Abstract

G-protein coupled receptors (GPCRs) are cell surface receptors which interact with a diverse range of external stimuli including hormones, neurotransmitters and many drugs and thereby mediate a wide range of intracellular responses. Their importance in health and disease has made them the subject of intense study, particularly in terms of understanding their structure-function relationships. Adenosine A<sub>2a</sub>R is a GPCR activated by the nucleoside, adenosine. It is present ubiquitously within the human body and has roles in immune regulation, sleep induction and neurological disorders making it an important target for structural work. The aim of the research outlined in this thesis was to develop reliable, reproducible protocols for the expression and purification of a stable form of human adenosine A<sub>2a</sub>R. WT and C-terminally truncated adenosine A<sub>2a</sub>R constructs were expressed in the methylotrophic yeast *Pichia pastoris*. Very high-level expression (11 mg/L, 222 pmol/mg) was obtained for an adenosine A<sub>2a</sub>R truncated at residue V334. The functional yield following both solubilisation and purification was monitored by radioligand binding analysis using [³H] ZM241385 as a means of optimising recovery of receptor. Solubilisation trials revealed that enrichment of functional recovery of V334 adenosine A<sub>2a</sub>R was only possible in the presence of cholesterol. A two-step purification protocol including Flag tag affinity chromatography followed by TEV cleavage and reverse His-trap resulted in a highly homogenous and pure sample with a specific activity of 21 nmol/mg, close to the theoretical maximum of 24 nmol/mg. The purified receptor is functionally stable and degradation resistant for a period of 15 days at 4°C and 20°C. G-protein peptides were generated as potential tools for co-crystallisation. The protein was submitted to crystallisation trials in the presence and absence of the peptides using the lipidic cubic phase technique.
Declaration

I declare that the work detailed in this thesis has not been accepted in any previous application for a higher degree, that unless otherwise stated this work has been carried out by myself and that all sources of information have been specifically acknowledged.
Acknowledgements

I would like to thank my supervisor Dr. Bernadette Byrne for this immense opportunity and her undoubted support and encouragement. A sincere thanks to Prof. Philip Strange for being a source of wisdom and for adopting me as one of his own. A special thanks to Diana, Elodie, Hong, Helen and Doug at Reading University for accepting me as a member of their team. Thanks to Dr. Benjamin Bax for his guidance and finding the right people at GSK, especially Emma Jones who made the arduous protein purifications fun. Thanks to Prof. So Iwata for providing the expertise and infrastructure at MPC and supporting me in the final stages of the PhD. I thank Dr. Momi Iwata and Tian for making the last day of my PhD lab work truly memorable. I want to express my gratitude to Dr. Stuart Haslam, Dr. Alana Trolley, Dr. Tadeusz Skarzynski, Dr. Pil Seok-Chae and Dr. Renaud Wagner for their valuable collaborations. I also want to thank James Mansfield, Nien-Jen, Yilmaj, James, Rohini, Adrien, Jonathan, Magda, Yang, Sam, Ying, Manami, Mutsuko and all my colleagues over the years for your fellowship. Thanks to Dr. Julia Hubbard for providing the opportunity to continue working on GPCRs.

A heartfelt thanks to my extended family at Wilson; Mike, Gwendolyn, the little stars James and Verity, Dom W, Tim, Alex, Giudi, Dom B, Laura, Dave and Matina for all the love and good times! Megha and Mansi I cherish your friendship. Suyash my little brother, the love that you have for me and the sincerity that you show towards your work inspires me to do my best. Ma and Dad you have unconditionally loved and encouraged me to pursue my dreams and invested your lives in me and bhai. I hope we can always give you reasons to be proud parents. Vinayak, I couldn’t have seen this through without your love, support and patience, I must be blessed to have found you.
# Table of Contents

1 INTRODUCTION .............................................................................................................. 15

1.1 OVERVIEW .................................................................................................................. 16

1.2 GPCR CLASSIFICATION ............................................................................................... 18

1.3 GPCR SIGNALING MECHANISMS .................................................................................. 19

1.3.1 Mechanistic models of ligand binding ...................................................................... 20

1.3.2 G-proteins and GPCR activation .............................................................................. 23

1.3.3 Intracellular events following GPCR activation ....................................................... 29

1.4 MUTAGENESIS STUDIES ............................................................................................. 33

1.5 GPCR Oligomerisation ................................................................................................ 34

1.6 CHALLENGES OF GPCR STRUCTURE DETERMINATION ........................................... 35

1.6.1 Comparison between different recombinant expression systems for production of GPCRs .................................................................................................................. 36

1.6.2 Challenges with solubilisation and purification ....................................................... 40

1.7 STRUCTURAL INSIGHTS ............................................................................................... 44

1.7.1 Rhodopsin structure ................................................................................................. 44

1.7.2 The β2 adrenergic receptor structures ...................................................................... 50

1.7.3 The β1 adrenergic receptor structure ....................................................................... 55

1.7.4 The adenosine A2a receptor structure ...................................................................... 57

1.8 THE ADENOSINE A2A RECEPTOR .............................................................................. 61

1.8.1 Adenosine and the adenosine receptors .................................................................. 61

2 EXPRESSION AND CHARACTERISATION OF MEMBRANE BOUND ADENOSINE A2A RECEPTOR CONSTRUCTS ................................................................. 67

2.1 INTRODUCTION ............................................................................................................. 68

2.2 MATERIALS AND METHODS ....................................................................................... 72

2.2.1 Materials .................................................................................................................. 72
2.2.2 Expression Constructs........................................................................................................... 72
2.2.3 Small Scale cultures............................................................................................................. 75
2.2.4 Bioreactor cultures............................................................................................................... 75
2.2.5 Small scale membrane preparations ................................................................................... 77
2.2.6 Large scale membrane preparations .................................................................................. 77
2.2.7 Protein concentration determination................................................................................... 78
2.2.8 SDS-PAGE and Western blotting analysis.......................................................................... 78
2.2.9 Saturation radioligand binding assays................................................................................ 79
2.2.10 Data Analysis: Saturation binding assay........................................................................... 80
2.2.11 Competition radioligand binding assays.......................................................................... 80
2.2.12 Data analysis: Competition radioligand binding assay...................................................... 81

2.3 RESULTS.................................................................................................................................. 83
2.3.1 Bioreactor culture allows controlled growth........................................................................ 83
2.3.2 Comparison of Flask and Bioreactor cell growth............................................................... 85
2.3.3 Comparison of expression levels of the three adenosine $A_{2a}$R forms in a bioreactor ........................................................................................................................................ 89
2.3.4 Pharmacological characterisation of receptor forms expressed in P. pastoris......................... 90

2.4 DISCUSSION............................................................................................................................ 93

3 DEVELOPING A SOLUBILISATION AND PURIFICATION PROTOCOL FOR THE ADENOSINE $A_{2a}$ RECEPTOR........................................................................................................... 97

3.1 INTRODUCTION ...................................................................................................................... 98
3.2 MATERIALS AND METHODS................................................................................................. 100
3.2.1 Materials............................................................................................................................ 100
3.2.2 Solubilisation...................................................................................................................... 100
3.2.3 Modified protein concentration determination assay....................................................... 101
3.2.4 Analysing column efficiency............................................................................................. 101
3.2.5 Radioligand binding assay on solubilised and purified receptor....................................... 102
3.2.6 Analysing data for solubilised and purified receptor ........................................ 103
3.2.7 Solubilisation condition screening ........................................................................ 103
3.2.8 Determining T_m of detergent solubilised receptor ............................................. 103
3.2.9 Purification ........................................................................................................... 104
3.2.10 SDS-PAGE and Western blot ............................................................................. 106
3.3 RESULTS ................................................................................................................. 107
3.3.1 Analysing column efficiency ................................................................................ 107
3.3.2 Developing an assay for determining functional adenosine A_2aR activity in detergent solution .................................................................................................................. 107
3.3.3 Presence of CHS significantly increases the functional yield and stability of detergent-solubilised adenosine A_2aR ........................................................................................................ 109
3.3.4 Assessment of binding characteristics of solubilised receptor ............................ 109
3.3.5 The T_m values for V334 adenosine A_2aR are comparable to the conformational thermostable mutants .......................................................................................................................... 113
3.3.6 Development of a purification protocol .............................................................. 114
3.3.7 Second step of purification .................................................................................. 118
3.3.8 TEV cleavage and subsequent protein purification optimisation ....................... 121
3.4 DISCUSSION ............................................................................................................ 125

4 BIOPHYSICAL CHARACTERISATION OF THE PURIFIED V334 ADENOSINE A_2A RECEPTOR FORM ................................................................................................................................. 129

4.1 INTRODUCTION ................................................................................................... 130
4.2 MATERIALS AND METHODS ............................................................................... 133
4.2.1 Materials ............................................................................................................. 133
4.2.2 Analysing long term stability of purified receptor .............................................. 133
4.2.3 CPM analysis ..................................................................................................... 134
4.2.4 Circular Dichroism analysis .............................................................................. 135
4.2.5 Analytical SEC ................................................................................................. 135
4.2.6 Glycosylation profiling ................................................................. 136
4.2.7 Amphiphiles .................................................................................. 138
4.2.8 SEC of receptor in different detergents ......................................... 139
4.2.9 Crystallisation trials ..................................................................... 140

4.3 RESULTS ............................................................................................ 142
4.3.1 The purified V334 adenosine A\textsubscript{2a}R is highly stable .......... 142
4.3.2 Ligands help stabilise the receptor ................................................ 143
4.3.3 CHS is important to maintain the receptor in a folded state .......... 144
4.3.4 The purified protein is monodisperse ........................................... 145
4.3.5 The purified protein is glycosylated .............................................. 146
4.3.6 Amphiphiles improve the thermostability of receptor .................... 148
4.3.7 The receptor state in solution is different in different amphiphiles and detergents ......................................................................................... 149
4.3.8 Crystallisation trials set up through vapour diffusion and Lipid Cubic Phase ................................................................. 151

4.4 DISCUSSION ......................................................................................... 153

5 ALTERNATIVE APPROACHES TO PRODUCING AN ADENOSINE A\textsubscript{2a} RECEPTOR FORM SUITABLE FOR STRUCTURAL STUDIES ........................................... 157

5 ............................................................................................................. 158
5.1 INTRODUCTION ................................................................................. 158
5.1.1 GFP adenosine A\textsubscript{2a}R fusion construct .................................... 158
5.1.2 Adenosine A\textsubscript{2a}R and the G-protein ...................................... 161

5.2 MATERIALS AND METHODS ............................................................ 164
5.2.1 Materials ........................................................................................ 164
5.2.2 Generation and assessment of adenosine A\textsubscript{2a}R GFP fusion construct .......... 164
5.2.3 G-protein fusion constructs ........................................................... 168
5.2.4 Saturation radioligand binding studies ........................................ 169
List of Figures

Figure 1.1 Schematic representation of the mammalian phospholipid bilayer........................................ 16
Figure 1.2 Schematic representation of a class A GPCR structure.......................................................... 18
Figure 1.3 Ternary complex model......................................................................................................... 22
Figure 1.4 Extended ternary complex model........................................................................................... 22
Figure 1.5 Level of GPCR activation in the presence of different ligands.................................................. 23
Figure 1.6 Structure of the heterotrimeric G-protein complex..................................................................... 24
Figure 1.7 GPCR signaling mechanism via the heterotrimeric G-protein...................................................... 26
Figure 1.8 Topology of the Gα subunit of a heterotrimeric G-protein......................................................... 28
Figure 1.9 GPCR internalisation and recycling mechanisms........................................................................ 32
Figure 1.10 Illustration of membrane protein solubilisation and crystallisation.......................................... 42
Figure 1.11 The structure of rhodopsin..................................................................................................... 46
Figure 1.12 The ligand binding pocket of rhodopsin.................................................................................. 47
Figure 1.13 The ionic lock of rhodopsin.................................................................................................... 48
Figure 1.14 The opsin structures................................................................................................................ 49
Figure 1.15 The three β2AR structures....................................................................................................... 52
Figure 1.16 The ligand binding pocket of β2AR......................................................................................... 54
Figure 1.17 Structure of the β1 adrenergic receptor................................................................................... 56
Figure 1.18 The β1AR cytoplasmic surface structure.................................................................................. 57
Figure 1.19 The adenosine A2βR structure................................................................................................. 58
Figure 1.20 Ligand binding pocket of adenosine A2βR.............................................................................. 60
Figure 1.21 Schematic representation of metabolism of adenosine............................................................. 62
Figure 1.22 Ligands of adenosine A2βR used in this study.......................................................................... 63
Figure 2.1 Schematic representation of the adenosine A2βR constructs used in this study....................... 73
Figure 2.2 Sequence alignment of adenosine A2βR and bovine rhodopsin................................................. 74
Figure 2.3 Representation of key bioreactor parameters............................................................................ 85
Figure 2.4 Expression profile of adenosine A2βR construct over time...................................................... 87
Figure 2.5 Functional expression of adenosine A2βR constructs over time.............................................. 88
Figure 2.6 Saturation analysis of [3H]ZM241385 to the membrane bound adenosine A2βR............... 89
FIGURE 5.2 SCHEMATIC REPRESENTATION OF THE GFP FUSION EXPRESSION VECTOR USED IN THE STUDY......165
FIGURE 5.3 SCHEMATIC REPRESENTATION OF ADENOSINE A₂₃R G-PROTEIN FUSION CONSTRUCTS ............169
FIGURE 5.4 SEQUENCE ALIGNMENT FOR TRANSDUCIN AND Gₑα ..............................................170
FIGURE 5.5 LOCATION OF Gₑα PEPTIDES WITHIN THE G-PROTEIN .............................................172
FIGURE 5.6 SCHEMATIC REPRESENTATION OF THE PRINCIPLE OF THE cAMP FUNCTIONAL ASSAY ..174
FIGURE 5.7 CELL DENSITY OPTIMISATION FOR FUNCTIONAL ASSAY ........................................175
FIGURE 5.8 A STANDARD CURVE FOR FREE cAMP CONCENTRATION ........................................176
FIGURE 5.9 SATURATION ANALYSIS FOR [³H]ZM241385 TO THE MEMBRANE BOUND GFP ADENOSINE A₂₃R
FUSION CONSTRUCT ..........................................................................................................................178
FIGURE 5.10 SATURATION ANALYSIS OF [³H]ZM241385 TO THE MEMBRANE BOUND G-PROTEIN FUSION
ADENOSINE A₂₃R CONSTRUCTS .......................................................................................................179
FIGURE 5.11 DETECTING ABILITY OF PEPTIDES TO INHIBIT THE ADENOSINE A₂₃R G-PROTEIN INTERACTION...181
FIGURE 5.12 CRYSTALS OBSERVED IN MEMGOLD SCREENS FOR V334 ADENOSINE A2AR WITH NECA AND
PEPTIDE 736 ........................................................................................................................................182
List of Tables

TABLE 2.1 ESTIMATE OF RECEPTOR YIELD IN A BIOREACTOR CULTURE THROUGH SATURATION ASSAY ............ 90
TABLE 2.2 COMPETITION ASSAY PARAMETERS FOR MEMBRANE BOUND RECEPTOR .................................. 92
TABLE 2.3 LIGAND BINDING AFFINITIES FOR ADENOSINE A2aR IN DIFFERENT EXPRESSION SYSTEMS .............. 92
TABLE 3.1 OPTIMISATION OF LIGAND BINDING ASSAY FOR SOLUBLE RECEPTOR .................................. 108
TABLE 3.2 SATURATION ASSAY PARAMETERS FOR ADENOSINE A2aR FORMS SOLUBILISED IN DM OR DDM + CHS. .................................................................................................................. 111
TABLE 3.3 COMPETITION ASSAY PARAMETERS FOR ADENOSINE A2aR SOLUBILISED IN DM+CHS .......... 113
TABLE 3.4 FUNCTIONAL RECEPTOR YIELDS FROM CO2+ AND Ni2+ IMAC AS THE FIRST STEP OF PURIFICATION.. 117
TABLE 3.5 SINGLE POINT BINDING ANALYSIS PARAMETERS FOR V334 ADENOSINE A2aR PURIFIED THROUGH
DIFFERENT PROCEDURES .................................................................................................................. 123
TABLE 3.6 COMPETITION ASSAY PARAMETERS FOR PURIFIED V334 ADENOSINE A2aR .......................... 124
TABLE 4.1 THE CRITICAL MICELLE CONCENTRATION OF NOVEL AMPHIPHILES ........................................ 139
TABLE 5.1 SATURATION ASSAY PARAMETERS FOR G-PROTEIN FUSION ADENOSINE A2aR CONSTRUCTS .... 180
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2aR</td>
<td>Adenosine A2a Receptor</td>
</tr>
<tr>
<td>7 TM</td>
<td>Seven transmembrane</td>
</tr>
<tr>
<td>A316 A2aR</td>
<td>Adenosine A2A Receptor truncated at Alanine 316</td>
</tr>
<tr>
<td>AOX</td>
<td>Alcohol Oxidase</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine Triphosphate</td>
</tr>
<tr>
<td>BCA</td>
<td>Bicinchoninic Acid</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic Adenosine Mono Phosphate</td>
</tr>
<tr>
<td>CCD</td>
<td>Charge Couple Device</td>
</tr>
<tr>
<td>CD</td>
<td>Circular Dichroism</td>
</tr>
<tr>
<td>CHO</td>
<td>Chinese Hamster Ovary</td>
</tr>
<tr>
<td>CHS</td>
<td>cholesteryl hemisuccinate</td>
</tr>
<tr>
<td>CMC</td>
<td>Critical Micellar Concentration</td>
</tr>
<tr>
<td>CPM</td>
<td>N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide dye</td>
</tr>
<tr>
<td>CV</td>
<td>Column Volume</td>
</tr>
<tr>
<td>DDM</td>
<td>n-Dodecyl-β-D-Maltoside</td>
</tr>
<tr>
<td>DM</td>
<td>n-Decyl-β-D-Maltoside</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl Sulphoxide</td>
</tr>
<tr>
<td>dO2</td>
<td>Dissolved Oxygen</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>ECL</td>
<td>Extracellular loop</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic Acid</td>
</tr>
<tr>
<td>Fab</td>
<td>Fragment antigen binding</td>
</tr>
<tr>
<td>G Protein</td>
<td>Guanine Nucleotide Binding Protein</td>
</tr>
<tr>
<td>GDP</td>
<td>Guanine Diphosphate</td>
</tr>
<tr>
<td>GFP</td>
<td>Green Fluorescent Protein</td>
</tr>
<tr>
<td>GPCR</td>
<td>G Protein Coupled Receptor</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanine Triphosphate</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)-1-piperazinethanesulfonic acid</td>
</tr>
<tr>
<td>IC50</td>
<td>Half maximal inhibitory concentration</td>
</tr>
<tr>
<td>ICL</td>
<td>Intracellular Loop</td>
</tr>
<tr>
<td>IMAC</td>
<td>Immobilised Metal Affinity Chromatography</td>
</tr>
<tr>
<td>Kd</td>
<td>Dissociation Constant</td>
</tr>
<tr>
<td>Ki</td>
<td>Inhibition Constant</td>
</tr>
<tr>
<td>LCP</td>
<td>Lipid Cubic Phase</td>
</tr>
<tr>
<td>mAb</td>
<td>Monoclonal antibody</td>
</tr>
<tr>
<td>MES</td>
<td>2-[N-morpholino]ethanesulfonic acid</td>
</tr>
<tr>
<td>NECA</td>
<td>5'-N-Ethylcarboxamidoadenosine</td>
</tr>
<tr>
<td>NTA</td>
<td>Nitrilotriacetate</td>
</tr>
<tr>
<td>OD600</td>
<td>Optical Density at 600nm</td>
</tr>
<tr>
<td>OG</td>
<td>n-Octyl-β-D-Glucopyranoside</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate Buffered Saline</td>
</tr>
<tr>
<td><strong>Abbreviation</strong></td>
<td><strong>Full Form</strong></td>
</tr>
<tr>
<td>------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein Data Bank</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethylene Glycol</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene Fluoride</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium Dodecyl Sulfate</td>
</tr>
<tr>
<td>SDS PAGE</td>
<td>SDS Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>SEC</td>
<td>Size Exclusion Chromatography</td>
</tr>
<tr>
<td>TEV</td>
<td>Tobacco Etch Virus</td>
</tr>
<tr>
<td>TM</td>
<td>Transmembrane</td>
</tr>
<tr>
<td>V334 A2R</td>
<td>Adenosine A2A Receptor truncated at Valine 334</td>
</tr>
<tr>
<td>WT</td>
<td>Wild Type</td>
</tr>
</tbody>
</table>
1 INTRODUCTION
1.1 OVERVIEW

The plasma membrane is an essential component of the cell. It separates the cell from the external environment maintaining discrete intracellular conditions. It is also a very dynamic system, composed of a bilayer of mainly phospholipid molecules, containing a central hydrophobic core and outward facing polar head groups, along with other lipids such as cholesterol in mammals (Paila et al., 2009) or ergosterol in yeast (van der Rest et al., 1995) (Figure 1.1). It also houses a wide range of membrane proteins with key cellular functions. These integral membrane proteins have diverse roles in respiration, photosynthesis, uptake of essential nutrients and export of waste products. Membrane proteins mediate the cellular responses to a range of bioactive molecules such as hormones, neurotransmitters and a large number of drugs and thus allow the cell to respond to changes in the external environment in a controlled manner. It is estimated that integral membrane proteins account for 25-30% of the open reading frames encoded by the human genome (Wiener, 2004).

![Figure 1.1 Schematic representation of the mammalian phospholipid bilayer](image)

The phospholipid bilayer constituting the plasma membrane is presented. The hydrophilic head groups (orange spheres) are located on the extracellular and intracellular surfaces of the membrane, in contact with the aqueous environment while the nonpolar tails (blue line) self-assort to form the hydrophobic central core. Lipids such as cholesterol are an essential part of the membrane providing it with the required rigidity and as a lipid partner for proteins. The polar head group of cholesterol is shown as red circle and non polar steroid ring and hydrocarbon tail are shown in yellow. A molecule of the integral membrane protein the β2AR (grey) in complex with the ligand cyanopindolol (2VT4) is also shown.
G-protein coupled receptors (GPCRs) are the largest class of membrane proteins with at least 799 GPCR encoding genes within the human genome (Gloriam et al., 2007). These mediate cellular responses to a wide range of biologically active molecules including organic odorants, nucleotides, nucleosides, peptides, lipids and drugs. Ligand binding causes activation of the receptor, which allows binding and activation of an intracellularly located cognate G-protein. Activation and subsequent dissociation of the G-protein causes stimulation of an intracellular signaling cascade which in turn effects a cellular change (discussed in more detail below).

It is estimated that 50% of all modern drugs and almost 25% of the top 200 best selling drugs target GPCRs (Klabunde et al., 2002; Overington et al., 2006; Sarramegna et al., 2003). These drugs have a wide range of pharmacological effects and are involved in the treatment of cancer, cardiovascular disease, gastrointestinal disease as well as central nervous system (CNS) and immunological disorders (Klabunde et al., 2002).

All GPCRs share a common secondary structure with a hydrophobic central core domain comprised of 7 transmembrane helices (TM I through to TM VII) connected by three intracellular loops ICL (1,2,3) and three extracellular loops ECL (1,2,3) (Figure 1.2). A short helical segment (helix VIII) is also present in the C-terminal tail of class A GPCRs as visualised first with rhodopsin (Palczewski et al., 2000) and subsequently in the β2 adrenergic receptor (Rasmussen et al., 2007) and β1 adrenergic receptor (Warne et al., 2008) as well as the adenosine A2a receptor (Jaakola et al., 2008).
A basic overview of a class A GPCR structure is presented. The transmembrane helices are labeled I-VII, the amino (NH$_2$) and carboxy (COOH) terminus and intracellular (ICL 1-3) and extracellular loops (ECL 1-3) are indicated. A helix VIII is also present within the cytoplasmic C-terminal domain. Conserved residues and their approximate locations are shown in white circles. Cysteine residues (C) involved in disulphide linkages within the ECL1 and ECL2 and the palmitoylation site at the C-terminus end are also indicated.

Despite a very similar overall architecture GPCRs share comparatively low sequence identity overall. The receptors belonging to the rhodopsin subfamily have a reasonable sequence identity within the transmembrane domain with some highly conserved residues throughout GPCRs such as the DRY motif present in TM III and the NPXXY motif in TM VII (Rhee et al., 1996) however the loop regions differ greatly in size and sequence (Mirzadegan et al., 2003).

1.2 GPCR CLASSIFICATION

Many attempts have been made to phylogenetically classify the diverse members of the GPCR superfamily. Attwood and Findlay in 1993 developed a sequence based
fingerprinting method (Attwood et al., 1993) which was subsequently used to identify sequences belonging to GPCRs. This was followed by a classification in 1994 by Kolakowski (Kolakowski, 1994) a modified version of which is also used by the International Union of Pharmacology, Committee on Receptor Nomenclature and Classification (NC-IUPHAR). Kolakowski classified the GPCRs that had been experimentally proven to bind to a G-protein into A-F families while all the other 7-TM spanning proteins were grouped into the O (other) family. The 6 different classes included Class A Rhodopsin-like, Class B Secretin-like, Class C Metabotropic glutamate / Pheromone, Class D Fungal pheromone, Class E cAMP receptors and the frizzled/smoothed family. Bockaert and Pin further classified the GPCRs into 5 families based on both their structural morphology and ligand-binding properties. The members of each family possess conserved amino acid motifs and similar ligand binding sites (Bockaert et al., 1999). The GRAFS nomenclature provided by Fredriksson and colleagues divided the GPCRs into five major families including Glutamate, Rhodopsin, Adhesion, Frizzled/Taste2 and Secretin (Fredriksson et al., 2003). The total number of known and verified or deorphanised human GPCRs continues to grow and the database is regularly updated. The work in this thesis focuses on the human adenosine A2aR, a class A rhodopsin-like GPCR.

1.3 GPCR SIGNALING MECHANISMS

GPCRs are versatile signaling proteins, activated by a remarkably wide spectrum of structurally diverse ligands such as biogenic amines, peptides, glycoproteins, hormones, ion and proteases. The ligands bind at a specific site on the extra cellular face of the GPCR. The location of the ligand binding domains for many GPCRs has been determined (Ji et al., 1998) and can include sites within the TM segments, the
amino terminus and/or extracellular sequences joining the TM domains (Bockaert et al., 1999).

1.3.1 Mechanistic models of ligand binding

Compounds that activate the receptor and initiate downstream signaling by activation of G-proteins are called agonists. Numerous models have been proposed describing the activation of GPCRs. The classical form of the ‘ternary complex model’ (De Lean et al., 1980) describes the interaction between the receptor (R), G-protein (G) and an agonist (A) (Figure 1.3). This model was extended (Samama et al., 1993) to accommodate the active (R*) and inactive state (R) of the receptor. An active receptor here refers to a receptor able to cause activation of downstream effector molecules and thus this model also accounts for the constitutive activity of the receptor.

According to the model the binding of an agonist to a receptor promotes the formation of the receptor/G-protein complex hence shifting the equilibrium $R + G \leftrightarrow RG$ to the right. Partial agonists induce submaximal activation of the G-protein even at saturating concentration. An antagonist does not affect the $R + G \leftrightarrow RG$ equilibrium or the basal activity of the receptor but inhibits binding of ligand to the receptor. An inverse agonist binds to the receptor and inhibits its basal activity thereby shifting the $R + G \leftrightarrow RG$ equilibrium to the left (Figure 1.3, Figure 1.5).

The ‘extended ternary complex’ model (Figure 1.4) includes the intermediate R* form. Agonists have higher affinity for R* while inverse agonist bind preferentially to R. Partial agonist owing to their affinity for both R and R* are less effective at shifting the equilibrium towards R*.

The molecular mechanisms involved during the formation of each state of the receptor and its interaction with the G-protein have been studied extensively through ligand binding and mutagenesis studies. The structures of G-protein complexes and
GPCRs have provided new insights into molecular mechanisms involved in the GPCR signaling.
Figure 1.3 Ternary complex model

The model involves interaction between agonist (A), GPCR (R) and G-protein (G). The equilibrium constant K defines binding of agonist to the receptor while the constant M defines the binding of GPCR to G-protein. The α constant defines the differential affinity of the compound to bind to the active versus the inactive form of the GPCR (De Lean et al., 1980).

Figure 1.4 Extended ternary complex model

In addition to the parameters defined in the ternary complex model this extended model defines an inactive GPCR (R) and an active GPCR (R*). The receptor undergoes a conformational change from R to R* defined by the constant J. The effect of agonist binding is defined by the constant β (Samama et al., 1993).
Figure 1.5 Level of GPCR activation in the presence of different ligands

GPCR activation is measured based on the amount of second messenger molecule produced. A few GPCRs have been reported to be constitutively active and hence show a basal level of activation (blue line) even in the absence of ligands. With increasing concentration of agonist increase in GPCR activation is observed (red line) whereas the opposite is true for inverse agonists (green line) which inhibit the basal level of activation of receptor. An antagonist (brown line) binds to the receptor and inhibits an agonist from binding and activating the receptor however it has no effect on the basal secondary molecule production by a GPCR.

1.3.2 G-proteins and GPCR activation

GPCRs regulate cellular activity by the activation of a heterotrimeric GTP-binding regulatory protein (G-protein) (Oldham et al., 2006; Sprang, 1997). The heterotrimeric G-protein is composed of three different chains namely \( \alpha \), \( \beta \) and \( \gamma \). There are a total of sixteen different genes encoding for \( G\alpha \) subunits, five for \( G\beta \) and twelve genes encoding for \( G\gamma \) subunits (Downes et al., 1999). The G-proteins are classified into four main families and referred to by their \( \alpha \) subunit namely the \( G\alpha_s \), \( G\alpha_i \), \( G\alpha_q \) and \( G\alpha_{12} \) (Simon et al., 1991). The \( G\alpha \) subunit is made up of two domains with different functions. A GTPase domain is responsible for binding and hydrolysis of GTP while a unique helical domain containing a six \( \alpha \)-helix bundle acts as a cap over the nucleotide binding pocket and is responsible for burying the GDP within the protein. Thus release of the GDP from the binding pocket acts as a rate-limiting
step for G-protein activation (Hamm, 1998). In the inactive, GDP bound form the Gα contains a hydrophobic region. The Gβ and Gγ subunits associate to form the heterotrimeric complex by direct interaction with this hydrophobic region (Figure 1.6).

Figure 1.6 Structure of the heterotrimeric G-protein complex
The structure of Gαi (green), Gβt (blue) and Gγt (red) is shown. Gβγ bind to the hydrophobic region present within the Gα bound to GDP (pink stick). The structure (PDB 1GOT) shows the α helix and β sheets within the Gα and Gβ subunits and the interaction between the α, β and γ subunits.

Currently structures of the active (Coleman et al., 1994; Noel et al., 1993), inactive and transition state (Lambright et al., 1994; Sondek et al., 1994) conformations of Gαi are known. Structures of an inactive heterotrimeric complex have also been determined (Lambright et al., 1996; Wall et al., 1995). The G-protein acts as a molecular switch based on the ability of the Gα subunit to cycle between a GDP bound inactive state and a GTP bound state capable of activating downstream effectors. When an extracellular ligand binds to a GPCR the receptor changes its conformation allowing the G-protein to associate with the intracellular regions of the receptor. A constitutively active receptor can interact with the G-protein in the absence of an activating ligand. The conformational change of the Gα subunit upon
interaction with the GPCR leads to dissociation of GDP followed by binding of GTP and subsequent G-protein activation. The GPCR-G-protein complex associated with GTP is highly unstable and rapidly dissociates into free GPCR, free Gα and free Gβγ complex. Once the Gα dissociates from the Gβγ dimer it couples to different effector molecules and initiates an intracellular signaling cascade. G-proteins have specific effector molecules and downstream effects such as the Gs and Gi couple to adenylate cyclase and cause stimulation and inhibition of adenylate cyclase respectively. Gi also causes activation of G-protein coupled inwardly rectifying potassium (GIRK) channels. The Gq is responsible for activation of phospholipase Cβ and G12 for activation of Rho guanine nucleotide exchange factors (GEFs) (Simon et al., 1991). The Gβγ complex has also been reported to bring about some downstream signaling (Figure 1.7). An alternate pathway involves β-arrestins acting as scaffolds and adapters that localise signalling molecules such as Erk (extracellular-regulated kinase) to ligand activated GPCRs and result in activation of MAP kinase (Miller et al., 2001).
The GPCR in its inactive form is unable to bind to a heterotrimeric G-protein coupled to GDP (red oval). Agonist (yellow oval) binding in the transmembrane region in class A GPCRs causes activation of the receptor allowing it to bind to the Gα protein (green oval) of the heterotrimeric G-protein. The conformational change of the Gα subunit upon interaction with the GPCR leads to dissociation of GDP followed by binding of GTP (blue oval) and subsequent G-protein activation. The GTP bound G-protein complex is highly unstable and dissociates into Gα and Gβγ (orange oval). The four families of Gα protein stimulate different effector molecules and trigger different signaling cascades which culminate in a cellular response. The Gβγ subunit has also been shown to cause cellular response through alternate pathways.

The bound GTP is hydrolysed to GDP mainly through the autocatalytic GTPase activity of the α-subunit. Hydrolysis of GTP to GDP inhibits the interaction
of Gα protein with the effector molecules and allows reconstitution of the Gαβγ heterotrimeric complex.

Structures of all major G-protein conformations except for the GPCR-G-protein complex are known and have provided an understanding of the G-protein at the molecular level. The Gα subunit consists of a mixture of α helices and β sheets within the GTPase domain (Oldham et al., 2006). A helical domain as mentioned earlier plays a key role in trapping nucleotides within the protein. The α3β5, α4β6, αNβ1 and C-terminus of the Gα protein have been reported as interaction sites with the receptor. A TCAT motif on the β6α5 loop is located on the opposite end of the receptor binding site and is involved in stabilising the binding of GDP to the loop connecting the β1 and α1 helix (Figure 1.6, Figure 1.8) (Oldham et al., 2006). The GTPase domain consists of the flexible loops designated molecular switches I, II, III. Comparison between the GDP bound (Lambright et al., 1994; Mixon et al., 1995) and GTPγS bound (Coleman et al., 1994; Noel et al., 1993) structures of transducin reveals a significant structural difference within this domain with a more ordered structure for the GTP-bound active form. The binding sites for GPCR, Gβγ dimer and effector molecules are also located on the GTPase domain. The Gβ subunit consists of a β-propeller structure with seven WD-40 repeats. The N-terminus of the Gγ subunit interacts with the N-terminus of the Gβ subunit through a coiled-coil interaction while the C-terminus of the Gγ subunit binds to the base of the Gβ subunit (Sondek et al., 1996; Wall et al., 1995). The Gβγ dimer is a functional unit that does not dissociate under physiological conditions (Fig 1.6). The membrane localisation and protein-protein interaction of G-proteins is regulated by post-translational modifications such as myristoylation of the Gα subunit (Chen et al., 2001; Smotrys et al., 2004). Acylation
of the N and C terminus of Gα and Gγ respectively may increase the affinity of the Gα and Gβγ dimer for each other (Iiri et al., 1996; Iniguez-Lluhi et al., 1992).

**Figure 1.8 Topology of the Gα subunit of a heterotrimeric G-protein**

The secondary structure of the Gα subunit is shown by numbered cylinders for α-helices and arrows for β-sheets. The helical region is depicted by lettered brown cylinders. The linkers L1 and L2 connect the helical and GTPase domains. The molecular switch regions (yellow) and receptor contact loops (red) are depicted. The TCAT motif and P loop involved in nucleotide binding are shown. The figure was adapted from Oldham et al., 2008.

In addition to the information revealed by X-ray structures, studies such as mutation of specific residues, generation of peptides corresponding to a particular domain or production of antibodies recognising a specific sequence have also provided information about GPCR/G-protein interactions. The C-terminus of Gα, in particular the last five residues, has been established to be important for GPCR/G-protein interactions (Bourne, 1997; Martin et al., 1996; Wess, 1997). Peptides corresponding to the C-terminus of Gα protein have been shown to stabilise the G-protein interacting form of the β adrenergic receptor (Rasenick et al., 1994) and inhibit GPCR/G-protein interaction (Hamm et al., 1988). A structure has also been published of opsin in complex with a peptide corresponding to the C-terminal region of transducin and is discussed in detail later (Scheerer et al., 2008). In addition, an
antibody specific to the Gα C-terminus has been shown to inhibit GPCR/G-protein interaction (McFadzean et al., 1989).

1.3.3 Intracellular events following GPCR activation

Activation of GPCRs leads to an alteration in the concentration of one or more small intracellular signaling molecules termed the second messengers. Two of the most common intracellular mediators are cyclic AMP (cAMP) and inositol (1,4,5)-trisphosphate (IP₃). Adenylate cyclase is responsible for producing cAMP. The Gα stimulates while Gαᵢ family of proteins inhibits adenylate cyclase thereby effecting the level of cAMP production and initiates a cascade of reactions that bring about a change within the cell (Cabrera-Vera et al., 2003). PKA (cAMP dependant protein kinase A) consists of two catalytic and two regulatory subunits in its inactive state. cAMP binds to the regulatory subunit causing the subunits to dissociate. The activated catalytic subunit can then phosphorylate substrate proteins. These can range from enzymes which act as activating or inhibitory factors, ion channels or may also regulate gene promoters causing increased expression of a specific gene. PKC (protein kinase C) also activates substrate protein through phosphorylation however it is activated by diacylglycerol and Ca²⁺.

Phospholipase C is regulated by the Gαq family of G-proteins. Phospholipase C hydrolyses phosphatidylinositol 4,5 bisphosphate (PIP₂) found primarily on the inner surface of the plasma membrane. Hydrolysis of PIP₂ forms the two important second messengers namely IP₃ and diacylglycerol. IP₃ activates the calcium channel leading to release of Ca²⁺ from intracellular stores into the cytosol (Bockaert et al., 1999) while diacylglycerol activates protein kinase C. The Gβγ complex has been shown to directly activate Phospholipase A and Phospholipase C upon activation by some receptors and also directly activate ion channels and several kinases thereby
regulating signaling (Milligan et al., 2006). Other proteins are also involved in GPCR signaling. Other key processes in the GPCR signaling cycle are the regulation pathways including receptor desensitisation and internalisation. Desensitisation is a process of reducing the signaling process and serves as a regulation mechanism. Three families of molecules are reported to control GPCR desensitisation; the second messenger dependant protein kinases, GPCR kinases (GRKs) and arrestins (Ferguson et al., 1996; Krupnick et al., 1998; Lefkowitz, 1993; Pitcher et al., 1998). GRKs and β-arrestins also contribute to GPCR endocytosis, intracellular trafficking and receptor recycling (Ferguson, 2001; Tsuga et al., 1994). Desensitisation can occur through various mechanisms including receptor phosphorylation and uncoupling of heterotrimeric G-proteins, internalisation of cell surface receptor into intracellular membrane compartments, down regulation of cellular components of the GPCR signaling following reduced mRNA and protein synthesis and degradation of pre-existing receptor (Pierce et al., 2002).

cAMP-dependant protein kinase (PKA) and protein kinase C (PKC) and GRKs directly uncouple G-proteins by modifying the GPCR as a consequence of phosphorylation by intracellular kinases. It is believed that serine and threonine residues in ICL3 and the C-terminus of GPCRs are selectively phosphorylated by PKA and PKC (Pierce et al., 2002; Pitcher et al., 1998).

Two types of desensitisation; heterologous and homologous can occur upon receptor phosphorylation (Klaasse et al., 2008). The heterologous form of desensitisation occurs by phosphorylation by PKA and PKC. Receptor not activated by agonists can also be phosphorylated and subsequently desensitised (Lohse et al., 1990). Kinase activated by one type of receptor can also lead to phosphorylation and desensitisation of another receptor. Homologous desensitisation is specific for the
agonist bound form of receptor. This type of desensitisation is carried out by GRKs which phosphorylate agonist activated receptor form and allow binding by β-arrestins (Pei et al., 1994).

Previously it had been shown that GRK-mediated phosphorylation was unable to cause complete inactivation of rhodopsin and β2AR. Full inactivation required an additional component to arrest the signaling process. The first mediators of this process to be discovered were the visual arrestins bound to light activated rhodopsin (Pfister et al., 1985). Four arrestin genes have so far been identified of which two express exclusively in retina (visual and cone arrestin) while β-arrestin 1 and β-arrestin 2 are expressed ubiquitously and are thought to interact with all other GPCRs (Krupnick et al., 1998). Arrestins preferentially bind to agonist-activated and GRK phosphorylated GPCRs as opposed to receptor phosphorylated by PKA or PKC or non-phosphorylated receptor. The β-arrestins terminate the G-protein activation by inhibition of G-protein coupling or target the receptor to clathrin coated pits for internalisation.

Internalisation is considered an important aspect of regulation of GPCR signaling by promoting sequestration of GPCR, recycling of receptor and regulation of receptor signaling (Ferguson, 2001). Clathrin-coated pits are specialised regions of cell membrane involved in internalisation of GPCRs. The inner surface of the membrane contains a polygonal lattice composed of three heavy chain and three light chain clathrin molecules and associated proteins (Klaasse et al., 2008). The β-arrestin binds to phosphoinositides present on the inner surface of the cell membrane (Gaidarov et al., 1999) and heavy chain of clathrin (Goodman et al., 1996) and β2 adaptin subunit of clathrin adaptor protein AP-2 (Laporte et al., 1999) of the clathrin coated machinery. The β-arrestin thus mediates the incorporation of GPCRs at the
cell surface into the clathrin coated pits. GTPase dynamin pinches off these receptor containing pits leading to rapid recycling of the receptor, targeting to endosomes for slow recycling or degradation in lysosome (Figure 1.9).

![Figure 1.9 GPCR internalisation and recycling mechanisms](image)

**Figure 1.9 GPCR internalisation and recycling mechanisms**

A GPCR activated by agonist is phosphorylated (P) at the serine and threonine residues by GRKs. β-arrestins (β-Ar) are then able to bind to receptor activated by agonist and coupled to GRK. The β-arrestins can bind to AP-2 (salmon ring) and clathrin (blue line) within the clathrin coated pits and lead to internalisation of the receptor with the aid of Dynamin. The receptor is then dephosphorylated in the endosome and recycled back to the membrane. Alternatively, the GPCR may be trafficked to the lysosomes for degradation. Adapted from Pierce et al., 2002.

Resensitisation of a GPCR includes endocytosis of the receptor followed by dephosphorylation within endocytic vesicles prior to recycling back to the membrane.
1.4 MUTAGENESIS STUDIES

Mutagenesis studies have been utilised to identify key residues involved in various aspects of GPCR function. A detailed account of the use of mutagenesis approaches to understand the molecular mechanism of ligand binding, signaling and regulation is presented in (Edvardsen *et al.*, 2002; Kristiansen, 2004; Kristiansen *et al.*, 1996).

Chimeric receptors have been generated using closely related subtypes in order to highlight regions involved in ligand binding. The information obtained from chimeric receptors has been used to identify domains with a specific function and inform further mutagenesis studies (Schwartz, 1994). Pioneering work involved the generation of chimeric proteins of the human α₂a and β₂ adrenergic receptors (β₂AR). Through this study TM VII was identified as a site for receptor specific ligand binding with particular importance for antagonist binding (Kobilka *et al.*, 1988). The approach of random mutagenesis was used with the δ-opioid receptors leading to identification of 30 residues involved in receptor activation. This study showed that the ECL3 of the receptor is a key site for ligand binding, while the signal is propagated via TM III, VI and VII to an activation switch located at the intracellular face of TM VI and VII (Decaillot *et al.*, 2003).

Certain residues in ECL2 have been identified as key sites for ligand binding for GPCRs activated by small molecules such as the adenosine receptor through site directed mutagenesis (Kim *et al.*, 1996). Similarly His 251 and His 278 residues on TM VI and VII respectively have been shown to be important for ligand binding to adenosine A₁ receptor (Olah *et al.*, 1992). Site directed mutagenesis studies on the adenosine A₂a receptor combined with radioligand binding studies identified the role of Glu 13 in TM I and His 278 in TM VII to be important for ligand binding (Gao *et al.*, 2000). ECL1 and ECL3 have been shown to be indirectly involved in receptor
folding and conformation through mutagenesis studies on the serotonin receptor (Kroeze et al., 2002). Through alanine scanning mutagenesis studies on the histamine H3 receptor the residue Glu 206 was identified to be important for ligand binding (Uveges et al., 2002).

These studies show that such approaches have been useful for the identification of small regions or individual residues important in GPCR function. However a full understanding of GPCR structure–function relationship is only possible through the determination of high resolution structures.

1.5 **GPCR OLIGOMERISATION**

Some receptors have been found to exist natively as dimers. Much work on GPCR oligomerisation was stimulated by the discovery that the functional GABAb receptor is a heterodimer of two different GPCRs (Marshall et al., 1999a; Mohler et al., 1999). Popular experimental techniques allowing the study of GPCR dimerisation include co-immunoprecipitation, where differentially epitope-labelled GPCR gene products can be shown to coexist as multimers (Marshall, 2001).

Fluorescence resonance energy transfer (FRET) and bioluminescence resonance energy transfer (BRET) are both very powerful techniques enabling the observation of GPCR oligomerisation. Initial resonance energy transfer studies were performed on *S. cerevisiae* Ste2p (Overton et al., 2000), where receptor tagged with either cyan or yellow fluorescent proteins at the C-terminus showed Ste2p receptor oligomerisation. This study revealed that the equilibrium between oligomer and monomer forms of the receptor was unaffected by agonist or antagonist binding. The dopamine D2 receptor has also been shown to dimerise in a ligand independent manner using FRET (Wurch et al., 2001).
The formation of GPCR dimers is currently thought to occur at the biosynthesis stage (Terrillon et al., 2003). However, there is some evidence to show that dimerisation occurs upon ligand binding (Cornea et al., 2001; Kroeger et al., 2001; Zhu et al., 2002).

In some cases like that of the metabotropic γ-aminobutyric acid b receptor, dimerisation has been shown to have a primary role in receptor maturation allowing the correct trafficking of the receptor from the endoplasmic reticulum (ER) to the cell surface (Marshall et al., 1999b). Once at the plasma membrane, dimers may become the target for dynamic regulation by ligand binding. It has also been proposed that GPCR heterodimerisation leads to both positive and negative ligand binding cooperativity, as well as potentiating/attenuating signaling or changing G protein selectivity. Heterodimerisation can also promote co-internalisation following the stimulation of only one of the two receptors as, for example, heterodimers of the adenosine A2a and dopamine D2 receptors were found in internal membrane compartments (Hillion et al., 2002).

1.6 CHALLENGES OF GPCR STRUCTURE DETERMINATION

In order to gain a precise understanding of the interactions between specific GPCRs and their ligands and to obtain insights into the mechanisms of G-protein activation it is necessary to acquire high resolution 3D structures. The expression and purification of these receptors, in order to achieve sufficient protein for crystallisation and structural determination, remains a non-trivial task even after years of research. The scarcity of structural data is in part due to the naturally low abundance of these proteins and their highly hydrophobic and dynamic nature.

The first integral membrane protein structure, that of the photosynthetic reaction structure was reported in 1985; reviewed in (Deisenhofer et al., 1989) while
the first structure of a membrane protein expressed in a recombinant system was only reported in 1998 (Doyle et al., 1998). The first GPCR structure, of an inactive form of rhodopsin (Palczewski et al., 2000), which is naturally abundant in bovine eye, was solved in 2000. This was followed by a seven year gap before the structure of the β2 adrenergic receptor (Cherezov et al., 2007), a GPCR expressed as a recombinant protein, was determined. These examples illustrate the challenges associated with the overexpression of membrane proteins in recombinant systems. With years of research in this field different expression systems ranging from bacteria, yeast and insect to mammalian cell lines have been established (Sarramegna et al., 2003; Tate et al., 1996). It is now possible to produce sufficient quantities of functional protein for structural work for some GPCRs. However, there is no generic expression system for making all GPCRs yet (Alkhalfioui et al., 2009; Chiu et al., 2008).

1.6.1 Comparison between different recombinant expression systems for production of GPCRs

1.6.1.1 Escherichia coli

The major advantage of E. coli is the short doubling-time of the host cell. GPCRs have been overexpressed in E. coli as a fusion to a periplasmic maltose binding protein (MBP) which has been shown to improve expression of protein to 1000 copies per cell (Abrahmsen et al., 1986; di Guan et al., 1988). There are many examples in the literature of expression and successful purification of GPCRs in E. coli such as the adenosine A2aR (Weiss et al., 2002), serotonin 5-HT1A (Bertin et al., 1992) and neurotensin (Hanninen et al., 1994; Tucker et al., 1996) receptor. The lack of post-translational modification within E. coli does guarantee a homogeneous population of receptor however it is known that post-translational modifications such as
glycosylation may affect the trafficking of the receptor (Li et al., 2007) to the membrane and function of GPCRs. Absence of glycosylation may modify the folding of rhodopsin thereby affecting its function (Kaushal et al., 1994). The lipid composition of the E. coli membrane differs from that of eukaryotic cells and hence may affect the functionality of receptor. The lipid environment has been shown to have an effect on the quality of receptors such as µ-opioid receptor (Hasegawa et al., 1987; Lagane et al., 2000) and oxytocin receptor (Gimpl et al., 1995).

One of the disadvantages of the system is that E. coli lacks native GPCRs and thus there is an absence of endogenous G-proteins. This property makes E. coli unsuitable for pharmacological characterisation of GPCR binding states as high-affinity agonist binding sites can not be achieved. However addition of G-protein to membrane bound serotonin 5HT_{1A} receptor did restore high-affinity agonist binding (Bertin et al., 1992).

With a short generation time this system is preferable for quick screening of different protein constructs for stability or expression. This approach was used for selecting the conformationally thermostable turkey β_{1} adrenergic receptor developed through alanine scanning mutagenesis (Serrano-Vega et al., 2008). The selected mutant was then transferred to a eukaryotic insect cell line for production of protein for structural work. Although many GPCRs have been successfully expressed and purified in E. coli it has not yet been possible to obtain a structure from this system. It is possible that E. coli is best used as a screening system as described above.

1.6.1.2 Yeast: *Saccharomyces cerevisiae* and *Pichia pastoris*

Unicellular eukaryotic yeasts have long been used for large scale fermentation in industry, which has helped develop them as heterologous expression systems. A short
generation time and an ability to grow on simple media make the process of expression simple and inexpensive (Hollenberg et al., 1997; Stratton et al., 1998). In addition the eukaryotic yeast can perform post-translational modifications such as glycosylation. Fusion of the α-mating factor prepropeptide (α-factor) to the N-terminus of the protein enables trafficking of the receptor to the membrane and helps improve cell-surface receptor expression (King et al., 1990). The presence of G-proteins in yeast may help maintain the agonist bound or active state of the receptor (Stefan et al., 1998). Expression plasmids which can be maintained in episomal form or integrated into the host genome are readily available. These incorporate strong promoters such as GAL1 for \textit{S.cerevisiae} and AOX1 for \textit{P. pastoris} that drive overexpression where up to 80% of protein content can be the target protein (Sarramegna et al., 2003). The production of protease deficient strains such as SMD1163 for \textit{P. pastoris} has been shown to increase production of protein (Gleeson et al., 1998; Sander et al., 1994).

Expression levels of 115 pmol/mg in \textit{S.cerevisiae} (King et al., 1990) and 25 pmol/mg in \textit{P. pastoris} have been achieved for the β2 adrenergic receptor (Weiss et al., 1998). This highlights the variability in receptor expression possible in different yeast systems. Despite successful expression and purification of a number of GPCRs (Sarramegna et al, 2003) no structure of a GPCR expressed in yeast has been solved yet.

\subsection{1.6.1.3 Baculovirus/insect cell system}

The use of the insect cell line \textit{Spodoptera frugiperda} 9 (\textit{Sf9}) together with the recombinant baculovirus \textit{Autographa californica} is well established as a system for protein expression (Fraser, 1992). Indeed, all of the recently reported ligand-activated GPCR structures were obtained using protein expressed in \textit{Sf9} cells (Cherezov et al., 2007; Jaakola et al.,
The system has been used to express high levels of functional receptor however like most expression systems the success can be receptor specific. Non-functional receptor expression has been reported, probably as a result of improper post-translational modifications (Parker et al., 1991). Sf9 cells are deficient in endogenous GPCRs but contain a number of G-proteins (Butkerait et al., 1995). This property is useful for functional assays due to a low-background environment and ability of the heterologously expressed receptor to couple to G-proteins. Owing to their eukaryotic nature insect cells can perform several post-translational modifications. However, receptors may be poorly glycosylated and a heterogeneous population of receptor is often obtained (Reilander et al., 1991). Another issue with expression in insect cell is a relatively long generation time and requirement of complex media making expression in this system both much more expensive and time-consuming than expression in either yeast or E. coli. However the track record of insect cell expression in terms of high-resolution structures does make it the system of choice for the production of receptors for structural studies. However mammalian cell lines may be better suited for functional studies and yeast systems still have considerable potential as GPCR expression hosts.

1.6.1.4 Mammalian cells
Mammalian cells undoubtedly provide the most native background including similar lipid composition, post translational modifications such as glycosylation and native G-proteins for GPCR expression. Though ideally suited for functional work such as drug-screening assays due to the production of active protein the typically low yields achieved in such expression systems make them less suitable for structural work (Sarramegna et al., 2003). A large number of GPCRs have been expressed in mammalian cells lines and used for mutagenesis studies to identify key residues.
involved in ligand binding and receptor activation (Schwartz, 1994; Strader et al., 1994). High-level expression in mammalian cells can be achieved by transfecting the cells with a recombinant virus harboring the target GPCR gene (Lundstrom et al., 1994; Walker et al., 1993). Overexpression in stable cell lines is also possible (Lohse, 1992; Needham et al., 1995). However, the process of producing such a cell line is difficult in comparison to bacteria or yeast. To date no structures of GPCRs have been obtained using protein overexpressed in mammalian cell systems.

A range of different systems have been extensively utilised and developed for the expression of GPCRs for structural studies. The most notable successes so far have been achieved with protein produced in insect cell based system (Jaakola et al., 2008; Rasmussen et al., 2007; Rosenbaum et al., 2007; Warne et al., 2008). However it is possible that other expressions systems could also be successful following further optimisation.

1.6.2 Challenges with solubilisation and purification

The hydrophobic membrane-spanning domains of integral membrane proteins are shielded by the lipid bilayer in which they reside naturally and which provides a stable native environment. The isolation of membrane proteins from such a lipid environment involves replacing this with a solvent with similar properties, such as detergents (Ostermeier et al., 1997). Detergents are only effective at solubilising protein at or above their critical micelle concentration (CMC). Typically a large excess of detergent is added to membranes in order to effectively extract as much of the membrane protein as possible. The resultant protein-detergent micelle (Figure 1.10) can then be purified. It is rare that solubilisation efficiency of 100% can be achieved. The most commonly used detergent for membrane protein solubilisation is dodecyl-β-maltoside (DDM), a C12 detergent which forms a comparatively large micelle around
the protein effectively shielding the hydrophobic domains from the aqueous environment. Formation of a large detergent micelle around the protein may increase solubility however it can reduce the likelihood of the molecules forming protein-protein interactions essential for crystal formation (Ostermeier et al., 1997). Similarly a short aliphatic chain detergent may be best suited for crystallisation due to the comparatively small micelle size. However these can also be the least stabilising as the small micelle size leads to increased exposure of the hydrophobic regions of the protein and an increased likelihood of non-specific aggregation. Often successful crystallisation means finding the detergent or combination of detergents which maintains the protein in a stable state whilst maximising the regions of the protein available to form crystal-crystal contacts (Iwata, 2003).
Membrane proteins are solubilised and extracted from the lipid environment using amphipathic detergents. Type I 2D crystals can be formed by exchanging detergents with lipids. Type I 3D crystals are obtained in lipidic cubic phase. Type II 3D crystals are the most commonly obtained crystals. These are obtained in exactly the same way as soluble protein crystals but these tend to have a much higher solvent content due to the presence of the detergent micelles (Iwata, 2003).

Purification of membrane proteins serves as another major hurdle. The challenge is to obtain protein which is monodisperse, homogenous and functionally active with as few isolation steps as possible. A prolonged exposure to detergents and excessive washes during purification can lead to delipidation and aggregation or precipitation of the protein. Often isolation protocols need to be individually optimised for different proteins (Kobilka, 1995; Warne et al., 2003)
Another issue is the highly dynamic nature of GPCRs which equilibrate between multiple conformational states (Kobilka et al., 2007). Specific ligands with high affinity to a receptor are used to stabilise the receptor in a specific conformation and produce a homogenous population of the protein in order to improve the chances of crystallisation. The recently published structures of GPCRs are of receptor bound to either an antagonist (Jaakola et al., 2008; Warne et al., 2008) or partial inverse agonist (Cherezov et al., 2007; Hanson et al., 2008). One recent success in crystallisation has involved addition of antibodies to enlarge the soluble domain to increase crystal contacts (Rasmussen et al., 2007). An antibody was raised to recognise and bind to the N and C-terminal ends of the ICL3 of β2AR. Interaction with the antibody helped maintain the receptor in a stable conformation. A similar result was achieved by replacing ICL3 by T4 lysozyme (Jaakola et al., 2008; Rosenbaum et al., 2007). The known structures of GPCRs are of receptor bound to antagonists or inverse agonists which stabilise the receptor in its inactive conformation whereas a structure for the active form of receptors or one in complex with its cognate G-protein is also desirable. The recently described structures of a ligand free opsin (Park et al., 2008) and opsin in its G-protein interacting form (Scheerer et al., 2008) (see section 1.6.1) show it is possible to obtain crystals of GPCRs in different conformations and interacting with different molecules.

The ability to grow diffraction quality crystals for GPCRs is still a limiting step for GPCR structural biology. Previously workers have been able to express and purify highly functional GPCRs (Fraser, 2006; Grisshammer et al., 1995; Warne et al., 2003; Weiss et al., 2002; Weiss et al., 1995) but have been unable to obtain crystals. Crystallisation of GPCRs in lipid cubic phase has seen some recent successes of both non-GPCRs 7 TM proteins, Bacteriorhodopsin (Pebay-Peyroula et al., 1997) and
halorhodopsin (Kolbe et al., 2000) and GPCRs (Cherezov et al., 2007; Jaakola et al., 2008). The viscous nature of the lipids makes them difficult to manipulate and requires special equipment and expertise (Caffrey et al., 2009). However this technique does seem to be particularly suited to the 7-TMs. The crystals obtained from lipid cubic phase are usually very small and sensitive to radiation. However, recent developments in microdiffraction technology are facilitating data collection from such tiny, fragile crystals.

An alternative approach to the production of GPCR protein suitable for structural studies is the identification of thermostable mutants. It has been possible through alanine scanning mutagenesis to produce a β1AR protein which is highly stable even in the absence of lipids and can be crystallised using the vapour diffusion technique (Warne et al., 2008). The recent advances in structural work on GPCRs are mentioned below explaining these approaches in much more detail.

### 1.7 STRUCTURAL INSIGHTS

The last decade has witnessed the structure determination of four independent GPCRs with the majority of these reported in the past 3 years. The recently obtained structures of the human β2 adrenergic receptor (β2AR), the turkey β1 adrenergic receptor (β1AR), the human adenosine A2a receptor and bovine opsins the ligand free form of rhodopsin have provided unprecedented insight into the understanding of GPCR function at the molecular level.

#### 1.7.1 Rhodopsin structure

Rhodopsin is a class A GPCR localised in the rod cells of the retina, responsible for non-colour vision. Rhodopsin expression is also required for formation of rod outer
segments (Okada et al., 2001). Bovine retinas are large and contain mainly rod cells and are thus a rich source of native receptor. Rhodopsin was the GPCR of choice for initial structure studies mainly due to its ready availability (Filipek et al., 2003). Rhodopsin consists of the ligand free protein, (apoprotein) opsin, and the chromophore 11-cis-retinal covalently bound by a protonated Schiff base to Lys 296 in TM VII. 11-cis-retinal serves as an inverse agonist and stabilises the receptor in an inactive state. Absorption of a single photon of light causes a cis-trans isomerisation of retinal which in turn causes deprotonation of the retinal Schiff base linkage and formation of the active metarhodopsin II state via different photointermediate states (Knierim et al., 2007; Okada et al., 2001). This results in a change in the interaction between the retinal and polypeptide causing a re-packing of the transmembrane helices. This results in a conformational rearrangement of the intracellular surface of the receptor, providing a surface for interaction with the rhodopsin specific G-protein, transducin (Farrens et al., 1996). Following binding and activation of transducin, metarhodopsin II interacts with rhodopsin kinase and arrestin, leading to G-protein inactivation and subsequent hydrolysis of the retinal Schiff base. The all-trans retinal is released leaving the unliganded form of the receptor, opsin. Rhodopsin is then regenerated by the supply of new 11-cis-retinal produced metabolically (Lamb et al., 2004).

There are now fifteen bovine rhodopsin structures in the PDB including that of the receptor in the ground state, the ligand free opsin and photointermediates (Lodowski et al., 2009). The initial ground state rhodopsin structures included four tetragonal crystal forms (Okada et al., 2002; Okada et al., 2004; Palczewski et al., 2000; Teller et al., 2001) and one trigonal crystal form (Li et al., 2004).
The first high-resolution structure of rhodopsin (Palczewski et al., 2000) revealed a wealth of information and confirmed the projected topology and loop connectivity (Baldwin et al., 1997; Unger et al., 1997). The seven transmembrane spanning domains followed by a short helix VIII running parallel to the cytoplasmic surface of the membrane were observed (Figure 1.11).

![Figure 1.11 The structure of Rhodopsin](image)

**Figure 1.11 The structure of Rhodopsin**

The high resolution structure of Rhodopsin is presented (1F88, (Palczewski et al., 2000)). A single molecule from the crystallographic dimer in the asymmetric unit is shown in purple revealing the position of helix VIII relative to the other transmembrane domains and the membrane. The bound 11-cis-retinal which acts as an inverse agonist is shown in pink stick model. The region corresponding to the integral membrane domain of the protein is bounded by the grey square.

ECL 2 between TM IV and V contains two short β-strands stabilised by a disulphide bond between Cys110 and Cys187 (Figure 1.13). This structure forms a cap which stabilises the retinal chromophore located within the binding site (Yan et al., 2002).
As mentioned above, 11-cis retinal acts as an inverse agonist inhibiting basal activity. It is thought that the presence of the 11-cis retinal induces formation of the so-called ionic lock (Yao et al., 2006) (Figure 1.14) between Glu134 and Arg135 in TM III (D/ERY) motif and Glu247 and Thr251 in TM VI as revealed by the ground state structure of rhodopsin (Palczewski et al., 2000). The presence of the ionic lock constrains TM VI within the helix bundle.
The ionic lock has been observed only in rhodopsin of all class A GPCR structure solved. The transmembrane helices of rhodopsin (1F88) are depicted in grey. The residues Glu134 (blue stick model) and Arg135 (purple stick model) present on TM III and Glu247 (green stick model) and Thr251 (yellow stick model) on TM VI involved in the ionic lock are shown in stick model.

A more recent structure of light-activated, deprotonated rhodopsin at 4.15 Å (Salom et al., 2006) shows a highly flexible and disordered cytoplasmic domain and absence of an ionic lock. It was suggested that this flexibility within the cytoplasmic loop could allow coupling of transducin, an interaction not possible in the ground state of the receptor with the ionic lock.

The crystal structure of ligand free native opsin at 2.9 Å (Park et al., 2008) and opsin in its active G-protein interacting conformation at 3.2 Å (Scheerer et al., 2008) have also been published (Figure 1.14).
Figure 1.14 The Opsin structures

The ligand free (A) and G-protein interacting structure (B) of the ligand free form of rhodopsin, opsin is presented. The structure shown in blue is of the ligand free form of opsin (3CAP). The structure shown in purple is of opsin bound to a peptide (3DQB) corresponding to the residues of the C-terminal of the cognate G-protein (orange).

The most prominent feature revealed by the structures of the G-protein interacting form of receptor (Scheerer et al., 2008) was the outward tilt of the cytoplasmic end of the TM VI bringing it in much closer proximity to TM V. This is the only reported GPCR structure where outward tilting of TM VI is observed. Biophysical studies such as mutagenesis (Sheikh et al., 1996) and labeling studies (Altenbach et al., 2008; Farrens et al., 1996) have suggested that this movement is essential for GPCR activation. Tilting of TM VI causes the ionic lock to be broken as the distance between Arg135 and Glu247 increases from 2.7 Å in the ground state to 15.3 Å in the active opsin structure (Lodowski et al., 2009). New interactions are formed between Arg135 and Tyr223 and Glu247 with Thr251/Lys 231A stabilising the closed conformation of TM VI and TM V. These rearrangements are also suggested to cause formation of two openings in the binding pocket which may act as selective gates for uptake of 11 cis-retinal, the inverse agonist, and release of all-trans retinal, an agonist for rhodopsin. The cytoplasmic ends of TM V and TM VI in the G-protein interacting form of opsin provide a hydrophobic surface for interaction with
the Gα C-terminal (CT) peptide. Hydrophobic contacts are also provided by residues in ICL 2 and ICL 1. The GαCT peptide forms a helical structure and is shown to connect both the E/DRY and NPXXY motifs in the G-protein interacting structure of opsin (Scheerer et al., 2008). Thus it is clear that the same set of conserved residues, E/DRY and NPXXY, stabilise the inactive rhodopsin through the ionic lock and are also involved in mediating the interaction between transducin and opsin. (Scheerer et al., 2008).

Rhodopsin differs from other GPCRs in that its initial signal does not occur due to the binding of a small molecule ligand. The photon induced conformational changes in retinal are replaced by a ligand-binding event in other GPCRs. Undoubtedly the structures of rhodopsin, have provided a template for studies on other GPCRs as well as for drug design (Ballesteros et al., 2001). However, the very features of rhodopsin that make it amenable to structural studies mean that it is of limited use for understanding the ligand-receptor interactions for all other GPCRs. In order to gain a precise understanding of the interactions between specific GPCRs and their ligands and to facilitate rational design of novel drugs it is necessary to obtain high resolution structures of a number of these receptors in different conformations. Similar to the role of the rhodopsin structure in the understanding of GPCRs the structures of the opsin in the G-protein interacting form may also act as a stepping stone on the way to a comprehensive understanding of the signaling mechanism of GPCRs until structures of other activated GPCRs are elucidated.

1.7.2 The β2 adrenergic receptor structures

Despite extensive efforts for several years structural information for only one member of the eukaryotic GPCR family, the non-signaling conformation of the bovine rhodopsin bound to 11-cis retinal was known. In contrast to rhodopsin most other
GPCRs are expressed naturally at low levels, bind diffusible high affinity ligands and exhibit a high level of functional and structural variability (Kobilka et al., 2008).

The human β2AR is a class A GPCR activated by epinephrine (adrenaline) and norepinephrine. β2AR is mainly expressed in pulmonary and cardiac tissues and plays an important role in their physiology (Takeda et al., 2002). β blockers, β-adrenergic receptor antagonists, have been used to treat cardiac arrhythmias and hypertension for many years (Kolb et al., 2009; Overington et al., 2006). The β2AR was the first non-rhodopsin GPCR to be cloned (Dixon et al., 1986) and has been studied extensively over the last two decades. Fittingly, the first two non-rhodopsin GPCR structures obtained were both of the human β2AR (Fig 1.16). The first of these structures was of the β2AR bound to the inverse agonist carazolol and in complex with a Fab fragment which binds to ICL3 (Figure 1.16 B) solved at a resolution of 3.4 Å /3.7 Å (Rasmussen et al., 2007). The second is of a β2AR-T4 lysozyme fusion protein (Figure 1.16 A) bound to the inverse agonist carazolol solved at 2.4 Å resolution (Cherezov et al., 2007). In this case the ICL3 was deleted, removing a highly mobile region of the protein and replaced with 160 residues of T4-lysozyme which forms a highly stable compact domain.
Figure 1.15 The three β2 AR structures
The (A) β2AR-T4 lysozyme fusion structure (2RH1), (B) β2AR bound to Fab (2R4R) and (C) β2AR-T4 lysozyme fusion structure showing the position of cholesterol binding sites (3D4S) are shown. The transmembrane helices are coloured using a rainbow spectrum from TM I-TM VII for the β2AR-Fab structure. The heavy (green) and light (yellow) chains of the antibody interact with the ICL3 on the cytoplasmic surface. The ligand carazolol (blue) bound to β2AR-T4 lysozyme fusion protein (grey), timolol (black) and cholesterol (red) bound to β2AR depicting cholesterol binding site (green) is presented. The T4 lysozyme (magenta) is incorporated within the ICL3.

As mentioned previously the biggest problem for structure determination of GPCRs is their inherent instability in detergent solution. Resistance to proteolytic degradation to improve the chances of obtaining well-diffracting crystals of the β2AR was improved by truncating the receptor at position 365 where it aligns approximately with the position of the rhodopsin C-terminus (Rosenbaum et al., 2007). β2AR-T4L was crystallised in LCP, a technique which has had significant success for the structure determination of a number of proteins with 7TMs. LCP provides a more native, lipid environment as compared to other crystallisation techniques (Caffrey, 2003). The receptor was expressed in Sf9 cells, solubilised in 1% DDM and purified by sequential
antibody and ligand affinity chromatography. Saturation and competition binding studies revealed that the modified receptor form had a similar pharmacological profile to that of the WT receptor. The inverse agonist carazolol has a high affinity and slow dissociation kinetics, with both characteristics helping to stabilise $\beta_2$AR against extreme pH and temperature thereby facilitating crystallisation. The final model of $\beta_2$AR-T4 lysozyme included 442 amino acids with well-defined structure from 29-342 amino acids, the N-terminus (1-28) and C-terminus (343-364) were disordered and not visible in the structure (Cherezov et al., 2007).

Within the structure, the helices are well defined with proline-induced kinks present in TM II, V, VI and VII thought to enable structural rearrangement for G-protein effectors. In contrast to rhodopsin and, not predicted by computational analyses of secondary structure, there is a short helical segment in the middle of ECL 2. This short $\alpha$-helix is locked in position by two disulphide bonds and a number of hydrophobic interactions making the ligand binding site accessible to diffusible ligands. In the rhodopsin structure (Palczewski et al., 2000) the N-terminus interacts with the ECL 2 and is clearly defined, in contrast the N-terminus of the $\beta_2$AR is highly disordered which may be attributed to the presence of the Flag tag at the N-terminus. Carazolol is a partial inverse agonist capable of reducing basal activity unlike cis-retinal a full inverse agonist which suppresses all basal activity. This suggests that the $\beta_2$AR-T4 lysozyme may not be completely inactive and thus provide an alternate conformational state of the receptor for homology modelling. Through homology modelling with rhodopsin as the template the structural divergence observed in the crystal structure could not be predicted highlighting the fact that structures of GPCR from different families are needed to provide a more reliable template for homology modelling of receptors.
The ligand binding pocket of β2AR (2RH1) is shown. Unlike rhodopsin the ECL 2 of β2AR has a helix (shown in yellow) stabilised by disulphide linkages (red stick model). The transmembrane helices involved in the binding are shown as grey lettered chains and the bound ligand carazolol as blue stick model. The hydrophobic residues Val 114, Phe 193 and Phe 290 interacting with the bound ligand are shown as green stick models.

The high degree of similarity between the two β2AR structures (Fig 1.15 RMSD = 0.8 Å) demonstrate that radical changes made in the construction of the β2AR-T4 lysozyme fusion did not significantly perturb the conformation of the receptor (Rosenbaum et al., 2007).

A third structure of the same receptor was obtained at a resolution of 2.8 Å by combining the T4 lysozyme fusion protein with a thermostabilising mutation, Glu122Trp (Hanson et al., 2008). In this case, the structure is of β2AR bound to the inverse agonist timolol, which occupies a similar binding site to carazalol as predicted by earlier docking experiments (Topiol et al., 2008). The key structure also revealed important insights into the role of cholesterol in the structural integrity of the protein. Cholesterol was shown to improve the thermal stability of the receptor and gave preference towards a high affinity agonist binding site. A consensus motif for cholesterol binding was identified and has been predicted to be present in 44% of class A GPCRs.
1.7.3 The $\beta_1$ adrenergic receptor structure

Inverse agonists and antagonists which inactivate the $\beta$ adrenergic receptor subtype $\beta_1$ and $\beta_2$ are used for modulating heart function. Drugs acting on $\beta_1$ adrenergic receptor ($\beta_1$AR) are preferred due to fewer side effects as compared to $\beta_2$AR (Warne et al., 2008). Structural determination of the $\beta_1$AR would improve the understanding of specificity of ligands to different receptor subtypes and help in better drug design.

The $\beta_2$AR structures were obtained by modifying the ICL 3 to reduce the dynamic nature of the protein. Presence of a soluble domain and addition of lipids also helped improve the stability of the protein. An alternative approach is that of conformational stabilisation which has to date been applied to $\beta_1$AR (Serrano-Vega et al., 2008), adenosine $A_{2a}$R (Magnani et al., 2008) and neurotensin NTS1 (Shibata et al., 2009). This approach involves alanine scanning mutagenesis in order to identify mutants which increased thermostability and reduced conformational flexibility of the receptor. In the case of the turkey $\beta_1$AR (Serrano-Vega et al, 2008) this process identified six mutations which when combined resulted in a receptor with significantly higher thermostability ($T_m=52^\circ\text{C}$) than the WT ($T_m=32^\circ\text{C}$) and a preference for the antagonist bound conformation. The $\beta_1$AR construct included six thermostabilising mutations along with two additional mutations which increased functional expression and eliminated the palmitoylation site (Serrano-Vega, 2008). The initial 35 amino acids at the N-terminal and 102 residues at the C-terminal were also removed to reduce degradation and maintain a compact structure. The structure is that of the receptor interacting with a high affinity antagonist cyanopindolol (Warne et al., 2008). The resultant mutant was also much more stable in smaller chain detergents than the WT receptor. This approach produced a receptor construct which yielded well-
diffracting crystals in octylthioglucoside (OG) and ultimately led to the high resolution structure at 2.8 Å of the turkey β₁AR (Warne et al, 2008) (Figure 1.17).

![Image of β₁ adrenergic receptor](image)

**Figure 1.17 Structure of the β₁ adrenergic receptor**

The structure of β₁AR (2VT4) bound to the antagonist cyanopindolol is presented. The TM5s for the molecule B in the asymmetric unit are shown in rainbow spectrum with the bound antagonist in a black stick model. ICL3 was not well resolved in the structure.

The β₁AR structure is very similar to the β₂AR structures validating the approach of introducing mutations as a means to producing stable protein suitable for structural studies (Warne et al., 2008). However, the binding sites for the ligands for β₂AR and β₁AR are also very similar making it difficult to explain the ligand selectivity of the different receptor types. The turkey β₁AR structure, unlike those of β₂AR and rhodopsin has a short α-helix in the ICL 2 parallel to the membrane surface. The ICL 2 has been shown to be involved in G-protein activation (Wong et al., 1994). In the β₁AR structure the α-helix of ICL 2 interacts directly with the DRY motif in TM III. The Tyr149 located in ICL 2 interacts with Asp138 in TM III (Figure 1.18). In contrast the corresponding Tyr141 is embedded within the TM III, TM IV and TM VI cavity within the β₂AR structure for both the antibody and T4 lysozyme structures. The β₂AR-T4 lysozyme and β₂AR Fab crystal structure can not
possibly accommodate an $\alpha$-helix in ICL 2 because of the lattice contacts within the adjacent molecules. However in $\beta_1$AR the ICL 2 makes lattice contacts which are different between the four molecules suggesting a physiological role of the structure (Warne et al., 2008).

**Figure 1.18 The $\beta_1$AR cytoplasmic surface structure**

The structure of $\beta_1$AR is presented (2VT4) indicating the $\alpha$-helix structure of the ICL 2. The residues involved in the interaction between the ICL 2 and TM III are shown. The Tyr 149 present on ICL 2 is shown as a black stick model and Asp 138 present on TM III is shown as a blue stick model.

The salt bridge between arginine in TM III and glutamate in TM VI characteristic of an ionic lock is not present in the $\beta_1$AR structure bound to an antagonist. Thus it is likely that an ionic lock is not essential for GPCR inactivation.

### 1.7.4 The adenosine $A_{2a}$ receptor structure

As for the $\beta_2$AR, the adenosine $A_{2a}$R was made as a fusion protein with T4 lysozyme in Sf9 cells and was crystallised using LCP (Jaakola et al., 2008). The structure published just a year after the $\beta_2$AR structure indicated that the GPCR-T4 lysozyme chimera approach could be successfully applied to other GPCRs. The construct had been truncated at Ala316 removing the last 96 residues and hence the majority of the intracellular C-terminus. The ICL3 was truncated and the T4 lysozyme incorporated.
The receptor contains a predicted N-glycosylation site at Asn154 and an enzymatic deglycosylation step was used to remove the sugar group during purification (Jaakola et al., 2008). Stability analysis revealed that the addition of up to 800 mM NaCl during purification increased the Tₘ of the A₂aR-T4 lysozyme construct. The earlier study by Hanson et al, 2008 suggested that there was a specific cholesterol binding site in the adenosine A₂aR (Hanson et al., 2008). The protein was therefore purified in the presence of 500 mM NaCl and cholesterol hemisuccinate. The protein was crystallised in LCP containing monoolein and cholesterol in the presence of the high affinity antagonist ZM241385 (Figure 1.19).

The adenosine A₂aR structure has some significant differences compared to the β-adrenergic and rhodopsin structures. In adenosine A₂aR the ECL2 is comprised of random coil which forms three disulphide linkages with the ECL1 and one within ECL2. This extensive disulphide bond network contributes to a open structure which
exposes the ligand binding pocket to the extracellular environment. In addition, the binding pocket is somewhat different than those found in the rhodopsin and $\beta_2$AR structures with ZM241385 binding in closer proximity to TM VI and VII compared to retinal or carazolol which bind nearer to TM III. ZM241385 binds perpendicular to the membrane plane and is stabilised by $\pi$ stacking interactions with Phe168, hydrophobic interactions with Trp246, Leu 249, His 250, Leu 267, Met 270, Ile 274 and H-bonds with Glu169, Asn253 (Figure 1.20). Several of these residues had previously been identified as having key roles in agonist and/or antagonist binding (Kim et al, 1995). Interestingly, it was not possible to predict the binding position of ZM241385 using the rhodopsin and $\beta$ adrenergic structures (Yuzlenko et al., 2009). This highlights the significant limitations of modeling GPCRs using a small number of structures. Thus it is clear that there is a need for structures of different receptors in complex with a range of different ligands.
Figure 1.20 Ligand binding pocket of adenosine A\textsubscript{2a}R

The ligand binding cavity of adenosine A\textsubscript{2a}R (3EML) bound to antagonist ZM241385 shown as a green stick model is presented. The transmembrane helices involved in ligand binding are shown in grey and ECL 1 and ECL 2 are shown in blue. The residues involved in disulphide linkages are shown as red stick model. The disulphide linkage between Cys 77 and Cys 166 is conserved through Class A GPCRs. Adenosine A\textsubscript{2a}R has an additional disulphide linkage in the ECL 3 between Cys 262 and Cys 259. The residues involved in stabilising the ZM241385 ligand binding are shown as orange stick model. The ligand binding is stabilised through π stacking with Phe 168 and hydrogen bonding with Glu 169 and Asn 253. Hydrophobic interactions with Trp 246, Leu 249, His 250, Leu 267, Met 270 and Ile 274 are also involved. Adapted from Jaakola et al., 2008.

The ICL 2 of adenosine A\textsubscript{2a}R is similar in structure to β\textsubscript{1}AR containing a α-helix and forming interactions between Trp112 and Asp101 in TM III. An ionic lock is thus absent from this structure. The stabilisation of the inactive conformation of the receptor is likely to be due at least in part to the direct interaction of ZM241385 with Trp246, the so-called “toggle-switch” tryptophan proposed to have a key role in controlling switching of the receptor between active and inactive states.

In the case of each of the non-rhodopsin GPCR structures major modifications were required in order to obtain protein which ultimately yielded well-diffracting crystals. In addition, to date only a single conformation of each of the receptors has so
far been obtained. Indeed, while tremendous advances have been made in our understanding of the mechanism of action of these important proteins, many questions remain including the precise molecular details of the interaction of the receptors with their cognate G-proteins. A structure of adenosine A$_{2a}$R bound to an agonist or including ICL3 would facilitate understanding of the signaling mechanism at the molecular level and help improve drug design.

1.8 THE ADENOSINE RECEPTOR

1.8.1 Adenosine and the adenosine receptors

Adenosine is the ubiquitously present natural ligand of the adenosine receptors (Fredholm et al., 2001). Through the four known subtypes of the adenosine receptor, A$_1$, A$_{2a}$, A$_{2b}$, and A$_3$, adenosine has modulatory effects in the central nervous system (Benarroch, 2008) and roles in pain perception (Sawynok et al., 2003), sleep induction (Basheer et al., 2004) and immune regulation (Hasko et al., 2008) among others.

The production of adenosine is not localised and it can be released from any cell type due to metabolic stress or regulation of intracellular adenosine (Fredholm et al., 2007). Equilibrative nucleoside transporters (ENT) maintain a tight and dynamic equilibrium between intracellular and extracellular adenosine concentrations (King et al., 2006). In an event of metabolic stress such as hypoxia, ischemia, trauma or seizures the concentration of extracellular adenosine increases (Boison, 2006; Dunwiddie et al., 2001; Fredholm et al., 2005; Jacobson et al., 2006; Ribeiro, 2005; Ribeiro et al., 2002). Adenosine triphosphate (ATP) is dephosphorylated to adenosine monophosphate (AMP) and subsequently converted to adenosine by 5’ nucleotidase and transported outside the cell via the ENT (Figure 1.21). The rate of energy
utilisation and lack of substrates such as oxygen during hypoxia or ischemia increases the rate of breakdown of intracellular ATP and down regulates adenosine kinase activity within astrocytes (Boison, 2006) which has the effect of increasing intracellular adenosine concentrations. During seizure adenosine may also be produced through extracellular conversion of ATP to adenosine through a cascade of ectonucleases (Zimmermann, 2000).

Figure 1.21 Schematic representation of metabolism of adenosine
Extracellular adenosine is derived through two sources. Adenosine can be transported from the intracellular compartments via equilibrative nucleoside transporter (ENT). Adenosine monophosphate (AMP) is broken down by 5'-nucleotidase (5'NTase) into adenosine. ATP may appear in the extracellular region via release of ATP upon cell damage and death or be transported from the intracellular compartments. Extracellular ATP is metabolised into adenosine via ectonucleotidases (EctoN).

Numerous other agonist and antagonists act via the adenosine receptors. Xanthines such as caffeine and theophylline extracted from natural sources are antagonists of the adenosine receptors and these together with adenosine have provided a template for the design of adenosine analogues. NECA and CGS21680 are non-receptor specific ligands which bind to all adenosine receptor subtypes with varying affinity. Derivatisation of the xanthine antagonists has produced receptor specific high affinity
antagonists (Katritch et al., 2010). The molecule ZM241385 is an artificially synthesised ligand which is a highly selective adenosine A$_{2a}$R antagonist (Figure 1.22). Since the determination of the structure of adenosine A$_{2a}$R the binding pocket can now be studied in greater detail and ligands with higher specificity and selectivity can be designed.

Figure 1.22 Ligands of adenosine A$_{2a}$R used in this study

Adenosine (A) is a natural ligand for all the adenosine receptor subtypes. Agonists NECA (B) has similar affinities for the A$_1$, A$_2a$, and A$_3$ receptor subtypes while CGS21680 (C) has a higher affinity for A$_{2a}$ than A$_{2b}$ receptor. The antagonist theophylline (D) is a naturally occurring compound with similar affinities for all the adenosine receptor subtypes. ZM241385 (E) is a highly selective antagonist for adenosine A$_{2a}$R.

1.8.1.1 The adenosine A$_{2a}$R

The adenosine A$_{2a}$R was first identified and characterised as a GPCR in 1990 (Maenhaut et al., 1990). The adenosine A$_{2a}$R is a Class-A GPCR constituting 412
amino acids. The N-terminus of the adenosine A2aR is shorter than rhodopsin and adrenergic receptors whereas the C-terminus is considerably longer. The majority of the C-terminus is not involved in ligand binding or G-protein coupling (Klinger et al., 2002a; Klinger et al., 2002b) however it appears to recruit other accessory proteins and may have roles in internalization of receptor upon agonist binding (Burgueno et al., 2003).

The activated adenosine A2aR predominantly couples to G\( \alpha_s \) causing stimulation of adenylate cyclase leading to an increase in cAMP concentration (Bruns et al., 1986; Hide et al., 1992). However activated adenosine A2aR may also couple to G\( \alpha_{ol} \) in striatum where it is present in high density (Kull et al., 2000). The adenosine A2aR bears the highest sequence similarity of 59% to adenosine A2bR which also signals via the G\( \alpha_s \). The adenosine A1b and A3 receptor subtype signal via G\( _i \) and G\( _o \) respectively and cause decrease in cAMP concentration.

Like many other GPCRs a vast amount of literature is available emphasising the biomedical relevance of the adenosine A2aR and its importance as a drug target (Jacobson et al., 2006). The adenosine A2aR has been implicated in heart diseases, diabetes, asthma and sepsis among other medical conditions.

The A\( \beta \) peptide is a toxic peptide which accumulates in the brain of patients suffering from Alzheimer’s disease. Blockade of adenosine A2aR may reduce the damage caused by this peptide (Ritchie et al., 2007). During aging the ability of A1 receptor antagonist to inhibit neurotransmitter release and synaptic transmission is reduced while that of adenosine A2aR agonist to facilitate neurotransmission is increased. Understanding of the mechanism of such an unbalanced response could aid in development of modulators to decrease the symptoms of aging disorders (Takahashi et al., 2008).
The immune system has developed the ability to rapidly heal wounds and defend against pathogens in response to threats such as trauma and infection in healthy cells. However, increased immune cell-mediated destruction may result in collateral tissue damage and immunopathologies. It has been shown *in vivo* that adenosine is physically involved in inhibiting activated immune cells and preventing tissue damage. The adenosine A\(_{2a}\)R senses oxygen depletion via extracellular adenosine accumulation in damaged tissue. Intracellular cAMP generated by adenosine A\(_{2a}\)R then inhibits activated immune cells and prevents tissue damage (Sitkovsky *et al.*, 2004).

Adenosine A\(_{2a}\)R activation also has a suggested role in asthmatic airway inflammation. Elevated adenosine levels are observed during asthmatic attack. The adenosine liberated in the airway may act on the different subtypes of adenosine receptor and cause airway inflammation. Studies in rats have shown that adenosine A\(_{2a}\)R suppresses activation of T-lymphocyte, degranulation of human mast cells and secretion of interleukin-12 and suppresses T-cell effector function (Luijk *et al.*, 2008). Theophylline, an antagonist for the adenosine A\(_{2a}\)R, is widely used in treatment of asthma.

Various agonists and antagonists of the adenosine receptor have been developed as potential drug targets. The adenosine A\(_{2a}\)R thus remains an attractive target for drug development and structural studies.
**Aims**

GPCRs constitute the largest family of proteins in the human genome and mediate cellular responses to a wide range of bioactive molecules including hormones, neurotransmitters and the vast majority of currently available drugs. High resolution structural studies of multiple receptors in different conformations is essential in order to gain precise insights into their mechanisms of action as well as provide better models for rational drug design. The adenosine A_2aR is a GPCR implicated in immune regulation and neurological disorders. The last two years have seen a mini-explosion of GPCR structures obtained by a variety of means. The underlying principle in each case was to stabilise the receptor in order to facilitate crystallisation. In each case the receptors were significantly altered by truncation, mutation, deletion/insertion or combinations of these. In addition only a single conformation of each receptor has so far been obtained. Thus there remain significant gaps in our understanding of the molecular mechanisms of GPCRs. In addition, all the structures obtained thus far are from receptor produced in Sf9 cells. However, it would be useful to explore whether other systems could be used. Thus the initial aim of this study was to express a human adenosine A_2aR construct with minimal sequence modifications in the *P. pastoris* heterologous expression system. A further aim of the study was to functionally and biophysically characterise the adenosine A_2aR constructs during the isolation process in order to identify constructs and conditions which yielded the most active and thermostable receptor. We also aimed to explore the use of G-protein peptides as tools to facilitate crystallisation of a receptor construct incorporating ICL3. Finally we aimed to submit our highly pure, highly stable receptor to crystallisation trials using a variety of techniques.
2 Expression and characterisation of membrane-bound adenosine $A_{2a}$ receptor constructs

Part of the work presented in this chapter is published in:

Microbial cell factories 2008 Oct 10;7:28

Large-scale functional expression of WT and truncated human adenosine $A_{2A}$ receptor in *Pichia pastoris* bioreactor cultures

Shweta Singh, Adrien Gras, Cédric Fiez-Vandal, Jonathan Ruprecht, Rohini Rana, Magdalena Martínez, Philip Strange, Renaud Wagner, Bernadette Byrne
2.1 INTRODUCTION

The large-scale production of G-protein coupled receptors (GPCRs) for functional and structural studies has until recently been a challenge mainly due to low endogenous expression levels. With the exception of rhodopsin, all GPCRs are expressed at very low levels endogenously, thus requiring the development of recombinant over expression systems. Recent studies have screened both expression systems and receptors in order to optimise production levels (Griesshammer, 2006; Sarramegna et al., 2006; Sarramegna et al., 2003; Tate et al., 1996), which has led to *Pichia pastoris* becoming a well characterised system for protein production (Clare et al., 1998; Gleeson et al., 1998; Reilander et al., 1998; Stratton et al., 1998). *P. pastoris* has advantages over other systems for large scale expression in that it is comparatively inexpensive and easy to grow to high cell densities making the volumes required for large scale protein expression easier to process. It is preferred over prokaryotic expression systems for producing functional receptor mainly due to the similarity in translational, folding and insertion machinery. *P. pastoris* has the ability to glycosylate expressed receptors, albeit in a modified form compared to higher eukaryotes, a process essential for the proper functioning and membrane targeting of many receptors (Reilander et al., 1998). *P. pastoris* is comparatively easy to work with in terms of genetic manipulation and readily adapts to high density fermentation cultures growth in the bio-reactor. Recombinant protein production is driven by the strong, tightly regulated, methanol inducible alcohol oxidase (AOX1) promoter. The yeast can utilise methanol as its sole carbon source and grow to high cell density thus making the process cheap and highly reproducible (Fraser, 2006; Hollenberg et al., 1997; Sarramegna et al., 2002; Weiss et al., 1995; Wiener, 2004). In addition, there are
several examples of membrane proteins expressed heterologously in *P. pastoris* which have yielded high resolution structures including plant aquaporin from *Arabidopsis thaliana* (Verdoucq et al., 2008), the human leukotrienes C4 synthase to 2 Å (Martinez Molina et al., 2007), both ligand free and drug bound P-glycoprotein structure (Aller et al., 2009) and the rat membrane protein K+ channel (Long et al., 2005). High resolution structures of these membrane proteins demonstrate that *P. pastoris* is indeed a valuable tool for production of protein for structural work.

Considerable effort has been made in establishing a *P. pastoris* expression system for GPCRs for the production of functional receptor at sufficiently high levels for structural studies. The basic system uses a pPIC9K vector (Invitrogen) where GPCR expression is under the control of the AOX1 promoter. Protease deficient expression strains, such as the SMD1163 strain, and the use of the α-factor leader sequence have improved receptor expression levels (Weiss et al., 1995). Modifications to the growth media including addition of histidine, receptor specific ligands and dimethyl sulfoxide (Andre et al., 2006), which facilitates phospholipid biosynthesis and membrane proliferation in yeast (Murata et al., 2003) have been shown to increase the expression levels of 20 different GPCRs in *P. pastoris* (Andre et al., 2006). This system has been successfully used to express a number of GPCRs including the mouse 5-HT3A serotonin receptor (Weiss et al., 1995), the endothelin ETB receptor (Schiller et al., 2000) and dopamine D2 receptor (de Jong et al., 2004). Different GPCRs express at different levels under similar conditions. Despite the vast amount of work done to optimise expression conditions for GPCRs it has proved necessary to establish specific expression conditions for each receptor. One particular issue is the amount of methanol added for induction as very high levels of methanol can induce cytotoxic effects which reduce cell viability and thus expression (Sinha et al., 2003). Methanol
sensors, which detect the level of unmetabolised methanol have been key to reducing these cytotoxic effects and improving the reproducibility of the cultures. Another important issue to consider is the osmotic stress induced during high cell density culturing which is known to be responsible for adaptative cell response mechanisms, such as changes in the membrane lipid content (Mattanovich et al., 2004). This may not be desirable in the process of membrane protein production hence regulating the biomass through medium cell density culturing and determining the length of the induction period appear essential.

A number of studies have previously shown heterologous expression of the adenosine A\textsubscript{2a}R in a range of expression hosts including \textit{E. coli} (Weiss et al., 2002), \textit{P. pastoris} (Fraser, 2006), \textit{S. cerevisae} (O’Malley et al., 2007) and \textit{Sf}9 cells (Jaakola et al., 2008). The adenosine A\textsubscript{2a}R has been reported to undergo C-terminal degradation (Weiss et al., 2002). Recent studies have shown the importance of removal of the unstructured C-terminal tail, among other modifications, in obtaining well diffracting crystals and ultimately structure determination of GPCRs (Cherezov et al., 2007; Jaakola et al., 2008; Warne et al., 2008). The adenosine A\textsubscript{2a}R has a particularly long C-terminal tail, comprised of approximately 96 amino acid residues (Jaakola et al., 2008), likely to reduce the chances of obtaining well diffracting crystals. Constructs were generated that are truncated at the C-terminal domain of the adenosine A\textsubscript{2a}R in order to prevent this C-terminal degradation.

Several studies have described the expression of GPCRs in \textit{P. pastoris} using shaker flasks, which allow culturing of small volumes (500 ml) with moderate cell densities (\textit{OD}\textsubscript{600} ~15). A small scale expression trial is useful for determining the best expression conditions. Once that has been established expression in a bioreactor is preferred because structural studies require large amounts of protein. However in
most cases scale up in the bioreactor requires some optimisation of both time required to reach optimal OD$_{600}$ for induction and methanol induction time period for each individual receptor. It is also essential to maintain reproducibility between different expression runs and to have a stringent control on all the parameters involved.

The aims of the study were threefold; we wanted to compare the levels of expression through a protocol developed in the lab for the human adenosine A$_{2a}$R in \textit{P. pastoris} under control of a methanol-inducible promoter in both flask and bioreactor cultures. The study also involved a detailed functional analysis of the human adenosine A$_{2a}$R produced in \textit{P. pastoris} with a view to assessing the suitability of a protein expressed in a recombinant system for structural studies. The functional analysis enabled us to assess the integrity of the receptor at each stage of the production procedure. We also wished to assess the effects of generating a C-terminal truncation on the functional expression, pharmacological profile and stability of the receptor.
2.2 MATERIALS AND METHODS

2.2.1 Materials

Yeast nitrogen base and yeast extract were purchased from Difco UK. Peptone, L-histidine and mouse M2 anti-FLAG antibody were from Sigma-Aldrich Gillingham UK and dimethyl sulfoxide (DMSO) was from Acros Organics Leicestershire UK. Complete EDTA-free protease inhibitor cocktail tablets were purchased from Roche Welwyn UK and the bicinchoninic acid assay (BCA) kit was purchased from Thermo-Fisher Leicestershire UK. Scintillation cocktail (Ultima Gold MV) was obtained from PerkinElmer USA. [\(^3\)H] ZM241385 was obtained from American Radiolabelled Chemicals Inc. USA while ZM241385 was obtained from Tocris Bristol UK and theophylline from Fisher Scientific. Nitrocellulose membrane was obtained from Millipore and GF/B filters were from Whatman-GE Buckinghamshire UK. The sheep anti-mouse IgG-horseradish peroxidase (HRP) conjugate from GE Healthcare Buckinghamshire UK. All other chemicals were obtained from Sigma-Aldrich Gillingham UK.

2.2.2 Expression Constructs

The WT adenosine A\(_{2a}\)R was expressed using a modified pPIC9K (Invitrogen) vector incorporating a Flag tag followed by a 10 His tag and a TEV cleavage site all upstream of the gene coding for the receptor as described by (Andre et al., 2006). This vector was further modified to remove the Bio-tag located downstream of the gene coding for the receptor.

The Flag tag is an octapeptide sequence recognised by commercially available antibodies. The epitope tag has the following sequence N-DYKDDDDK-C. This tag
is followed by ten consecutive histidine codons and the coding region for the tobacco etch virus protease cleavage site. The Bio-tag is the coding region for the biotinylation domain of the transcarboxylase from *Propionibacterium shermanii* (Figure 2.1).

**Figure 2.1 Schematic representation of the adenosine A$_{2a}$R constructs used in this study**

(A) The AOX1 promoter (arrow) is followed by the signaling peptide from yeast alpha factor (grey; αF). The two tags used for purification and identification are an octa peptide Flag tag (yellow FLAG) and a deca His tag (orange; HIS 10). A TEV protease cleavage sequence (green; TEV) is utilised as the final polishing step to separate all the tags from the receptor. The WT adenosine A$_{2a}$R (blue) form is 412 aa long while the mutants have been truncated at alanine 316 (lilac) and valine 334 (plum). (B) The stars in the schematic representation of the receptor indicate the approximate truncation points on the C-terminus.

The forward primer 5’-GGT GGA TCC ATG CCC ATC ATG GGC TCC TCG GTG TAG-3’ together with the reverse primer either 5’-CAT GGA ATT CAC TAG TTG CCT TGA AAG GTT CTT GCT GCC TCA GGA CGT GGC TG-3’ introducing a stop codon after alanine 316 or 5’-CAT GGA ATT CAC TAG TGA CCT GCT CTC CGT CAC TGC CAT GAG CTG CCA AG-3’ introducing a stop codon after residue valine 334 was used to amplify the truncated adenosine A$_{2a}$R gene fragments. The Expand long Template PCR system (Roche) was used for amplification of gene. The reaction consisted of 500µM dNTPs, 300nM each of forward and reverse primer, buffer containing MgCl$_2$, 100ng template DNA and 0.5µl expand long polymerase. The cycling conditions involved an initial denaturation step of 2 minutes at 95°C followed by 35 cycles of denaturation at 95°C for 10 seconds, annealing at 55°C for 30 seconds and extension at 68°C for 2 minutes followed by a
final elongation at 72°C for 5 minutes. The PCR sample was run on a gel and desired band extracted using the gel filtration kit from Qiagen. These amplified gene fragments were sequenced and cloned into the BamHI and SpeI restriction sites of the modified pPIC9K (Invitrogen) vector. Expression plasmids were transformed by electroporation into SMD1163 P. pastoris cells. Colonies containing the gene were selected on Yeast extract Peptone Dextrose (YPD) plates [1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) dextrose, 2% (w/v) agar] containing 0.1 mg/ml geneticin.

**Figure 2.2 Sequence alignment of adenosine A2aR and bovine Rhodopsin**

The amino acid sequence highlighted in grey are the helices for rhodopsin H-I(35 to 64), H-II(107-139), H-IV(151-173), H-V(200-225), H-VI(247-277), H-VII(286-306) (Palczewski et al., 2000). The residue highlighted in blue (Alanine) marks the beginning for C-terminal truncation for A-316 constructs, residue highlighted in red (Valine) marks the beginning for C-terminal truncation for V-334 construct.
2.2.3 Small Scale cultures

Single *P. pastoris* colonies containing the plasmid of interest were selected on YPD plates containing 0.1 mg/ml geneticin. Cells from a single colony were used to inoculate 5 ml of Buffered Glycerol complex media (BMGY) [100 mM potassium phosphate pH 6.0, 1% (w/v) yeast extract, 2% (w/v) peptone, 1.34% (w/v) yeast nitrogen base without amino acids, 0.00004% (w/v) biotin, 1% (w/v) glycerol]. The culture was grown overnight at 30°C to an OD$_{600}$ of 12-15. The cells were centrifuged at 3,000 g for 5 min and the cell pellet was typically resuspended in 10 ml Buffered Methanol Complex Media (BMMY) [similar to BMGY with the following changes: phosphate buffer at pH 8.0, 2.5% (v/v) dimethyl sulfoxide, 0.04% (w/v) histidine, and 0.5% (v/v) methanol instead of 1% glycerol] to achieve a starting OD$_{600}$ of 5 for all cultures. The culture was incubated for 18 h at 22°C and then the cells harvested by centrifugation at 5000 g for 5 mins. The highest expressing clones were selected by Western blot for further study. In case of time course experiments samples were taken at regular hours during induction.

2.2.4 Bioreactor cultures

Single *P. pastoris* colonies containing the plasmid of interest were selected on YPD plates containing 0.1 mg/ml geneticin. Cells from a single colony were used to inoculate a 150 ml starter culture of MGY medium (0.1M potassium phosphate pH 6, 1.34% yeast nitrogen base, 0.0004 % biotin, 1% glycerol). The culture was grown at 30°C with aeration for 20-22 hours to an OD$_{600}$ ~10-15. The entire inoculum was added to a 5 L fermenter vessel at t=0. Cultures were grown in a 5 L bioreactor vessel controlled by an ADI 1010 bio-controller connected to a PC running BioExpert software (all from Applikon Biotechnology). A methanol sensor (Raven Biotech) was used to monitor and control methanol levels in the vessel. The internal regulation
software of the sensor was used to control the external peristaltic pump and adjust methanol concentration to the set point. Agitation was set at 1000 rpm, pH at 5.0 and dO$_2$ at 35% for the entire run. For the initial cell growth phase the temperature was set at 30°C. In this initial phase, the culture was allowed to consume all the glycerol contained in the 4 L of FM22 culture medium (4.3% (w/v) monobasic potassium phosphate, 0.5% (w/v) ammonium sulphate, 0.1% (w/v) calcium sulphate, 1.43% (w/v) potassium sulphate, 1.17% (w/v) magnesium sulphate) containing 2% (w/v) glycerol and 1 ml/L PMT4 trace elements solution (0.2% (w/v) copper sulphate, 0.008% (w/v) sodium iodide, 0.3% (w/v) manganese sulphate, 0.02% (w/v) sodium molybdate, 0.002% (w/v) boric acid, 0.05% (w/v) calcium sulphate, 0.05% (w/v) cobalt chloride, 0.7% (w/v) zinc chloride, 2.2% (w/v) iron sulphate, 0.02% (w/v) biotin, 1 ml/L sulphuric acid). During a second growth phase the culture was fed 50% glycerol at a rate of 3.5 ml/L/h until the OD$_{600}$ reached approximately 75 (~3-4h). The temperature of the vessel was then reduced to 22°C and the cells treated with a 0.1% v/v aliquot of 100% methanol supplemented with 4 ml/L PMT4, 3% DMSO, 0.04% histidine and the methanol sensor signal allowed to stabilise. The signal output was recorded in mV and the methanol feeding system set point was set to this value. The culture was allowed to adapt to 0.1% methanol for 1 hour. For the induction stage, the methanol concentration was raised to 0.3% v/v in steps of 0.1% v/v every 2 to 3 hours. At the end of the induction period (~18 h) the cells were harvested by centrifugation at 6,000 g for 20 min, the temperature was maintained at 4°C. To monitor the induction procedure through time course experiments samples (1 ml) were taken at various time points during the culture (6, 4 and 2 hours pre induction and 0, 3, 6, 9, 12, 15, 18, and 21 hours post induction). The cells were harvested and stored at -80°C for further analysis.
2.2.5 Small scale membrane preparations

Cells were resuspended in ice-cold breaking buffer (50 mM HEPES-NaOH pH 7.4, 100 mM NaCl, 10% (w/v) glycerol, 2 mM EDTA, (1/100 ml) protease inhibitor PI). The cell suspension along with glass beads was transferred to 2 ml safe lock eppendorf tubes and broken using a tissue lyser (Qiagen) set to 30 MHz for 15 mins. The beads were spun down by centrifugation at 1500 g for 1 min. Cell debris, including intact cells, were removed by a low speed spin (3,000 g) for 10 min. The supernatant was retained, and membranes isolated by centrifugation at 100,000 g for 30 min. Each membrane pellet was resuspended in membrane buffer (50 mM HEPES-NaOH pH 7.4, 100 mM NaCl, 10% (w/v) glycerol) and flash frozen in liquid nitrogen for storage.

2.2.6 Large scale membrane preparations

Cell pellet was resuspended in five volumes of buffer. Typically 25 grams of cells from a fermentor culture were resuspended in 125 ml of ice-cold breaking buffer. The mixture was then passed three times through a constant cell disruptor (Constant Systems) at 40 kpsi. Cell debris including intact cells, if any, were removed by a low spin at 6000 g for 20 minutes. The supernatant was retained and membranes were isolated by centrifuging at 170,000 g for 1 hour using 45ti rotor in an ultra XP centrifuge (Beckmann Coulter). The membranes were resuspended in 10 ml of membrane buffer. The membrane was homogenised and either divided in 300 µl aliquots and stored in 0.5 ml microcentrifuge tube at -80°C or used immediately for other downstream processing steps. All steps were carried out at 4°C.
2.2.7 Protein concentration determination

All protein concentrations were determined using bovine serum albumin as standard and based on the protocol developed by Lowry et al (1951). Protein standard at concentration ranging from 10-100 µg were added to tubes in triplicates. Protein samples at varying concentration were also added to tubes. 1 ml of 10% Trichloro acetic acid (TCA) was added to all tubes and vortexed immediately. The tubes were centrifuged for 15 minutes at 2200g. The supernatant was discarded and the precipitate was allowed to dry. 1 ml of solution containing Solution A (2% anhydrous Na₂CO₃, 0.1M NaOH) and Solution B (0.5% CuSO₄.5H₂O, 1% NAK tatrante) at a ratio of 50:1 v/v was added to all tubes, vortexed immediately and incubated at room temperature for 10 minutes. 100 µl of Folin-Ciocalteau reagent diluted with an equal volume of water was added to each tube vortexed immediately and incubated at room temperature for 10 minutes. A volume of 2 ml of distilled water was added to each tube and incubated for 20 minutes. Absorbance was measured at 760 nm on a spectrophotometer. Data were analysed using GraphPad prism software.

2.2.8 SDS-PAGE and Western blotting analysis

The proteins were separated on a 4-12% Bis-Tris NuPAGE gel prior to transfer to a PVDF membrane using the X-Cell system. The membrane was blocked with 3% w/v skimmed milk powder in phosphate buffer saline (PBS) at 4°C for 12-14 hours. The membrane was washed three times with PBS containing 0.03% Tween20. The blot was probed for 2 hours with 1:10,000 dilution of primary anti-FLAG M2 antibody. The unbound antibody was washed with PBS containing 0.03% Tween 20 followed by a 2 hour incubation with 1:10,000 dilution of secondary goat anti-mouse horseradish peroxidase antibody. The incubations with antibody were carried out at
room temperature. Protein bands were detected through chemiluminescence using ECL substrate.

2.2.9 Saturation radioligand binding assays

Saturation radioligand binding assays were carried out to determine the expression level and functionality of the adenosine A$_{2a}$R. Saturation radioligand binding experiments measure the specific radioligand binding at equilibrium at various concentration of the radioligand. The receptor is incubated with increasing concentrations of radioligand and bound ligand is determined. With an increasing radioligand concentration all the available receptor binding sites are occupied and hence the curve starts to plateau after a level. Membranes expressing WT, A316 and V334 adenosine A$_{2a}$R were incubated with its antagonist $[^{3}H]$ ZM241385 (0.036-4 nM) in binding buffer (20mM HEPES pH 7.4, 1mM EDTA, 1 mM EGTA, 0.1% BSA). Non-specific binding was defined in the presence of 30 µM ZM241385 at each reaction concentration. Total binding was defined in the absence of the unlabelled antagonist. Reactions were performed in triplicates in a final volume of 1000 µl containing radioligand and unlabelled competing ligand where needed. Each reaction concentration contained 2 µg of membrane protein. The reactions were incubated for 3 hours at 25° C following incubation receptor bound radioligand was collected by filtration through Whatman glass micro-fibre GF/C filters using a Brandel harvester. The filters were washed with 5 ml of ice-cold phosphate buffered saline (137mM NaCl, 2.7mM KCl, 10mM Na$_2$HPO$_4$, 2mM KH$_2$PO$_4$, pH 7.4). The filter discs were cut out and placed in scintillation vials along with 2 ml of scintillation fluids. The filters were allowed to soak for at least 6 hours before the amount of bound radiolabel was determined using a LS 6500 scintillation counter (Beckman Coulter).
2.2.10 Data Analysis: Saturation binding assay

The assay produced a set of values as disintegration per minute (DPM) for the total and non-specific reactions. Data was analysed using the Graphpad prism using non-linear regression. The specific [³H] ZM241385 bound was calculated as the difference between bound ligand in the total and non-specific reaction i.e. in the absence and presence respectively of the antagonist for each radioligand concentration. The specific binding so obtained in DPM was converted to the concentration of radioligand bound (pmol) by using the specific activity of the [³H] ZM241385 used in the assay. The free radioligand concentration for each reaction was calculated by subtracting the total [³H] ZM241385 bound in each reaction from the DPM obtained for the radioligand in the absence of protein. The average specific binding for each radioligand concentration was then plotted against the free radioligand added and the data analysed by non-linear regression and by fitting the equation for either one-site or two-site binding model. The best fit was determined by an F test. The two parameters obtained were the equilibrium dissociation constant (Kd) and Bmax. The Bmax corresponds to the maximum number of receptor sites occupied by the ligand therefore represents the number of receptor binding sites. The specific activity or Bmax values are presented as pmol of receptor per mg of membrane protein. Kd denotes the concentration of ligand required to occupy half of the receptors binding sites hence the affinity of the ligand to the receptor.

2.2.11 Competition radioligand binding assays

Agonist and inverse agonist profiles for binding to the adenosine A₂aR was determined by competition radioligand binding assays against [³H] ZM241385. In each reaction condition 2 µg of membrane protein expressing WT, A316 or V334 adenosine A₂aR were incubated with a fixed concentration of [³H] ZM241385 (1 nM) and varying
concentration of competing ligand ranging (from $10^{-4}$ to $10^{-10.5}$ M for NECA, $10^{-3}$ to $10^{-9.5}$ M for theophylline) in binding buffer containing 10 mM MgCl$_2$. Non-specific binding was defined in the presence of ZM241385 (3 µM) and absence of competing ligand. Reactions were performed in triplicates in a final volume of 1 ml. Initiation, incubation and termination procedure were similar to as described for the saturation binding assay.

2.2.12 Data analysis: Competition radioligand binding assay

Specific binding was calculated as the difference between total and non-specific binding. Binding in the presence of competitive ligand at different concentrations was calculated as a percentage of the maximum specific binding. Non-specific binding was subtracted from each condition. The data were analysed by non-linear regression and fitted to one site or two site binding equation and the best fit determined by an F test. The best fit equation was used to generate the IC$_{50}$ value. IC$_{50}$ represents the concentration of competing ligand in the reaction required to inhibit 50% of the specific binding for the radioligand. The IC$_{50}$ values were used to calculate the $K_i$ values by the Cheng and Prusoff equation (Cheng et al., 1973).

$$K_i = \frac{IC50}{1 + \left(\frac{L}{Kd}\right)}$$

$L$ = Concentration of radiolabelled ligand

$K_d$ = dissociation constant of the radiolabelled ligand binding to the receptor
Binding assay parameters such as $K_d$ and $K_i$ are presented as negative log of $K_d$ ($pK_d$) or negative log of $K_i$ ($pK_i$). The concentrations of competing ligand are equally spaced on a log scale. Hence values computed by Prism such as $IC_{50}$ and its SE and 95% CI are symmetrical on the log scale. These turn asymmetrical when converted to a linear scale by taking an antilog. Thus to find a mean from several experiments it is better to average the log of $K_i$ ($pK_i$) rather than the $K_i$. 
2.3 RESULTS

2.3.1 Bioreactor culture allows controlled growth

The bioreactor allows tight control of the growth parameters improving reproducibility of the culture. Factors such as dissolved oxygen (dO$_2$) response on addition of methanol, pH, temperature and foaming of the culture were monitored throughout the run.

Figure 2.3 shows the changes in two parameters: dO$_2$ and OD$_{600}$ during a sample bioreactor culture for both WT (A) and V334 (B) adenosine A$_2a$R constructs. The initial glycerol batch phase (phase 1) showed a gradual decrease in dO$_2$ levels followed by a stabilisation at 35%, the configured set point. As the cells multiply and consume glycerol, the culture’s requirement for oxygen increased, lowering the observed dO$_2$ in the vessel. Once the 35% set point was reached, the system maintained dO$_2$ levels by modulating the airflow. This phase lasted approximately 22 hours with minor differences between runs depending on the OD$_{600}$ of the starting inoculum. A sharp rise in dO$_2$ after ~19 h cultivation indicated that the cells were starving, indicating that the glycerol within the media had been completely consumed.

The glycerol fed-batch phase (phase 2) was then initiated. The culture was fed 50% glycerol at 3.5 ml/L/hour until the OD$_{600}$ reached approximately 75. This caused dO$_2$ levels to drop and stabilise again at 35%. This phase lasted between 3 and 4 hours. The culture was then allowed to consume any remaining glycerol before the start of the induction phase. The temperature was reduced from 30°C to 22°C and DMSO and histidine added before induction. Protein expression was induced by the addition of methanol. An adaptation period of 3-5 hours was required during which
the cells alter their metabolism to accept methanol as the new carbon source. This was observed as a steady decrease in dO₂ levels as the cells slowly adapted to methanol utilisation. During this phase the OD₆₀₀ remained relatively stable at about 75 for the WT adenosine A₂aR construct while typically a small increase was observed for the V334 adenosine A₂aR (from 75 to ~90). The highly similar dO₂ and OD₆₀₀ traces for the two different constructs demonstrate the reproducibility of the bioreactor culture protocol described in this study. A four litre bioreactor culture generated 300-320 grams of biomass. The final culture of adenosine A₂aR had an OD₆₀₀ typically ranging between 75-90.
2.3.2 Comparison of Flask and Bioreactor cell growth

Samples were collected at regular time points during the glycerol fed batch stage when the cells are growing in number by utilising glycerol as an energy source and after the conditions had been changed for induction. Expression of the two different receptor forms was monitored through radioligand binding assays and Western blots.
Western blot analysis of the flask cultures expressing WT adenosine A\textsubscript{2a}R showed a gradual increase in expression over the course of the induction phase with receptor protein first detectable at 14 hours post induction (Figure 2.4) as a band of approximately 38 kDa. Due to the large gap in sampling times it is likely that receptor is expressed before 14 hours, however it is clear that no receptor is expressed at 4 hours post induction. The amount of WT adenosine A\textsubscript{2a}R did not seem to significantly increase after this. The V334 adenosine A\textsubscript{2a}R expressed in flasks was first detected as a band of approximately 28 kDa at a similar time point however the amount of protein increased significantly after this to the 20 hour time point when the culture was harvested. Expression of the WT and V334 adenosine A\textsubscript{2a}R in the bioreactor also increased with time although the expression was detectable by Western blot at 3 hours. There were some low molecular weight bands visible for the WT adenosine A\textsubscript{2a}R, indicating that there is proteolytic degradation of the expressed receptor. The fact that these bands are being detected using the N-terminally positioned Flag tag suggests that the bands observed are due to C-terminal degradation. This is not observed in the flask culture for either construct, or the bioreactor culture expressing the V334 adenosine A\textsubscript{2a}R.

Interestingly there are distinctive higher molecular weight bands at 45 and 62 kDa visible for the WT and V334 adenosine A\textsubscript{2a}R samples from the bioreactor. The various species of protein that appeared on the gel were also visible on a Western blot conforming that they were the adenosine A\textsubscript{2a}R however they could not be fully characterised at this stage. Further work to characterise the different populations of receptor was carried out on purified receptor and is described in Chapter 3.
Figure 2.4 Expression profile of adenosine $A_{2a}$R construct over time

Western blots for WT (A), V334 (B) expressed in flask, WT (C) and V334 (D) expressed in bioreactor is presented. Membranes were prepared from samples taken at regular intervals during induction. The M2 antibody selective to flag tag located at the N-terminal was used. The molecular weight markers are labeled on the left and time in hours on the top of the blot. The band corresponding to the monomeric adenosine $A_{2a}$R is denoted by WT (*) and V334 (**).

Radioligand binding analysis indicated a similar trend in increasing expression of functional receptor from the flask cultures to that shown by Western blot analysis. There was no measurable functional receptor for the first 4 hours post induction (Figure 2.5) while after this time functional expression increased rapidly, to a maximum of 23 pmol/mg after 14 hours. The V334 adenosine $A_{2a}$R expresses to a much higher level after 14 hours (53 pmol/mg) and increases significantly after that to a final expression level of 72 pmol/mg after 20 hours.

The expression profile is slightly different in the case of the bioreactor cultures. Derepression of the methanol promoter due to depletion of the glycerol during the second growth phase, induces expression from the AOX1 promoter resulting in low level but measurable receptor production prior to induction. For both the WT and the V334 adenosine $A_{2a}$R constructs there is a steady increase in expression up to 6 hours.
post induction. After this time the expression of the WT adenosine A$_{2a}$R continues to increase steadily peaking at approximately 125 pmol /mg at the end of the culture (18 h) while V334 adenosine A$_{2a}$R increases much faster peaking at 250 pmol/mg at 18 h. The culture for the V334 adenosine A$_{2a}$R construct was continued for a further 3 hours however the data suggests that this results in a slight drop in functional expression. It was concluded that 18 h post-induction is the optimal time for harvesting the bioreactor cultures.

Figure 2.5 Functional expression of adenosine A$_{2a}$R constructs over time.
Expression of WT (■) and V334 (▲) adenosine A$_{2a}$R in flask culture (A) and bioreactor culture (B) is depicted. Regular time points were taken before and during the methanol induction phase. Radioligand binding assays at saturating concentrations of [H] ZM241385 (4nM) were done to estimate the functional receptor present at each stage. The data is representative of three independent experiments.
2.3.3 Comparison of expression levels of the three adenosine A2aR forms in a bioreactor

Saturation radioligand binding assays were performed on membrane bound WT, A316 and V334 adenosine A2aR in order to fully characterise the binding kinetics of the WT and C-terminal truncated forms of the receptor following optimised bioreactor expression. The results of three typical saturation binding assays are shown in Figure 2.6.

Saturation radioligand binding analysis for the WT and the C-terminal truncated adenosine A2aR forms revealed a similar pKd values indicating that removal of the C-terminus did not affect ligand binding activity. The functional expression levels were significantly higher for the C-terminal truncated constructs in comparison to the WT receptor form (Figure 2.6; Table 2.1). The Bmax values achieved for all three receptor constructs were high giving estimated yields of 11 mg/L for V334 adenosine A2aR (Table 2.1) which is more than adequate for further studies.

![Figure 2.6 Saturation analysis of [3H]ZM241385 to the membrane bound adenosine A2aR](image)

Saturation curve representative of three independent experiments for the WT (■) with (A) A316 (●) and (B) V334 (▲) to membrane bound adenosine A2aR is presented. Bmax and Kd values were derived by non-linear regression and mean values ± SEM are presented in Table 2.1
Table 2.1 Estimate of receptor yield in a bioreactor culture through saturation assay

Table below provides an estimate of the amount of receptor obtained from a litre of bioreactor culture. $B_{\text{max}}$ and $K_d \pm \text{SEM}$ from three independent experiments are also presented. Membranes were prepared from a bioreactor culture and measured by three independent experiments each containing triplicates.

<table>
<thead>
<tr>
<th>RECEPTOR</th>
<th>$B_{\text{max}} \pm \text{SEM}$ (pmol/mg)</th>
<th>$pK_d \pm \text{SEM}$ ($K_d$ nM)</th>
<th>RECEPTOR (mg/l)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>124.2 ± 25.6</td>
<td>8.9 ± 0.1</td>
<td>4.3</td>
</tr>
<tr>
<td>A316</td>
<td>179.5 ± 16.4</td>
<td>8.7 ± 0.1</td>
<td>9</td>
</tr>
<tr>
<td>V334</td>
<td>222.2 ± 35.6</td>
<td>8.4 ± 0.2</td>
<td>11.1</td>
</tr>
</tbody>
</table>

\* Typically 100gm of membrane containing 4.5-5 g of protein is obtained from a 4-litre bioreactor culture. The molecular weight (kDa) is WT- 44, A316- 35, V334-37.

2.3.4 Pharmacological characterisation of receptor forms expressed in P. pastoris

Competition radioligand binding experiments using an agonist NECA and an antagonist theophylline were also performed for the three receptor forms. Competition curves for the ligands ZM241385, NECA and theophylline are presented for the WT adenosine A$_{2a}$R (Figure 2.7). The top of the curve represents a value equal to radioligand binding in the absence of the competing unlabeled drug. The bottom of the curve is equal to non-specific binding. With a decrease in the concentration of competing ligand the binding moves towards total binding i.e. 100% of the available binding sites are occupied by the radioligand. The ligands used in the study were chosen as they had affinities that were appreciably different from one another and allowed comparison of the binding over a range of $K_d$ values. The pK$_i$ values obtained (Table 2.2) indicated that NECA binds the receptor with higher affinity than
theophylline but lower affinity than ZM241385 (Table 2.1) and is similar to that reported by other workers (Table 2.3). The pKᵢ values obtained for all these ligands demonstrated that neither expression in a heterologus system nor truncations of the C-terminus affects the receptor’s ligand binding ability. The lower end of the curve was utilised to judge the suitability of theophylline as a ligand to define non-specific binding and the amount of ligand required. Similar level of inhibition in binding is achieved for a micromolar concentration of ZM241385 and millimolar concentrations of theophylline which was used henceforth for defining non-specific binding.

![Figure 2.7 Competition radioligand binding profile for the WT adenosine A₂aR form](image)

**Figure 2.7 Competition radioligand binding profile for the WT adenosine A₂aR form**

A typical curve for the WT adenosine A₂aR form for binding of theophylline (●), NECA (■) and ZM241385 (▲) in competition versus [³H] ZM241385 binding is presented. Similar curves were obtained for the C-terminal truncated constructs through independent experiments done under similar conditions presented in Table 2.2. Data were best fitted to a one binding site model.
Table 2.2 Competition assay parameters for membrane bound receptor

The average pKᵢ ± SEM (Kᵢ) values from 3 independent experiments for the adenosine A₂aR forms for binding of NECA, theophylline and ZM241385 in competition assays is presented.

<table>
<thead>
<tr>
<th>RECEPTOR</th>
<th>NECA pKᵢ ± SEM (Kᵢ µM)</th>
<th>Theophylline pKᵢ ± SEM (Kᵢ µM)</th>
<th>ZM241385 pKᵢ± SEM (Kᵢ nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>6.2 ± 0.3 (0.63)</td>
<td>5.2 ± 0.1 (5.7)</td>
<td>8.6±0.5 (0.1)</td>
</tr>
<tr>
<td>A316</td>
<td>6.1 ± 0.4 (0.79)</td>
<td>5.0 ± 0.1 (9.1)</td>
<td>n.d.</td>
</tr>
<tr>
<td>V334</td>
<td>6.0 ± 0.3 (0.95)</td>
<td>5.0 ± 0.1 (9.8)</td>
<td>n.d.</td>
</tr>
</tbody>
</table>

Table 2.3 Ligand binding affinities for adenosine A₂aR in different expression systems

pKi or pKd values for the ligands as reported in the literature for different expression systems is presented.

<table>
<thead>
<tr>
<th>Expression system</th>
<th>ZM241385</th>
<th>NECA</th>
<th>Theophylline</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td>9.7</td>
<td>7.1</td>
<td>5.6</td>
</tr>
<tr>
<td>(Weiss et al., 2002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P. pastoris</td>
<td>8.4</td>
<td>5.9</td>
<td>4.6</td>
</tr>
<tr>
<td>(Fraser, 2006)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHO cells</td>
<td>9.1</td>
<td>7.7</td>
<td>5.3</td>
</tr>
<tr>
<td>(Dionisotti et al., 1997; Klotz et al., 1998)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rat brain</td>
<td>8.8</td>
<td>6.5</td>
<td>4.7</td>
</tr>
<tr>
<td>(Alexander et al., 2001)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.4 DISCUSSION

One of the aims of the study was to develop a reliable, large-scale and well expressing system for the production of adenosine A2aR. To this aim we compared the expression of adenosine A2aR in flask and bioreactor based cultures. The flask conditions used in this study were similar to, or the same, as those which have previously been used for expression of the adenosine A2aR in P. pastoris (Andre et al., 2006; Fraser, 2006). Flask cultures enable the exploration of several expression conditions in parallel and require a limited set-up compared to bioreactor based cultures. However there are a number of drawbacks to flask cultures, which the bioreactor is particularly well suited to overcome. The principal advantage of the bioreactor set up is that it enables large volumes of air to be circulated within the culture to satisfy the high oxygenation requirements of P. pastoris cultivation to high OD600. The ability to regulate pH and dissolved oxygen (dO2) was shown to greatly improve batch-to-batch reproducibility while glycerol and methanol feeding lines enable accurate control over the amount of nutrients added to the culture. P. pastoris synthesises alcohol oxidase, which metabolises methanol by conversion to formaldehyde. However, if high concentrations of methanol are introduced without pre-conditioning of the culture to this carbon source, excess formaldehyde can accumulate to toxic levels in the cells (Stratton et al., 1998). It is therefore very important to allow the culture to adapt to methanol utilisation using low methanol concentration. It is also essential to experimentally determine the optimum induction period as over time the production of receptor saturates. It is possible that due to the large amount of receptor produced the cellular machinery is unable to cope with the translational modification required
to form a functional receptor and hence stops production. It is thus desirable to harvest the cells before the production reaches saturation level.

The cell density achieved in the bioreactor cultures \( \text{OD}_{600} \sim 75-90 \) was routinely six times higher than that achieved in flask cultures \( \text{OD}_{600} \sim 15 \). In addition radioligand binding indicated that there was an approximately three and a half times higher functional expression for the bioreactor culture compared to the flask culture (Figure 2.5). Even though the protein amount loaded onto the gel was normalised peculiarly a stronger band intensity was observed for Western blots of flask cultures than bioreactor cultures suggesting that a large proportion of protein expressed in flask cultures was non-functional. Thus the expression level in the bioreactor is increased in two ways firstly the biomass is increased and secondly the production of functional receptor is increased per milligram of membrane protein. These data clearly demonstrate the high efficiency receptor production that can be achieved by the careful regulation and optimisation of the growth parameters possible in bioreactor cultures. A \( B_{\text{max}} \) of 125 pmol/mg for the WT and 222 pmol/mg for the V334 is the highest yield reported for adenosine A\(_2\)R in \( P. \text{pastoris} \) (Andre et al., 2006; Sarramegna et al., 2003). This is lower than that achieved for expression of the adenosine A\(_2\)R in both \( S. \text{cerevisae} \) (480 pmol/mg; (O'Malley et al., 2007)) and mammalian cells (287 pmol/mg; (Hassaine et al., 2006)). However the higher cell densities achieved using \( P. \text{pastoris} \) mean that higher expression levels per litre are obtained.

Western blot analysis of the WT adenosine A\(_2\)R expressed in the bioreactor indicated the presence of degradation bands likely to be the result of C-terminal degradation. Previous studies expressing the WT adenosine A\(_2\)R in \( E. \text{coli} \) demonstrated C-terminal degradation upon solubilisation (Weiss et al., 2002). This is a
major concern since the purification tags are N-terminally located thus purification would not differentiate between intact and degraded receptor. Such a heterogeneous population of receptor is unlikely to crystallise. In order to answer the issue of proteolytic degradation of the receptor C-terminal truncations were made. Radioligand binding analysis indicated that higher levels of receptor were produced for the truncated A316 and V334 compared to the WT adenosine A2aR. The V334 adenosine A2aR expresses at approximately double the amount of the WT adenosine A2aR in both flask (Figure 2.5) and bioreactor culture (Table 2.1). Interestingly, unlike earlier reported results where the full length adenosine A2aR gave higher yields than the A316 adenosine A2aR (Weiss et al., 2002) a higher expression is achieved for both the V334 and A316 adenosine A2aR truncated forms when expressed in P. pastoris. It is not entirely clear why an opposite effect is observed but this may in part be due to the degradation observed for the WT receptor construct during induction. Unlike the WT the V334 adenosine A2aR does not exhibit this degradation during expression, indicating an intrinsic instability in the WT adenosine A2aR. This degradation has not been previously reported for expression of the receptor in relatively small scale P. pastoris cultures (Fraser, 2006) and this is supported by the findings of this study since no degradation bands were observed for either construct in flask cultures. It may be that the very high expression levels achieved in the bioreactor increases the processing and targeting time of the receptor making it more vulnerable to degradation by intracellular proteases. The lack of degradation of the V334 A2aR indicates that this is more stable and therefore a more suitable candidate for further downstream structural and functional studies.

A prerequisite for structural studies is the production of receptor that has a native like pharmacological profile while expressing it in a heterologous system such as
*P. pastoris*. The results presented here show that all the three receptor forms have highly similar pharmacological profiles with affinities similar to that of receptor expressed in mammalian cells (Dionisotti *et al.*, 1997; Klotz *et al.*, 1998) and native tissues (Alexander *et al.*, 2001) (Table 2.3) The values are also similar to that reported in earlier work done on the full length non-glycosylated adenosine A2aR expressed in *P. pastoris* (Fraser, 2006).

In conclusion a comparison between flask and bioreactor expression ability was performed and a protocol for high level expression of the adenosine A2aR in bioreactors established. This high level production was achieved through an optimised expression protocol and generation of C-terminally truncated constructs which were resistant to proteolytic degradation. A detailed functional characterisation of the adenosine A2aR receptor expressed in *P. pastoris* revealed similar affinities to that of receptor expressed in mammalian cell lines. A high expression level and native pharmacological characterisation is a promising start for the next downstream processing steps of solubilisation and purification.
3 Developing a solubilisation and purification protocol for the adenosine $A_{2a}$ receptor

Part of the work presented in this chapter is published in:

Protein Expression and Purification 2010 May 10

A purified C-terminally truncated human adenosine A(2A) receptor construct is functionally stable and degradation resistant

Shweta Singh, Diana Hedley, Elodie Kara, Adrien Gras, So Iwata, Jonathan Ruprecht, Philip G. Strange and Bernadette Byrne
3.1 INTRODUCTION

Successful characterisation of GPCRs involves achieving high level of expression followed by efficient solubilisation and purification which allows removal of all contaminants while maintaining the functional and conformational stability of the receptor (Kobilka et al., 2008).

The work described in the previous chapter demonstrated that the adenosine A$_{2a}$R forms with their C-terminus terminating at residue valine 334 and alanine 316 express to exceptionally high level and are more stable than the WT receptor when expressed under conditions of high cell density in bioreactors. Despite a number of reports in the literature (Fraser, 2006; Niebauer et al., 2006; Warne et al., 2003; Weiss et al., 2002) describing the successful expression of functional GPCRs, crystallisation and structure determination remains extremely difficult. One issue is that it is usually necessary to develop a protein-specific solubilisation and purification protocol in order to maximise the recovery of high-quality functional protein.

It is essential to have proper detection strategies in place to follow the quality of protein at each stage. One such technique is the radioligand binding assay which can be used to monitor recovery of functional receptor at each stage in the isolation process. In addition, radioligand binding analysis can also provide a measure of the long term functional stability of the receptor which can also be used as an assessment of conformational integrity. Binding affinity values obtained through competition assays can be used to compare the receptors pharmacological characteristics to the receptor expressed in mammalian cell lines.

However a highly functional receptor population may not always be an indication of a protein suitable for structural work in terms of homogeneity of protein.
Hence SDS-PAGE and SEC profiles were also used during the course of the study to assess the purity of the receptor and the monodispersity respectively. One of the limitations of ligand binding studies is that for an accurate measurement of functional receptor the protein needs to be prepared in the absence of ligands. Ligands especially ones that lock the receptor in an inactive conformation such as antagonist (Jaakola et al., 2008) or inverse agonists (Cherezov et al., 2007; Rasmussen et al., 2007) are routinely used for GPCR structural work as they help to stabilise the protein. Often there is a requirement to include these during solubilisation and purification complicating the functional characterisation of the pure material.

The aim of this study was to perform a detailed functional analysis of human adenosine A<sub>2a</sub>R expressed in <i>P. pastoris</i> at each step in the preparation process with a view to assessing forms of the receptor with minimal modifications (truncated C-terminus) in terms of suitability for structural studies. Such a study would also enable us to develop an optimised purification protocol which would keep the adenosine A<sub>2a</sub>R stable over time and hence help in recovering and maintaining a functional receptor for crystallisation. In the case of the adenosine A<sub>2a</sub>R it proved possible to isolate the protein in the absence of ligand facilitating the functional characterisation process.
3.2 MATERIALS AND METHODS

3.2.1 Materials

Complete EDTA-free protease inhibitor cocktail tablets were purchased from Roche Welwyn UK. The mouse M1 and M2 anti-FLAG antibody, M2 affinity resin, Flag peptide and Cholesteryl hemisuccinate (CHS) were from Sigma-Aldrich Gillingham UK. Dodecyl β-D maltopyranoside (DDM) and decyl β-D maltopyranoside (DM) were obtained from Anatrace USA. The 100 kDa molecular weight cut off concentrator was from Amicon-Millipore Watford UK and the Imperial stain was from Thermo Fisher UK. The ECF substrate, G-50 microspin columns, Superose 6 and the MonoQ columns were obtained from GE biosciences Buckinghamshire UK. The Co\textsuperscript{2+}-TALON resin was from Clontech and the Ni\textsuperscript{2+}-NTA resin was from Qiagen Crawley UK. The anti-His antibody was from Dianova Germany. All other chemicals were either as described before or obtained from Sigma-Aldrich Gillingham UK.

3.2.2 Solubilisation

Membranes were prepared as described in (2.2.6) and solubilised in Solubilisation Buffer (20 mM HEPES 7.4, 50% glycerol, 250 mM NaCl, 1 Protease Inhibitor tablet/100ml buffer) supplemented with either 2.5% dodecyl β-D maltoside (DDM) or decyl β-D maltoside (DM) ± 0.5% Cholesteryl hemisuccinate (CHS) at a concentration of 12-15 mg membrane protein per ml of solubilisation buffer. The samples were incubated for 2 hr at 4°C with gentle agitation followed by centrifugation for 45 minutes at 100,000 g to remove the insoluble fraction. The
solubilised material was either immediately purified or used for radioligand binding analysis.

3.2.3 Modified protein concentration determination assay

The solubilisation of membranes was carried out at a high concentration of non-ionic detergents which can cause precipitate formation during protein analysis and hence render the protein estimation inaccurate. In order to avoid inconsistencies, a modification to the protein analysis mentioned in Section 2.2.7 was made. The protein was precipitated with 1 ml of 25% TCA and centrifuged at 2200 g for 15 minutes. The supernatant was discarded and a mixture of Solution A and B at 50:1 v/v was added to each pellet. To prevent any precipitation of protein 1 ml of 10% SDS was added to each tube and the solutions incubated overnight at room temperature. Folin-Ciocalteau reagent was diluted with an equal volume of water and 100 µl of this solution was added to each tube. The tubes were vortexed immediately and incubated at room temperature for 10 minutes. A 1 ml volume of distilled water was added to each tube and the samples incubated for a further 20 minutes. Absorbance was measured at 760 nm on a spectrophotometer. Data were analysed using GraphPad prism software.

3.2.4 Analysing column efficiency

In order to separate the receptor bound radioligand from free radioligand in soluble protein samples the principle of size exclusion was applied. Pre-packed G-50 micro-spin columns containing 500 µl Sephadex resin were used for separation. It was important to verify that the initial eluate only contained ligand bound to the protein and unbound ligand was not co-eluting. Three different reaction types were assayed. Control tubes contained buffer and radioactive ligand but no protein. The “total” and
“non-specific” reaction tubes were similar to that used in the saturation ligand binding assays i.e. total contained the radioligand and the protein while non-specific contained the protein along with the radioligand and competing ligand. A 40 µl volume from a 200 µl volume of sample was loaded on to a column and centrifuged for 2 minutes at 1000 g. The eluate was retained for analysis. The columns were then washed with 40µl buffer and the fractions analysed through radioligand binding studies.

### 3.2.5 Radioligand binding assay on solubilised and purified receptor

All radioligand binding assays on soluble receptor samples were performed in binding buffer supplemented with 0.05% DDM or 0.5% DM, 0.01% CHS and 100 mM NaCl in an assay volume of 800 µl. Theophylline at 10⁻² M was used to define non-specific binding. The assay was incubated at 4°C for 1 hour. Individual 40 µl aliquots were removed and loaded onto columns and centrifuged at 1000 g for 2 minutes. Before loading the assay mixture, the columns were equilibrated with at least 100 µl of binding buffer. The eluant containing the receptor bound radioligand was added to 2 ml of scintillation fluid and counted immediately. All measurements were carried out in duplicate. For competition binding assays 2 nM [³H] ZM241385 was used in competition with NECA and theophylline. Assay conditions as described for the membrane (Section 2.2.9) were used for both solubilised and purified receptor samples.

Binding conditions for purified receptor were similar to that used for soluble receptor. As the protein at this stage was highly enriched a very high binding was expected, 100 times less protein (0.01µg per assay tube) was used. Such low protein concentration was also important to prevent ligand depletion in the assay.
3.2.6 Analysing data for solubilised and purified receptor

Data analysis was similar to that for membrane bound receptor with the exception that the DPM values had to be adjusted for the fact that only 1/20 of the reaction mixture was used for determining the radioactive count.

3.2.7 Solubilisation condition screening

In order to achieve an optimised solubilisation protocol various factors were assessed. All the factors were assessed for V334 adenosine A\textsubscript{2a}R in binding buffer containing 0.05% DDM and 0.01%CHS unless specified. The incubation time for the protein and detergent mixture was varied between 2-16 hours to assess for extraction ability. In order to assess the time required to reach equilibrium for the ligand protein reaction at 4°C incubation time of 1 and 16 hours was assessed. Theophylline an antagonist at 10^{-2} M was used for determining non-specific binding vs [\textsuperscript{3}H] ZM241385. Adenosine deaminase at 1 unit/ ml of buffer for every reaction tube was used to remove any endogenous adenosine present. Different concentration of protein 12-25 µg per 200 µl assay volume was incubated with adenosine deaminase followed by the ligand binding assay. The binding buffer was supplemented with 150mM NaCl.

3.2.8 Determining T\textsubscript{m} of detergent solubilised receptor

V334 adenosine A\textsubscript{2a}R was solubilised in DM+CHS and DDM+CHS as described in Section 3.2.2. The solubilised receptor was incubated over a range of temperature 28°C-80 °C in binding buffer containing 300nM ZM241385, 0.5% DM or 0.05% DDM + 0.01% CHS. The samples were incubated at specific temperature for 30 minutes following which the samples were quenched on ice and passed through a GE microspin column to remove ZM241385 present in the buffer. Specific binding was
measured by incubation at 4°C for 1 hour at saturating concentration (4nM) of [³H] ZM241385. The assay was repeated three times and the average percentage of binding remaining at each temperature in comparison to the binding for the receptor at 4°C was plotted and the curve fitted to non-linear regression. The temperature corresponding to 50% binding was defined as the Tₘ.

### 3.2.9 Purification

The following purification methods were used during the development of a purification protocol for the adenosine A₂aR. All optimisation experiments were done on the V334 adenosine A₂aR. The purified fractions were assessed by radioligand binding analysis. All the protein samples were also further analysed using SDS-PAGE and Western blotting. In case of a multiple purification-step protocol isolated fractions were immediately used for the next step. All samples were stored at 4°C during the course of the study.

#### 3.2.9.1 Immobilised metal affinity chromatography

The deca His-tag present on the N-terminal of the receptor construct was utilised for immobilised metal affinity chromatography (IMAC). An aliquot of Co²⁺-TALON resin was batch equilibrated with IMAC base buffer (20 mM HEPES pH 7.4, 10% Glycerol, 250 mM NaCl, 0.05% DDM, 0.01% CHS) The soluble protein sample was batch-bound to the resin at 4°C for 3 hr. Typically 15-20 mg of solubilised protein was bound to 1 ml resin. The resin was then transferred to a drip column and the flow through containing unbound contaminants was collected. High and low imidazole purification protocols were compared. The high imidazole protocol included 30 column volume (CV) washes at 40 mM imidazole and elution at 400 mM imidazole. The imidazole concentrations were reduced to 15 mM for washes and 150 mM for
elution in the low imidazole purification protocol. In the case of isolation on Ni²⁺ resin, IMAC base buffer was supplemented with 30 mM imidazole during washing and 300 mM imidazole during elution.

3.2.9.2 Flag Tag affinity chromatography
An octa peptide Flag tag is also present on the N-terminus of the construct to which the M2 antibody can bind. The M2 resin was equilibrated with 3 sequential CV of 0.1 M glycine-HCl pH 3.5 followed by 5 CV washes with the Flag tag buffer (20mM Tris-HCl pH-8, 10% glycerol, 250mM NaCl, 0.05% DDM, 0.01% CHS). The protein typically at 15-20 mg/ml resin was batch bound to the resin for 2 hours at 4°C if volume was larger than 20 ml. The resin slurry was then loaded in to a drip column and washed with 30 CV of Flag tag buffer. In the case of analytical scale purification of 1-5 ml, soluble sample was passed three times through the resin bed of 1 ml to allow binding of the Flag tag protein to M2 resin. The bound Flag tag fusion protein was eluted by competitive elution with 4 CV of a Flag tag buffer solution containing 150 µg/ml Flag peptide.

3.2.9.3 Anion exchange chromatography
Anion exchange chromatography through MonoQ 5/5 gl (GE) was used as a second step of purification after IMAC or Flag tag purification. The protein was buffer exchanged on a 100 kDa molecular weight cut-off filter into MonoQ buffer (20mM Tris-HCl pH-8, 10% glycerol, 0.05% DDM, 0.01% CHS) with at least 25 times the protein volume to wash off the NaCl and imidazole. The column was equilibrated with MonoQ buffer and typically 500 µl of protein at 5-10 mg/ml was loaded onto a 1 ml column. The column was then washed with 2 CV of MonoQ buffer. The proteins were eluted using a step gradient of 0-1000 mM NaCl.
3.2.9.4 Size exclusion chromatography

Size exclusion was used as an analytical step. Superose 6 was equilibrated with SEC buffer (20mM Tris-HCl pH-8, 5% glycerol, 250mM NaCl, 0.05% DDM, 0.01% CHS). The protein at 5-10 mg/ml was loaded onto the column, depending on the concentration of the protein 500-1000 µl of protein was loaded on to the column. The separation was run at 1 ml/min and 1 ml fractions were collected.

3.2.9.5 TEV proteolysis and reverse Ni\(^{2+}\) purification

In order to remove the tags TEV proteases was added to pure receptor protein following Flag tag affinity chromatography at a 1:1 molar ratio of pure protein. The mixture was incubated at 4°C for 12-14 hours. In case of the time course time samples were taken at intervals of 0,2,4,6,8 and 20 hours after addition of TEV proteases. The TEV protease and purified adenosine A\(_{2a}\)R protein mixture was passed through a Ni\(^{2+}\)-NTA resin bed 3 times and the unbound fraction collected for analysis.

3.2.10 SDS-PAGE and Western blot

SDS-PAGE and Western blot analysis were performed as described in Section 2.2.8. Western blots were performed simultaneously with equal amounts of protein using primary antibodies M1 and M2 followed by the secondary goat anti-mouse horseradish peroxidase antibody to differentiate between the processed and unprocessed receptor forms. To detect TEV protein anti-his antibody at a dilution of 1:5000 was used followed by secondary goat anti-mouse alkaline phosphatase antibody at a dilution of 1:5000. In all cases protein bands were visualised using ECL substrate by the chemiluminescence method.
3.3 RESULTS

3.3.1 Analysing column efficiency

Experiments to rule out any leakage of free ligand from the G-50 microspin columns used for separation of bound from unbound radioligand were carried out (Figure 3.1). The highest levels of radioactivity were detected in the samples containing radioligand along with the protein. A much lower signal was measured for the non-specific binding samples indicating that the majority of the free ligand was trapped in the resin and hence the counts from total and non-specific mixtures were reliable.

![Column efficiency test](image)

**Figure 3.1 Column efficiency test**

Amount of radioligand (DPM values) present in eluates from free ligand (green), total (plum) and non-specific (blue) for the initial elution and washes thereafter from a G-50 microspin column is presented.

3.3.2 Developing an assay for determining functional adenosine $A_{2a}$R activity in detergent solution

In order to develop a robust method for radioligand binding analysis on detergent solubilised adenosine $A_{2a}$R different variables within the assay were tested. A summary of the optimisation experiments is presented in (Table 3.1).
Table 3.1 Optimisation of ligand binding assay for soluble receptor

A summary of conditions analysed for binding studies for receptor solubilised in detergent is presented. All optimisation trials were done with V334 adenosine A2aR. The condition highlighted in grey was used for further assays.

<table>
<thead>
<tr>
<th>CONDITIONS</th>
<th>RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low protein concn (0.05-5 µg/200 µl assay)</td>
<td>Low DPM hence low B_max</td>
</tr>
<tr>
<td>High protein concn (5-25 µg/200 µl assay)</td>
<td>High DPM but ligand depletion</td>
</tr>
<tr>
<td>Increased solubilisation/incubation time</td>
<td>No significant effect</td>
</tr>
<tr>
<td>Presence of Adenosine Deaminase</td>
<td>No significant effect</td>
</tr>
<tr>
<td>Theophylline as a competing ligand</td>
<td>Tighter triplicates (DPM)</td>
</tr>
<tr>
<td>Solubilisation in presence of DM +/- CHS</td>
<td>Further decrease in B_max</td>
</tr>
<tr>
<td>8 µg /400 µl assay + 150mM NaCl in binding buffer</td>
<td>High DPM, decrease in ligand depletion</td>
</tr>
<tr>
<td>2/4 µg /800 µl assay + 150mM NaCl in binding buffer</td>
<td>High DPM, Ligand depletion reduced</td>
</tr>
</tbody>
</table>

Optimisation of the protein concentration and overall assay volume to 2 µg/800 µl gave a high radioactivity value and less than 25% ligand depletion. The incubation times was reduced to one hour which was enough to reach equilibrium, as supported by the literature (Palmer et al., 1995). A total solubilisation time 2-3 hours was identified as optimal as no significant binding was observed in the insoluble fraction at this time. The other conditions that made the assay reproducible and reliable were addition of 150 mM NaCl in the binding buffer, use of theophylline as the ligand to define non-specific binding and increasing the volume of the assay to 800 µl.
### 3.3.3 Presence of CHS significantly increases the functional yield and stability of detergent-solubilised adenosine A2aR

The V334 adenosine A2aR was solubilised in either DM or DDM in the presence and absence of CHS. Recovery of functional protein following solubilisation in DDM or DM was less than 80 pmol/mg. Solubilisation with the detergent alone led to a significant loss in functional binding as compared to that in the membrane (Bmax = 222 pmol/mg). Addition of 0.05% CHS to the solubilisation buffer had a dramatic effect, more than doubling the recovery of functional receptor (Figure 3.2).

![Figure 3.2 Comparison of different buffer conditions for solubilisation](image)

**Figure 3.2 Comparison of different buffer conditions for solubilisation**

The V334 adenosine A2aR was solubilised in the presence of DM or DDM +/- of CHS to determine functional recovery of receptor. Single point binding assay at saturating concentrations of [³H] ZM241385 were carried out to assess the recovery of functional receptor. Data is representative of three independent experiments.

### 3.3.4 Assessment of binding characteristics of solubilised receptor

Saturation assays were performed under optimised solubilisation and assay conditions to obtain saturation curves and a Bmax for detergent-solubilised receptor. Saturation binding analysis performed on the WT and adenosine A2aR constructs with truncated
C-terminus solubilised in DDM + CHS (Figure 3.3) revealed an increase in $B_{\text{max}}$ compared to the membrane bound receptor indicating a slight enrichment of receptor at this stage (Table 3.2). The $B_{\text{max}}$ values for the solubilised receptors revealed a similar trend to the membrane bound samples with both the truncated C-terminus constructs yielding significantly higher levels of functional receptor than the WT. Similar results were obtained for the A316 and V334 adenosine $A_2a$R forms solubilised in DM + CHS however under these conditions the $B_{\text{max}}$ value for the solubilised WT receptor was lower than that obtained for the membrane bound WT sample (Figure 3.4, Table 3.2). This further supports the fact that the WT construct is less stable than the C-terminal truncated constructs.

![Figure 3.3 Saturation curves for receptor solubilised in DDM+CHS](image)

**Figure 3.3 Saturation curves for receptor solubilised in DDM+CHS**

Saturation curve representative of three independent experiments for the WT (■) and V334 (▲) adenosine $A_2a$R solubilised in DDM+CHS is presented. The $B_{\text{max}}$ values for WT and V334 adenosine $A_2a$R were 155 and 424 pmol/mg respectively. The standard error means for triplicates from an experiment are shown.

Interestingly V334 adenosine $A_2a$R solubilised in both DM and DDM has a significantly ($P<0.05$) higher binding affinity for the radioligand $[^3H]ZM241385$ compared to both the membrane bound and purified receptor protein (Figure 3.4, Table 3.2)
Figure 3.4 Saturation curves for receptor solubilised in DM+CHS

Saturation curve representative of three independent experiments for the WT (■) and (A) A316 (○) and (B) V334 (▲) adenosine A₂aR solubilised in DM+CHS is presented. The Bₘₐₓ values for WT, A316 and V334 adenosine A₂aR were 97.8, 384.2 and 420.7 pmol/mg respectively. The standard error means for triplicates from an experiment are shown.

Table 3.2 Saturation assay parameters for adenosine A₂aR forms solubilised in DM or DDM + CHS.

The three receptor forms were solubilised in DM or DDM + CHS. The Bₘₐₓ ± SEM and pKₐ ± SEM (Kₐ nM) for at least three independent experiments is presented. *P<0.05 relative to membrane bound or purified receptor (ANOVA with Bonferroni’s post test). n.d. = not determined

<table>
<thead>
<tr>
<th>RECEPTOR TYPE</th>
<th>DM+CHS</th>
<th>DDM+CHS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bₘₐₓ ± SEM (pmol/mg)</td>
<td>pKₐ ± SEM (Kₐ nM)</td>
</tr>
<tr>
<td>WT</td>
<td>75 ± 22</td>
<td>9.7 ± 0.1* (0.21)</td>
</tr>
<tr>
<td>A316</td>
<td>282 ± 51</td>
<td>9.2 ± 0.2 (0.61)</td>
</tr>
<tr>
<td>V334</td>
<td>405 ± 20</td>
<td>9.1 ± 0.1* (0.72)</td>
</tr>
</tbody>
</table>

The amount of receptor in milligrams in the membrane and that extracted after solubilisation indicated 95% solubilisation efficiency. The insoluble pellet did not contain any measurable binding activity.
Competition binding analysis of the solubilised receptors (Figure 3.5, Table 3.3) reveals overall similar pKᵢ values to those obtained for membrane bound receptor for NECA and theophylline, indicating that although the affinity of ZM241385 is slightly higher, the binding site has not been greatly perturbed upon solubilisation.

Figure 3.5 Competition binding profile for the V334 adenosine A₂₅R form solubilised in DDM+CHS
A typical curve for the V334 adenosine A₂₅R form for binding of theophylline (●) and NECA (■) in competition versus [³H] ZM241385 binding is presented. The kᵣ value for NECA was 3.7 µM and theophylline was 35 µM. Similar curves were obtained for the WT and A316 adenosine A₂₅R forms through three independent experiments done under similar conditions. Data were best fitted to a one binding site model. The standard error means for triplicates from an experiment are shown.

The exception to this is the binding affinity of the solubilised A316 A₂₅R for theophylline which is significantly higher than that obtained for the membrane bound and purified receptor (P<0.05).
Table 3.3 Competition assay parameters for adenosine A₂aR solubilised in DM+CHS

The average $pK_i \pm$ SEM ($K_i \mu M$) values from 3 independent experiments for the adenosine A₂aR forms for binding of NECA and theophylline in competition assays is presented. *$P<0.05$ relative to membrane bound or purified receptor (ANOVA with Bonferroni’s post test).

<table>
<thead>
<tr>
<th>RECEPTOR</th>
<th>NECA $pK_i \pm$ SEM ($K_i \mu M$)</th>
<th>theophylline $pK_i \pm$ SEM ($K_i \mu M$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>6.4 ± 0.7 (0.36)</td>
<td>5.1 ± 0.1 (7.35)</td>
</tr>
<tr>
<td>A316</td>
<td>5.9 ± 0.7 (1.20)</td>
<td>4.4 ± 0.1* (73.48)</td>
</tr>
<tr>
<td>V334</td>
<td>6.1 ± 0.2 (0.49)</td>
<td>5.0 ± 0.2 (10.15)</td>
</tr>
</tbody>
</table>

3.3.5 The $T_m$ values for V334 adenosine A₂aR are comparable to the conformational thermostable mutants

The $T_m$s for V334 receptor solubilised in DM+CHS and DDM+CHS were 50°C and 49°C respectively (Figure 3.6). Conformational thermostable mutants have been generated through alanine scanning mutagenesis for the adenosine A₂aR (Magnani et al., 2008), β₁ adrenergic receptor (Serrano-Vega et al., 2008) and the NT1 neurotensin receptor (Shibata et al, 2009). The $T_m$ values for A₂aR mutants were in the range of 46°C - 49°C (Magnani et al., 2008) which are comparable to that obtained for the V334 adenosine A₂aR.
Figure 3.6 Determining the $T_m$ for V334 adenosine $A_2a$R

Radioligand binding analysis of V334 adenosine $A_2a$R solubilised in DDM + CHS (filled circles, dashed line) or DM + CHS (filled triangles, solid line). The solubilised receptor was incubated for 30 minutes over a range of temperatures 28°C - 80°C in binding buffer containing 300 nM ZM241385. Specific radioligand binding was measured by incubation at 4°C for 1 hour at saturating concentrations (4nM) of $[^3H]$ZM241385. The percentage of binding remaining at each temperature in comparison to the binding for the receptor at 4°C is shown. The specific activity measured at 4°C was taken as 100% activity (for clarity this value is not shown on the graph). The $T_m$ was estimated as the temperature at which 50% of the starting activity remained. The data is an average of three independent experiments.

3.3.6 Development of a purification protocol

High imidazole purification protocol

Previous work in the laboratory had developed an isolation protocol based on the procedure which gave the highest purity as assessed by SDS-PAGE analysis. This protocol was therefore initially used for isolation of the V334 $A_2a$R affinity chromatography using Co$^{2+}$ resin and an elution buffer containing 400 mM imidazole protocol gave a highly homogeneous preparation. Two clear bands are visible on an SDS-PAGE gel; a 32 kDa band corresponding to the receptor monomer and a higher molecular band (~90 kDa) likely to correspond to a higher oligomeric form (Figure 3.7). These bands had previously been identified to be the adenosine $A_2a$R through Western blot and mass spectrometry. Radioligand binding assays were performed to measure the quantity of functional protein present after the first step of purification.
Even after much optimisation of the assay, the maximum binding values achieved were no higher than 1 nmol/mg.

**Figure 3.7 SDS gel profile for high imidazole Co\(^{2+}\) IMAC purification**

Adenosine A\(_2a\)R V334 purified through Co\(^{2+}\) IMAC as the first step of purification using a imidazole purification protocol where the bound protein was eluted using 400 mM imidazole in the elution buffer. Two bands corresponding to a monomer and probably a higher oligomeric form can be observed. The molecular weight marker (M), sample bound to the resin (L), the flow through (FT) and washes (W) followed by the elution fraction (E1-4) are shown.

**Optimisation of IMAC purification protocol**

In order to improve the functional receptor recovery, modifications to the purification protocol were made. A comparison was made between Ni\(^{2+}\) and Co\(^{2+}\) slurry for IMAC (Figure 3.8). The Co\(^{2+}\) resin was washed with 15 mM imidazole instead of 40 mM as in the high imidazole protocol. The washed sample was then divided and the protein eluted at different concentrations of imidazole. The manufacturers protocol and previous work in the lab suggested that Ni\(^{2+}\) binds with higher affinity and lower sensitivity than Co\(^{2+}\) resin. The Ni\(^{2+}\) slurry was thus supplemented with 10 mM imidazole during incubation with the protein to reduce non-specific binding. The Ni\(^{2+}\) resin was washed with 40 mM imidazole. The elution
fractions from both the purifications were collected and analysed by SDS-PAGE (Figure 3.8 A), Western blot (Figure 3.8 B) and radio-ligand binding assay (Table 3.4). The fraction collected from the Co\(^{2+}\) column was cleaner than that obtained by Ni\(^{2+}\) purification. Interestingly for the Ni\(^{2+}\) fractions no band correlating to a monomeric form was observed. However both the preparations from Co\(^{2+}\) and Ni\(^{2+}\) were not as clean as the high imidazole preparation (Figure 3.6)

![Image of SDS-PAGE and Western blot](image)

**Figure 3.8 Comparison between Co\(^{2+}\) and Ni\(^{2+}\) IMAC as first step of purification**

Concentration of imidazole for elution and resin type were assessed as first step of purification. The protein was washed with 40 mM imidazole for Ni\(^{2+}\) and 15mM imidazole for Co\(^{2+}\) and eluted at the given concentrations. The samples were assessed through commassie stained SDS-PAGE (A) and Western blot (B). The blot was probed with primary M2 antibody and visualized though ECL substrate.

Functional values obtained for the fractions purified through Co\(^{2+}\) were higher than that of the Ni\(^{2+}\) resin (Table 3.4). Radioligand binding studies on the flow through and washes did not give a significant binding value suggesting that the protein bound well to the resin. The binding values were higher than that obtained from the earlierpurifications suggesting an enrichment and maintenance of functionality from solubilisation to first step of purification.
The sensitivity, efficiency and functional yields from the two purification protocols were assessed on the basis of functional receptor at each stage of purification. Functional receptor in unbound fractions, washes and elution was checked through single point binding analysis at saturating concentrations of [3H] ZM241385.

### Table 3.4 Functional receptor yields from Co\(^{2+}\) and Ni\(^{2+}\) IMAC as the first step of purification.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Ni(^{2+}) purification</th>
<th>Conditions</th>
<th>Co(^{2+}) purification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Load</td>
<td>Binding nmol/mg</td>
<td>Load</td>
<td>Binding nmol/mg</td>
</tr>
<tr>
<td>Flow Through</td>
<td>0.5</td>
<td>Flow through</td>
<td>0.5</td>
</tr>
<tr>
<td>Wash 40mM</td>
<td>0.3</td>
<td>Wash 15mM</td>
<td>0.3</td>
</tr>
<tr>
<td>Elution 200mM</td>
<td>3.4</td>
<td>Elution 100mM</td>
<td>4.7</td>
</tr>
<tr>
<td>Elution 300mM</td>
<td>2.5</td>
<td>Elution 150mM</td>
<td>3.1</td>
</tr>
<tr>
<td>Elution 400mM</td>
<td>2.4</td>
<td>Elution 200mM</td>
<td>2.5</td>
</tr>
<tr>
<td>Elution 500mM</td>
<td>2.4</td>
<td>Wash 15mM; 500mM NaCl</td>
<td>0.6</td>
</tr>
<tr>
<td>Elution 500mM; 500mM NaCl</td>
<td>1.7</td>
<td>Elution 200mM; 500mM NaCl</td>
<td>1.8</td>
</tr>
</tbody>
</table>

Assessment of the pharmacological characteristics of the protein purified by the Co\(^{2+}\) IMAC low imidazole protocol confirmed the enrichment of sample with a $B_{\text{max}} \pm \text{SEM}$ of 3.62 ± 0.51 nmol/mg. In addition the protein exhibited a binding constant $K_d \pm \text{SEM}$ of 9.1 ± 0.1 similar to that of membrane bound and solubilised V334 adenosine A$_{2a}$R (figure 3.9).
Figure 3.9 Saturation analysis of [3H]ZM241385 to the Co^{2+} IMAC purified V334 adenosine A_{2a,R}

Data are representative of three independent experiments. A $B_{\text{max}}$ of 3.4 nmol/mg and $K_d$ of 0.4 nM were derived by non-linear regression.

3.3.7 Second step of purification

In order to separate the three different species as observed on the SDS-PAGE/Western blot analysis (Fig 3.8) after Co^{2+} IMAC, a further SEC step through Superose 6 was carried out. The chromatogram for the protein mixture suggested a homogenous population of receptor with a single mono disperse peak. Fractions from the peak were loaded onto a SDS gel and three bands were observed (Figure 3.10). A 24 ml Superose 6 column was unable to fully separate the three species of protein.
Radioligand binding studies on the SEC fractions indicated that all the forms were functional. Single point radioligand binding assays done on the SEC fraction at saturating concentration of ligand gave a binding value of 3.2 nmol/mg. This value was similar to that from the first step of purification suggesting that SEC was not a good second step of purification as no enrichment occurred.

Anion exchange chromatography through a MonoQ column was assessed as an alternative second step of purification. A two step purification of the V334 adenosine A2aR by Co²⁺ affinity chromatography and anion exchange chromatography produced highly pure protein with a yield of ~ 2 mg /L of cell culture. Saturation binding of [³H] ZM241385 gave a Bₘₐₓ of 20.5 ± 0.7 nmol/mg (Figure 3.11) corresponding to 83% of the theoretical Bₘₐₓ.
The purified receptor had very similar binding affinities to the membrane bound receptor for all the ligands tested demonstrating retained pharmacological properties. However, three bands were still observed on an SDS gel (Figure 3.12).
3.3.8 TEV cleavage and subsequent protein purification optimisation

V334 adenosine A_2aR from the MonoQ column was incubated with TEV followed by reverse IMAC. This produced a homogenous protein population with a single band corresponding to the monomeric receptor form (Figure 3.14) however radioligand binding studies indicated a loss of activity compared to the sample prior to TEV treatment (Table 3.5).
Figure 3.13 Time course for TEV protease cleavage of purified V334 adenosine A2aR.
The receptor was purified by Co²⁺ IMAC followed by MonoQ prior to incubation with TEV proteases. A 1:1 molar ratio of protein to TEV proteases was used. The samples taken at different time points (0-20 hours) were mixed with SDS loading dye and frozen at -20°C. The samples were then submitted to SDS-PAGE (A) and Western blot (B) analysis. TEV protease was also loaded onto the gel as a positive control for the blot. The blot was probed with anti-His primary antibody.

It was possible that the addition of a third chromatographic step caused significant delipidation of the receptor resulting in the loss of function. An alternative purification protocol was developed which included purification through a M2 antibody affinity resin followed by incubation with TEV proteases for 14 hours at 4°C followed by reverse IMAC and analysed by SDS-PAGE (Fig 3.15). The protein ran as a single band (Figure 3.15) of about 30 kDa corresponding to a monomeric receptor form with the optimal time for cleavage 14 h. Receptor purified in this way had a $B_{\text{max}} \pm \text{SEM}$ of 20.3 ± 2.0 nmol/mg and a $pK_d \pm \text{SEM}$ of 8.5 ± 0.1. The theoretical $B_{\text{max}}$ for the glycosylated V334 adenosine A2aR is 24 nmol/mg demonstrating that this protocol produced highly pure, highly functional receptor protein.
Figure 3.14 SDS-PAGE and functional analysis of the purified V334 A2aR

(A) Saturation binding of [3H]ZM241385 to the pure V334 A2aR obtained following reverse IMAC. The receptor was isolated in the presence of DDM+CHS. Data are representative of three independent experiments. B<sub>max</sub> and K<sub>d</sub> values were derived by non-linear regression. The B<sub>max</sub> was 18.2 nmol/mg and K<sub>d</sub> 3.3 nM. (B) Coomassie-stained SDS-PAGE gel of protein samples obtained following anti-Flag tag affinity chromatography (Lane 2) and following reverse IMAC (Lane 3). Molecular weight markers are shown in Lane 1.

<table>
<thead>
<tr>
<th>Purification method</th>
<th>Specific activity nmol/mg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flag affinity chromatography</td>
<td>4.6</td>
</tr>
<tr>
<td>Flag + TEV + Ni2+ reverse IMAC</td>
<td>9.3</td>
</tr>
<tr>
<td>Flag + MonoQ</td>
<td>18.7</td>
</tr>
<tr>
<td>Flag+MonoQ+TEV+Ni2+reverse IMAC</td>
<td>5.8</td>
</tr>
</tbody>
</table>
Figure 3.15 Competition binding profile for purified V334 adenosine A2aR

A typical curve for the V334 adenosine A2aR form for binding of theophylline (○) and NECA (■) in competition versus [3H]ZM241385 binding is presented. Data were best fitted to one binding site model. The $K_i$ for theophylline was 22 µM and NECA was 0.52 µM.

The binding affinities obtained through competition radioligand binding assay indicate a similar pharmacological profile of the solubilised receptor to membrane bound receptor.

Table 3.6 Competition assay parameters for purified V334 adenosine A2aR

The average $pK_i \pm$SEM ($K_i$ µM) values from 3 independent experiments for the V334 adenosine A2aR for binding of NECA and theophylline in competition assays is presented.

<table>
<thead>
<tr>
<th>PURIFICATION TYPE</th>
<th>NECA $pK_i \pm$SEM ($K_i$ µM)</th>
<th>Theophylline $pK_i \pm$SEM ($K_i$ µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMAC + monoQ</td>
<td>6.5 ± 0.3 (0.35)</td>
<td>4.8 ± 0.2 (15.5)</td>
</tr>
<tr>
<td>M2 + reverse IMAC</td>
<td>6.3 ± 0.2 (0.52)</td>
<td>4.7 ± 0.1 (22.2)</td>
</tr>
</tbody>
</table>
3.4 DISCUSSION

Solubilisation and isolation of pure, functionally active protein remains a major challenge to the structure determination of GPCRs. Solubilisation must be as efficient as possible while maintaining structural and functional integrity (Ostermeier et al., 1997). The standard detergent for membrane protein solubilisation and purification is DDM, which produces a large micelle and is thus an effective mimic of the native membrane environment (Prive, 2007; Iwata, 2003). The β2 AR (Cherezov et al., 2007) and the adenosine A2aR (Jaakola et al., 2008) structures were obtained using protein solubilised in DDM. Membrane proteins tend to be relatively stable in this detergent. In contrast membrane proteins tend to be much more unstable in smaller detergents such as DM, which leave larger regions of the molecules exposed to the aqueous environment. The challenge for membrane protein structural biologists is to balance the long term stability of the membrane protein, usually requiring large chain detergents, with an optimised likelihood of obtaining diffracting crystals, often requiring shorter chain detergents. A comparison between DM and DDM ± CHS as solubilising agents for the V334 adenosine A2aR (Figure 3.2) indicated an essential role for CHS in functional recovery. In the absence of CHS the binding level for the receptor in both the detergents was lower than that in the membranes indicating that solubilisation by detergent alone rendered the receptor inactive however addition of CHS led to a significant increase in the binding levels. Cholesterol has been shown to increase the functional recovery of GPCRs (Hanson et al., 2008; Jaakola et al., 2008; Weiss et al., 2002). The study by Hanson et al also identified a cholesterol binding motif in the β2 adrenergic receptor. Analysis of amino acid sequences for a range of GPCRs identified a potential binding motif for cholesterol in a large number of
GPCRs including the adenosine $A_2a$R. This information along with the data presented in this study suggests a critical role for cholesterol in maintaining functional and structural integrity of the receptor. The role of cholesterol is possibly more important that that of the detergent itself since little difference in functional recovery was observed between sample solubilised in DM or DDM.

A high efficiency of solubilisation was achieved for both the V334 and A316 receptor forms. Saturation and competition radioligand binding analysis revealed similar pharmacological characteristics to the membrane bound and native receptor. The WT adenosine $A_2a$R unlike the two C-terminal truncated constructs was not stable in DM+CHS and had poor functional recovery. This supports our earlier results (Chapter 2) demonstrating that that the C-terminal truncated constructs are more robust or resistant to degradation than the WT construct. Among the three constructs studied the highest level of expression and functional recovery was achieved for the V334 adenosine $A_2a$R which was taken further for subsequent purification and functional analysis. The observation that the profile for V334 adenosine $A_2a$R is similar in the presence of DM and DDM+ CHS is promising as it indicates that exchange into the smaller detergent in the final stages of purification may be possible (see Chapter 4 for further comparative analysis of these detergents on the stability of the purified receptor).

Thermostability analysis of the V334 adenosine $A_2a$R gave similar $T_m$ values to that reported for the adenosine $A_2a$R following conformational stabilisation by alanine scanning mutagenesis ($46\text{-}49^\circ\text{C}$; Serrano-Vega et al., 2008). In the case of the mutagenised construct, the receptor was solubilised in the absence of stabilising lipid (Serrano-Vega et al., 2008). Whilst the V334 $A_2a$R construct is dependent on CHS for effective solubilisation and long term stability, the high resolution structure of the
$A_2^aR + T4L$ construct obtained using protein isolated in the presence of CHS illustrates that a such a dependence is not necessarily a barrier to crystallisation. However it should be noted that in the case of the $A_2^aR + T4L$ construct, the amount of CHS used for preparation was an order of magnitude less than that used here (Jaakola et al, 2008).

Several purification protocols were attempted in order to produce a population of protein that was highly functional, homogenous and conformationally stable. The initial purification protocol was developed based on the SDS-PAGE profile of the protein. As the eluate from the IMAC step did not show any functional enrichment it was suspected that the high (40 mM) imidazole wash step was causing elution of monomeric, comparatively loosely bound receptor. It is possible that the high imidazole concentration used for elution (400 mM) displaced a population of partially unfolded, highly aggregated and much more tightly bound receptor which was mainly non-functional.

A comparison between Co$^{2+}$ and Ni$^{2+}$ IMAC purification provided the ability to assess a range of imidazole concentrations in order to ascertain the optimum wash and elution concentrations. A low imidazole wash on both columns led to a higher yield of functional receptor however an increased number of bands were observed on an SDS-PAGE. The purest preparation with the highest level of enrichment was obtained from the Co$^{2+}$ column. A combination of this step with MonoQ gave a functional value close to the theoretical $B_{\text{max}}$. Despite this, the SDS gel profile for the protein purified through this protocol appeared less promising as a heterogeneous population of receptor could be observed.

Even though they were highly functional such a heterogeneous population is not suitable for crystallisation trials. The process of TEV cleavage helped to remove
all the tags and produce a homogenous population of receptor. A highly functional and homogenous receptor population could be achieved by a two-step purification involving M2 affinity chromatography where the protein was eluted in the presence of Flag peptide rather than the high imidazole concentration making it a milder process. This was followed by TEV cleavage and reverse IMAC which made the purification short and hence less destabilising to the protein by reducing any possible delipidation, forced aggregation or precipitation. In addition, processes such as concentration of protein, binding and purification through a high-pressure SEC column increase the physical stress on the protein and may cause partially unfolded receptors to aggregate or precipitate.

Since the data presented show that the affinity for antagonists ZM241385 and theophylline and agonist NECA are similar for membrane-bound, solubilised receptor and pure protein purification we can be confident that the processing has not compromised the pharmacological characteristics of the protein. The affinity values obtained are also similar to those reported for the receptor expressed in a mammalian cell line (Fredholm et al., 2001).

In summary, a detailed functional characterisation of the WT and C-terminal truncated forms of the human adenosine A2aR expressed in P. pastoris has been carried out. Optimisation of the purification protocol in order to obtain highly pure protein that is also highly functional was achieved. In this case it was vital to look at both the factors in order to produce a protein suitable for further studies. Such a protein population can be used for further structural and biophysical characterisation as described in Chapter 4.
4 Biophysical characterisation of the purified V334

adenosine $A_{2a}$ receptor form
4.1 INTRODUCTION

The data presented in Chapters 2 and 3 describes the large-scale expression and purification of functional adenosine A\(_{2a}\)R. One of the receptor constructs tested, V334 adenosine A\(_{2a}\)R with a truncated C-terminus consistently expressed at the highest level and yielded milligram quantities of pure and highly functional protein. In order to further assess the suitability of the V334 adenosine A\(_{2a}\)R for crystallisation trials it was essential to perform further analysis to determine its long term structural and conformational stability.

One of the prerequisites for producing diffraction quality crystals is to obtain milligram quantities of protein which can maintain its native folded state over a long period of time. This is particularly important as the process of expression, purification followed by crystallisation screening and time required for crystals to develop often lasts several days or even weeks. The protein needs to be tested for its aggregation and degradation state as an unstable or heterogeneous population of receptor is likely to reduce the likelihood of forming crystals. Similarly maintenance of the native folded state of the receptor is also essential for crystallisation. Post-translational modifications such as glycosylation can be an extra potential source of heterogeneity of the receptor population, hence the glycosylation profiling of the protein is useful.

The research presented in this chapter describes the analysis of the purified V334 A\(_{2a}\)R by a number of biophysical techniques; SDS-PAGE, size exclusion chromatography (SEC), circular dichroism (CD) and mass spectrometry. The long term functional stability of the purified V334 adenosine A\(_{2a}\)R was also assessed through radioligand binding assays.
The thermostability of the protein in a number of different buffer conditions was also analysed since pre-screening can facilitate the identification of conditions suitable for crystallisation. The thermostability of the receptor was assessed using a fluorescent thermal stability assay (Alexandrov et al., 2008). (Hanson et al., 2008). The assay is based on the reporter dye, a thiol specific fluorochrome N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl] maleimide (CPM dye), used in the assay. The fluorochrome binds and reacts readily with the free thiols of cysteine residues. As the membrane protein unfolds, the normally buried cysteine residues become available to interact with the dye and the level of fluorescence gives an indication of the rate of unfolding. This assay is thus a rapid, high throughput way of comparing the relative stability of protein in different buffer conditions.

As previously discussed one of the key parameters for obtaining membrane protein crystals is the detergent choice (Iwata, 2003; Ostermeier et al., 1997). Several groups are attempting to design new detergent-like molecules which confer a higher degree of stability to membrane proteins and therefore make them more likely to crystallise. As part of the analysis of the V334 adenosine A2aR we tested the stability of the protein in a number of novel amphiphiles designed by Dr. Pil Seok Chae (University of Wisconsin-Madison) compared to currently available detergents. These amphiphiles have a much more rigid chemical structure which may be more suited to the crystallisation process. The MNG amphiphiles used in this study are so called due the their hydrophilic group being derived from Maltose-Neopentyl Glycol (MNG) whereas the lipophilic chain is made up of n-decyl chains. The MNG class of amphiphiles has a highly branched structure with two hydrophilic and hydrophobic groups projecting from the central quaternary carbon core (Figure 4.1 A and B). A branched hydrophilic portion has been reported to improve protein extraction (Chae
The core quaternary carbon imparts a conformational restriction to the molecule compared to a linear alkyl chain present in conventional detergents (McQuade et al., 2000). The MonoPod Amphiphile (MPA-12) also tested has a structure closer to that of the commercially available detergents (Figure 4.1 C). It is a linear molecule with a hydrophilic head and a lipophilic tail but contains the core quaternary carbon. The results presented in this chapter have confirmed that the V334 adenosine A2aR is extremely suitable for structural studies and identified a number of buffer conditions for crystallisation trials.

Figure 4.1 Structures of the amphiphiles used in this study
The structure of MNG-1 (A), MNG-14 (B) and MPA-12 (C) is presented. The MNG molecules have two hydrophilic and two hydrophobic chains projecting from the central carbon atom whereas the MPA is a linear molecule with single hydrophilic and hydrophobic group.
4.2 MATERIALS AND METHODS

4.2.1 Materials

CPM dye was obtained from Invitrogen Paisley UK, 96 well plates were purchased from Greiner Gloucestershire UK, amphiphiles were a gift from Dr. Pil-Seok Chae (University of Wisconsin-Madison), MemGold and MemSys/Memstart were purchased from Molecular Dimensions Suffolk UK, the Superdex 200 10/300 GL column was obtained from GE Biosciences Buckinghamshire UK. All other chemicals were either purchased from Sigma-Aldrich Gillingham UK or as previously described.

4.2.2 Analysing long term stability of purified receptor

4.2.2.1 Analysis of susceptibility to degradation

The susceptibility of the protein to proteolytic degradation was assessed at typical crystallisation temperatures of 4°C and 20°C. Purified receptor was incubated at 4°C and 20°C for a period of 15 days with aliquots collected at 0, 4, 8 and 15 days. The fractions were stored in SDS running buffer at -20°C prior to SDS PAGE analysis on 4-12% Bis-Tris gels.

4.2.2.2 Analysis of functional stability

The purified V334 adenosine A2aR was incubated at 4°C. Samples were assayed by radioligand binding assay at a saturating concentration (6nM) of [3H] ZM241385 on day 0 and 30 (Section 3.2.5).
4.2.3 CPM analysis

The CPM dye was dissolved in DMSO at 4 mg/ml and stored at -80°C. The dye stock was diluted 1:40 in dye dilution buffer (20mM Tris-HCl (pH 7.5), 150mM NaCl, 0.03% DDM, 5mM EDTA) and incubated for 5 minutes at room temperature prior to use. Care was taken to protect the dye from light to prevent photo bleaching. The thermal denaturation assay was performed in Greiner 96 well plate with the purified V334 adenosine A$_2a$R protein at 10mg/ml diluted in test buffer conditions (1:150) to give a final concentration of 10 µg of protein per assay condition. A volume of 3µl of diluted CPM dye was added to each condition to give the final reaction volume of 153 µl. The reaction was monitored for 130 minutes at 40°C using a microplate spectroflurometer set at an excitation wavelength of 387 nm and an emission wavelength of 463 nm. Readings were taken every 5 minutes after automatic agitation of the plate.

Radioligand binding studies (Chapter 3) indicated that 0.05% DDM (5xCMC) in the presence of 0.01% CHS maintains the functional activity of the receptor hence this condition was compared through the CPM assay to DDM or amphiphiles at 10xCMC values in test buffer (25mM Tris-HCl pH 7.5, 150mM NaCl, 10% glycerol). The effect of ligands $10^{-6}$ M ZM241385, $10^{-5}$ M NECA and $10^{-5}$ M theophylline were compared in test buffer containing 0.05% DDM and 0.01% CHS.

The highest raw fluorescence count within the data set is used as the value for 0% folded protein remaining and was used to calculate the percentage of relative folded protein remaining at a regular 5 minute interval up to 130 minutes at 40°C. Data were fitted to one-phase decay curves using GraphPad Prism. The percentage of folded protein remaining after 130 minutes of incubation at 40°C was used to assess the effect of different agents on the thermostability of the receptor.
4.2.4 Circular Dichroism analysis

V334 adenosine A2aR purified in 0.05% DDM + 0.01% CHS was concentrated to 10mg/ml for analysis in different buffer conditions. The detergents were used at ten times their critical micelle concentration (CMC) 0.05% DDM and 0.5% DDM in comparison to the standard 0.05% DDM+ 0.01% CHS. Typically 4 µl of concentrated protein was added to 250 µl of buffer to achieve a final concentration of ~0.2mg/ml. Such a dilution allowed an exchange of protein into the buffer condition being tested. CD spectra were obtained with a ChiraScan™ spectrometer in a 1 mm path-length cuvette at a scanning speed of 0.5 nm/s and a bandwidth of 1 nm over a range of 190nm to 260 nm. The spectra for each condition were measured from 10°C -100°C at an interval of 5°C.

The CD signal from 221 nm to 223 nm was averaged for each temperature and fitted to a non-linear regression curve using GraphPad Prism to obtain a transition midpoint T<sub>m</sub> corresponding to the temperature at which 50% of the receptor is in folded state.

4.2.5 Analytical SEC

Size exclusion was used as an analytical step and was similar to as described in Chapter 3. Superdex 200 was equilibrated with SEC buffer (20mM Tris-HCl pH-8, 5% glycerol, 250mM NaCl, 0.05% DDM, 0.01% CHS). A volume of 500-1000µl of protein at 5-10 mg/ml was loaded onto the column. The separation was run at 1 ml/min and 1 ml fractions were collected.
4.2.6 Glycosylation profiling

4.2.6.1 Sample preparation
V334 adenosine A2aR was purified as described in Chapter 3. Two different samples were passed through glycosylation profiling. The receptor was either purified through Co\(^{2+}\) metal affinity chromatography (IMAC) and monoQ chromatography containing the α factor signaling peptide and purification tags or through M2-Flag tag affinity chromatography and reverse IMAC following TEV cleavage to remove all the tags.

Protein was concentrated at 10mg/ml and exchanged into ammonium bicarbonate through dialysis o/n. A 50 kDa molecular weight cut off cassette was used for the dialysis. The sample was passed on for profiling in the lab of Dr. Stuart Haslam (Imperial College London). The protocol used for analysis is presented below.

4.2.6.2 Reduction and Carboxymethylation
The sample was lyophilized and then reduced in 200 µl of 50 mM Tris-HCl buffer, pH 8.5, containing 2 mg/ml dithiothreitol. Reduction was performed under a nitrogen atmosphere at 37°C for 1 h. Carboxymethylation was carried out by the addition of iodoacetic acid (6-fold molar excess over dithiothreitol), and the reaction was allowed to proceed under a nitrogen atmosphere at room temperature in the dark for 2 h. Carboxymethylation was terminated by dialysis against 4 × 4.5 l of 50 mM ammonium bicarbonate, pH 8.5, at 4°C for 48 h. After dialysis, the sample was lyophilized.

4.2.6.3 Digestion with chymotrypsin
The reduced carboxymethylated protein was digested with TLCK pre-treated bovine pancreas chymotrypsin (EC 3.4.21.1, Sigma-Aldrich), for 16 h at 37°C in 50 mM
ammonium bicarbonate buffer (pH 8.4). The products were purified by C₁₈-Sep-Pak (Waters Corp.) as described (Dell et al., 1994).

### 4.2.6.4 Sep-Pak* separation of released glycans from peptides

The reverse-phase C₁₈ Sep-Pak cartridge was primed sequentially with 5 ml methanol, 5 ml 5 % acetic acid (v/v) and 5 ml propan-1-ol before being re-equilibrated with 15 ml 5 % acetic acid (v/v). The sample was then dissolved in a minimum volume of 5 % acetic acid (v/v) and loaded directly onto the Sep-Pak. Elution was achieved using 5 ml of 5 % acetic acid (v/v), followed by 4 ml each of 20 % and 40 %, propan-1-ol in 5 % acetic acid (v/v). The 20 % and 40 %, propan-1-ol elutions were collected, had their volumes reduced in a Speed Vac, then were combined and lyophilised (Dell et al., 1994).

### 4.2.6.5 PNGase F digestion of glycopeptides

PNGase F (EC3.5.1.52, Roche Molecular Biochemicals, Lewes, UK) digestion was carried out in 200 µl ammonium bicarbonate (50 mM, pH 8.5) for 16 h at 37 °C using 5 U of enzyme. The reaction was terminated by lyophilisation and the released N-glycans were separated from peptides and O-glycopeptides by Sep-Pak C₁₈ (Waters, Elstree, UK) as described (Dell et al., 1994). Only the 5 ml of 5 % acetic acid elution containing N-glycans was collected.

### 4.2.6.6 Derivatisation for MALDI-TOF and tandem mass spectrometry analysis

Permethylation was performed using the sodium hydroxide procedure, as described previously (Dell et al., 1994). 1 g of sodium hydroxide pellets were crushed in a glass mortar with 3 ml distilled, anhydrous DMSO. A volume of 1 ml of the resulting slurry was mixed with 500 µl of methyl iodide and added to the lyophilised sample. The
mixture was then shaken for 15 min before the reaction was quenched by dropwise addition of water. The permethylated sample was then extracted into 1 ml of chloroform and washed four times with 1 ml of water. The chloroform was then removed under a stream of nitrogen.

4.2.6.7 Sep-Pak® separation of permethylated glycans
The reverse-phase C18 Sep-Pak cartridge was primed sequentially with 5 ml methanol, 5 ml water and 5 ml acetonitrile before being re-equilibrated with 15 ml of water. The lyophilised permethylated oligosaccharide sample was then dissolved in a minimum volume of methanol:water (1:1 v/v) and loaded directly onto the Sep-Pak. Elution was achieved using 5 ml of water followed by 3 ml each of 15%, 30%, 50% and 75% acetonitrile in water (v/v). Each elution step was collected, reduced in volume on a Speed Vac and lyophilised (Dell et al., 1994).

4.2.6.8 Mass spectrometric analysis
MALDI-TOF data were acquired on a Voyager-DE STR mass spectrometer (Applied Biosystems, Foster City, CA) in the reflectron mode with delayed extraction. Permethylated samples were dissolved in 10 µl of methanol and 1 µl of dissolved sample was premixed with 1 µl of matrix (20 mg/ml 2,5-dihydroxybenzoic acid (DHB) in 70% (v/v) aqueous methanol), spotted onto a target plate and dried under vacuum.

4.2.7 Amphiphiles
The critical micelle concentrations (CMC) of the amphiphiles tested are given in Table 4.1. The amphiphiles were dissolved in water to make stock solutions and stored at 4°C. Stock solutions were made at 150xCMC for MPA12 and 300xCMC for the MNG amphiphiles. Previous work (Pil Seok Chae et al Manuscript in progress)
evaluated the amphiphiles at 10xCMC and higher and established that 10xCMC was optimum for stabilising a range of membrane proteins. The stability of V334 adenosine A$_{2a}$R was assessed in buffer containing amphiphiles at 10xCMC. These were compared to the protein in the standard buffer conditions of 0.05% DDM (5xCMC) and 0.01% CHS.

<table>
<thead>
<tr>
<th>Amphiphile</th>
<th>CMC (mM)</th>
<th>CMC (wt %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNG 1</td>
<td>0.017</td>
<td>0.0019</td>
</tr>
<tr>
<td>(MNG Am-10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MNG 14</td>
<td>0.010</td>
<td>0.0010</td>
</tr>
<tr>
<td>(MNG Alk-10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MPA 12</td>
<td>0.660</td>
<td>0.0357</td>
</tr>
<tr>
<td>(L-Et-3-10)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 4.1 The critical micelle concentration of novel Amphiphiles**

The CMC values for amphiphiles used in the study are presented in mM and wt% (amount required in gm/100 ml).

4.2.8 **SEC of receptor in different detergents**

V334 adenosine A$_{2a}$R at 10mg/ml was diluted (1:100) in test buffer containing the specific detergent at 10xCMC. The samples were incubated at 40°C for 120 minutes. Protein purified in 0.05% DDM + 0.01% CHS was used as control sample. A Superdex 200 10/300 column was pre-equilibrated with test buffer (20mM Tris-HCl (pH 7.5), 150mM NaCl, 5% glycerol) supplemented with the same detergent as the sample. A volume of 500 µl of each sample was loaded onto the column following separation at 0.5ml/min and 500 µl fractions were collected.
4.2.9 Crystallisation trials

Sitting drop vapour diffusion crystallisation trials were set up for the purified V334 adenosine A\textsubscript{2a}R. The protein was purified in the presence of 0.05% DDM + 0.01% CHS and concentrated to 13 mg/ml in the presence of 30\textmu M ZM241385 and screened against MemGold (Appendix 7.2) and MemStart/MemSys (Appendix 7.3) crystallisation kits. The plates were prepared using the Mosquito robot with a volume of 0.1 \mu l of reservoir and 0.7 \mu l of protein solution. The plates were incubated at 20°C and observed regularly under a microscope for a period of 15 days.

The purified V334 adenosine A\textsubscript{2a}R at 13 mg/ml containing 200 \mu M ZM241385 was used for lipidic cubic phase (LCP) crystallisation trials. Screens were set up with MemGold, MemStart/MemSys and a grid screen (Appendix 7.4) based on the conditions used to obtain well diffracting crystals of the adenosine A\textsubscript{2a}R T4 lysosome constructs (Jaakola \textit{et al.}, 2008). The screen contained PEG 400 (range of 28-32\% (w/v)), lithium sulphate (range of 180 to 220 mM) and sodium citrate (pH 6.5) (Range of 5.5 to 6.5).

The LCP mixture contained lipids monoolein: cholesterol: receptor 54\%: 6\%: 40\% (w/w) respectively (Figure 4.2). The crystallisation setup was carried out at ~22°C on glass sandwich plates. The Flexsus robot (Anachem) dispensed 50 nl of LCP and 90 nl of reservoir onto the glass plates. The plate was incubated at 20°C and observed for a period of 20 days however no crystals could be observed in any of the screens tested.
Figure 4.2 A lipid mixing device for lipid cubic phase setup

A coupled syringe device is used to mix lipid solution containing monoolein and cholesterol with the protein. The lipid (yellow) is loaded into one syringe and protein (purple) to the other at a ratio of 60:40 by weight. The syringes are sealed by Teflon and connected by a syringe connector. The syringes are held in horizontal position between the thumb and finger. Force is applied by moving the piston by thumb about 20 times to allow formation of a homogenous protein lipid mixture (lilac). The mixture is then used for setting up lipid cubic phase crystallisation screens (Caffrey et al., 2009).
4.3 RESULTS

4.3.1 The purified V334 adenosine A₂aR is highly stable

The long term stability of the purified V334 adenosine A₂aR was compared using SDS-PAGE analysis of protein incubated at 20°C or 4°C for 15 days. On day 0 the V334 adenosine A₂aR is observable as a prominent band at approximately 28 kDa (Figure 4.3). No bands of smaller size are visible over the course of the 15 day incubation period demonstrating that the V334 adenosine A₂aR does not undergo any detectable degradation at either temperature. An equal amount of protein was loaded in each well however the band intensity for the samples incubated at 20°C is less than that of 4°C. It is possible that this apparent decrease in band intensity is due to aggregation of the protein.

![Figure 4.3 Long term stability analysis of purified V334 adenosine A₂aR](image)

*Comassie-stained SDS-PAGE gel of purified V334 A₂aR incubated at 4°C and 20°C. Lane 1 contains the molecular weight markers while Lane 2 contains pure V334 A₂aR on day 0. Lanes 3,4 and 5 contain protein stored at 4°C for 4, 8 and 15 days respectively. Lanes 6, 7 and 8 contain protein stored at 20°C for 4, 8 and 15 days respectively.*
Radioligand binding assay measurements carried out on the purified V334 adenosine A2aR incubated at 4°C over a period of approximately 30 days indicated that the receptor exhibited a high level of functional stability. The specific ligand binding achieved at saturating concentration of ligand on Day 0 was 21.7 nmol/mg which had only dropped to 17.6 nmol/mg, ~84% of the starting activity, after 30 days.

4.3.2 Ligands help stabilise the receptor

Thermostability analysis of the protein revealed that the purified V334 adenosine A2aR was highly unstable when exchanged into buffer containing only DDM in the absence of CHS (Figure 4.4). After an incubation of 130 minutes at 40°C, ten times more folded protein remained in the buffer containing DDM+ CHS as compared to DDM alone. These results are also in agreement with the functional recovery of the V334 adenosine A2aR upon solubilisation in DDM ± CHS (Section 3.3.).

The effect of ligands on stability of the protein was tested at a concentration ten times higher than their K_i to allow saturation of receptor binding sites. Addition of all three ligands tested resulted in an increase in stability compared to DDM+CHS (Figure 4.4). Approximately 60-75% of folded receptor was present in the presence of ligands after 130 minutes of incubation as compared to 14% (Table 4.2) for the receptor in the same buffer conditions lacking the ligands. The high affinity antagonist ZM241385 maintained the highest proportion of protein in folded state among all the conditions tested.
Figure 4.4 Assessing the effect of ligands on stability of receptor

V334 adenosine A2aR was purified in DDM+CHS (blue) and exchanged into DDM (maroon) and DDM+CHS buffer containing 10^{-6} M ZM241385 (green), 10^{-4} M theophylline (salmon), 10^{-5} M NECA (plum) for CPM analysis. The unfolding of protein in different conditions was monitored at 40°C for 130 minutes by a spectrophotometer set at an excitation wavelength of 387 nm and emission wavelength of 463 nm. Relative fraction of folded protein remaining plotted. The amount of relative folded protein remaining in each condition at the end of the experiment is presented as a table.

<table>
<thead>
<tr>
<th></th>
<th>DDM</th>
<th>DDM+CHS</th>
<th>NECA</th>
<th>theophylline</th>
<th>ZM241385</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative folded protein (%)</td>
<td>1.4</td>
<td>13.9</td>
<td>58.3</td>
<td>66.1</td>
<td>76.0</td>
</tr>
</tbody>
</table>

4.3.3 CHS is important to maintain the receptor in a folded state

It has been established (Chapter 3) through radioligand binding analysis that CHS is essential to maintain the functionality of receptor. The effect of CHS on the structural integrity of the receptor was assessed through CD analysis. A CD spectrum for α helical protein (Figure 4.5A) and a T_m of 60°C (Figure 4.5B) was obtained for V334 adenosine A2aR in DDM+CHS. Upon exchange into buffer containing either DM or DDM without CHS the CD spectra suggested a disordered protein (Figure 4.5C). It was not possible to obtain a T_m in this case.
Figure 4.5 CD analysis of V334 adenosine A₂aR to assess the role of CHS

The CD spectra (A) obtained for purified V334 adenosine A₂aR suggested a α-helical protein structure. A T_m of 60°C was obtained from the melting curve of protein purified in DDM+CHS (B). Purified protein exchanged into DM (C) or DDM (D) lacked a typical α-helical protein curve.

4.3.4 The purified protein is monodisperse

SEC of the purified protein revealed a minor aggregation peak (retention volume ~8ml) followed by a major peak at ~14ml. This suggests that the receptor is largely a monodisperse.
The V334 adenosine A$_2$R was purified in the presence of DDM+CHS by Flag tag affinity chromatography and reverse IMAC prior to loading onto a pre-equilibrated Superdex 200 column. A volume of 500 µl containing 100 µg of protein was loaded onto the column. The peak at 8 ml is likely to correspond to an aggregated protein based on the molecular weight. The major monodispersed peak at 14 ml corresponds to the molecular weight of a monomeric form of protein.

4.3.5 The purified protein is glycosylated

The mass spectrometry analysis of V334 adenosine A$_2$R purified through IMAC and monoQ anion exchange chromatography containing the purification tags and α factor signaling peptide indicated that the protein was glycosylated. The N-glycan structures observed were high-mannose type, which is a common feature for protein expressed in *P. pastoris*. As indicated by the partial MALDI-TOF MS glycosylation profile (Figure 4.7) a series of [M+Na]$^+$ molecular ions were observed fully consistent with a range of high mannose N-glycans (m/z 2192.4-3212.4, Man$_{8-13}$GlcNAc$_2$). Hence the recombinant protein is heterogeneously glycosylated. As reported in Chapter 3 the expressed receptor was unprocessed and contained the α-factor sequence which is predicted to have three potential N-glycosylation sites. It was thus impossible from this data alone to establish whether all of these sugar groups were present on the receptor itself or also the signaling peptide.
The purified V334 adenosine A$_2$R containing the signaling peptide was assessed for N-glycans. A heterogeneous population of high-mannose sugars typical for *P. pastoris* was observed.
To address this issue analysis of the purified V334 adenosine A$_{2a}$R lacking the α-factor signal peptide was performed. [M+Na]$^+$ molecular ions were observed at m/z 2192.6 and 2396.8 consistent with compositions of Man$_{8-9}$GlcNAc$_2$ (data not shown). Thus providing convincing evidence that the V334 adenosine A$_{2a}$R is N-glycosylated.

4.3.6 Amphiphiles improve the thermostability of receptor

We also compared the thermostability of the V334 adenosine A$_{2a}$R in a range of amphiphiles. All the amphiphiles tested improved the thermostability of the protein in comparison to DDM±CHS. In the case of MPA12 a dramatic increase in thermostability was observed with 94% of the receptor population remaining in a folded state following incubation at 40°C for 130 minutes (Figure 4.8).

The other amphiphiles MNG 1 and MNG 14 are slightly better than DDM + CHS with 20% to 30% folded protein remaining after 130 minutes of incubation as compared to 14% for DDM+CHS. Almost no folded protein (1.4%) remains in DDM in the absence of CHS.
V334 adenosine A2aR was purified in DDM+CHS (orange) and exchanged into DDM (maroon), MNG1 (green), MNG 14 (pink), MPA 12 (lilac) for CPM analysis. The unfolding of protein in different conditions was monitored at 40°C for 130 minutes by a spectrofluorometer set at an excitation wavelength of 387nm and emission wavelength of 463 nm. The relative fraction of folded protein remaining was fitted to a one phase decay curve. The amount of relative folded protein remaining in each condition at the end of the experiment is presented as a table.

<table>
<thead>
<tr>
<th>DDM</th>
<th>DDM+CHS</th>
<th>MNG 1</th>
<th>MNG 14</th>
<th>MPA12</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0</td>
<td>10.0</td>
<td>4.0</td>
<td>20.0</td>
<td>96.0</td>
</tr>
</tbody>
</table>

Figure 4.8 Screening for amphiphiles to improve protein stability

The receptor state in solution is different in different amphiphiles and detergents

To establish the effect of conventional detergents and amphiphiles on the structural integrity of the receptor, SEC with selected agents was carried out. The SEC profile obtained at t=0 for the receptor in 0.05% DDM and 0.01% CHS (Figure 4.6) was routinely obtained and is used as a control.

After heating the sample at 40°C for 120 minutes in buffer containing 5xCMC (0.05%) DDM + 0.01% CHS a profile similar to the control profile with a slightly
larger aggregation peak can be observed (Figure 4.9 A). An additional peak at ~ 20 ml could be a degradation product however due to low protein concentration SDS – PAGE analysis could not be performed.

The profiles obtained for receptor incubated in buffer containing 10xCMC MPA12 (Figure 4.9 D) or DDM (Figure 4.9 B) after incubation at 40°C for 120 minutes are much narrower than that obtained for receptor in DDM+CHS. The elution profile for the receptor incubated at 40°C for 120 minutes in buffer containing MNG 14 reveals a much higher proportion of the protein is aggregated evidenced by the presence of a large void peak eluting at with a retention volume of ~8ml. In addition the major protein peak elutes much earlier than in the other conditions tested, at ~11ml is observed (Figure 4.9 C). It is possible that the presence of MNG-14 encourages higher oligomer formation.
Figure 4.9 Comparison of dispersity of protein in commercially available detergents and amphiphiles

Purified V334 adenosine A$_{2a}$R was exchanged into DDM+CHS, DDM, MNG 14 and MPA 12 and incubated at 40°C for 120 minutes. Protein samples after incubation in each agent were loaded onto a Superdex 200 gel filtration column pre-equilibrated in buffer containing the appropriate agent. The void volume for the column is marked with an arrow.

4.3.8 Crystallisation trials set up through vapour diffusion and Lipid Cubic Phase

A highly functional and stable protein population could be obtained routinely for the V334 adenosine A$_{2a}$R. Crystallisation screens were set up using vapour diffusion and lipid cubic phase techniques with optimised crystallisation screens. These were checked routinely for crystals, but only precipitation in several conditions was observed over a period of 15-25 days.

The CPM and SEC analysis suggested that MPA 12 had the ability to maintain the receptor in a more conformationally stable state than DDM+CHS. MPA 12 was thus tested as an additive for crystallisation screens through the vapour diffusion
screen. Precipitation under several conditions was observed in the plates over a period of 7 days however no crystals were obtained.
4.4 DISCUSSION

The aim of the analyses described in this chapter was to further characterise the purified V334 adenosine A$_{2a}$R and assess the quality and thus suitability of the protein for structural studies. In addition we aimed to screen some buffer conditions using a variety of means in order to submit the optimised samples for crystallisation trials using a variety of methods. Thermostability analysis as described in Chapter 3 to determine the $T_m$ of the V334 adenosine A$_{2a}$R solubilised in DDM+CHS ($T_m = 49^\circ$C) indicated that the receptor was highly stable. However we also wished to assess the stability of the purified receptor. The WT adenosine A$_{2a}$R has been reported (Chapter 2) to undergo C-terminal proteolytic degradation (Weiss et al., 2002). In contrast the purified V334 A$_{2a}$R demonstrated a high level of resistance to degradation even following incubation at 20$^\circ$C for extended periods of time, further proving the superiority of this construct for downstream studies. In addition, long term functional analysis of the protein sample indicated that the unliganded receptor retained 80 % functionality after 30 days storage at 4$^\circ$C. This shows a remarkable functional stability for the purified receptor in the absence of a stabilising ligand.

The CPM assay results where a higher percentage of folded protein was present in buffer containing DDM+CHS as compared to DDM alone were in agreement with the results presented in Chapter 3 for the solubilised receptor. The importance of CHS for V334 adenosine A$_{2a}$R stability was further supported by CD analysis which yielded a $T_m$ of 59$^\circ$C measured through CD analysis for receptor in buffer containing DDM+CHS. In contrast there was a complete lack of ordered structure for the receptor in DM or DDM alone. This indicated that CHS not only improves the functional recovery of receptor but also helps to maintain the secondary
structure. This could in part be due to the lateral pressure applied by the bulky lipids present in the protein detergent micelle (Paila et al., 2009). A bigger micelle due to CHS may also help protect the hydrophobic ends better thus reducing aggregation of the receptor.

The radioligand binding assay suffers from the limitation of not being able to assess receptor functionality when bound to a ligand. The CPM assay has allowed assessment of the affect of ligands on stability indicating an improvement in thermostability in the presence of ligands. In the presence of a high affinity antagonist, ZM241385, five times more folded protein was present after incubation at 40°C in comparison to buffer containing DDM+CHS only. Such a result implies that the long-term stability of the receptor may be improved further in the presence of ligands. As a stable protein in the absence of ligands is routinely produced which is capable of binding both agonist and antagonist with native affinities, this can be used to set up crystallisation screens for both antagonist or agonist bound form.

Whilst many groups have focused on the modification of individual membrane proteins in order to improve the stability and thus increase the probability of obtaining well diffracting crystals an alternative approach could be of screening alternative detergents or novel amphipathic molecules (Seddon et al., 2004). Recent work has characterised a group of new amphiphiles (Chae et al. manuscript in progress), which have been shown to improve the thermostability and conformational stability of a range of membrane proteins. The CPM assay indicated that the receptor was more thermostable in the amphiphiles MPA 12 and MNG 14 than in commercially available detergents such as DDM. As the receptor was more thermostable in the amphiphiles tested in comparison to DDM+CHS it is possible
that these novel molecules could perhaps be useful as additives or alternatives to standard detergents during purification and crystallisation.

The increase in thermostability was strikingly evident in the case of MPA 12 but it was necessary to examine whether this was a result of aggregation. Aggregation of the protein may make the cysteine residues unavailable to the CPM dye and the apparently highly folded protein may thus be an artifact. The SEC profiles obtained for the V334 adenosine A2aR in MPA 12 showed a more monodisperse peak in comparison to that obtained for the protein in DDM+CHS indicating that the homogeneity of the receptor in MPA 12 is slightly better than in DDM+CHS. One curious result was that the SEC profile for the receptor in DDM alone was very similar to that obtained in MPA 12. This is surprising since the CD spectroscopy suggested that the V334 adenosine A2aR in DDM had little secondary structure, a finding supported by the thermostability and functional results. MPA 12 has a linear structure similar to DDM and it is possible that it may not be as stabilising an agent for the receptor as it appears from CPM assay. A highly branched MNG 14 molecule may be flexible and able to accommodate the hydrophobic protein surface better and help in improving the thermostability of the receptor. However the presence of a heterogeneous population as seen through the SEC profile will serve as a hindrance in the crystallisation process. It remains to be tested if these forms of receptor are more stable in terms of their functional or conformational integrity. Further analysis through radioligand binding assay or CD spectroscopy may improve our understanding of the effects of these novel molecules on the receptor.

The adenosine A2aR has a predicted N-glycosylation site (Jaakola et al., 2008). Mass spectrometry analysis of the receptor molecule with and without the α factor signal peptide indicated that the protein was glycosylated. Previous studies have
described expression and purification of a non-glycosylated mutant adenosine A2aR (Fraser, 2006). The recent structure of the adenosine A2aR-T4 lysozyme construct was obtained for a protein which was expressed in the glycosylated form in insect cells and then deglycosylated by enzymatic treatment using PNGaseF following purification (Jaakola et al., 2008). It is likely that any heterogeneity of sugar groups present or their large and disordered structure may interfere with crystal lattice formation. However structures obtained for rhodopsin (Okada et al., 2000; Palczewski et al., 2000) indicate that there are examples where the presence of sugar groups does not preclude three-dimensional crystal lattice formation.

The screening process in order to obtain well-diffracting crystals is complex, time-consuming and expensive. In addition, from the GPCR structures recently published it has been established that stability is key for successful crystallisation. A highly functional and structurally stable preparation of V334 adenosine A2aR could be routinely produced. Preliminary crystallisation screens were carried out however no crystals could be obtained in these conditions. Even though a stable and properly folded receptor is a prerequisite for successful crystallisation it is not the only factor involved. An attempt to address this issue while improving the chances of crystallising the protein is discussed in the following chapter.
5 Alternative approaches to producing an adenosine $A_{2a}$ receptor form suitable for structural studies
5.1 **INTRODUCTION**

The successful crystallisation and subsequent structure determination of GPCRs has involved overcoming several practical obstacles such as expression of large quantities of functional receptor (Sarramegna *et al.*, 2003), stabilisation of receptor in detergents (Tate *et al.*, 2009), stabilisation of the flexible domains through protein engineering (Rosenbaum *et al.*, 2007) or interaction with molecules such as antibodies (Rasmussen *et al.*, 2007) and peptides (Scheerer *et al.*, 2008) and in some cases combination with non-standard crystallisation techniques such as lipid cubic phase (Caffrey, 2003) or bicelle crystallisation (Faham *et al.*, 2002). In order to increase the chances of obtaining a protein suitable for crystallisation trials it is often necessary to try several approaches in parallel. This chapter describes a few approaches additional to those detailed in the previous chapters for the production of protein suitable for structural studies.

### 5.1.1 GFP adenosine A2aR fusion construct

The use of GFP fusion proteins has been successfully utilised for production (Sarramegna *et al.*, 2002; Wedekind *et al.*, 2006) and pre-crystallisation screening (Kawate *et al.*, 2006; Newstead *et al.*, 2007) of a number of integral membrane proteins including GPCRs.

One well characterised system expressing the target protein as a fusion with GFP (Drew *et al.*, 2008) involves expression of the target protein in *S.cerevisiae* as a fusion with C-terminal GFP-His tag. Production of GPCRs suitable for obtaining diffraction quality crystals is still a bottleneck for structural studies and involves assessment at each stage of the protein production. The system has been designed and
used as a pipeline for rapid screening and selection of targets suitable for structural studies.

In order to select a membrane protein that is optimal for structural determination, many expression constructs bearing carboxy and/or amino terminal truncations or mutations to stabilise the receptor are often generated. The ability of *S. cerevisiae* to undergo homologous recombination is utilised and considerably speeds up generation of expression constructs compared to standard restriction digest cloning. An expression vector encoding for a galactose inducible promoter, a TEV cleavage site, GFP and an octa Histidine tag is transformed in *S. cerevisiae* along with an amplified target protein gene (WT adenosine A2aR). Homologous recombination within the *S. cerevisiae* cells results in a construct with the target protein fused to a C-terminally located GFP. The presence of a TEV cleavage site downstream of the protein allows for removal of the purification tags and GFP at the final stage of the protein processing allowing crystallisation of the target protein only. The target protein can be expressed in 10 ml analytical scale cultures where measurement of whole cell fluorescence provides a rapid assessment of expression levels. A 10 ml culture provides sufficient amount of protein to allow analysis of the profile of the expressed receptor by in-gel fluorescence. The presence of GFP allows selective detection of target protein in crude cell lysates, at low concentrations or in the presence of large quantities of contaminating material (Drew *et al.*, 2006; Drew *et al.*, 2005).

As GPCRs are particularly sensitive to precipitation or degradation while in detergent solution the presence of GFP enables measurement of fluorescence before and after solubilisation to select for detergents which allow maximum extraction of protein from the membrane and maintain it in soluble form. The GFP allows the use
of fluorescence size exclusion chromatography (FSEC) to assess the stability of the crude solubilised target protein in a range of detergents (Kawate et al., 2006). The FSEC profile allows the detection of aggregates indicated by a large void peak while presence of a peak corresponding to free GFP is a marker for proteolytic degradation. This system thus provides an efficient and rapid method of selecting the best expression and solubilisation conditions (Figure 5.1). An adenosine A2aR-GFP construct was generated and submitted to preliminary characterisation.
**Figure 5.1 A flowchart representing the GFP screening workflow**

(A) The gene of interest is amplified through PCR by primers with ends coding for homologous recombination sites and transformed into competent *S.cerevisiae* cells along with linearised GFP plasmid. Through the process of homologous recombination the target gene is integrated within the GFP plasmid. (B) Cultures of 10 ml are grown up from the single colonies. The whole cell fluorescence of cultures is used to calculate the amount of target protein expressed. (C) Membranes are prepared from the cell pellet by breaking with glass beads and ultracentrifugation. The in-gel fluorescence profile of membrane samples confirms the integrity of the protein. (D) The membrane sample can be solubilised in different detergents to identify the detergents which provide high extraction yields of protein. The FSEC profile further confirms the state of the receptor in solution. Figure adapted from Drew et al., 2008.

### 5.1.2 Adenosine A$_2a$R and the G-protein

As described in detail in Chapter 1 activation of a GPCR by binding of an agonist leads to conformational changes within the receptor allowing binding of a
heterotrimeric G-protein. An activated G-protein leads to stimulation of downstream effector molecules such as adenylate cyclase which leads to production of second messenger molecule and initiates an intracellular signaling cascade. GPCR/Gα fusion proteins were first used as a means of studying receptor G-protein signaling more than fifteen years ago (Bertin et al., 1994). Since then they have been used to study many aspects of the signaling pathway such as the role of post-translational modification and the effect of mutations in both receptor and G-protein (Milligan et al., 2007).

The key properties of the GPCR/Gα fusion proteins which have made them particularly useful for functional studies is the 1:1 stoichiometric ratio of the receptor and G-protein, the close proximity of the two proteins and the close association of the Gα protein with the membrane (Seifert et al., 1999). The use of a fusion protein prevents the release of the G-protein into the cytosol upon activation by GTP thus maintaining the GPCR/G-protein complex. It has been shown that a β2AR/Gα fusion protein was more efficient at stabilising high-affinity agonist binding and stimulating adenylate cyclase in comparison to a non-fusion β2AR construct in Sf9 cells (Bertin et al., 1994). In addition a recent study has reported a 20 fold increase in expression level for a κ-opioid receptor/Gαi1 fusion protein in comparison to the κ-opioid receptor alone. The κ-opioid receptor/Gαi1 fusion protein was also much more degradation-resistant than the κ-opioid receptor alone. These data suggest that GPCR/G-protein constructs may be useful tools for structural biology.

An additional aim of this study was to characterise adenosine A2aR/Gαs or Gαi fusion proteins generated by linking the intracellular C-terminus of a GPCR in frame with the N-terminus of the naturally occurring long (Gαi) and short (Gαs) isoforms of the stimulatory type Gα-protein (provided by Dr. Renaud Wagner, University of Strasbourg).
Such GPCR G-protein fusion complexes have been studied for about two decades but so far it has proven impossible to obtain a structure of a receptor in complex with a complete G-protein. However recent success was achieved with the structure of opsin being determined in the presence of a peptide from the C-terminus of its cognate G-protein (Scheerer et al., 2008). In order to investigate such an approach for the crystallisation of the V334 adenosine A$_2$R, peptides based on the C-terminus of G$\alpha$s were designed. Preliminary characterisation of these peptides and co-crystallisation attempts with the receptor are presented in this chapter.
5.2 MATERIALS AND METHODS

5.2.1 Materials
Amino acids, glucose, PEG 3000, 425-600 µm glass beads and single stranded carrier DNA salmon sperm were obtained from Sigma-Aldrich Gillingham UK, the cAMP kit was from Cisbio France, CGS21680, rolipram and cilostamide were purchased from Tocris Bristol UK, Dulbecco Modified Eagle’s Media (DMEM) was obtained from Invitrogen Paisley UK, 96 well black optical bottom plate and 384 well plate were purchased from Nunc Rochester USA, 50 ml aerated capped tubes were supplied by Helena Biosciences Tyne UK and the LAS-300 CCD imaging system was from Fujifilm. All other reagents were either from Sigma-Aldrich Gillingham UK or as described previously.

5.2.2 Generation and assessment of adenosine A$_{2a}$R GFP fusion construct

5.2.2.1 PCR amplification of the adenosine A$_{2a}$R gene
The adenosine A$_{2a}$R gene was amplified by PCR from the pPIC9K vector harboring the WT adenosine A$_{2a}$R sequence. Primers contained a gene sequence specific for the adenosine A$_{2a}$ gene and a 30 bp region complementary to the homologous recombination sites (shown in bold) on the vector. Forward and reverse oligonucleotide sequences were as follows 5’ACCCCGGATT CTAGAACTAG TGGATCCCCC TACGGGTAGT ACCCGAGGAG3’ and 5’AAATTGACCT TGAAAATATA AATTTTCGCC GACACTCCTG CTCCATCCTG GG3’ respectively.
5.2.2.2 Plasmid design and preparation

The vector pDDGFP-2 (pRS426GAL1-GFP) (Newstead et al., 2007) contains the homologous recombination site, a *Sma*I restriction site between the homologous recombination sites a TEV cleavage site and a GFP tag followed by an 8-His tag (Figure 5.2). Linearization of the plasmid is required for the homologous recombination and thus integration of the GPCR gene into the vector. The plasmid was cleaved by *Sma*I at 30°C for 3 hours.

![Figure 5.2 Schematic representation of the GFP fusion expression vector used in the study](image)

The protein is expressed under the inducible galactose promoter (Gal 1; sky blue). The homologous recombination sites (HR; red and blue) and *Sma*I restriction site between the HR (grey) are required for integration of the target gene sequence through homologous recombination in *S.cerevisiae*. A GFP protein mutated for enhanced fluorescence in yeast (yEGFP; green) and an octa-His (8-His; yellow) tag fused to its C-terminus is present. A TEV cleavage site (TEV; purple) between the GFP and integration site for the target protein allow for removal of tags at the final stages of protein production. Selective markers for ampicillin resistance (Amp'; dark green) and expression in uracil deficient media (Ura 3; orange) is also present.

5.2.2.3 Transformation of *Saccharomyces cerevisiae*

Transformation was carried out using the Lithium acetate method (Gietz et al., 1995). This involved growing a single colony of the *S. cerevisiae* strain FGY217 (MATa, ura3-52, lys2 201 and pep4) (Kota et al., 2007) overnight in 10 ml synthetic complete (SC) media (0.67% YNB, 2% glucose) containing 2 g/L amino acid mix (18 mg adenine, 380 mg leucine, 8 mg p-aminobenzoic acid and 76 mg each of alanine, arginine,
asparagine, aspartic acid, cysteine, glutamic acid, glycine, histidine, myo-inositol, isoleucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, uracil and valine) in a shaking incubator at 30°C. A volume of 50 ml SC media was inoculated with the overnight culture to achieve an initial OD$_{600}$ of 0.1. The cells were grown for roughly 4 to 5 hours at 30°C to allow growth of culture to an OD$_{600}$ of 0.6 thus attaining maximum cells at log phase to improve transformation efficiency. The cells were pelleted through centrifugation at 3000 g, washed twice with sterile distilled water and resuspended in 1 ml of 100 mM lithium acetate. The cells were pelleted through centrifugation at 8000 g and resuspended in 400 µl lithium acetate. A volume of 50 µl was used for each transformation reaction. The transformation mix containing 240 µl of 50% w/v PEG 3350, 36 µl of 1 M lithium acetate, 25 µl of salmon sperm carrier DNA (2 mg/ml) was added to two tubes. One of the tubes contained 75 ng of plasmid digested with SmaI enzyme and 750 ng of PCR amplified adenosine A$_2$aR DNA. A no-DNA control was supplemented with sterile water instead of the DNA. The volume was made up to 360 µl with sterile water in both the cases. The cells were incubated at 30°C for 30 minutes followed by a heat shock at 42°C for 30 minutes. The cells were pelleted by centrifugation at 5000 g and resuspended in 100 µl of sterile water. The cell suspension was spread on SC plates lacking uracil and incubated at 30°C for 2-3 days. Colonies obtained were streaked onto SC plates lacking uracil to select for single colonies.

### 5.2.2.4 Small scale expression

Single colonies were inoculated in 10 ml synthetic complete media lacking uracil (SC-Ura) and incubated overnight in 50 ml aerated cap tubes at 30°C, 250 rpm. The overnight cultures were back diluted in 10 ml of SC-Ura media with 0.1% glucose to give an initial OD$_{600}$ of 0.12. The culture was grown at 30°C, 290 rpm for 6-7 hours
to reach an OD_{600} of 0.6. To induce expression of protein 2% galactose was added to
the medium and the culture incubated for 22 hours at 30°C. The cells were harvested
by centrifugation at 3000 g for 5 minutes. The cell pellet was resuspended in 200 µl of
GFP buffer (50mM Tris-HCl pH-7.6, 50mM EDTA, 10% v/v glycerol).

The cell suspension was transferred to a clear 96 well optical plate and whole
cell fluorescence was determined using a SpectraMax M2e microplate reader
(Molecular devices) with emission at 512 nm and excitation at 488 nm. The cell
suspension from the plates was transferred back to 2 ml safe lock eppendorf tubes
containing 250 µl of acid washed 425-600 µm glass beads. The cells were lysed by
shaking at 13 Hz for 8 minutes in a tissue lyser. The beads were pelleted by
centrifugation at 5000 g for 5 seconds at 4°C. The supernatant was transferred to 1.5
ml microcentrifuge tubes. The glass beads were washed again with 200 µl GFP buffer
and the above steps repeated. The membrane pellet was obtained by centrifugation of
the combined supernatant at 21,000 g for 1 hour at 4°C pellet. The pellet was
resuspended in 200 µl GFP buffer and used for radioligand binding assays.

5.2.2.5 Estimation of receptor produced

The amount of membrane protein expressed was estimated from the whole cell
fluorescence count using the procedure described by Drew et al., 2008. The amount of
GFP in the sample was calculated based on the relative fluorescence unit (RFU)
obtained against a standard curve of yeast-enhanced GFP (yEGFP) (Drew et al., 2006).
The molecular weight of protein was then used to calculate the final amount of
receptors in solution (Drew et al., 2008).

Cell pellet from a 10 ml culture was resuspended in 200 µl to give a final
volume of 250 µl. It has been established that 200 µl of 0.03 mg/ml of yEGFP solution
gives a fluorescence value of 11,300 (Drew et al., 2006) which was used to calculate the amount of GFP for the RFU obtained.

\[
\frac{\text{RFU - background}}{\text{GFP count}} \times \text{amount of GFP} = \text{amount mg/ml} \quad \frac{19200 - 3000}{11300} \times 0.03 = 0.043 \text{ mg/ml}
\]

\[
\frac{\text{amount} \times \text{GFP recovery}}{\text{volume correction}} = \text{GFP mg/ml} = \text{GFP mg/L,} \quad \frac{0.43 \times 0.6}{40} = 0.000645 \text{ mg/ml} = 0.645 \text{ mg/L}
\]

\[
\frac{A_{2a}R \text{ MW}}{\text{GFP MW}} \times \text{GFP mg/L} = \text{protein mg/L} \quad \frac{40}{28} \times 0.645 = 0.9 \text{ mg/L}
\]

As only 200 µl from a 250 µl suspension is used the amount of GFP calculated is equivalent to protein in 8 ml out of a 10 ml culture. In order to calculate the concentration of GFP in mg/ml in 8 ml of culture the amount obtained for 200 µl was divided by 40. As only 60% of the GFP can be recovered after membrane preparation the actual protein concentration was corrected by multiplying by 0.6.

Finally the amount of GFP produced was converted to the amount of membrane protein expressed by adjusting for the molecular weight of the receptor.

5.2.3 G-protein fusion constructs

Membrane samples expressing the WT and G-protein fusion adenosine A₂aR constructs were provided by Dr. Renaud Wagner (Strasbourg, MePNet). The expression and membrane preparation for the WT and G-protein fusion adenosine A₂aR constructs was carried out under conditions similar to that of small scale expression as described in Chapter 2. The adenosine A₂aR gene was followed by the G-protein sequence with a six base pair sequence coding for SpeI restriction site present between them (Figure 5.3).
**Figure 5.3 Schematic representation of adenosine A\(_{2a}\)R G-protein fusion constructs**

The two naturally occurring G\(_{\alpha}\) protein isoforms short (G\(_{\alpha_s}\)) and long (G\(_{\alpha_l}\)) are fused to the C-terminal of the full length (WT) adenosine A\(_{2a}\)R with a SpeI restriction site in between. The signal peptide (\(\alpha_F\)), purification tags (Flag, His 10) and TEV cleavage site are present on the N-terminal of the receptor sequence. The protein is expressed under the methanol inducible AOX1 promoter.

5.2.4 Saturation radioligand binding studies

Radioligand binding studies on the GFP fusion and G-protein fusion constructs were carried out as described in Chapter 2, section 2.8. Saturation radioligand binding assays for the GFP fusion construct and G-protein fusion constructs were carried out with 5 µg and 2 µg membrane protein per ml assay volume respectively.

5.2.5 G-protein peptide generation

The design strategy for the G-protein peptides was based on the successful crystallisation of the active form of opsin in complex with a peptide based on the C-terminus of its cognate G-protein transducin (Scheerer et al., 2008). As described in detail in Chapter 1 the adenosine A\(_{2a}\)R interacts with the stimulatory type G-protein (G\(_s\)). The peptides were designed by Dr. Benjamin Bax (GSK) using the C-terminal sequence of the stimulatory type \(\alpha\) subunit of the G-protein (G\(_s\alpha\)). The adenosine A\(_{2a}\)R has eight extra residues in the ICL3 as compared to that of opsin. The peptides
used in this study were longer in length than that used for opsin in order to try to increase the contacts between the loop and peptide (Figure 5.4).

\[ \text{Peptide 736 is 22 amino acids long (RRARNDDRDIQRMHLRQYELL-acid) most of which correspond to residues at the extreme C-terminus of the G}_\alpha\text{-protein (Figure 5.5 A). The extreme C-terminus of G}_\alpha\text{ has been shown to be important for receptor G-protein interaction (Bourne, 1997; Martin et al., 1996; Wess, 1997). The three underlined hydrophobic residues (Figure 5.4) were largely buried in the core of the G}_\alpha\text{ subunit. These three residues were mutated to an} \]
alanine and two charged residues on the same side of the helix to try to increase the solubility of the peptide and stabilise it in a helical conformation.

The design for peptide 737 (RWLRTISGGGIQRMLRQYELL-acid) was slightly different with the underlined residues corresponding to the $\alpha_3\beta_5$ loop of the Gs protein connected by a glycine linker to the extreme C-terminus of the Gs$\alpha$ protein (Figure 5.5 B). The $\alpha_3\beta_5$ loop has been proposed to interact with the receptor hence may help in stabilising the receptor further (Grishina et al., 2000).
Figure 5.5 Location of $G_\alpha$ peptides within the G-protein

A structure of the human $G_\alpha$ protein showing its secondary structure including alpha helix and beta sheet is presented (PDB1AZS). The residues for peptide 736 (A) are indicated in blue. The peptide 737 (B) was a complex of residues from two regions within the $G_\alpha$ protein. It included residues from the $\alpha_3\beta_5$ loop indicated in green and C-terminus indicated in blue linked together by glycine residues in the final peptide.

Peptides were purchased from Cambridge Research Biochemicals and stock solutions (10mM) were made in water and stored at -20°C.
5.2.6 Functional assay

A cAMP cell based assay kit from Cisbio based on HTRF technology (Homogenous Time-Resolved Fluorescence) was used to measure the level of cAMP that accumulates in the cell upon activation of the receptor (Figure 5.6). The method is a competitive immunoassay between native cAMP produced by cells and the cAMP labeled with the dye d2 (cAMP d2). The tracer binding is assessed by a mAb anti-cAMP labeled with Europium cryptate (mAb-cAMP). As depicted in figure 5.6 when cAMP d2 interacts with mAb-cAMP the two fluorophores d2 and cryptate are close enough so that a transfer of energy (FRET) can occur from cryptate to d2. The emission spectrum of d2 is measured by a spectrometer. Native cAMP produced by the cell competes with the cAMP d2 leading to mAb-cAMP being unavailable for energy exchange hence emission spectra for d2 cannot be observed.
A membrane bound GPCR upon activation by an agonist interacts with the α subunit of the G-protein and leads to production of the effector molecule cAMP. The cAMP labelled with dye d2 interacts with the mAb-cAMP cryptate conjugate and produces FRET. However in the presence of native cAMP no FRET is produced due to competitive inhibition of interaction between cAMP d2 and mAb-cAMP by native cAMP.

A HEK293 stable cell line expressing the full length human adenosine A2aR was used for these assays. Diana Hedley (University of Reading) kindly provided the cells and assisted during these assays. Cell density optimisation was carried out such that the level of cAMP produced by the cells was within the linear range of the standard curve (Figure 5.7). The aim was to obtain the highest signal difference possible between control and activated cells. The cells were stimulated with forskolin.
an activator of adenylate cyclase and the amount of cAMP produced was compared with cells in the absence of forskolin or any other stimulating agents such as an agonist. Based on this analysis 4000 cells per well were used for the assay.

![Graph showing cell density optimisation for functional assay](image)

**Figure 5.7 Cell density optimisation for functional assay**

A specific number of cells were treated with 1 µM and 10 µM forskolin, which activates adenylate cyclase leading to production of cAMP. The mean ratio of emission at 665/620 nm was used to calculate delta F values. The highest delta F difference between activated and inactivated cells was observed at 4000 cells per well, this number was used for future experiments.

A standard curve (Figure 5.8) was obtained using standards provided in the kit. Solutions containing cAMP (concentration ranging between 0-700 nM) were prepared in DMEM medium. The free cAMP concentrations were then plotted against the experimental delta F value obtained (the mean ratios of the emissions from both fluorophores taking into account standard error and the signal generated from the negative control, see delta F calculations below). This was required to determine the cAMP concentration for each of the test conditions.

**5.2.6.1 Delta F calculations**

Each delta F value was obtained though the following steps:
1. The mean ratio between emission at 665 nm and emission 620 nm was calculated at each given concentration of cAMP.

2. Delta F was calculated as the difference between mean ratio obtained at each condition and mean ratio generated by the negative control.

3. Delta F value were converted to percentage value.

Figure 5.8 A standard curve for free cAMP concentration
A known amount of cAMP present in the assay was plotted against the delta F values obtained from the mean ratio of 665/620 nm for each concentration. The amount of cAMP in each test condition was extrapolated from the standard curve.

A negative control included the DMEM medium, compound buffer (DMEM, phosphodiesterase inhibitor Rolipram 10 µM and Cilostamide 10 µM), conjugate and cell lysis buffer. The conjugate and cell lysis buffer were provided with the kit.

A 2.5 µl volume of DMEM medium containing 8000 cells was incubated with a range of peptide concentrations (10^-2 M to 10^-11 M) for 1 hour at 37°C. A pre-incubation of the cells with peptide 736 and 737 for 1 hour before stimulation with ligand was essential to allow the peptide to enter the cell and interact with the GPCR to cause inhibition of cAMP production by the cell.
An adenosine A$_2a$R specific agonist CGS21680 at 10$^{-6}$ M was added to the wells containing the cells and the plate incubated at room temperature for 30 minutes to allow stimulation of the adenosine A$_2a$R receptor. Conjugate and lysis buffer (5 µl) were added to the negative controls. cAMP-d2 was added to the wells containing the free cAMP for obtaining the standard curve and cells containing peptides. MAb-cAMP cryptate conjugate which selected for the cAMP antibody was added to all the wells to make up the final volume of the assay to 20 µl. The plate was sealed and incubated for 1 hour at room temperature. The emission spectra at 665 nm and 620 nm were read on a Tecan 384 plate reader. The mean ratio of 665/620 nm was calculated to obtain a delta F value. Using this value the cAMP concentration for each test condition was extrapolated from the standard curve using Graphpad Prism.

5.2.7 Setting up LCP with G$_{\alpha}$ peptide 736

Purified V334 adenosine A$_2a$R at 13 mg/ml (0.34 mM) was mixed with 10$^{-5}$ M NECA and 10$^{-6}$ M peptide to a final volume of 24 µl. MemGold and Memstart/MemSys screens were tested with a crystallisation setup as described in section 4.2. The plates were incubated at 20°C and checked periodically for crystals.

Crystals appeared after 3-5 days of incubation. These were tested for tryptophan fluorescence by Dr. Yilmaz Alguel at MPL Diamond in an attempt to differentiate between protein and salt, detergent or peptide crystals. At 280 nm the tryptophan residues are excited and emit light of lower energy which is observed as fluorescence. As the salt, detergent and peptides do not contain any tryptophan residues any fluorescence observed in the crystals at 280 nm could be attributed to the presence of protein which contains tryptophan residues.
5.3 **RESULTS**

5.3.1 **Functional expression of adenosine A<sub>2a</sub>R–GFP fusion construct**

An RFU of 19000 corresponding to an expression level of 0.9 mg/L was obtained for a one litre flask culture of WT adenosine A<sub>2a</sub>R–GFP fusion protein.

Saturation radioligand binding analysis was performed on membranes expressing the adenosine A<sub>2a</sub>R-GFP fusion protein (Figure 5.9). The maximal specific activity obtained was 10.6 pmol/mg. The pK<sub>d</sub> of 8.2 for ZM241385 is similar to that obtained for WT adenosine A<sub>2a</sub>R expressed in *P. pastoris* (Fraser, 2006) (Chapter 2) and mammalian cell lines (Alexander *et al.*, 2001) indicating that the receptor maintains native pharmacological properties. A further optimisation of the saturation assay conditions is required as the curve does not reach saturation and the maximal value presented is an estimation.

![Figure 5.9 Saturation analysis for [H] ZM241385 to the membrane bound GFP adenosine A<sub>2a</sub>R fusion construct](image)

Saturation binding analysis of [H]ZM241385 to the membrane bound WT adenosine A<sub>2a</sub>R GFP constructs expressed in *S. cerevisiae*. Preliminary results from a single saturation assay experiment are shown. B<sub>max</sub> and K<sub>d</sub> values were derived by non-linear regression analysis.
5.3.2 Functional expression of adenosine $A_2aR/G$-protein fusion constructs

Saturation radioligand binding assays were carried out on membranes prepared from small-scale flask cultures (Fig 5.10). Radioligand binding assay parameters from three independent experiments are presented in Table 5.1. A significantly higher $B_{\text{max}}$ was obtained for the G-protein fusion constructs in comparison to the WT adenosine $A_2aR$ ($p<0.05$). Fusion of the receptor to the G-proteins resulted in a higher expression level compared to the WT receptor as the specific activity increased by 7 times for the WT adenosine $A_2aR/G\alpha_s$ fusion construct and by 4 times for the WT adenosine $A_2aR/G\alpha_l$ fusion construct. A similar binding affinity is observed for both WT and G-protein fusion adenosine $A_2aR$ constructs indicating similar pharmacological properties.

![Figure 5.10 Saturation analysis of $[^3H]$ ZM241385 to the membrane bound G-protein fusion adenosine $A_2aR$ constructs](image)

Saturation binding of $[^3H]$ ZM241385 to WT (■), G$\alpha_s$ (○), and G$\alpha_l$ (▲) adenosine $A_2aR$ from flask cultures of *P. pastoris* membranes was determined. Data are representative of 3 independent experiments. $B_{\text{max}}$ and $K_d$ values were determined by non-linear regression.
Table 5.1 Saturation assay parameters for G-protein fusion adenosine A2aR constructs

The average $B_{\text{max}} \pm \text{SEM}$ from three independent experiments for the WT and G-protein fusion forms of the adenosine A2aR are presented. The average pK$_d \pm \text{SEM}$ from three independent experiments is presented for the three constructs.

<table>
<thead>
<tr>
<th>Receptor type/Parameters</th>
<th>WT</th>
<th>Gαs</th>
<th>Gαi</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_{\text{max}} \pm \text{SEM}$ (pmol/mg)</td>
<td>10.7 ± 1.2</td>
<td>70.1 ± 13.9</td>
<td>33.4 ± 2.5</td>
</tr>
<tr>
<td>pK$_d \pm \text{SEM}$</td>
<td>9.0 ± 0.2</td>
<td>8.7 ± 0.2</td>
<td>8.8 ± 0.1</td>
</tr>
</tbody>
</table>

5.3.3 Characterisation of the Gα peptides by a cAMP functional assay

A functional assay was used to assess the ability of the peptides to compete with native G-protein interactions with WT adenosine A2aR activated by an agonist in an in vivo mammalian cell based assay. In principle, the peptide should compete with the G-protein for binding to the activated receptor and inhibit G-protein activation. This in turn would reduce the level of cAMP production.

The amount of cAMP produced after activation with different concentrations of CGS21680 was measured (Figure 5.11 inset). As indicated by the curve only approximately 20% of the maximal cAMP production possible could be achieved upon activation with $10^{-6}$M CGS21680. The receptor was stimulated with CGS21680 at $10^{-6}$ M concentration and the amount of cAMP produced at varying concentration of peptides was assessed. A reduction in the amount of cAMP produced was observed at higher concentrations of peptide 736 (Figure 5.11) suggesting that the peptide was competing with the G-protein for binding to the receptor. No detectable binding of peptide 737 was observed (Fig 5.11).
Detecting ability of peptides to inhibit the adenosine A$_2$R G-protein interaction

The amount of cAMP produced by HEK293 cells expressing adenosine A$_2$R in the presence increasing concentrations of the ligand CGS21680 only (inset; blue), CGS21680 and peptide 736 (black) and CGS21680 and peptide 737 (red) is presented. The blue dashed line indicates the maximum amount of cAMP that can be produced on stimulation with $10^{-6}$ M CGS21680. An inhibition in level of cAMP detected is observed for the peptide 736 indicating the ability of the peptide to interact with the GPCR and inhibit the G-protein interaction and thus production of cAMP. A curve could not be fitted to the data obtained from cells incubated with peptide 737.

5.3.4 LCP crystallisation of V334 adenosine A$_2$R and peptide 736 mixture

Crystalline material was observed for protein in the presence of peptide 736 (Figure 5.12). Crystals were observed in conditions 1.1 (0.08M sodium citrate pH 5.2, 2.2 M (NH$_4$)$_2$ SO$_4$), 1.17 (0.1M MgCl$_2$, 0.03M Tris-HCl pH 8.2 and 32% v/v PEG 400)
and 2.9 (0.1M NaCl/0.1M CdCl₂ 0.1M Tris-HCl pH 8 and 33% v/v PEG 400) of
the MemGold screen (Appendix 7.2). The crystals were tested by tryptophan
fluorescence for presence of protein. As no fluorescence could be detected it suggests
that the crystals may not be protein but rather peptide, salt or detergent crystals.
However an absolute conclusion can not be drawn from these results hence it is
necessary to collect X-ray diffraction images of crystals.

**Figure 5.12 Crystals observed in Memgold screens for V334 adenosine A2aR with NECA
and peptide 736**
Crystallisation screen with LCP were set up for the V334 adenosine A₂aR incubated with NECA and peptide 736.
Crystals appeared after 5 days of incubation at 20°C. The wells containing the crystalline material were screened
for protein through Tryptophan fluorescence.

The protein mixture containing peptide 737 did not form a lipid cubic phase
even after repeated attempts at different concentration of monoolein hence
crystallisation screening could not be done with this peptide.
5.4 DISCUSSION

In addition to the V334 adenosine $A_{2a}$R expressed in *P. pastoris* as detailed in Chapters 2, 3 and 4 a number of other constructs were assessed for their suitability as alternatives for structural studies.

The WT adenosine $A_{2a}$R GFP fusion construct was designed with an aim to explore the *S. cerevisiae* as an alternative to the *P. pastoris* system. The *S. cerevisiae* GFP system is a rapid and efficient screening system for testing expression and solubilisation of different protein constructs in a range of different conditions. Very high level expression (4 mg/L) and purification of a human adenosine $A_{2a}$R GFP fusion protein in *S. cerevisiae* has been reported previously (O'Malley *et al.*, 2007; Wedekind *et al.*, 2006). The construct generated as part of this study expressed at a much lower level as determined by the fluorescence of the expression culture; 0.9 mg/L of adenosine $A_{2a}$R could be produced from a flask culture after incubation at 30°C for 20 hours. The specific activity of the protein was 10 pmol/mg. It is not clear why our expression level is less than the reported value however it may be possible to further optimise the expression level by; presence of a signaling peptide, addition of additives such as histidine or DMSO to the growth medium or reducing the induction temperature (Drew *et al.*, 2008). In light of the results obtained for the WT and C-terminal truncated adenosine $A_{2a}$R expressed in *P. pastoris* it would be sensible to design a V334 adenosine $A_{2a}$R GFP fusion construct and to assess if that improves expression and stabilises the receptor further. There is some evidence to suggest that *S. cerevisiae* is not a good host for production of GPCRs for structural studies (Prof So Iwata, personal communication) but rather it is a useful system for rapid preliminary
screening a range of targets prior to expression of suitable constructs using either *P. pastoris* or insect cell based systems.

Radioligand binding analysis of both adenosine $A_2a$R/G-protein fusion constructs indicated a higher yield of functional receptor in comparison to the WT adenosine $A_2a$R. Both the WT and the adenosine $A_2a$R/G-protein fusion constructs were expressed in a small scale flask culture which could account for a maximal specific binding value much lower than that reported in Chapter 2 for WT adenosine $A_2a$R (100 pmol/mg) which was expressed in the more controlled environment of a bioreactor. It remains to be seen if the increase in functional receptor observed for the G-protein fusion constructs in comparison to the WT adenosine $A_2a$R in small scale cultures would translate to a bioreactor culture. The binding affinities obtained for the G-protein fusion constructs for ZM241385 derived from saturation assay are similar to that of adenosine $A_2a$R constructs expressed in *P. pastoris* and mammalian cell lines (Chapter 2) indicating that the receptor G-protein fusion constructs have pharmacological properties similar to that of the receptor expressed in native tissues. A further assessment of binding affinities of different ligands including an agonist would help characterise the adenosine $A_2a$R/G-protein fusion constructs better.

Large scale expression and purification needs to be performed in order confirm the suitability of this construct for structural studies. As described in Chapter 2 the adenosine $A_2a$R undergoes proteolytic degradation of the exposed C-terminus. The presence of a G-protein at the C-terminus of the receptor may protect this region of the protein from proteolytic degradation thereby increasing the amount of functional receptor present at the cell surface in comparison to a WT adenosine $A_2a$R. This hypothesis could be tested by assessing for long term stability of the purified receptor forms through SDS-PAGE gels as described in Chapter 4.
The addition of the soluble G-protein may help in increasing the solubility or stability of the receptor/G-protein constructs facilitating crystallisation. The crystal structures of G-proteins are also known (Lambright et al., 1996) which may help in structure determination of the GPCR. However it should be pointed out that although the GPCR/G-protein fusion constructs have been used for pharmacological studies for more than 15 years they have not so far proved suitable for structural studies. However an alternative approach using G-protein peptides has been successfully used to structurally characterise the unliganded form of rhodopsin, opsin (Scheerer et al, 2008). G₆α peptides were designed as a means of stabilising ICL3 of V334 adenosine A₂aR, increasing the likelihood of obtaining crystals and obtaining more information on receptor/G-protein interactions. Peptide 736 was designed using the peptide obtained from transducin, the cognate G-protein of opsin as a guideline and corresponds to the C-terminal region of the G₆α. Peptide 737 included portions of the C-terminal and the α₅β₃ loop of the G₆α protein linked together by glycine residues.

Preliminary analysis indicated that peptide 736 did indeed interact with the adenosine A₂aR however a more complete kinetic analysis is required to confirm and further characterise the interaction. Lipidic cubic phase crystallisation trials using the V334 adenosine A₂aR in complex with the agonist NECA and peptide 736 were performed. Crystalline material was obtained from some of the conditions tested however it was not possible to screen the crystals through X-ray diffraction. Further optimisation to produce crystals that are larger in size and less fragile could be carried out. The structure of opsin interacting with the G-protein peptide was obtained with a four times molar excess of peptide to protein. A similar or range of molar ratios of ligand:protein:peptide could not be tried out for the adenosine A₂aR due to
insufficient amount of peptide. In the case of the opsin G-protein peptide interacting structure a range of synthetic peptides designed using the C-terminal of transducin as a template were screened for affinity to the activated receptor. A peptide with lysine substituted to leucine was used for structural work as it had a higher affinity for the activated receptor (Herrmann et al., 2006). Once the affinity of the interaction between peptide 736 and the adenosine A2aR is established a range of other peptides could be generated by mutating specific residues. The affinities of these could be tested though the earlier described functional assays in order to identify an optimal peptide for further work.

In contrast to the results obtained for peptide, no interaction between adenosine A2aR and peptide 737 was detected. It is possible that the peptide 737 was either unable to diffuse across the lipid bilayer or was unable to interact with the receptor and hence could not inhibit the production of cAMP within the cells. It proved impossible to form a lipid cubic phase with the peptide 737 for crystallisation trials which suggests unfavorable chemical properties making peptide 737 incompatible for lipid cubic phase at the concentrations tested. A glycine linker used between the peptide may make the peptide too flexible and unable to form a stable interaction with the receptor. A more rigid linker structure could be tested.

Though promising preliminary results have been obtained, forming a stable ternary complex between a receptor, agonist and peptide is highly challenging due to the dynamic nature of the interactions between the different molecules. Screening a range of peptides and crystallisation conditions in order to select for one which leads to a stable complex formation is likely to be necessary.
6 Conclusion and future outlook
GPCRs comprise a large class of structurally similar cell surface receptors that mediate cellular responses to a wide range of hormones, neurotransmitters and environmental stimulants. For many years a lack of high resolution structural information has limited understanding of the molecular basis of GPCR function. X-ray diffraction structures can provide information at the molecular level, of changes brought about within the receptor upon ligand binding or activation of a receptor leading to G-protein binding. The recent structures solved of GPCRs are the result of decades of intensive scientific research (reviewed in (Kobilka et al., 2008)). Heterologous overexpression of GPCRs followed by purification to obtain milligram quantities of protein while maintaining functional and structural integrity was a bottleneck for structural studies for a long time. Establishing overexpression in different heterologous expression systems (reviewed in (Andre et al., 2006; Grisshammer, 2006; McCusker et al., 2007; Sarramegna et al., 2003)) developing purification protocols (Chiu et al., 2008; Sarramegna et al., 2006), mutagenesis of the receptor (Kobilka, 1995; Rosenbaum et al., 2007; Serrano-Vega et al., 2008; Tate et al., 2009) and developments in the field of membrane protein structural biology including crystallisation in lipidic cubic phase (Caffrey, 2003) and the use of microfocus beamline (Riekel et al., 2005) are some of the many factors that have contributed to the elucidation of GPCR structures. The recently published structures (Cherezov et al., 2007; Jaakola et al., 2008; Rasmussen et al., 2007; Warne et al., 2008) are an important milestone in the field of GPCR biology. Nevertheless to improve our understanding of the diverse GPCR signaling mechanisms and to facilitate efficient structure based drug design many more structures of GPCRs bound to different ligands and/or interacting with their cognate G-proteins are required.
The adenosine A\(_{2a}\)R is an important drug target with implications in medical conditions such as asthma (Luijk et al., 2008), Parkinson’s disease (Schapira et al., 2006; Schwarzschild et al., 2006) and immunosuppression (Hasko et al., 2008; Sitkovsky et al., 2004). The recent A\(_{2a}\)R-T4 lysozyme structure has provided a wealth of information about the antagonist bound receptor. However a structure of adenosine A\(_{2a}\)R bound to ligands with different efficacies, such as theophylline, a ligand widely used in the treatment of asthma, would facilitate drug design (Barnes et al., 1994). Similarly structures of GPCRs in complex with agonists or cognate G-proteins would aid understanding of the molecular mechanism of signal transduction.

The aim of this study was to develop an expression and purification protocol for the adenosine A\(_{2a}\)R in order to produce a protein suitable for structural work. Since previous studies from both our laboratory (Chapter 2) (Singh et al., 2008) and that of others (Weiss et al., 2002) show that the adenosine A\(_{2a}\)R is susceptible to C-terminal degradation, constructs with truncated C-terminal were generated. These C-terminally truncated constructs, A316 and V334 adenosine A\(_{2a}\)R, were shown to be resistant to proteolytic degradation. A detailed pharmacological characterisation, in terms of ligand binding properties, of full length and C-terminal truncated adenosine A\(_{2a}\)R constructs has been performed at each stage of protein production. The previous literature supports \textit{P. pastoris} as a successful heterologous expression system for GPCRs (Abdulaev et al., 1997; Andre et al., 2006; Feng et al., 2002; Kim et al., 2005; Weiss et al., 1995) including the adenosine A\(_{2a}\)R (Fraser, 2006). As the process of crystallisation requires milligram quantities of protein a large-scale expression protocol using a bioreactor was established for production of the adenosine A\(_{2a}\)R. Such an approach improved both reproducibility of the expression culture and yield of functional receptor. The V334 adenosine A\(_{2a}\)R was the best expressing clone with about 11
mg/L of protein being produced from a 4 litre bioreactor culture. The V334 adenosine A$_{2a}$R expressed to the highest level reported for the receptor in *P. pastoris*. The recent GPCR structures are of constructs with truncated C-terminus demonstrating that the removal of the long and disordered C-terminal tail is highly advantageous to crystallisation.

In the present study, mg quantities of the highly pure V334 adenosine A$_{2a}$R were obtained in buffer containing DDM+CHS. The V334 adenosine A$_{2a}$R solubilised in DDM+CHS has a $T_m$ comparable to that of the conformationally thermostable adenosine A$_{2a}$R mutant solubilised in the absence of CHS (Magnani *et al.*, 2008). The purified V334 adenosine A$_{2a}$R is highly functional with a specific activity close to the theoretical maximum for the construct and is resistant to both degradation and aggregation as revealed by the SDS-PAGE analysis and SEC profile respectively. The purified V334 adenosine A$_{2a}$R has pharmacological properties similar to that of the adenosine A$_{2a}$R expressed in native tissues (Fredholm *et al.*, 2001). An increase in thermostability of the protein was observed in the presence of the agonist NECA or antagonists ZM241385 and theophylline as compared to the unliganded receptor.

Even though a highly stable and functional protein could be routinely purified no crystals could be obtained. Recent work has shown that maintaining the structural and functional integrity of the protein are key to crystallisation but they are not the only factors involved. Meticulously developed expression and purification protocols formed the basis of structural studies on the β$_2$AR (Kobilka, 1995). However, despite the availability of this optimised purification protocol it took another decade and major modifications of the receptor (Rosenbaum *et al.*, 2007) or the presence of a Fab
fragment (Rasmussen et al., 2007) to further improve conformational stability enough to yield a high resolution structure.

As mentioned, an important factor in the structure determination of GPCRs has been the development and optimisation of new tools. Several groups have invested significant amounts of effort into the design, generation and characterisation of new amphipathic molecules to increase the stability of membrane proteins (reviewed in (Seddon et al., 2004)). A class of novel amphiphiles was screened as potential additives or as alternatives to conventional detergents to enhance thermostability of the receptor. Initial results indicate that V334 adenosine A\textsubscript{2a}R is more thermostable in these amphiphiles than DDM+CHS. A further characterisation of these molecules to test their ability to maintain receptor function needs to be carried out. It has proven possible to obtain well-diffracting crystals of membrane proteins in both of the amphiphiles tested (Dr Pil-Seok Chae, personal communication; Dr Bernadette Byrne, personal communication). These therefore represent an exciting possibility for the future of structure determination of GPCRs and other membrane proteins.

One thing that seems clear from the available GPCR structures and the data presented here is that stabilising ICL 3 is crucial. The recent structure of ligand free opsin in complex with a G-protein peptide (Scheerer et al., 2008) provided not only insights into receptor/G-protein interactions but also outlined an alternative method for stabilising ICL3. In this study peptides based on the C-terminal end of the G\textsubscript{q}\alpha protein were designed. Preliminary characterisation studies indicated that one peptide, 736, interacted with the receptor in cells. Crystalline material was obtained from trials in lipidic cubic phase using V334 adenosine A\textsubscript{2a}R in complex with the agonist NECA and peptide 736 however these couldn’t be confirmed as protein crystals. Co-crystallisation of a receptor-agonist-peptide complex is challenging due to
the dynamic nature of the interactions between the molecules. In order to produce well diffracting crystals further screening of crystallisation conditions needs to be carried out. The initial optimisation could involve testing different molar ratios of peptide and protein. The peptides can be further modified to increase area of contact and affinity to the receptor thereby providing greater conformational stability to the complex. Other possible approach includes the generation of Fab fragments which bind to and therefore stabilise the third intracellular loop.

Other bilayer methods of crystallisation can also be tested. The lipidic cubic phase is very viscous and as a result crystal produced are small and difficult to see and extract from the lipidic phase (Caffrey et al., 2009). An alternative is the bicelle crystallisation method which has been successfully used for crystallisation of bacteriorhodopsin (Faham et al., 2002) and the β2AR-Fab complex (Rasmussen et al., 2007). The bicelles have an advantage over lipidic cubic phase in being less viscous at low temperatures enabling the use of standard screening methodology and make them much easier to handle.

A detailed characterisation of the adenosine A2aR forms expressed in P. pastoris has led to the development of a protocol for overexpression and purification of a highly stable and functional receptor. The protein produced is an excellent candidate for further studies.
References


Cheng, Y, Prusoff, WH (1973) Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem Pharmacol* **22**(23): 3099-3108.


# Appendix

## 7.1 MEMGOLD SCREEN

<table>
<thead>
<tr>
<th>Tube #</th>
<th>Salt</th>
<th>Buffer</th>
<th>pH</th>
<th>Precipitant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>None</td>
<td>0.08 M sodium citrate</td>
<td>5.2</td>
<td>2.2 M ammonium sulfate</td>
</tr>
<tr>
<td>1.2</td>
<td>None</td>
<td>0.01 M Tris</td>
<td>8.0</td>
<td>1.2 M tri-sodium citrate</td>
</tr>
<tr>
<td>1.3</td>
<td>None</td>
<td>0.015 M tricine</td>
<td>8.5</td>
<td>24% w/v PEG 4000</td>
</tr>
<tr>
<td>1.4</td>
<td>0.36 M sodium chloride/0.1% w/v sodium azide</td>
<td>0.015 M sodium phosphate</td>
<td>7.0</td>
<td>9.9% w/v PEG 4000</td>
</tr>
<tr>
<td>1.5</td>
<td>0.3 M sodium chloride</td>
<td>0.01 M Tris</td>
<td>8.0</td>
<td>27.5% w/v PEG 4000</td>
</tr>
<tr>
<td>1.6</td>
<td>None</td>
<td>0.225 M MES/bis-tris</td>
<td>6.6</td>
<td>6.6% w/v PEG 6000</td>
</tr>
<tr>
<td>1.7</td>
<td>0.1 M ammonium sulfate</td>
<td>0.1 M HEPES</td>
<td>7.5</td>
<td>12.0% w/v PEG 4000/22% w/v glycerol</td>
</tr>
<tr>
<td>1.8</td>
<td>0.02 M calcium chloride/0.01 M magnesium sulfate/0.02 M sodium chloride</td>
<td>0.02 M MES</td>
<td>6.5</td>
<td>7.7% w/v PEG 1500</td>
</tr>
<tr>
<td>1.9</td>
<td>None</td>
<td>0.05 M HEPES</td>
<td>7.5</td>
<td>2.5 M ammonium sulfate</td>
</tr>
<tr>
<td>1.10</td>
<td>None</td>
<td>0.0665 M HEPES</td>
<td>7.5</td>
<td>1.1 M tri-sodium citrate</td>
</tr>
<tr>
<td>1.11</td>
<td>None</td>
<td>0.15 M potassium phosphate</td>
<td>6.5</td>
<td>3.3 M ammonium sulfate</td>
</tr>
<tr>
<td>1.12</td>
<td>0.1 M magnesium acetate</td>
<td>0.1 M sodium citrate</td>
<td>5.8</td>
<td>14% w/v PEG 5000 MME</td>
</tr>
<tr>
<td>1.13</td>
<td>0.1 M sodium chloride</td>
<td>0.02 M sodium citrate</td>
<td>5.6</td>
<td>11% w/v PEG 3350</td>
</tr>
<tr>
<td>1.14</td>
<td>0.1 M sodium chloride</td>
<td>0.02 M sodium citrate</td>
<td>5.6</td>
<td>5.5% w/v PEG 3350</td>
</tr>
<tr>
<td>1.15</td>
<td>0.05 M calcium chloride/0.05 M barium chloride</td>
<td>0.1 M Tris</td>
<td>8.2</td>
<td>32% w/v PEG 400</td>
</tr>
<tr>
<td>1.16</td>
<td>0.05 M sodium chloride</td>
<td>0.1 M sodium phosphate</td>
<td>6.2</td>
<td>16% w/v PEG 4000</td>
</tr>
<tr>
<td>1.17</td>
<td>0.1 M magnesium chloride</td>
<td>0.03 M Tris-hydrochloride</td>
<td>8.2</td>
<td>19% w/v PEG 4000</td>
</tr>
<tr>
<td>1.18</td>
<td>0.2 M sodium chloride</td>
<td>0.025 M HEPES</td>
<td>7.5</td>
<td>13% w/v PEG 4000</td>
</tr>
<tr>
<td>1.19</td>
<td>None</td>
<td>0.1 M HEPES</td>
<td>7.5</td>
<td>11% w/v PEG 3350</td>
</tr>
<tr>
<td>1.20</td>
<td>0.1 M sodium chloride</td>
<td>0.02 M KMES</td>
<td>6.7</td>
<td>6.6% w/v PEG 4000</td>
</tr>
<tr>
<td>1.21</td>
<td>0.1 M potassium chloride</td>
<td>0.02 M Tris</td>
<td>7.0</td>
<td>20% w/v PEG 4000</td>
</tr>
<tr>
<td>1.22</td>
<td>0.05 M magnesium chloride/0.1% w/v sodium azide</td>
<td>0.1 M sodium cacodylate</td>
<td>6.7</td>
<td>6.6% w/v PEG 3350</td>
</tr>
<tr>
<td>1.23</td>
<td>0.2 M potassium chloride</td>
<td>0.1 M sodium citrate</td>
<td>5.5</td>
<td>37% w/v pentaerythritol propoxylate (5/4 PO/ OH)</td>
</tr>
<tr>
<td>1.24</td>
<td>None</td>
<td>0.1 M Tris</td>
<td>8.0</td>
<td>5.5% w/v PEG 4000</td>
</tr>
<tr>
<td>Tube #</td>
<td>Salt</td>
<td>Buffer</td>
<td>pH</td>
<td>Precipitant</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------------------------</td>
<td>-----------------</td>
<td>-----</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>1.25</td>
<td>0.1 M sodium chloride</td>
<td>0.02 M Tris</td>
<td>7.0</td>
<td>7.7 % w/v PEG 4000</td>
</tr>
<tr>
<td>1.26</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M Tris</td>
<td>7.5</td>
<td>22 % v/v PEG 400</td>
</tr>
<tr>
<td>1.27</td>
<td>0.04 M sodium chloride</td>
<td>0.04 M Tris</td>
<td>8.0</td>
<td>27 % v/v PEG 350 MME</td>
</tr>
<tr>
<td>1.28</td>
<td>0.05 M sodium chloride/0.02 M magnesium chloride</td>
<td>0.1 M sodium citrate</td>
<td>6.0</td>
<td>22 % v/v PEG 400</td>
</tr>
<tr>
<td>1.29</td>
<td>None</td>
<td>0.1 M sodium acetate</td>
<td>5.5</td>
<td>8.8 % w/v PEG 2000 MME</td>
</tr>
<tr>
<td>1.30</td>
<td>None</td>
<td>0.4 M ammonium acetate</td>
<td>8.0</td>
<td>13 % w/v PEG 2000 MME</td>
</tr>
<tr>
<td>1.31</td>
<td>None</td>
<td>0.02 M bis Tris</td>
<td>7.0</td>
<td>15 % w/v PEG 2000</td>
</tr>
<tr>
<td>1.32</td>
<td>0.1 M sodium chloride/0.1 M magnesium chloride</td>
<td>0.02 M Tris</td>
<td>7.5</td>
<td>11 % w/v PEG 1500</td>
</tr>
<tr>
<td>1.33</td>
<td>0.1 M sodium chloride/0.1 M magnesium chloride</td>
<td>0.1 M HEPES</td>
<td>8.0</td>
<td>11 % w/v PEG 1500</td>
</tr>
<tr>
<td>1.34</td>
<td>0.2 M sodium acetate/0.2 M Potassium Chloride</td>
<td>0.1 M HEPES</td>
<td>7.0</td>
<td>22 % w/v PEG 3000</td>
</tr>
<tr>
<td>1.35</td>
<td>0.02 M nickel sulfate</td>
<td>0.01 M HEPES</td>
<td>7.0</td>
<td>33 % v/v Jeffamine-M600</td>
</tr>
<tr>
<td>1.36</td>
<td>0.15 M sodium chloride</td>
<td>0.1 M Tris</td>
<td>8.0</td>
<td>13 % w/v PEG 6000</td>
</tr>
<tr>
<td>1.37</td>
<td>0.2 M calcium chloride</td>
<td>0.1 M HEPES</td>
<td>7.5</td>
<td>53 % v/v PEG 400</td>
</tr>
<tr>
<td>1.38</td>
<td>0.05 M magnesium acetate</td>
<td>0.05 M sodium acetate</td>
<td>5.0</td>
<td>28 % v/v PEG 400</td>
</tr>
<tr>
<td>1.39</td>
<td>None</td>
<td>0.05 M HEPES</td>
<td>7.5</td>
<td>22 % v/v PEG 4000</td>
</tr>
<tr>
<td>1.40</td>
<td>0.2 M calcium chloride</td>
<td>0.1 M Tris hydrochloride</td>
<td>8.0</td>
<td>44 % v/v PEG 400</td>
</tr>
<tr>
<td>1.41</td>
<td>0.05 M magnesium acetate</td>
<td>0.05 M sodium acetate</td>
<td>5.4</td>
<td>24 % v/v PEG 400</td>
</tr>
<tr>
<td>1.42</td>
<td>0.2 M calcium chloride</td>
<td>0.1 M MES</td>
<td>6.5</td>
<td>26 % w/v PEG 350 MME</td>
</tr>
<tr>
<td>1.43</td>
<td>0.1 M potassium chloride</td>
<td>0.1 M Tris</td>
<td>8.5</td>
<td>39 % w/v PEG 400</td>
</tr>
<tr>
<td>1.44</td>
<td>0.05 M magnesium chloride</td>
<td>0.1 M glycine</td>
<td>9.0</td>
<td>22 % v/v PEG 400</td>
</tr>
<tr>
<td>1.45</td>
<td>0.1 M ammonium sulfate</td>
<td>0.1 M glycine</td>
<td>3.8</td>
<td>28 % w/v tri-ethylene glycol</td>
</tr>
<tr>
<td>1.46</td>
<td>0.15 M sodium formate</td>
<td>0.1 M HEPES</td>
<td>7.2</td>
<td>18 % w/v PEG 3350</td>
</tr>
<tr>
<td>1.47</td>
<td>None</td>
<td>0.2 M sodium acetate</td>
<td>6.8</td>
<td>8.8 % w/v PEG 6000</td>
</tr>
<tr>
<td>1.48</td>
<td>0.2 M potassium chloride</td>
<td>0.1 M MES</td>
<td>6.5</td>
<td>18 % w/v PEG 6000</td>
</tr>
<tr>
<td>Tube #</td>
<td>Salt</td>
<td>Buffer</td>
<td>pH</td>
<td>Precipitant</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------------</td>
<td>----------</td>
<td>-----</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>2.1</td>
<td>0.22 M sodium citrate</td>
<td>0.1 M Tris</td>
<td>8.0</td>
<td>35 % v/v PEG 400</td>
</tr>
<tr>
<td>2.2</td>
<td>None</td>
<td>0.1 M sodium acetate</td>
<td>4.5</td>
<td>17 % v/v PEG 400</td>
</tr>
<tr>
<td>2.3</td>
<td>None</td>
<td>0.02 M Tris</td>
<td>8.5</td>
<td>1.0 M lithium sulfate/1.8 % w/v PEG 8000</td>
</tr>
<tr>
<td>2.4</td>
<td>None</td>
<td>0.02 M Tris</td>
<td>7.5</td>
<td>22 % v/v PEG 550 MME</td>
</tr>
<tr>
<td>2.5</td>
<td>0.05 M sodium chloride</td>
<td>0.02 M glycine</td>
<td>10.0</td>
<td>33 % w/v PEG 1000</td>
</tr>
<tr>
<td>2.6</td>
<td>0.2 M magnesium chloride</td>
<td>0.1 M Tris</td>
<td>8.5</td>
<td>25 % w/v PEG 4000</td>
</tr>
<tr>
<td>2.7</td>
<td>0.2 M magnesium chloride</td>
<td>0.1 M sodium cacodylate</td>
<td>6.5</td>
<td>31 % w/v PEG 2000</td>
</tr>
<tr>
<td>2.8</td>
<td>None</td>
<td>0.64 M sodium acetate</td>
<td>4.6</td>
<td>18 % w/v PEG 3350</td>
</tr>
<tr>
<td>2.9</td>
<td>0.1 M sodium chloride/0.1 M cadmium chloride</td>
<td>0.1 M Tris hydrochloride</td>
<td>8.0</td>
<td>33 % v/v PEG 400</td>
</tr>
<tr>
<td>2.10</td>
<td>None</td>
<td>0.1 M Bicine</td>
<td>8.9</td>
<td>31 % w/v PEG 2000</td>
</tr>
<tr>
<td>2.11</td>
<td>0.05 M sodium sulfate/0.05 M lithium sulfate</td>
<td>0.05 M Tris</td>
<td>8.5</td>
<td>35 % v/v PEG 400</td>
</tr>
<tr>
<td>2.12</td>
<td>0.1 M sodium chloride</td>
<td>0.05 M glycine</td>
<td>9.5</td>
<td>33 % v/v PEG 300</td>
</tr>
<tr>
<td>2.13</td>
<td>0.3 M magnesium nitrate</td>
<td>0.1 M Tris</td>
<td>8.0</td>
<td>23 % w/v PEG 2000</td>
</tr>
<tr>
<td>2.14</td>
<td>0.12 M lithium sulfate</td>
<td>0.02 M Tris/0.1 M sodium citrate</td>
<td>7.5/5.0</td>
<td>20 % v/v PEG 300</td>
</tr>
<tr>
<td>2.15</td>
<td>0.1 M sodium chloride</td>
<td>0.12 M Tris</td>
<td>9.4</td>
<td>20 % v/v PEG 400</td>
</tr>
<tr>
<td>2.16</td>
<td>0.2 M sodium chloride</td>
<td>0.1 M HEPES</td>
<td>7.0</td>
<td>22 % v/v PEG 550 MME</td>
</tr>
<tr>
<td>2.17</td>
<td>0.1 M sodium chloride/0.325 M sodium acetate</td>
<td>0.1 M Tris</td>
<td>8.0</td>
<td>21 % v/v PEG 400</td>
</tr>
<tr>
<td>2.18</td>
<td>0.02 M sodium citrate</td>
<td>0.08 M sodium phosphate</td>
<td>6.2</td>
<td>18 % w/v PEG 2000</td>
</tr>
<tr>
<td>2.19</td>
<td>0.02 M potassium nitrate</td>
<td>0.03 M potassium citrate</td>
<td>6.5</td>
<td>7.7 % w/v PEG 4000</td>
</tr>
<tr>
<td>2.20</td>
<td>0.1 M sodium chloride/0.005 M magnesium chloride</td>
<td>0.1 M Tris</td>
<td>8.5</td>
<td>30 % w/v PEG 2000 MME</td>
</tr>
<tr>
<td>2.21</td>
<td>0.2 M calcium chloride</td>
<td>0.1 M HEPES</td>
<td>7.0</td>
<td>33 % v/v PEG 400</td>
</tr>
<tr>
<td>2.22</td>
<td>0.1 M calcium chloride</td>
<td>0.1 M Tris</td>
<td>6.5</td>
<td>13 % v/v PEG 2000 MME</td>
</tr>
<tr>
<td>2.23</td>
<td>0.2 M ammonium sulfate/0.02 M sodium chloride</td>
<td>0.02 M sodium acetate</td>
<td>4.0</td>
<td>33 % v/v PEG 200</td>
</tr>
<tr>
<td>2.24</td>
<td>0.07 M sodium chloride</td>
<td>0.05 M sodium citrate</td>
<td>4.5</td>
<td>22 % v/v PEG 400</td>
</tr>
<tr>
<td>Tube #</td>
<td>Salt</td>
<td>Buffer</td>
<td>pH</td>
<td>Precipitant</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------------------------</td>
<td>----------------</td>
<td>-----</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>2.25</td>
<td>0.2 M ammonium sulfate</td>
<td>0.1 M sodium acetate</td>
<td>4.6</td>
<td>28 % v/v PEG 550 MME</td>
</tr>
<tr>
<td>2.26</td>
<td>None</td>
<td>0.05 M glycine</td>
<td>9.0</td>
<td>55 % v/v PEG 400</td>
</tr>
<tr>
<td>2.27</td>
<td>0.1 M magnesium chloride/0.1 M sodium chloride</td>
<td>0.1 M Tris</td>
<td>8.5</td>
<td>33 % v/v PEG 400</td>
</tr>
<tr>
<td>2.28</td>
<td>0.1 M lithium sulfate/0.05 M di-sodium hydrogen phosphate</td>
<td>0.05 M citric acid</td>
<td>None</td>
<td>19 % w/v PEG 1000</td>
</tr>
<tr>
<td>2.29</td>
<td>0.2 M magnesium chloride/0.1 M potassium chloride</td>
<td>0.025 M sodium citrate</td>
<td>4.0</td>
<td>33 % v/v PEG 400</td>
</tr>
<tr>
<td>2.30</td>
<td>0.05 M zinc acetate</td>
<td>0.05 M MES</td>
<td>6.1</td>
<td>11 % w/v PEG 8000</td>
</tr>
<tr>
<td>2.31</td>
<td>0.3 M magnesium nitrate</td>
<td>0.1 M Tris</td>
<td>8.0</td>
<td>22 % w/v PEG 8000</td>
</tr>
<tr>
<td>2.32</td>
<td>0.1 M sodium chloride/4% v/v ethylene glycol</td>
<td>0.1 M MES</td>
<td>6.5</td>
<td>33 % v/v PEG 400</td>
</tr>
<tr>
<td>2.33</td>
<td>0.05 M sodium chloride</td>
<td>0.1 M sodium citrate</td>
<td>5.5</td>
<td>26 % v/v PEG 400</td>
</tr>
<tr>
<td>2.34</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M glycine</td>
<td>9.3</td>
<td>30 % v/v PEG 400</td>
</tr>
<tr>
<td>2.35*</td>
<td>0.15 M potassium citrate/0.05 M lithium citrate</td>
<td>0.1 M sodium phosphate</td>
<td>-</td>
<td>22 % w/v PEG 6000</td>
</tr>
<tr>
<td>2.36</td>
<td>0.001 M zinc sulfate</td>
<td>0.05 M HEPES</td>
<td>7.8</td>
<td>28 % v/v PEG 600</td>
</tr>
<tr>
<td>2.37</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M sodium phosphate</td>
<td>7.0</td>
<td>33 % v/v PEG 300</td>
</tr>
<tr>
<td>2.38</td>
<td>0.1 M sodium chloride</td>
<td>0.05 M Bicine</td>
<td>9.0</td>
<td>33 % v/v PEG 300</td>
</tr>
<tr>
<td>2.39</td>
<td>0.05 M zinc acetate/6% v/v ethylene glycol</td>
<td>0.1 M sodium cacodylate</td>
<td>6.0</td>
<td>6.6 % w/v PEG 8000</td>
</tr>
<tr>
<td>2.40</td>
<td>0.2 M lithium sulfate</td>
<td>0.1 M sodium citrate</td>
<td>3.5</td>
<td>28 % v/v PEG 400</td>
</tr>
<tr>
<td>2.41</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M Tris</td>
<td>7.5</td>
<td>11 % w/v PEG 4000</td>
</tr>
<tr>
<td>2.42*</td>
<td>0.05 M lithium sulfate</td>
<td>0.1 M tricine</td>
<td>7.4</td>
<td>7 % w/v PEG 3000</td>
</tr>
<tr>
<td>2.43*</td>
<td>0.2 M calcium chloride</td>
<td>0.1 M MES</td>
<td>6.5</td>
<td>33% v/v PEG 400</td>
</tr>
<tr>
<td>2.44*</td>
<td>1 M sodium chloride</td>
<td>0.1 M sodium citrate</td>
<td>6.0</td>
<td>28% w/v PEG 4000</td>
</tr>
<tr>
<td>2.45*</td>
<td>None</td>
<td>0.1 M HEPES</td>
<td>7.5</td>
<td>11 % w/v PEG 4000</td>
</tr>
<tr>
<td>2.46*</td>
<td>0.002 M zinc sulfate</td>
<td>0.08 M HEPES</td>
<td>7.0</td>
<td>25 % v/v Jeffamine</td>
</tr>
<tr>
<td>2.47*</td>
<td>0.001 M cadmium chloride/0.03 M magnesium chloride</td>
<td>0.1 M MES</td>
<td>6.5</td>
<td>30 % v/v PEG 400</td>
</tr>
<tr>
<td>2.48*</td>
<td>None</td>
<td>0.1 M bis-tris-propane</td>
<td>7.0</td>
<td>3.0 M sodium chloride</td>
</tr>
</tbody>
</table>
### 7.2 MEMSYS/MEMSTART SCREEN

<table>
<thead>
<tr>
<th>Tube #</th>
<th>Salt</th>
<th>Buffer</th>
<th>pH</th>
<th>Precipitant</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>None</td>
<td>0.1 M sodium acetate</td>
<td>4.6</td>
<td>2 M ammonium sulfate</td>
</tr>
<tr>
<td>A2</td>
<td>None</td>
<td>0.1 M ADA</td>
<td>6.5</td>
<td>1 M ammonium sulfate</td>
</tr>
<tr>
<td>A3</td>
<td>None</td>
<td></td>
<td></td>
<td>2 M ammonium sulfate</td>
</tr>
<tr>
<td>A4</td>
<td>None</td>
<td>0.1 M Tris</td>
<td>8.5</td>
<td>2 M ammonium sulfate</td>
</tr>
<tr>
<td>A5</td>
<td>None</td>
<td>0.1 M Na HEPES</td>
<td>7.5</td>
<td>1.5 M lithium sulfate</td>
</tr>
<tr>
<td>A6</td>
<td>None</td>
<td>0.1 M sodium acetate</td>
<td>4.6</td>
<td>1 M magnesium sulfate</td>
</tr>
<tr>
<td>A7</td>
<td>None</td>
<td>0.1 M tri-sodium citrate</td>
<td>5.6</td>
<td>1 M magnesium sulfate</td>
</tr>
<tr>
<td>A8</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M ADA</td>
<td>6.5</td>
<td>1 M magnesium sulfate</td>
</tr>
<tr>
<td>A9</td>
<td>None</td>
<td>0.1 M ammonium dihydrogen phosphate</td>
<td>6.5</td>
<td>None</td>
</tr>
<tr>
<td>A10</td>
<td>0.1 M ammonium sulfate</td>
<td>0.5 M di-potassium hydrogen phosphate/ 0.5 M di-sodium hydrogen phosphate</td>
<td>7.5</td>
<td>None</td>
</tr>
<tr>
<td>A11</td>
<td>0.1M lithium sulfate</td>
<td>0.1 M sodium acetate</td>
<td>4.6</td>
<td>1 M ammonium dihydrogen phosphate</td>
</tr>
<tr>
<td>A12</td>
<td>None</td>
<td>0.1 M tri-sodium citrate</td>
<td>5.6</td>
<td>1 M ammonium dihydrogen phosphate</td>
</tr>
<tr>
<td>B1</td>
<td>None</td>
<td>0.1 M Tris</td>
<td>8.5</td>
<td>2 M ammonium dihydrogen phosphate</td>
</tr>
<tr>
<td>B2</td>
<td>None</td>
<td></td>
<td></td>
<td>2 M sodium formate</td>
</tr>
<tr>
<td>B3</td>
<td>None</td>
<td></td>
<td></td>
<td>4 M sodium formate</td>
</tr>
<tr>
<td>B4</td>
<td>None</td>
<td>0.1 M MES</td>
<td>6.5</td>
<td>1.4 M sodium acetate</td>
</tr>
<tr>
<td>B5</td>
<td>None</td>
<td>0.1 M Na HEPES</td>
<td>7.5</td>
<td>1 M ammonium dihydrogen phosphate</td>
</tr>
<tr>
<td>B6</td>
<td>None</td>
<td>0.1 M Na HEPES</td>
<td>7.5</td>
<td>1 M potassium sodium tartrate</td>
</tr>
<tr>
<td>B7</td>
<td>None</td>
<td>0.1 M Na HEPES</td>
<td>7.5</td>
<td>2 % v/v PEG 400/ 2 M ammonium sulfate</td>
</tr>
<tr>
<td>B8</td>
<td>0.1M magnesium chloride</td>
<td>0.1 M sodium acetate</td>
<td>4.6</td>
<td>30 % v/v PEG 400</td>
</tr>
<tr>
<td>B9</td>
<td>0.1M sodium chloride</td>
<td>0.1 M tri-sodium citrate</td>
<td>5.6</td>
<td>30 % v/v PEG 400</td>
</tr>
<tr>
<td>B10</td>
<td>0.1M lithium sulfate</td>
<td>0.1 M tri-sodium citrate</td>
<td>5.6</td>
<td>30 % v/v PEG 400</td>
</tr>
<tr>
<td>B11</td>
<td>0.3 M lithium sulfate</td>
<td>0.1 M ADA</td>
<td>6.5</td>
<td>30 % v/v PEG 400</td>
</tr>
<tr>
<td>B12</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M Na HEPES</td>
<td>7.5</td>
<td>30 % v/v PEG 400</td>
</tr>
<tr>
<td>C1</td>
<td>0.1 M ammonium sulfate</td>
<td>0.1 M Na HEPES</td>
<td>7.5</td>
<td>30 % v/v PEG 400</td>
</tr>
<tr>
<td>C2</td>
<td>0.2 M tri-sodium citrate</td>
<td>0.1 M Tris</td>
<td>8.5</td>
<td>30 % v/v PEG 400</td>
</tr>
<tr>
<td>C3</td>
<td>0.1 M zinc acetate</td>
<td>0.1 M sodium acetate</td>
<td>4.6</td>
<td>12 % w/v PEG 4000</td>
</tr>
<tr>
<td>C4</td>
<td>0.2 M ammonium sulfate</td>
<td>0.1 M sodium acetate</td>
<td>4.6</td>
<td>12 % w/v PEG 4000</td>
</tr>
<tr>
<td>C5</td>
<td>None</td>
<td>0.1 M sodium acetate</td>
<td>4.6</td>
<td>12 % w/v PEG 4000</td>
</tr>
<tr>
<td>C6</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M tri-sodium citrate</td>
<td>5.6</td>
<td>12 % w/v PEG 4000</td>
</tr>
<tr>
<td>C7</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M tri-sodium citrate</td>
<td>5.6</td>
<td>12 % w/v PEG 4000</td>
</tr>
<tr>
<td>C8</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M ADA</td>
<td>6.5</td>
<td>12 % w/v PEG 4000</td>
</tr>
<tr>
<td>C9</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M Na HEPES</td>
<td>7.5</td>
<td>12 % w/v PEG 4000</td>
</tr>
<tr>
<td>C10</td>
<td>0.1 M ammonium sulfate</td>
<td>0.1 M Na HEPES</td>
<td>7.5</td>
<td>12 % w/v PEG 4000</td>
</tr>
<tr>
<td>C11</td>
<td>0.2 M magnesium chloride</td>
<td>0.1 M Tris</td>
<td>8.5</td>
<td>12 % w/v PEG 4000</td>
</tr>
<tr>
<td>C12</td>
<td>0.2 M lithium sulfate hydrate</td>
<td>0.1 M Tris</td>
<td>8.5</td>
<td>12 % w/v PEG 4000</td>
</tr>
<tr>
<td>D1</td>
<td>0.2 M ammonium sulfate</td>
<td>None</td>
<td></td>
<td>12 % w/v PEG 4000</td>
</tr>
<tr>
<td>D2</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M sodium acetate</td>
<td>4.6</td>
<td>12 % w/v PEG 6000</td>
</tr>
<tr>
<td>D3</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M sodium acetate</td>
<td>4.6</td>
<td>12 % W/v PEG 6000</td>
</tr>
<tr>
<td>D4</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M ADA</td>
<td>6.5</td>
<td>12 % w/v PEG 6000</td>
</tr>
<tr>
<td>D5</td>
<td>0.1 M di-ammonium hydrogen phosphate</td>
<td>0.1 M Tris</td>
<td>8.5</td>
<td>12 % w/v PEG 6000</td>
</tr>
<tr>
<td>D6</td>
<td>1 M lithium sulfate</td>
<td>None</td>
<td></td>
<td>2 % w/v PEG 8000</td>
</tr>
<tr>
<td>D7</td>
<td>0.2 M sodium acetate</td>
<td>0.1 M MES</td>
<td>6.5</td>
<td>10 % w/v PEG 8000</td>
</tr>
<tr>
<td>D8</td>
<td>0.2 M zinc acetate</td>
<td>0.1 M MES</td>
<td>6.5</td>
<td>10 % w/v PEG 8000</td>
</tr>
<tr>
<td>D9</td>
<td>0.2 M calcium acetate</td>
<td>0.1 M MES</td>
<td>6.5</td>
<td>10 % w/v PEG 8000</td>
</tr>
<tr>
<td>D10</td>
<td>None</td>
<td>0.1 M Tris</td>
<td>8.5</td>
<td>10 % w/v PEG 8000</td>
</tr>
<tr>
<td>D11</td>
<td>0.2 M ammonium sulfate</td>
<td>None</td>
<td></td>
<td>10 % w/v PEG 8000</td>
</tr>
<tr>
<td>D12</td>
<td>0.5 M lithium sulfate</td>
<td>None</td>
<td></td>
<td>10 % w/v PEG 8000</td>
</tr>
<tr>
<td>Tube #</td>
<td>Salt 1</td>
<td>Salt 2</td>
<td>Buffer</td>
<td>pH</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------</td>
<td>-------------------------</td>
<td>-------------</td>
<td>-----</td>
</tr>
<tr>
<td>E1</td>
<td>None</td>
<td>None</td>
<td>0.1 M Na citrate</td>
<td>5.5</td>
</tr>
<tr>
<td>E2</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M Na citrate</td>
<td>3.5</td>
</tr>
<tr>
<td>E3</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M Na citrate</td>
<td>4.5</td>
</tr>
<tr>
<td>E4</td>
<td>0.1 M sodium chloride</td>
<td>None</td>
<td>0.1 M Na citrate</td>
<td>5.5</td>
</tr>
<tr>
<td>E5</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M Na citrate</td>
<td>5.5</td>
</tr>
<tr>
<td>E6</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M Na citrate</td>
<td>5.5</td>
</tr>
<tr>
<td>E7</td>
<td>None</td>
<td>None</td>
<td>0.1 M MES</td>
<td>6.5</td>
</tr>
<tr>
<td>E8</td>
<td>None</td>
<td>None</td>
<td>0.1 M MES</td>
<td>6.5</td>
</tr>
<tr>
<td>E9</td>
<td>0.1 M sodium chloride</td>
<td>None</td>
<td>0.1 M MES</td>
<td>6.5</td>
</tr>
<tr>
<td>E10</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M MES</td>
<td>6.5</td>
</tr>
<tr>
<td>E11</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M MES</td>
<td>6.5</td>
</tr>
<tr>
<td>E12</td>
<td>None</td>
<td>None</td>
<td>0.1 M MOPS</td>
<td>7.0</td>
</tr>
<tr>
<td>F1</td>
<td>None</td>
<td>None</td>
<td>0.1 M Na HEPEs</td>
<td>7.5</td>
</tr>
<tr>
<td>F2</td>
<td>0.1 M sodium chloride</td>
<td>None</td>
<td>0.1 M MOPS</td>
<td>7.0</td>
</tr>
<tr>
<td>F3</td>
<td>None</td>
<td>None</td>
<td>0.1 M Na HEPEs</td>
<td>7.5</td>
</tr>
<tr>
<td>F4</td>
<td>0.1 M sodium chloride</td>
<td>None</td>
<td>0.1 M Na HEPEs</td>
<td>7.5</td>
</tr>
<tr>
<td>F5</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M Na HEPEs</td>
<td>7.5</td>
</tr>
<tr>
<td>F6</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M Na HEPEs</td>
<td>7.5</td>
</tr>
<tr>
<td>F7</td>
<td>None</td>
<td>None</td>
<td>0.1 M Tris</td>
<td>8.5</td>
</tr>
<tr>
<td>F8</td>
<td>0.1 M sodium chloride</td>
<td>None</td>
<td>0.1 M Tris</td>
<td>8.5</td>
</tr>
<tr>
<td>F9</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M Tris</td>
<td>8.5</td>
</tr>
<tr>
<td>F10</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M Tris</td>
<td>8.5</td>
</tr>
<tr>
<td>F11</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M CAPSO</td>
<td>9.5</td>
</tr>
<tr>
<td>F12</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M CAPSO</td>
<td>9.5</td>
</tr>
<tr>
<td>G1</td>
<td>None</td>
<td>None</td>
<td>0.1 M Na citrate</td>
<td>5.5</td>
</tr>
<tr>
<td>G2</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M Na citrate</td>
<td>3.5</td>
</tr>
<tr>
<td>G3</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M Na citrate</td>
<td>4.5</td>
</tr>
<tr>
<td>G4</td>
<td>0.1 M sodium chloride</td>
<td>None</td>
<td>0.1 M Na citrate</td>
<td>5.5</td>
</tr>
<tr>
<td>G5</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M Na citrate</td>
<td>5.5</td>
</tr>
<tr>
<td>G6</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M Na citrate</td>
<td>5.5</td>
</tr>
<tr>
<td>G7</td>
<td>None</td>
<td>None</td>
<td>0.1 M MES</td>
<td>6.5</td>
</tr>
<tr>
<td>G8</td>
<td>None</td>
<td>None</td>
<td>0.1 M MES</td>
<td>6.5</td>
</tr>
<tr>
<td>G9</td>
<td>0.1 M sodium chloride</td>
<td>None</td>
<td>0.1 M MES</td>
<td>6.5</td>
</tr>
<tr>
<td>G10</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M MES</td>
<td>6.5</td>
</tr>
<tr>
<td>G11</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M MES</td>
<td>6.5</td>
</tr>
<tr>
<td>G12</td>
<td>None</td>
<td>None</td>
<td>0.1 M MOPS</td>
<td>7.0</td>
</tr>
<tr>
<td>H1</td>
<td>None</td>
<td>None</td>
<td>0.1 M Na HEPEs</td>
<td>7.5</td>
</tr>
<tr>
<td>H2</td>
<td>0.1 M sodium chloride</td>
<td>None</td>
<td>0.1 M MOPS</td>
<td>7.0</td>
</tr>
<tr>
<td>H3</td>
<td>None</td>
<td>None</td>
<td>0.1 M Na HEPEs</td>
<td>7.5</td>
</tr>
<tr>
<td>H4</td>
<td>0.1 M sodium chloride</td>
<td>None</td>
<td>0.1 M Na HEPEs</td>
<td>7.5</td>
</tr>
<tr>
<td>H5</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M Na HEPEs</td>
<td>7.5</td>
</tr>
<tr>
<td>H6</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M Na HEPEs</td>
<td>7.5</td>
</tr>
<tr>
<td>H7</td>
<td>None</td>
<td>None</td>
<td>0.1 M Tris</td>
<td>8.5</td>
</tr>
<tr>
<td>H8</td>
<td>0.1 M sodium chloride</td>
<td>None</td>
<td>0.1 M Tris</td>
<td>8.5</td>
</tr>
<tr>
<td>H9</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M Tris</td>
<td>8.5</td>
</tr>
<tr>
<td>H10</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M Tris</td>
<td>8.5</td>
</tr>
<tr>
<td>H11</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M lithium sulfate</td>
<td>0.1 M CAPSO</td>
<td>9.5</td>
</tr>
<tr>
<td>H12</td>
<td>0.1 M sodium chloride</td>
<td>0.1 M magnesium chloride</td>
<td>0.1 M CAPSO</td>
<td>9.5</td>
</tr>
</tbody>
</table>

**Abbreviations:**
ADA: N-(2-Acetamido)iminoacetic acid, CAPSO: 3-(Cyclohexylamino)-2-hydroxy-1-propanesulfonic acid Sodium Salt, Na HEPEs: N-(2-hydroxyethyl)-piperazine-N'-2-ethanesulfonic acid, sodium salt MES: 2-(N-morpholino) ethanesulfonic acid, MME: Monomethylether, MOPS: 3-(N-Morpholino) propanesulfonic acid, PEG: Polyethylene glycol, Tris: 2-Amino-2-(hydroxymethyl) propane-1,3-diol.