54,** 989-993.

Toulme E, Soto F, Garret M & Boue-Grabot E. (2006). Functional properties of internalization-deficient P2X4 receptors reveal a novel mechanism of ligand-gated channel facilitation by ivermectin. *Mol Pharmacol* **69,** 576-587.

Traut TW. (1994). The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. *Eur J Biochem* **222,** 9-19.

Tsuda M, Shigemoto-Mogami Y, Koizumi S, Mizokoshi A, Kohsaka S, Salter MW & Inoue K. (2003). P2X4 receptors induced in spinal microglia gate tactile allodynia after nerve injury. *Nature* **424,** 778-783.

Turner CM, Ramesh B, Srai SK, Burnstock G & Unwin RJ. (2004). Altered ATP-sensitive P2 receptor subtype expression in the Han:SPRD cy/+ rat, a model of autosomal dominant polycystic kidney disease. *Cells Tissues Organs* **178,** 168-179.

Urano T, Nishimori H, Han H, Furuhata T, Kimura Y, Nakamura Y & Tokino T. (1997). Cloning of P2XM, a novel human P2X receptor gene regulated by p53. *Cancer Res* **57,** 3281-3287.

Valera S, Hussy N, Evans RJ, Adami N, North RA, Surprenant A & Buell G. (1994). A new class of ligand-gated ion channel defined by P2x receptor for extracellular ATP. *Nature* **371,** 516-519.

Valera S, Talabot F, Evans RJ, Gos A, Antonarakis SE, Morris MA & Buell GN. (1995). Characterization and chromosomal localization of a human P2X receptor from the urinary bladder. *Receptors Channels* **3,** 283-289.

van Calker D, Muller M & Hamprecht B. (1979). Adenosine regulates via two different types of receptors, the accumulation of cyclic AMP in cultured brain cells. *J Neurochem* **33,** 999-1005.

Vassort G. (2001). Adenosine 5'-triphosphate: a P2-purinergic agonist in the myocardium. *Physiol Rev* **81,** 767-806.

Vial C & Evans RJ. (2000). P2X receptor expression in mouse urinary bladder and the requirement of P2X1 receptors for functional P2X receptor responses in the mouse urinary bladder smooth muscle. *Br J Pharmacol* **131,** 1489-1495.

Vial C & Evans RJ. (2002). P2X1 Receptor-Deficient Mice Establish the Native P2X Receptor and a P2Y6-Like Receptor in Arteries. *Mol Pharmacol* **62,** 1438-1445.

Vial C, Roberts JA & Evans RJ. (2004). Molecular properties of ATP-gated P2X receptor ion channels. *Trends Pharmacol Sci* **25,** 487-493.

Vial C, Rolf MG, Mahaut-Smith MP & Evans RJ. (2002). A study of P2X1 receptor function in murine megakaryocytes and human platelets reveals synergy with P2Y receptors. *Br J Pharmacol* **135,** 363-372.

von Kugelgen I & Starke K. (1990). Evidence for two separate vasoconstriction-mediating nucleotide receptors, both distinct from the P2x-receptor, in rabbit basilar artery: a receptor for pyrimidine nucleotides and a receptor for purine nucleotides. *Naunyn Schmiedebergs Arch Pharmacol* **341,** 538-546.

Vulchanova L, Arvidsson U, Riedl M, Wang J, Buell G, Surprenant A, North RA & Elde R. (1996). Differential distribution of two ATP-gated channels (P2X receptors) determined by immunocytochemistry. *Proc Natl Acad Sci U S A* **93,** 8063-8067.

Vullo A & Frasconi P. (2004). Disulfide connectivity prediction using recursive neural networks and evolutionary information. *Bioinformatics* **20,** 653-659.

Waterman MS & Vingron M. (1994). Rapid and accurate estimates of statistical significance for sequence data base searches. *Proc Natl Acad Sci U S A* **91,** 4625-4628.

Webb TE, Simon J, Krishek BJ, Bateson AN, Smart TG, King BF, Burnstock G & Barnard EA. (1993). Cloning and functional expression of a brain G-protein-coupled ATP receptor. *FEBS Lett* **324,** 219-225.

Werner P, Seward EP, Buell GN & North RA. (1996). Domains of P2X receptors involved in desensitization. *Proc Natl Acad Sci U S A* **93,** 15485-15490.

Wieraszko A. (1996). Extracellular ATP as a neurotransmitter: its role in synaptic plasticity in the hippocampus. *Acta Neurobiol Exp (Wars)* **56,** 637-648.

Wiklund NP & Gustafsson LE. (1988). Indications for P2-purinoceptor subtypes in guinea pig smooth muscle. *Eur J Pharmacol* **148,** 361-370.

Wilkinson WJ, Jiang LH, Surprenant A & North RA. (2006). Role of ectodomain lysines in the subunits of the heteromeric P2X2/3 receptor. *Mol Pharmacol* **70,** 1159-1163.

Worthington RA, Smart ML, Gu BJ, Williams DA, Petrou S, Wiley JS & Barden JA. (2002). Point mutations confer loss of ATP-induced human P2X(7) receptor function. *FEBS Lett* **512,** 43-46.

Xiong K, Hu XQ, Stewart RR, Weight FF & Li C. (2005). The mechanism by which ethanol inhibits rat P2X(4) receptors is altered by mutation of histidine 241. *Br J Pharmacol.*

Xiong K, Stewart RR, Hu XQ, Werby E, Peoples RW, Weight FF & Li C. (2004). Role of extracellular histidines in agonist sensitivity of the rat P2X4 receptor. *Neurosci Lett* **365,** 195-199.

Yan Z, Liang Z, Obsil T & Stojilkovic SS. (2006). Participation of the Lys313-Ile333 Sequence of the Purinergic P2X4 Receptor in Agonist Binding and Transduction of Signals to the Channel Gate. *J Biol Chem* **281,** 32649-32659.

Yan Z, Liang Z, Tomic M, Obsil T & Stojilkovic SS. (2005). Molecular Determinants of the Agonist Binding Domain of a P2X Receptor Channel. *Mol Pharmacol* **67,** 1078-1088.

Yiangou Y, Facer P, Baecker PA, Ford AP, Knowles CH, Chan CL, Williams NS & Anand P. (2001). ATP-gated ion channel P2X(3) is increased in human inflammatory bowel disease. *Neurogastroenterol Motil* **13,** 365-369.

Zemkova H, He M-L, Koshimizu T-a & Stojilkovic SS. (2004). Identification of Ectodomain Regions Contributing to Gating, Deactivation, and Resensitization of Purinergic P2X Receptors. *J Neurosci* **24,** 6968-6978.

Zhou QY, Li C, Olah ME, Johnson RA, Stiles GL & Civelli O. (1992). Molecular cloning and characterization of an adenosine receptor: the A3 adenosine receptor. *Proc Natl Acad Sci U S A* **89,** 7432-7436.