Technology development for the over-expression, purification and crystallisation of human membrane proteins

Gillian Young
B.Sc.(Hons.) M.Res.

Submitted in part fulfilment for the degree of Doctor of Philosophy

Institute of Cellular, Molecular and Systems Biology

College of Medical, Vetinary and Life Sciences

University of Glasgow

2010

Author's Declaration

This thesis has been written in accordance with the University of Glasgow's regulations and is less then 50,000 words in length. This thesis is an original contribution, which describes work performed entirely by myself unless otherwise cited or acknowledged. Its contents have not previously been submitted for any other degree. The research for this thesis was performed between June 2007 and December 2010.

Gillian Young

Abstract

Currently, the field of mammalian membrane protein structural biology is in its infancy. Existing technologies and experiences have shown that it is possible to obtain the structures of mammalian membrane proteins if sufficient work and thought has been invested. However, there is still an urgent need to develop new methodologies and approaches to improve all aspects of this important area of biological research. Here, a series of novel technologies for the overproduction, purification and crystallisation of human membrane proteins are described which have been tested with a representative member from each of the G-protein coupled receptor (adenosine 2a receptor (A2aR)) and membrane enzyme (sterol isomerase (SI)) superfamilies.

The methylotrophic yeast *Pichia pastoris* is an excellent host cell for the overproduction of recombinant proteins including membrane proteins of mammalian origin. However, the commercially available expression vectors are far from what is required to maximise the production levels as well as simplify the detergent extraction and purification of human membrane proteins. Here, a series of related expression constructs were made that had different combinations of tags at both ends of the recombinant protein. The final optimised expression vectors had a C3 protease-iLOV-biotin acceptor-His10 (CLBH) tag fused to the C-terminus of the recombinant protein. The -CLBH vectors gave high level production of both test proteins (one N_{in} – hSI; one N_{out} – hA2aR) that could be rapidly purified to homogeneity using a generic protocol. The position of the His10 tag did not affect the expression level of the recombinant protein. In contrast, fusion of the biotin acceptor domain to the C-terminus of the recombinant protein increased its expression by a factor of between 2-4. The biotin acceptor domain could also be fully biotinylated *in vitro* using recombinantly expressed biotin ligase allowing purification/immobilisation of the target protein with streptavidin beads. Removal of the expression/ purification tags from the recombinant proteins with C3 protease occurred more efficiently than when TEV protease was used.

An optimised protocol was developed that gave maximal production of our target proteins in fermenter culture at an induction temperature of 22°C. Care was taken to find a methanol feed rate that gave the highest levels of protein production without causing the accumulation of excess methanol in the culture (which is known to be toxic to the yeast). Using this protocol it was possible to make both hSI and hA2aR with a production level >10 mg of recombinant protein per litre of culture.

As most MPs are colourless, target protein identification is usually performed by methods such as radioligand binding and/or Western blotting. However, these techniques can be time-consuming, use a lot of protein and do not give any information on the aggregation state of the protein in detergent solution. Previously, it has been shown that the processes of identifying and analysing membrane proteins in detergent solution can be accelerated by attaching green-fluorescent protein to the C-terminus of the recombinant MP. Here, the potential of the recently described iLOV fluorescence tag for membrane protein applications was assessed. iLOV was shown to be an useful tool for optimising processes such as yeast clonal selection, protein production in fermenter culture, detergent and construct screening as well as tracking recombinant MPs through the purification process. Of note, the iLOV tag allowed a direct assessment of the stability and dispersity state of both target MPs in a range of detergents by fluorescence size exclusion chromatography (FSEC). Using this approach, it was shown that wild-type hA2aR solubilised using a combination of dodecyl-βD-maltoside (DDM) and cholesteryl-hemisuccinate (CHS) aggregated during purification on a Ni^{2+} column. Furthermore, it was shown that the hA2aR agonist-conformationally-fixed mutant Rag23 is stable in DDM without any CHS present. Moreover, Rag23 was found to be monodisperse in a series of short-chain detergents (decyl-βD-maltoside, nonyl-βD-maltoside (NM) and β-octylglucoside) suggesting that this mutant is well-suited to structural studies. SI was remarkably robust in short chain detergents demonstrating a reasonable level of stability in the short chain detergent NM. The FSEC experiments showed that wild-type SI has considerably higher intrinsic stability than native hA2aR suggesting that

membrane enzymes will prove to be more amenable to structural analysis than GPCRs.

Rag23 and SI were both purified to homogeneity in a simple four-step procedure: i) Ni^{2+} purification, ii) cleavage with C3 protease, iii) reverse Ni^{2+} purification and iv) gel-filtration chromatography. A buffer/salt screen was devised that allowed those conditions where SI had maximal thermostability in detergent-solution to be identified. SI was found to have greatest stability in sodium phosphate buffer at acidic pH. Using this information, it was possible to purify monodisperse SI in DM suggesting that this protein may make an excellent candidate for structural studies too.

Crystallisation trials with SI were performed using the commercially available sparse matrix screen MemSys/MemStart. In addition, a lipidic-sponge phase sparse-matrix crystallisation screen that was developed in collaboration with Prof. Richard Neutze (University of Chalmers, Sweden) was tested using SI. Cholesterol could be incorporated into all of the sponges that make up the screen upto a concentration of 10%. (This is important as the activity of many mammalian membrane proteins is cholesterol-dependent). To date, no diffracting crystals of SI have been obtained with either the conventional or lipidic-sponge phase crystallisation approaches.

In short, a series of novel technologies/methodologies have been developed that will act as a platform for future efforts to solve the structures of a wide-range of human membrane proteins.

Acknowledgements

I would like to sincerely thank my supervisors Dr. Niall Fraser and Prof. Richard Cogdell FRS without whom I would have never started a Ph.D. Their belief in my abilities has helped me to complete my project and I want to thank them for the immense opportunities I have had. Additionally, the invaluable advice and guidance Niall has given me has helped to shape me as a scientist and I will always be grateful for his encouragement.

I would also like to thank Prof. Gordon Lindsay for helping me to obtain my 4-year BBSRC studentship.

Throughout my Ph.D. I have had the pleasure of working with a variety of lovely and talented people. My dear friend Alette Brinth has been a constant support to me, and I will always be grateful for her company in the lab as well as being able to bounce ideas around with her. Khuram Ashraf has been a wonderful help to me in the latter stages of my Ph.D. for which I will always be indebted to him. I want to thank Rachel Mulvaney for having me stay with her during the final stages of my Ph.D. These three students have become my close friends and I will be forever thankful for our tea-breaks, lab-dinners and laughs together!

My thanks also goes to the members of the Cogdell and PX groups (in particular Tatas Brotosudarmo, Mads Gabrielsen, Alastair Gardiner, Nichola Picken, Alan Riboldi-Tunnicliffe, Aleks Roszak and June Southall) for all their advice and help over the years. I want to specially mention Susan Kitson without whom labwork would have been considerably harder, lonelier and definitely less fun!

I have been extremely fortunate throughout my Ph.D. to have had the opportunity to travel to various labs around the world and, therefore, have a list of collaborators I would like to acknowledge. For help with yeast fermentations – Roslyn Bill, Richard Darcy and Mohammed Jamshad (Aston University); learning how to make lipidic sponges for crystallisation experiments - Linda Johansson, Anne-Marie Whöri and Richard Neutze (Chalmers University, Sweden); learning the CPM assay – VP Jaakola and Ray Stevens (The Scripps Research Institute, US).

I want to extend my sincere thanks to Dr. Chris Tate for his mentoring and invaluable advice on protein purification as well as access to his adenosine 2a receptor mutants.

I would never have embarked on the challenge of a post-graduate degree had it not been for the amazing support of my family. My mum in particular advised me to follow my dreams when faced with the decision of returning to University and I am eternally grateful to her for helping me to choose the more challenging option. I want to thank my dad for his words of encouragement as well as my sister Carole for her ability to cheer me up during those times when writing this thesis were particularly difficult. Without their love and support I wouldn't be where I am today.

Finally and most importantly, I want to thank my fiancé Mark for his unending support. He, more than anyone, understands the emotional challenges a Ph.D. entails. I am so lucky to have had someone so amazing and loving to support and encourage me through the hard times as well as being there to share the thrill of the good times with me also.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Definition
3D	3 dimensional
A_{280}	absorbance at 280 nm
A1R	adenosine 1 receptor
A2aR	adenosine 2a receptor
A2bR	adenosine 2b receptor
A3R	adenosine 3 receptor
ABC	ATP-binding cassette
ADP	adenosine diphosphate
AMP	adenosine monophosphate
ARNO	ADP-ribosylation factor nucleotide-binding-site opener
ATP	adenosine triphosphate
BCA	bicinchoninic acid
βOG	n-octyl-β-D-glucopyranoside
BSA	bovine serum albumin
cAMP	cyclic adenosine monophosphate
CD	circular dichroism
CHH	Conradi-Hünermann-Happle
CHS	cholesterol hemisuccinate
CLBH	C3-iLOV-biotin-His construct
CMC	critical micelle concentration
CPDX2	chondrodysplasia punctata type II
CPM	7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin
D2R	dopamine D2 receptor
Da	Dalton
DAG	diacylglycerol
DDM	n-dodecyl-β-D-maltoside
DM	n-decyl-β-D-maltoside
DMSO	dimethyl sulfoxide
dO_2	dissolved Oxygen
DRY motif	Aspartic acid, Arginine, Tyrosine motif
DTT	dithiothreitol
EBP	emopamil-binding protein

EDTA	ethylenediaminetetraacetic acid
EGFR	epidermal growth factor receptor
EL	extracellular loop
ERK	extracellular-signal related kinase
F_{Ab}	antibody binding fragment
FRET	fluorescence energy resonance transfer
FSEC	fluorescence size exclusion chromatography
FU	fluorescence units
G-protein	guanine nucleotide binding protein
GABA	γ-aminobutyric acid
GDP	guanine diphosphate
GEF	guanine exchange factor
GFP	green fluorescent protein
GIP	GPCR-interacting protein
GPCR	G-protein coupled receptor
GRK	G-protein coupled receptor kinase
GTP	guanosine-5'-triphosphate
GTPase	guanosine-5'-triphosphate hydolase
hA2aR	human adenosine 2a receptor
HDL	high-density lipoproteins
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMG-CoA	3-hydroxy-3-methylglutaryl co-enzyme A
HRV C3	human rhinovirus C3 protease
HT-TLB	His-TEV-protein-TEV-iLOV-biotin construct
IL	intracellular loop
iLOV	improved light, oxygen, voltage sensor
IMAC	immobilised metal affinity chromatography
IMM	inner mitochondrial membrane
IP_3	inositol trisphosphate
IPTG	isopropyl-β-D-thiogalactopyranoside
JAK	Janus kinase family
JNK	Jun N-terminal kinase
K_D	dissociation constant
LDAO	lauryldimethylamine-oxide
LDL	low-density lipoproteins

MAPK	mitogen-activated protein kinase
MES	2-(N-morpholino)ethanesulfonic acid
MFS	major facilitator superfamily
MMP	mammalian membrane protein
MO	mono-olein
MOPS	3-(N-morpholino)propanesulfonic acid
MP	membrane protein
MW	molecular weight
MWCO	molecular weight cut off
NECA	5'-*N*-ethylcarboxamidoadenosine
NM	n-nonyl-β-D-maltoside
OD_{600}	optical density at 600 nm
PCR	polymerase chain reaction
PDB	protein databank
PEG	polyethylene glycol
PS	photosystem
Rag23	agonist thermostabilized adenosine 2a receptor
Rant21	antagonist thermostabilized adenosine 2a receptor
RC	reaction centre
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SI	sterol isomerase
SRC	non-receptor tyrosine kinase
STAT	signal transducers and activators of transcription
TEV	Tobacco Etch Virus
TM	transmembrane
TRIS	tris(hydroxymethyl)aminomethane
UV	ultraviolet
VDAC	voltage dependent anion channel
VLDL	very low-density lipoproteins
WT	wild type
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
ZM 241385	4-(2-[7-amino-2-(2-furyl)[1,2,4]triazolo[2,3-a][1,3,5]triazin-5-ylamino]ethyl)phenol

1 Membrane protein structural biology

1.1 Biological membranes

Membranes are an essential component of all cells where they function as selectively-permeable barriers. They consist of a rich mixture of lipid molecules and proteins (Figure 1.1). In mammalian cells, the plasma membrane physically separates the intracellular components from the extracellular environment (Brown, 1998). Moreover, intracellular membranes are used to form distinct compartments within the cell that perform highly-specialised functions (e.g. mitochondria – oxidative phosphorylation; lysosome – protein degradation) (Brown, 1998). All biological membranes are primarily composed of a thin layer of amphipathic molecules called phospholipids (which have polar hydrophilic head groups and hydrophobic tails) (Singer and Nicolson, 1972). In an aqueous environment the phospholipids spontaneously form bilayers with the hydrophobic tails sandwiched between the two hydrophilic headgroups (Figure 1.1) (Singer and Nicolson, 1972; Rietveld and Simons, 1998).

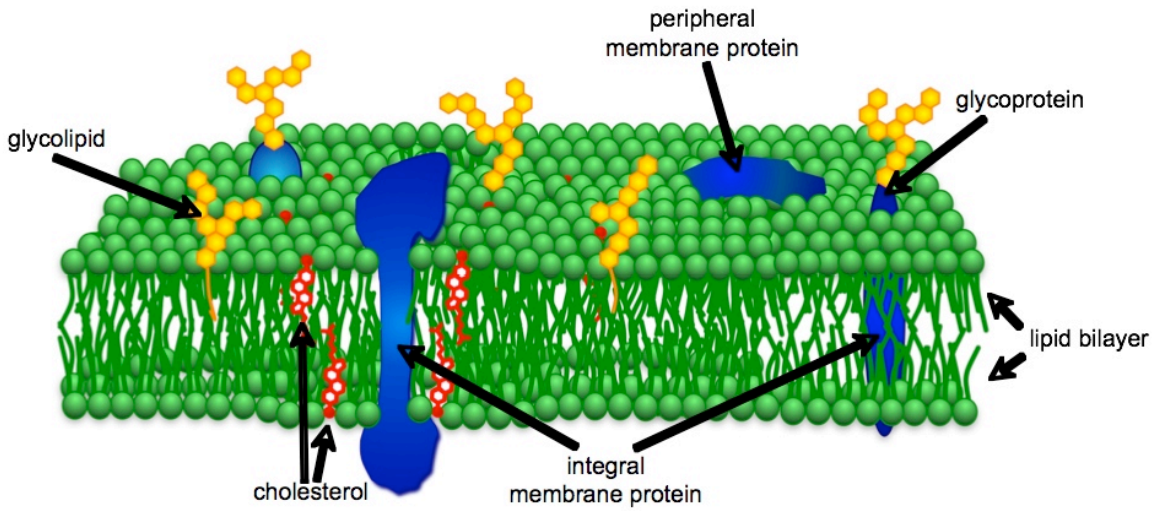


Figure 1.1 Schematic representation of a mammalian plasma membrane

Higher-eukaryotic membranes are a rich mixture of lipid molecules and proteins. Both peripheral and integral membrane proteins are located at and in the phospholipid bilayer, respectively. Mammalian membranes are distinct from those found in bacteria, lower eukaryotes and plants in that they contain cholesterol (which gives the membrane increased rigidity).

Although mammalian cell membranes are primarily composed of phospholipids, they also contain sphingolipids, glycolipids and cholesterol (Figure 1.2) (Rietveld and Simons, 1998). In their simplest form, phospholipids consist of a glycerol backbone to which two fatty acids and a phosphate group are connected. Mammalian cells make a range of related phospholipids (phosphatidyl-choline, -ethanolamine, -serine, -glycerol and -inositol) by attaching different headgroups to the core phosphate moiety (Figure 1.2). Many different fatty acids (including both saturated and unsaturated versions) can be added to the glycerol backbone. The biophysical properties of the fatty acid that are attached influence both the thickness and fluidity of the membrane (Becher and McIlhinney, 2005). Sphingolipids have a similar overall structure to phospholipids except that they have a sphingosine rather than a glycerol spine. Certain phospho- and sphingo-lipids located on the extracellular face of the plasma membrane have short carbohydrate chains attached to them and are called glycolipids (Figure 1.2). Cholesterol (Figure 1.2) is a major component of mammalian membranes and is exclusively found in higher eukaryotes. Cholesterol brings rigidity to membranes by restricting the movement of the fatty acids in neighbouring phospholipids through its bulky four-ring structure (Brown and London, 1998; Chini and Parenti, 2009).

Membranes are dynamic, and can be described as a “fluid mosaic” as they behave like a two-dimensional liquid in which all the lipid and protein molecules diffuse to a greater or lesser extent (Singer and Nicolson, 1972). There are regions in mammalian membranes that have high local concentrations of cholesterol and sphingolipids called lipid rafts (Brown and London, 1998; Brown, 1998). Rafts are more ordered and tightly packed than the surrounding bilayer and float freely in the membrane. Rafts function as specialised membrane microdomains that enable certain signalling and trafficking events at the cell’s surface to be compartmentalised (Rietveld and Simons, 1998; Simons and Ikonen, 2000; Becher and McIlhinney, 2005; Epand, 2006).

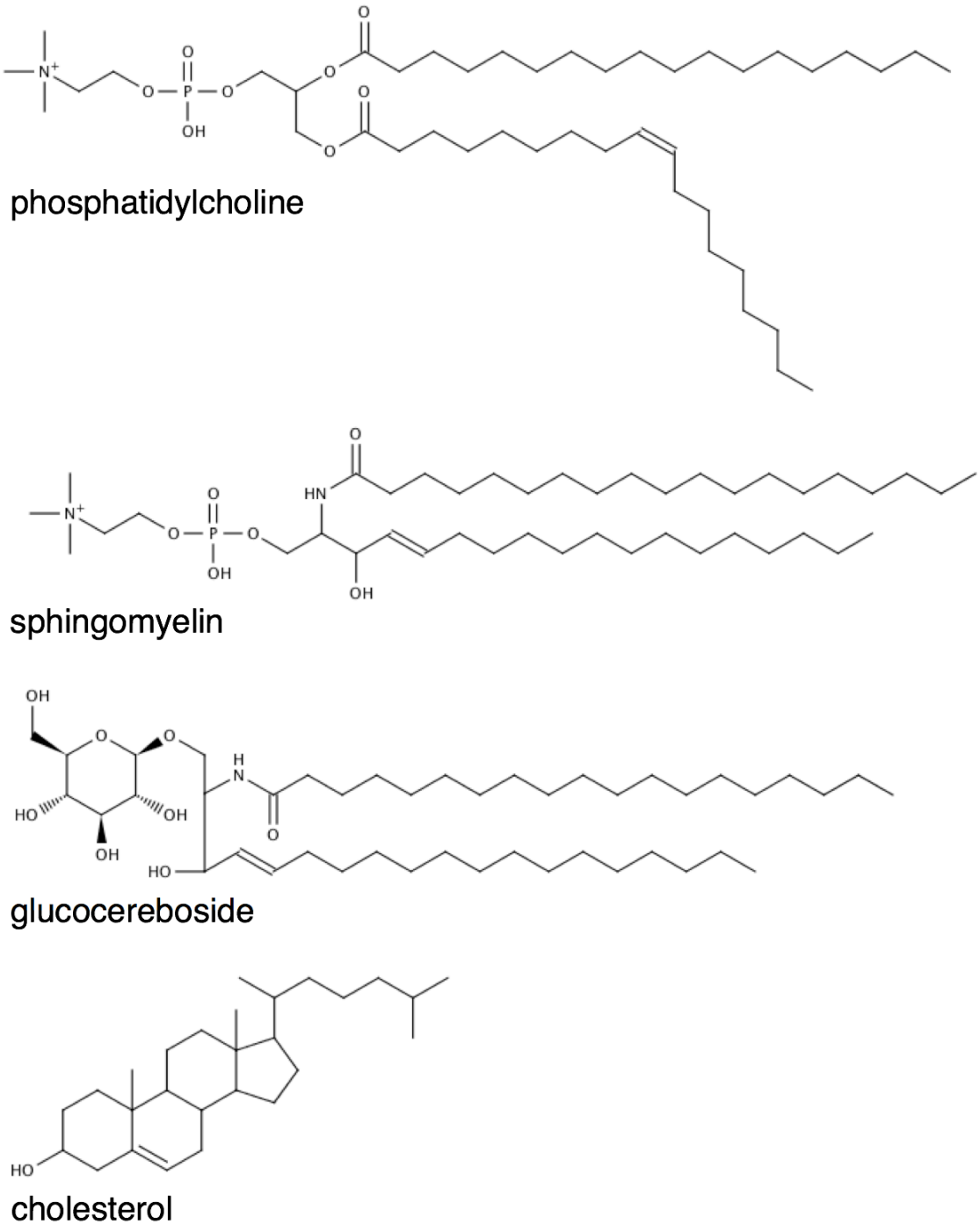


Figure 1.2 Chemical structures of selected lipids found in mammalian membranes

Higher-eukaryotic membranes contain four different kinds of lipid namely phospholipids (for example - phosphatidylcholine), sphingolipids (sphingomyelin), glycolipids (glucocereboside) and cholesterol.

1.2 Membrane proteins

Proteins found on/in membranes can be classed either as peripheral or integral membrane proteins according to the way that they attach to the membrane (Singer and Nicolson, 1972) (Figure 1.1). Peripheral membrane proteins (MPs) are located at the surface of membranes where they are anchored to them either directly by interacting with the lipid bilayer or indirectly via contact with an integral membrane protein. Peripheral MPs can usually be released from the membrane by washing with buffer of high ionic strength. In contrast, integral MPs (as their name suggests) span the entire membrane, and consist of two hydrophilic regions (that are exposed on either side of the membrane) separated by a hydrophobic region (that is buried within the membrane core). These proteins can only be released from the lipid bilayer by treating the membrane with detergent. As integral MPs need to be coated with detergent at all times in aqueous solution, they are considerably harder to work compared to either soluble of peripheral MPs. As even mild detergents are denaturants when integral MPs are removed from the membrane for structural studies it is not uncommon for the protein to lose its functional activity (Seddon *et al.*, 2004).

1.3 Membrane proteins and secondary structure

Following translation, proteins fold first into a series of defined small secondary structures from which a 3-dimensional shape is eventually formed. There are two main types of secondary structure found in all proteins called α-helix and β-sheet, both of which are formed by hydrogen bonding between the carbonyl and amine groups of the protein backbone (Figure 1.3) (Kennedy, 1978). α-helices are more common than β-sheet, and have a right-handed coiled conformation. In an α-helix, every backbone N-H group donates a hydrogen bond to the backbone C=O group of the amino acid four residues previous (Kabsch and Sander, 1983). There are *ca.* 3.6 amino acids per turn of helix (Pauling and Corey, 1951a; Pauling *et al.*, 1951). The other main type of secondary structure is β-sheet, and it consists of neighbouring β-strands (stretches of polypeptide where the backbone has an almost fully-extended conformation) that are hydrogen bonded

to each other (Figure 1.3) (Kabsch and Sander, 1983). Where the β-strands all have a similar orientation the β-sheet that is formed is denoted parallel. Where the β-strands have alternating orientations, the β-sheet is described as being anti-parallel (Pauling and Corey, 1951b).

The most common secondary structure found in integral MPs is the α-helix with a minimum of 17-20 residues required to span the bilayer (Kennedy, 1978; Nyholm *et al.*, 2007; McLuskey *et al.*, 2010). α-helices are energetically stable in the membrane as the amine and carbonyl groups of the peptide bonds are involved in hydrogen-bonding within the helix. The majority of the amino acids that occur within membrane-spanning helices tend to be hydrophobic in nature (Nyholm *et al.*, 2007). The presence of proline and/or glycine residues in a helix can cause it to kink (as was observed in the structure of rhodopsin) (Palczewski *et al.*, 2000). The number and orientation of the helices traversing the membrane, in combination with the variation of the connecting loop regions, creates a diverse range of membrane proteins including transporters, ion channels and receptors.

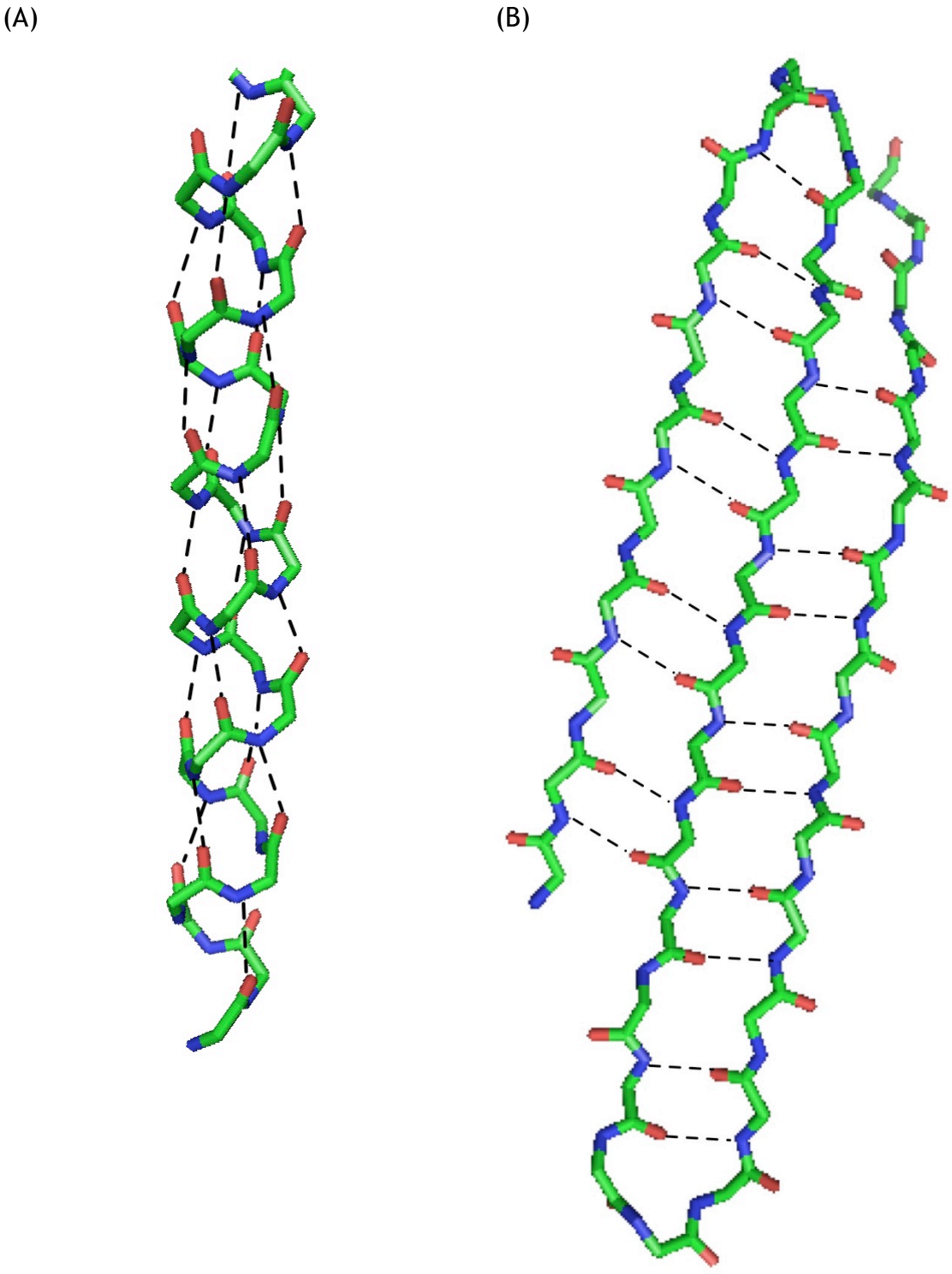


Figure 1.3 Representation of the two main types of protein secondary structure

(A) is an α-helix in which every backbone N-H group donates a hydrogen bond to the backbone C=O group of the amino acid four residues previous. There are *ca.* 3.6 amino acids per turn of

helix. (B) shows an anti-parallel β-sheet with hydrogen bond "bridges" between neighbouring β-strands. Hydrogen bonds are shown as black broken lines.

Alternatively, the plasma membrane can be traversed by a large β-sheet that has been twisted and coiled to form a closed structure called a β-barrel (Schulz, 2000). β-strands in β-barrels are arranged in an anti-parallel fashion with the first strand hydrogen bonded to the last (Kennedy, 1978; Cowan *et al.*, 1992; Buchanan, 1999). Unlike transmembrane α-helices that are largely composed of aliphatic amino acids, the β-sheets that make up β-barrels do not have obvious long hydrophobic regions which makes their identification from primary sequence analysis alone difficult (Schulz, 2000). Furthermore, whereas α-helical containing integral MPs have a ubiquitous distribution, proteins that have a β-barrel structure are only found in the outer membranes of gram-negative bacteria as well as mitochondria and chloroplasts (Schulz, 2000).

1.4 History of membrane protein structural biology

The PDB currently contains nearly 60,000 entries of which 700 are structures of integral membrane proteins. Of these, however, only 200-250 are regarded as unique structures (White, 2009; McLuskey *et al.*, 2010). The number of integral membrane protein structures solved to date does not reflect either their importance in biology or indeed their natural abundance (with 30% of the genes in a typical genome encoding MPs) (Arora and Tamm, 2001). The slow rate at which MPs have had their structures determined reflects the many challenges that have to be overcome before crystals that diffract to high resolution are obtained. These include the difficulties of making membrane proteins in a heterologous system on a milligram scale, preventing the protein from aggregating once it has been removed from its membrane environment as well as trying to find the optimal detergent for crystallisation. Furthermore, mammalian MPs have proved to be even more difficult to work with than bacterial MPs as they are harder to make in large amounts as well as being less stable in detergent solution. This is reflected in the PDB where there are <20 unique

mammalian membrane protein structures (White, 2009; McLuskey *et al.*, 2010). Having said this, membrane protein crystallisation is no longer considered as difficult, and with considerable advances in technology, *ca.* 30-40 new integral MP structures are being deposited each year in the PDB (Figure 1.4) (White, 2009). Here, a brief overview of the history of membrane protein structural biology is presented.

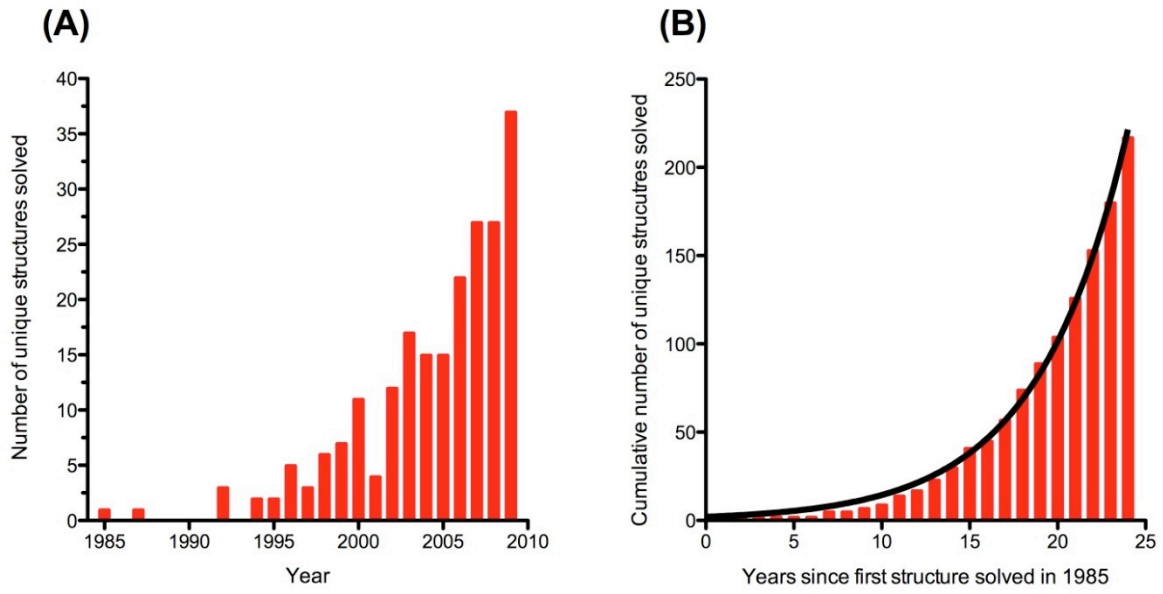


Figure 1.4 Progress in determining membrane protein structures

(A) The number of new structures reported each year since 1985. (B) A plot of the cumulative number of membrane protein structures solved (y-axis) versus the number of years since the first integral MP structure was solved (x-axis). The number of new membrane protein structures is increasing exponentially on a yearly basis. Figure adapted from (White, 2009).

In 1975, bacteriorhodopsin became the first membrane protein to have its structure elucidated. Using low-resolution (7 Å) electron microscopy data Richard Henderson and Nigel Unwin were able to unequivocally show that bacteriorhodopsin consisted of seven transmembrane α-helices (Henderson and Unwin, 1975). However, at this resolution, it was not possible to visualise the molecular details of bacteriorhodopsin's structure. It was another ten years before the first membrane protein structure at atomic resolution was obtained when Johann Deisenhoffer and Hartmut Michel solved the structure of the reaction centre (RC) from the purple photosynthetic bacteria

Rhodopseudomonas viridis to a resolution of 3 Å (Figure 1.5) (Deisenhofer *et al.*, 1984; Deisenhofer *et al.*, 1985). This landmark achievement was recognised when Michel, Deisenhoffer and Huber were awarded the Nobel Science prize for Chemistry in 1988. The RC is a protein-pigment complex of *ca.* 145 kDa consisting of three different sub-units (H, L, and M). The L and M subunits form the core of the RC and have five transmembrane α helices each (Michel, 1982). The H sub-unit has just one transmembrane helix which interacts with both L and M, and its globular domain is located in the bacteria's cytoplasm. The RC was solved in complex with a cytochrome that interacted with the periplasmic side of the core formed from the L and M sub-units (Deisenhofer and Michel, 1989). The protein acts as a scaffold holding the pigments at the correct geometry to allow efficient charge separation and electron transfer.

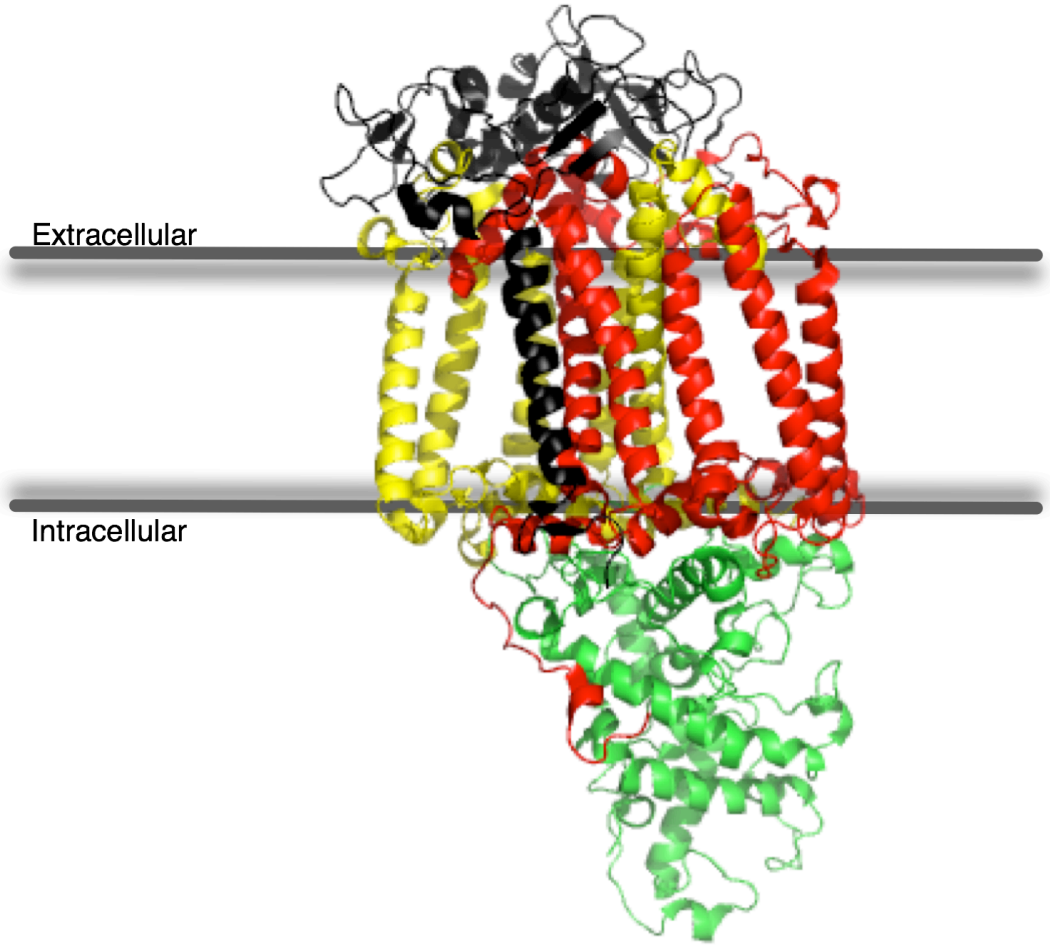


Figure 1.5 The first high-resolution structure of a membrane protein

The photosynthetic reaction centre from the purple bacteria *Rhodopseudomonas viridis* (accession code 1PRC) was the first membrane protein to have its structure solved to atomic resolution by X-ray crystallography (Deisenhofer *et al.*, 1984; Deisenhofer *et al.*, 1985). The heavy chain is coloured black, medium chains are red, light chains are yellow and the globular domain is coloured green.

Since this RC structure, several other photosynthetic integral MP complexes from purple bacteria have had their structures determined including the RC from *Rb. sphaeroides* (Yeates *et al.*, 1987; Stowell *et al.*, 1997), light harvesting complexes from *Rps. acidophila* (McDermott *et al.*, 1995; McLuskey *et al.*, 2001) and *Rs. molischianum* (Koepke *et al.*, 1996), as well as the "core" complex from *Rps. palustris* (Roszak *et al.*, 2003). In addition, structures of photosystem (PS) II (Zouni *et al.*, 2001; Kamiya and Shen, 2003; Ferreira *et al.*, 2004) and PSI (Krauss *et al.*, 1996; Schubert *et al.*, 1997) from cyanobacteria, as well as PSI (Ben-Shem *et al.*, 2003; Amunts *et al.*, 2007) and light-harvesting complex II (LHCII) (Liu *et al.*, 2004; Standfuss *et al.*, 2005) from plants (all integral membrane proteins) have been determined to atomic resolution. The structures of these photosynthetic complexes have given detailed insights into the molecular mechanisms of light capture, charge separation and water-splitting which occur in nature.

Arguably, the most important process that occurs within cells is the production of adenosine triphosphate (ATP) by oxidative phosphorylation. This process occurs in and across the inner mitochondrial membrane (IMM) (where both the electron carriers and ATP synthase are located). Electrons are transferred from either NADH or $FADH_2$ to molecular oxygen which is reduced to water. Electron transport through complexes I (NADH dehydrogenase), III (Cytochrome bc1 complex) and IV (Cytochrome oxidase) results in H^+ pumping across the IMM. The proton-gradient which results is used to drive the formation of adenosine-triphosphate (ATP) from inorganic phosphate and adenosine diphosphate (ADP) via the integral membrane protein ATP synthase. The structures of all four electron transfer complexes (I-IV) (Iwata *et al.*, 1995; Tsukihara *et al.*, 1996; Iwata, 1998; Iwata *et al.*, 1998; Yankovskaya *et al.*, 2003; Hunte *et al.*, 2010) and ATP synthase (Abrahams *et al.*, 1994; Stock *et al.*, 1999) (Figure 1.6) have

all been solved to atomic resolution. The structure of ATP synthase was solved by John Walker and co-workers, for which he was awarded the Nobel Prize for Chemistry in 1997. ATP synthase has three distinct domains: the transmembrane rotor domain, a globular catalytic domain and a rod subunit (comprising of two anti-parallel alpha helices) that connects the transmembrane and catalytic regions (Stock *et al.*, 1999). Proton translocation through the rotor domain causes the catalytic region to turn, which drives the formation of ATP.

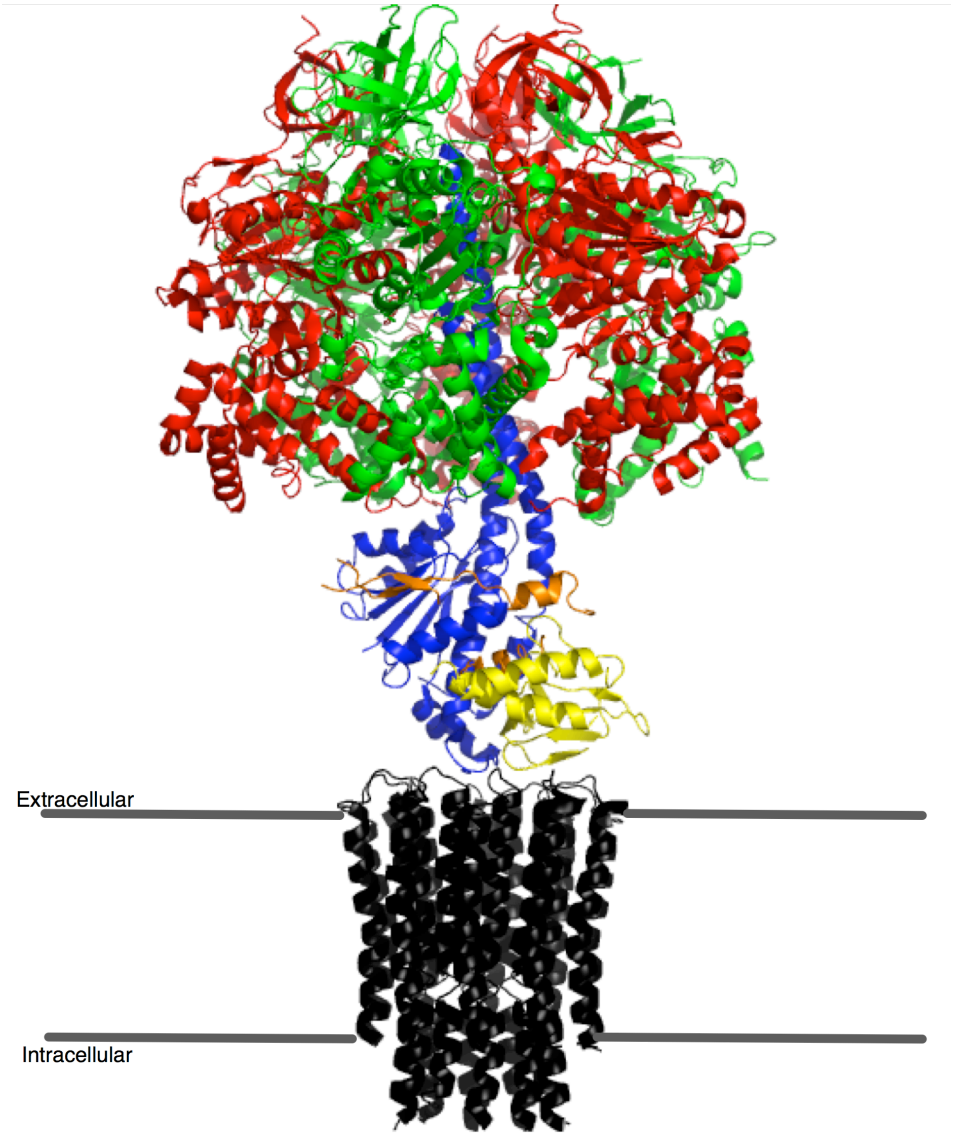


Figure 1.6 The 3-dimensional structure of F1-ATP synthase

ATP synthase (accession code 2XOK) is composed of a number of subunits forming the proton translocating transmembrane domain (coloured black) and the catalytic portion which is comprised of α (red), β (green), γ (blue) and ε (yellow/orange) subunits (Stock *et al.*, 1999).

β-barrel containing proteins are found in the outer membranes of gram-negative bacteria, mitochondria and chloroplasts where they have functions as diverse as ion transport, passive nutrient uptake, membrane anchors, membrane-bound enzymes as well as defense against attack proteins (Buchanan, 1999; Schulz, 2000). Membrane protein β-barrels contain between 8 and 22 strands, and the overall pore size depends on the number of β-strands that make up the barrel (Buchanan, 1999; Schulz, 2000). The first membrane protein with a β-barrel motif to have its structure solved was porin from *Rb. capsulatus* (Figure 1.7) (Weiss and Schulz, 1992). Porin is a 16-stranded β-barrel that associates to form homo-trimers, and carries out passive diffusion of small (*ca.* 600 Da) molecules. OmpA and OmpX both contain eight stranded barrels (Figure 1.7) (Pautsch and Schulz, 1998; Vogt and Schulz, 1999). These proteins are not thought to have any channel activity but rather function as transmembrane anchors. The structures of other bacterial β-barrel containing MPs are discussed in (Buchanan, 1999; Schulz, 2000). Recently, structures of the voltage-dependent anion channel (VDAC) (the major pathway for entry and exit of metabolites across the outer mitochondrial membrane) from human (Bayrhuber *et al.*, 2008) and mouse (Ujwal *et al.*, 2008) have been solved to atomic resolution. The VDACs contain a 19-stranded β-barrel.

Although membrane proteins that contain β-barrel structures are relatively rare, there are a large variety of protein families that have transmembrane α-helical domains. These include transporters, water and ion channels, enzymes as well as G-protein coupled receptors.

The major facilitator superfamily (MFS) is the largest family of active transporters, comprising one quarter of all transport proteins (Saier *et al.*, 1999). These proteins overcome the thermodynamically unfavourable process of transporting a substrate across the membrane against a concentration gradient by either utilising the free energy stored in an electrochemical ion (e.g. Na^+, H^+) gradient or coupling the primary transport event to the movement of a second substrate down a concentration gradient. There are three main types of MFS protein : i) uniporters - that translocate a single substrate across the membrane,

ii) symporters – that transport two or more substrates across the membrane in the same direction (e.g. the lactose/H^+ symporter LacY), and iii) antiporters – that transfer two or more substrates across the membrane in opposing directions (e.g. the inorganic phosphate/glycerol-3-phosphate antiporter GlpT)) (Law *et al.*, 2008).

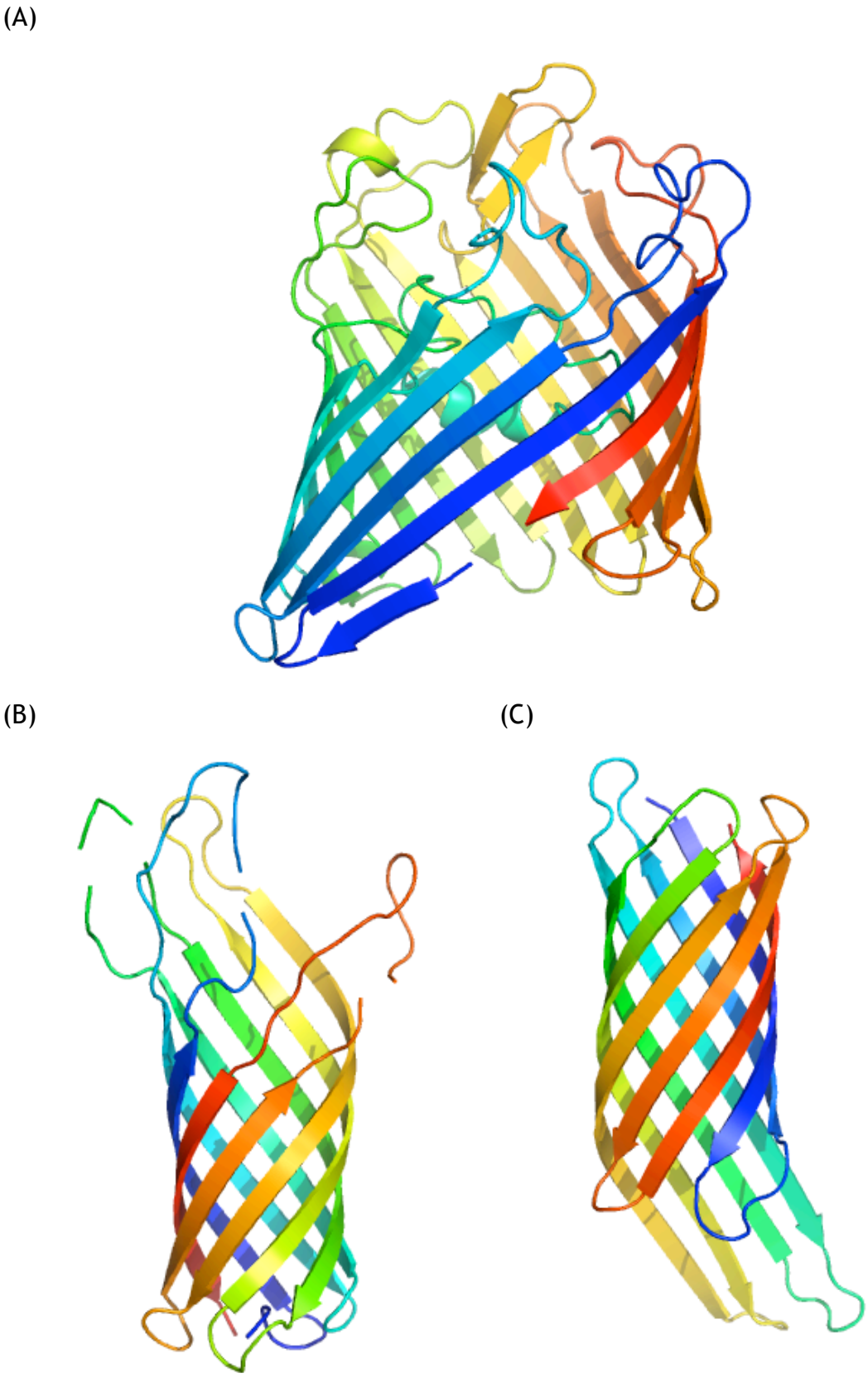


Figure 1.7 Examples of membrane proteins with a β-barrel type structure

(A) porin (2OMF), (B) OmpA (1BXW), and (C) OmpX (1QJ8). (Weiss and Schulz, 1992; Pautsch and Schulz, 1998; Vogt and Schulz, 1999)

The first member of the MFS family to have its structure solved was LacY from *E. coli* (to a resolution of 3.5 Å) (Figure 1.8(A)) (Abramson *et al.*, 2003 a,b). Since then, numerous bacterial MFS transporters have had their structures solved including GlpT (Huang *et al.*, 2003), the multidrug transporters EmrD (Yin *et al.*, 2006) and EmrE (Chen *et al.*, 2007), the Na^+/H^+ antiporter NhaA (Figure 1.8 (B)) (Hunte *et al.*, 2005), the sodium-dependent glutamate transporter GltPh (Yernool *et al.*, 2004), the leucine transporter LeuT (Yamashita *et al.*, 2005) and the zinc transporter YiiP (Lu and Fu, 2007). As yet, no mammalian MFS transporter has had its structure solved.

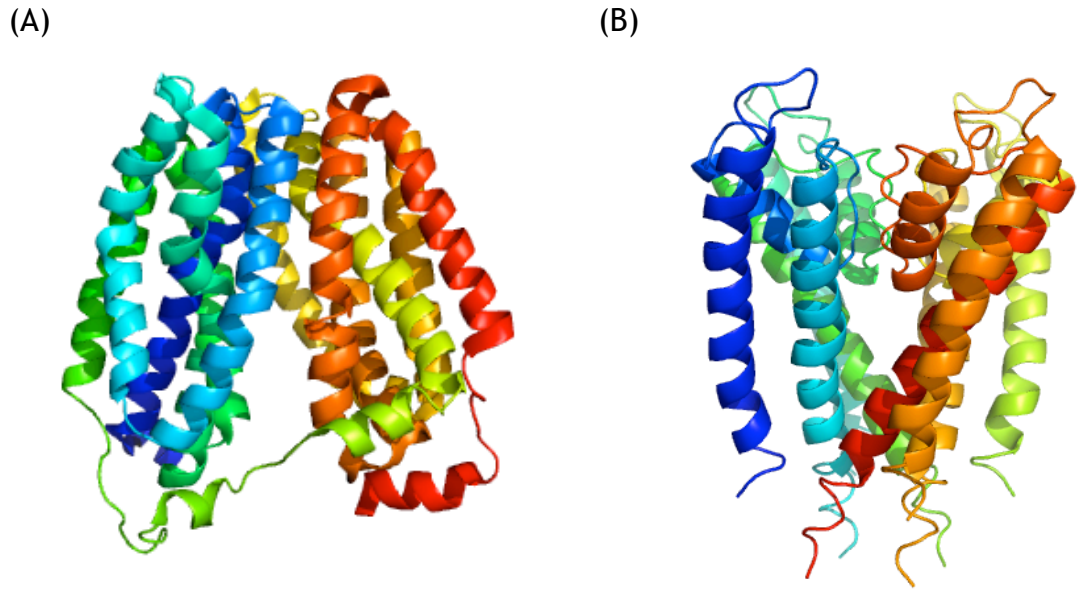


Figure 1.8 Structures of bacterial members of the MFS family
(A) the lactose/H^+ symporter, LacY (1PV7) has a large internal hydrophilic cavity which opens to the cytoplasm (Abramson *et al.*, 2003 a,b) and (B) the Na^+/H^+ antiporter, NhaA (1BL8) has a water pore on the outer side (Hunte *et al.*, 2005).

ATP-binding cassette (ABC) transporters use the energy released on ATP hydrolysis to transport specific substrates across the membrane. All ABC transporters have a four-domain structure in common consisting of two nucleotide binding domains (NBD), that hydrolyse ATP and drive the transport reaction, and two transmembrane domains (TMD) that form the translocation pathway (Higgins, 1992; Rees *et al.*, 2009). The first ABC transporter to have its

structure determined was the vitamin B12 importer from *E. coli* (BtuCD) to a resolution of 3.2 Å (Figure 1.9) (Locher *et al.*, 2002). Since then, a further five bacterial (including the multidrug resistance protein Sav1866) (Dawson and Locher, 2006; Hollenstein *et al.*, 2007 ; Oldham *et al.*, 2007; Pinkett *et al.*, 2007; Ward *et al.*, 2007) and one mammalian (P-glycoprotein) (Aller *et al.*, 2009) ABC transporters have had their structures solved.

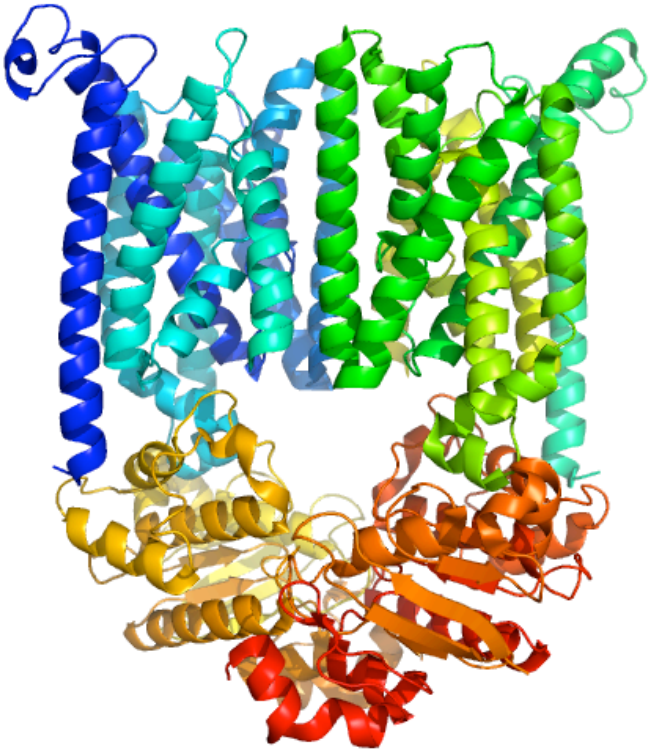


Figure 1.9 The bacterial ABC transporter that transfers vitamin B12 across the inner cytoplasmic membrane

The α helices (coloured blue through green) create the transmembrane region of BtuC domains while the BtuD regions (coloured yellow through red) are involved in ATP-binding. In combination these four domains complete the BtuCD structure (accession code 1L7V) (Locher *et al.*, 2002).

Aquaporins are water-specific membrane channels that allow water to move freely and bi-directionally across the membrane (Meinild *et al.*, 1998). Structures for bacterial (GlpF, AqzP) (Fu *et al.*, 2000; Savage *et al.*, 2003), plant (Tornroth-Horsefield *et al.*, 2006) and mammalian (AQP1, 4 and 5) (Murata *et al.*, 2000; Horsefield *et al.*, 2008; Ho *et al.*, 2009) aquaporins are known (Figure 1.10).

Ion channels allow the selective passage of particular ions (e.g Na^+, K^+, Cl^-) across membranes in response to signals such as ligand-binding (e.g. glutamate, GABA) or changes in transmembrane voltage. All ion channels have a similar architecture with multiple sub-units arranged around a central pore (where the ion conduction pathway is located). The first ion channel to have its structure determined was the K^+ channel from the bacteria *S. lividans* (KcsA) (Doyle *et al.*, 1998). Since this ground-breaking achievement, a plethora of other bacterial ion channels have had their structures solved including the Ca^{2+}-gated K^+ channel from *M. thermoautotrophicum* (MthK) (Jiang *et al.*, 2002a), the K_V from *A. pernix* (K_VAP) (Jiang *et al.*, 2002b), the inward-rectifying K^+ channel from *Burkholeria pseudomallei* (KirBac1.1) (Kuo *et al.*, 2003), the non-selective cation channel (Shi *et al.*, 2006), the large and small mechanosensitive channels from *M. tuberculosis* (MscL) (Chang *et al.*, 1998) and *E. coli* (MscS) (Bass *et al.*, 2002), respectively as well as the ClC Cl^- channels from *S. enterica* and *E. coli* (Dutzler *et al.*, 2002; Dutzler *et al.*, 2003). Several ion channels of mammalian origin have also been determined including rat K_V1.2 (a member of the K^+ Shaker channel family) (Long *et al.*, 2005), chicken Kir2.2 (Tao *et al.*, 2009), the human high-conductance voltage and Ca^{2+}-activated (BK) channel (Yuan *et al.*, 2010), the chicken acid sensing ion channel (Jasti *et al.*, 2007), the zebrafish ATP-gated P2X(4) ion channel (Kawate *et al.*, 2009) and the rat AMPA-sensitive ionotropic glutamate receptor (GluA2) (Sobolevsky *et al.*, 2009). The abundance of ion channel structures in the PDB suggests that they are considerably more tractable to work with than most other membrane protein superfamilies.

The major mammalian G-protein coupled receptor superfamily has proved remarkably resistant to crystallisation. However, several GPCR structures have recently been solved and deposited in the PDB. These include bovine rhodopsin (Palczewski *et al.*, 2000), the human β2-adrenergic receptor (Cherezov *et al.*, 2007; Rasmussen *et al.*, 2007), turkey β1 adrenergic receptor (Warne *et al.*, 2008) as well as the human adenosine 2a receptor (Jaakola *et al.*, 2008). These are discussed in more detail in Section 2.

In short, there are now a substantial (several hundred), and increasing, number of bacterial membrane proteins in the PDB with rather few (several tens) of mammalian origin. This reflects the extra difficulties in obtaining high quality, diffracting crystals of membrane proteins from mammalian cells as compared to those from bacteria.

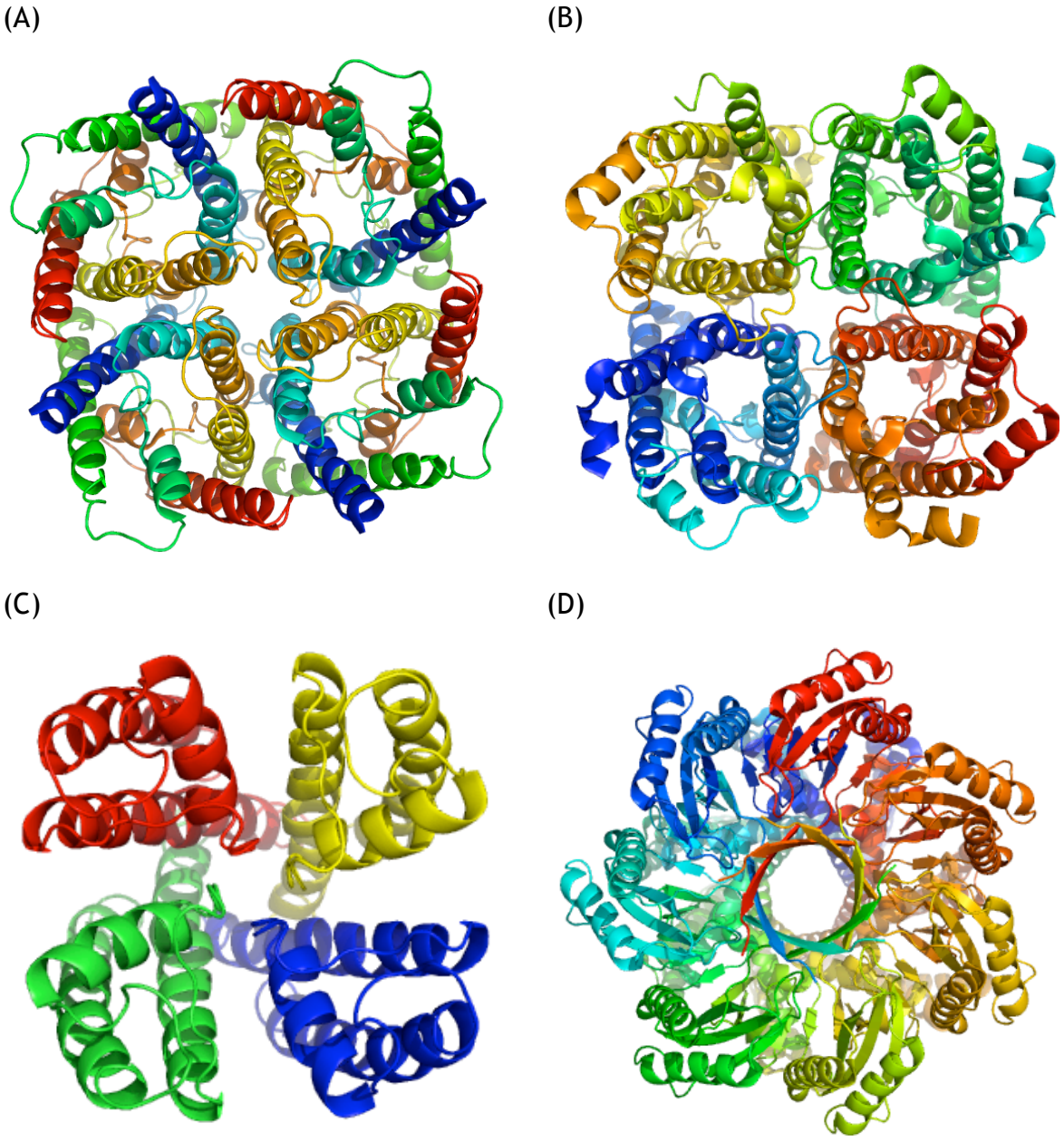


Figure 1.10 Structures of selected water and ion channels showing the ways that the aqueous pore can be organised

The human water channels (A) aquaporin 1, AQP1 (accession code 1FQY) and (B) aquaporin 5, AQP5 (3D9S) (Murata *et al.*, 2000; Ho *et al.*, 2009). (C) The potassium channel from *Streptomyces*

lividans, KcsA (1BL8) (Doyle *et al.*, 1998). (D) The *E. coli* mechanosensitive channel of small conductance, MscS (2OAU) (Bass *et al.*, 2002).

1.5 The problems associated with solving the structures of mammalian membrane proteins

Before the structure of any mammalian membrane protein can be solved, several major obstacles need to be overcome. Firstly, milligram quantities of the protein of interest are required for either crystallographic or nuclear magnetic resonance experiments (Sarramegna *et al.*, 2003). In certain instances, mammalian membrane proteins exist in native tissues at sufficiently high levels to allow their purification to homogeneity e.g. rhodopsin. However, this is not the case for the vast majority of mammalian membrane proteins which need to be overexpressed in an appropriate host cell system before they can be purified (Seddon *et al.*, 2004). Secondly, the mammalian MP needs to be isolated from its native membrane environment by detergent solubilisation. The choice of detergent is critical as mammalian MPs are very unstable and are readily denatured by detergent (Lacapere *et al.*, 2007). Thirdly, the protein has to be purified to homogeneity prior to structural analysis in a monodisperse state keeping its activity intact. However, given their poor stability, many mammalian membrane proteins tend to aggregate during the purification process (Gutmann *et al.*, 2007). Finally, mammalian membrane proteins (like all membrane proteins) have relatively small exposed polar surface areas making them considerably harder to crystallise than soluble proteins.

1.6 Engineering membrane proteins for crystallisation

Given that membrane proteins are so difficult to crystallise, mammalian ones in particular, it is imperative that the protein to be used in the crystallisation trials is largely rigid, thermostable, homogenous and has sufficient polar surface exposed to allow the formation of crystal contacts. As most wild-type mammalian membrane proteins do not fulfil all of these criteria, it is often

necessary to engineer the target protein to increase the probability of it crystallising. There are a variety of techniques that can be employed to increase the stability and rigidity of target proteins as well as their homogeneity. These include using deglycosylated mutants (Fraser, 2006), removing disordered regions from the protein (Cherezov *et al.*, 2007), introducing single point mutations to increase its stability (Standfuss *et al.*, 2007) as well as adding binding partners to "lock" the target protein into a particular conformation (often increasing the available polar surface area for lattice contacts) (Palczewski *et al.*, 2000; Rasmussen *et al.*, 2007). With all of these strategies, however, it is possible to engineer mutants that have non-physiological properties. Therefore, care must be taken when trying to create modified proteins better suited to crystallisation trials so that they retain their biological activity also.

1.6.1 Removing regions predicted to be disordered

It is not uncommon for proteins involved in signalling (eg p53, GPCRs) to have large regions that are intrinsically disordered. Well-ordered crystals cannot be obtained for proteins that have extensive flexible regions (Warne *et al.*, 2009). Utilising disorder predictor programs such as RONN (University of Oxford) likely site(s) of disorder within the target protein can be identified and removed by mutagenesis (Yang *et al.*, 2005). For example, the C-terminal tails of most GPCRs are highly disordered and are routinely removed by researchers trying to make a crystallisable construct (Warne *et al.*, 2003; Fraser, 2006; Rasmussen *et al.*, 2007; Rosenbaum *et al.*, 2007; Magnani *et al.*, 2008). Not unexpectedly, removal of disordered regions from a target protein can improve its expression by reducing the number of protease sensitive sites within it (Warne *et al.*, 2003).

1.6.2 Increasing the thermostability and therefore the crystallisability of membrane proteins

Most mammalian membrane proteins that function at the cell surface are glycosylated. Although not necessarily an impediment to crystallisation, it is often desirable to remove the glycosylation site(s) from the target protein by mutagenesis to ensure that the final purified sample has negligible heterogeneity. Although many glycosylation-deficient mutants retain wild-type properties (Fraser, 2006), this is not always the case (Tate, 2001).

The contribution that individual amino acids make to the thermostability of a membrane protein can be readily determined by combining alanine-scanning mutagenesis with an appropriate activity assay following partial heat-denaturation of the Ala mutants (for example DAG kinase (Gorzelle *et al.*, 1999), the β1-adrenergic (Serrano-Vega *et al.*, 2008) and adenosine 2a receptors (Magnani *et al.*, 2008)). Once the amino acids that are critical for stability have been identified, it is possible to combine several mutations to make a temperature- and detergent-resistant mutant suitable for crystallisation (Standfuss *et al.*, 2007). Proteins can also be selectively engineered to have increased stability, through the introduction of cysteine residues that form disulphide bonds fixing the protein's conformation (Blois and Bowie, 2009).

An alternative way that the thermostability of a membrane protein can be increased is through the addition of a specific ligand (eg inhibitor, agonist, antagonist). Indeed, the recent structures of the β1-, β2-adrenergic and adenosine 2a receptors were all solved in the presence of a high-affinity (pico- to nano-molar) inverse agonist (Cherezov *et al.*, 2007; Rasmussen *et al.*, 2007; Rosenbaum *et al.*, 2007; Hanson *et al.*, 2008; Jaakola *et al.*, 2008). One of the reasons that rhodopsin has proved considerably more tractable to crystallisation than ligand-activated GPCRs is that in the ground state its ligand, 11-cis-retinol, is covalently bound to it making rhodopsin substantially more thermostable compared to other GPCRs (Standfuss *et al.*, 2007).

To increase the probability of getting a membrane protein to crystallise strategies for increasing their polar surface area have been developed. For example, cytochrome oxidase (Hunte and Michel, 2002) and the β2 adrenoceptor (Rasmussen *et al.*, 2007) have been successfully crystallised in complex with an antibody binding fragment (F_{Ab}) (Figure 1.11 (coloured red)). The rationale behind this approach is that when the F_{Ab} binds to the target protein it creates a complex with a large polar surface area that can readily form lattice contacts (Hunte and Michel, 2002). Although this approach has been successfully used to obtain crystal structures of selected membrane proteins it is not widely practiced as generating F_{Ab} is both expensive and time-consuming.

A related approach has recently been developed by Stevens and co-workers (The Scripps Research Institute, California) where, to improve the crystallisability of GPCRs, they largely replaced the receptor's third intracellular loop (IL3) with T4-lysozyme. This GPCR-fusion protein has a large hydrophilic domain useful for forming crystal contacts. This approach has been used to solve the structures of the β2-adrenoceptor (Figure 1.11 (coloured blue)) (Cherezov *et al.*, 2007) and the adenosine 2a receptor (Jaakola *et al.*, 2008) to atomic resolution.

Figure 1.11 Overlay of the two structures of the ligand-activated GPCR, β2-adrenergic receptor, that have been solved using strategies to increase their polar surface area
Displayed in red is a complex of the human β2-adrenergic receptor in complex with a F_{ab} that binds to IL3 (accession code 2R4R); while blue shows a chimera consisting of the human β2-adrenergic receptor with T4-lysozyme instead of IL3 (accession code 2RH1).

1.7 Heterologous production of mammalian membrane proteins

Mammalian membrane protein expression can be performed in a variety of host cells [reviewed in (Sarramegna *et al.*, 2003)] including bacteria (Weiss and Grisshammer, 2002), yeast (Reilander and Weiss, 1998; Jahic *et al.*, 2006), insect (Cherezov *et al.*, 2007; Rasmussen *et al.*, 2007) and mammalian cells (Mize *et al.*, 2008). For any given protein, the choice of system will depend on a number of factors including the availability of specialised expression vectors and

strains, budget as well as any post-translational modifications required for maintaining the protein's functionality (Sarramegna *et al.*, 2003).

Although bacterial expression systems are routinely used to make soluble mammalian proteins their use with membrane proteins from higher-eukaryotes is rather more restricted. This is not surprising given that bacteria cannot perform many of the post-translational modifications that occur in mammalian cells (e.g. glycosylation, palmitylation) that are often required for full activity of the membrane protein. Furthermore, as the lipid composition of bacterial membranes is very different from that in higher eukaryotic cells it is not possible to functionally produce those mammalian membrane proteins that have specific lipid requirements in bacteria. This said, however, *E. coli* has been used to make both integral (including GPCRS) (Tucker and Grisshammer, 1996; Weiss and Grisshammer, 2002) and peripheral mammalian membrane proteins (e.g. 11β-hydroxylase) (Zollner *et al.*, 2008) in milligram amounts. The main advantages of making mammalian membrane proteins in bacteria are the short (*ca.* 20 min) doubling-time of the host cell and the use of cheap media (Sarramegna *et al.*, 2003).

Yeasts (in particular *Saccharomyces cerevisiae* and *Pichia pastoris*) are commonly used to make mammalian membrane proteins as they combine a short generation time (*ca.* 2 h) on cheap growth media with their ability to perform many of the post-translational modifications that occur in higher eukaryotes (Reilander and Weiss, 1998). The methylotrophic yeast *P. pastoris* has been shown to be particularly useful for making a wide-range of human membrane proteins including GPCRs, ion channels, transporters and pumps (Weiss *et al.*, 1995; Weiss *et al.*, 1998; Grunewald *et al.*, 2004; Andre *et al.*, 2006; Fraser, 2006; Shukla *et al.*, 2007; Horsefield *et al.*, 2008 ; Aller *et al.*, 2009 ; Ho *et al.*, 2009; Tao *et al.*, 2009). The main advantages of using *P. pastoris* are that it has a very strong, methanol-inducible promoter (Alcohol Oxidase 1 (AOX1)) which drives the production of high levels of recombinant protein (Hartner and Glieder, 2006; Jahic *et al.*, 2006) as well as its ability to grow to high cell densities in fermenter culture (>10^9 cells/ml) (Fantoni *et al.*, 2007). One drawback of using yeast, however, is that its membranes contain ergosterol (the principal fungal

sterol) instead of the mammalian one, cholesterol. This lack of cholesterol is thought to have prevented the functional expression of certain human membrane proteins in yeast including the serotonin transporter (Tate, 2001) and the μ-opioid receptor (Lagane *et al.*, 2000).

Insect cells that have been infected with recombinant baculovirus are useful for expressing a number of mammalian membrane proteins. Indeed, all of the ligand-activated GPCRs that have been crystallised to date have been made in insect cells (Cherezov *et al.*, 2007; Rasmussen *et al.*, 2007; Jaakola *et al.*, 2008; Warne *et al.*, 2008). Other advantages of this system are that they perform a simple form of glycosylation as well as growing to cell densities of 10^7-10^8 cells/ml in spinner flasks (Sarramegna *et al.*, 2003). Although the media for growing insect cells costs substantially more than that for either bacteria or yeast, it is also significantly less expensive than that for mammalian cells.

In many respects, mammalian cells are the ideal host system for making mammalian membrane proteins as the lipid composition of the membrane is exactly what the recombinant proteins require as well as the cells having all of the machinery to perform post-translational modifications to the target protein. However, mammalian cells have several drawbacks including a long generation time (>24h), they grow to low cell density in bioreactors (10^6-10^7 cells/ml) and the media can be prohibitively expensive (Sarramegna *et al.*, 2003; Tate *et al.*, 2003).

As there is a wide-range of expression systems available, each with their own advantages and disadvantages, the particular host cell chosen for heterologous expression will largely be determined by the particular mammalian membrane protein being over-expressed. Whichever system is chosen, it must deliver high yields of functional protein (Tate *et al.*, 2003).

1.8 Detergent solubilisation

Prior to their purification, membrane proteins are extracted from the lipid bilayer by adding an amphipathic substance usually detergent. Although detergents come in many shapes and forms they all have a hydrophilic head-group connected to a hydrophobic domain. The nature of the headgroup defines the class of detergent: anionic detergents have a negatively charged headgroup whereas cationic detergents are positively charged. Detergents with headgroups that have both positive and negative charges are termed zwitterionic whereas those detergents without a charged headgroup are classified as non-ionic (Seddon *et al.*, 2004).

The choice of detergent used for solubilisation is critical and the chosen one must retain the protein's stability and functionality in solution. Although strong ionic detergents (e.g. LDAO) are excellent at solubilising membranes they can also remove essential lipids from a protein's surface rendering it useless for structural studies (Johansson *et al.*, 2009). Therefore, mild non-ionic detergents (e.g. DDM) are commonly used for solubilisation as they often yield stable, active protein in aqueous solution (Misquitta and Caffrey, 2003).

The concentration at which discrete detergent monomers spontaneously forms micelles is known as the critical micelle concentration (CMC). Membranes can be readily solubilised by adding detergents at a concentration significantly (10 – 20 times) greater than their CMC. The detergent molecules replace the lipids surrounding the MP resulting in the formation of discrete MP-containing micelles (Figure 1.12). Following solubilisation, the target protein can be purified using conventional chromatographic techniques for crystallisation trials (Weiss and Grisshammer, 2002; Krueger-Koplin *et al.*, 2004). To prevent membrane proteins precipitating out of solution, however, all purification buffers must contain detergent with a concentration greater than the CMC.

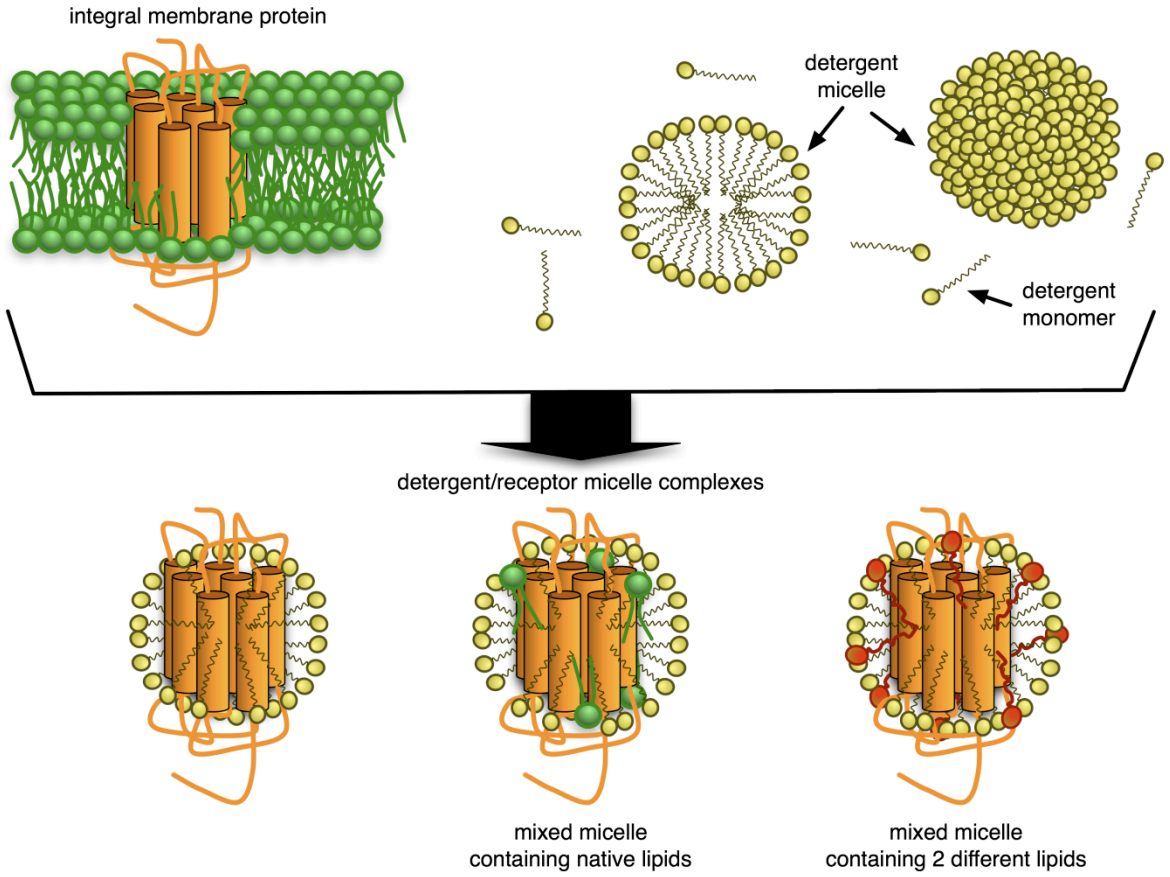


Figure 1.12 Schematic diagram showing solubilisation of an integral membrane protein with detergent

Following the addition of detergent at levels 10-20 fold greater than its critical micelle concentration (CMC), the lipids surrounding the membrane protein are stripped away and replaced by detergent. This results in the formation of discrete protein-detergent micelle complexes. If the solubilisation detergent is harsh all of the native lipids may be stripped away. When milder detergents are used, those lipids that bind tightly to the surface of the protein are often retained. If a second detergent or lipid is included within the solubilisation mixture mixed micelles will result.

When working with membrane proteins the particular detergent used in the crystallisation trials is often a critical determinant of whether or not any crystals are obtained. Large detergent micelles (e.g. DDM) can mask the hydrophilic (non-transmembrane) regions of membrane proteins, preventing the formation of protein-protein interactions necessary for crystal formation (Kobilka, 2007). In contrast, as short chain detergents have small micelles they are better suited to membrane protein crystallisation. However, this factor has to be weighed against the effect of the detergents' tail length on its ability to stabilise the target membrane protein (Prive, 2007). In general, detergents with long (C=12) alkyl chains tend to be stabilising whereas those with short (C=8) tails are not.

Therefore, for the vast majority of mammalian membrane proteins the detergents best suited for crystallisation can not be used because they are too destabilising (Gutmann *et al.*, 2007; Prive, 2007). One way in which this problem can be overcome is by engineering thermostable mutants of the target protein that retain their structure in short chain detergents. For example, considerable effort was invested by Tate and co-workers to make a thermostable mutant of the turkey β1-adrenergic receptor which was stable in a wider range of detergents than the wild type protein (Serrano-Vega *et al.*, 2008; Warne *et al.*, 2008) (Figure 1.13). Using their most thermostable mutant, it was possible to grow diffracting crystals of the β1AR in the detergent octyl-thioglucoside from which a structure to atomic resolution has been obtained (Warne *et al.*, 2008).

While modifications of protein sequence can aid protein stability in detergent solution, it is becoming increasingly apparent that the activities of many mammalian membrane proteins are influenced by specific lipids, in particular by cholesterol (Epand, 2006). By including specific lipids in the solubilisation mixture as well as the purification buffers it is possible to increase the stability of the target protein in solution (Lacapere *et al.*, 2007). For example, the cholesterol derivative cholesterol hemisuccinate (CHS) is essential to obtain radioligand binding by the human adenosine 2a receptor in detergent solution (Weiss and Grisshammer, 2002; Fraser, 2006).

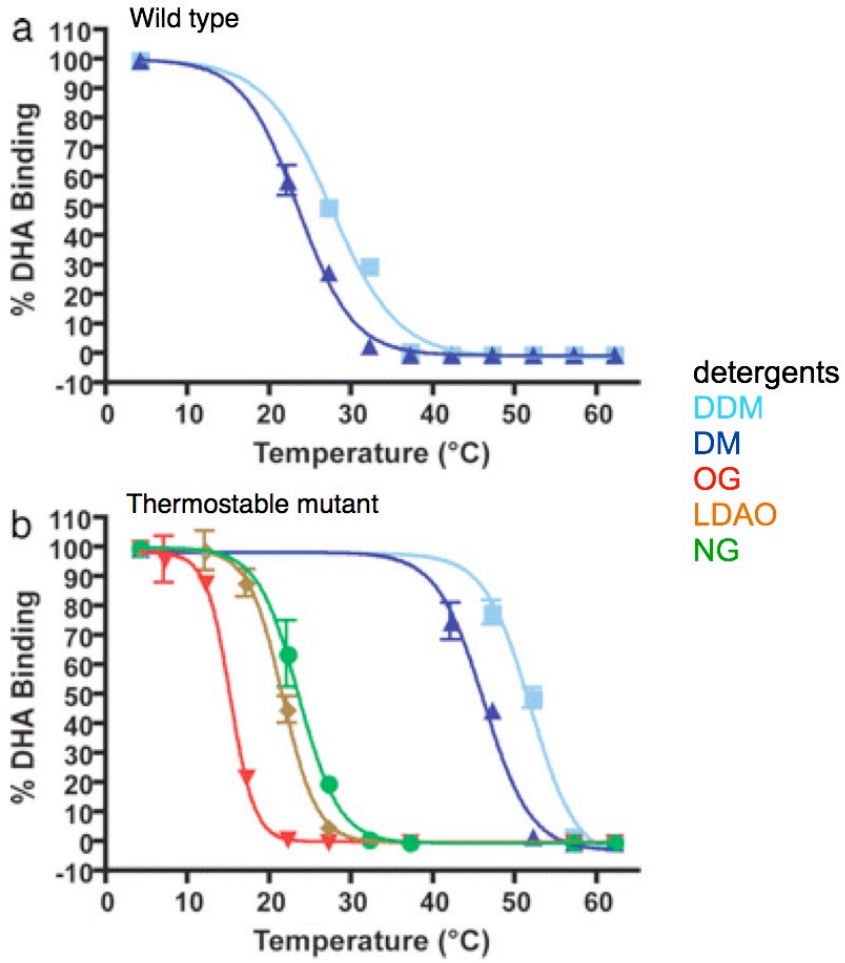


Figure 1.13 Stability of (a) wild-type and (b) a thermostable mutant of the turkey β1 adrenoceptor in various detergents

In contrast to that observed for the wild-type receptor, the thermostable mutant was able to bind radioligand even when it had been solubilised using the harsh detergents octyl-glucoside (OG), nonyl-glucoside (NG) and LDAO. Diffracting crystals of the thermostable mutant were obtained using the short-chain detergent octyl-thioglucoside. (Figure taken from (Serrano-Vega *et al.*, 2008)).

1.9 Membrane protein purification

Following solubilisation, the protein of interest must be purified to homogeneity prior to structural experiments. Two of the most common approaches to purifying soluble proteins are ion-exchange and gel-filtration chromatography. In ion-exchange chromatography, proteins are firstly bound to a charged resin and then selectively eluted by gradually increasing the ionic strength of the buffer (Seddon *et al.*, 2004). With gel-filtration, proteins are separated according to

their molecular weight by passage through a size exclusion chromatography column where large proteins pass through the column quickly and smaller proteins travel more slowly through the bead resin. These purification techniques can also be used with membrane proteins but the presence of detergent limits their effectiveness.

More recently, recombinant membrane proteins have been made as fusions with affinity tags located at either their N- or C- termini. All affinity tags bind to specific chromatographic resin, allowing the protein to be purified rapidly and efficiently (Lundstrom, 2005). For example, a deca-histidine tag (His10) binds to immobilized metal (usually Ni^{2+} but also Co^{2+} or Zn^{2+}) affinity chromatography (IMAC) resin. Non-specific proteins can be removed from the column by washing it with a buffer of high ionic strength containing a small (30 mM) concentration of imidazole. The recombinant protein can be eluted using a high (300 mM) concentration of imidazole (which competes with the His tag for binding to the metal ion). Affinity chromatography is the best approach for purifying membrane proteins as long as the detergent micelle does not interfere with binding of the tag to the resin.

1.10 Membrane protein crystallisation

1.10.1 Aggregation

For crystallisation to occur, the protein must be monodisperse in solution (Krueger-Koplin *et al.*, 2004). However, due to their hydrophobic nature membrane proteins have a propensity to aggregate in detergent solution rendering efforts to crystallise them meaningless (Sanders and Oxenoid, 2000; Gutmann *et al.*, 2007). Factors that affect membrane protein aggregation include the protein to detergent ratio, glycerol and salt content of the buffer as well as the type of detergent used. However, caution must be exercised when changing any of these parameters as they have to be compatible with crystallisation trials.

1.10.2 Crystal growth by vapour-diffusion

Protein crystallisation typically requires a highly concentrated (>10 mg/ml) pure protein sample which on the addition of a precipitant (and possibly small amphiphile molecules) starts to form crystals (Garavito *et al.*, 1996). The precipitant drives the formation of nucleation sites which, as the protein concentration within the drop increases due to water loss, forms the site of crystal growth (Caffrey, 2003). The formation of well-ordered crystals requires slow, regulated crystal growth which is highly influenced by the protein's concentration and solubility. As the range of possible precipitants is large, the process of identifying which will be best for crystal growth can be particularly time consuming. However, the development and use of sparse matrix screens for initial crystallisation trials has greatly helped to overcome this rate-limiting step. Crystal hits are then optimised through a series of small variations to the original crystallisation condition with a view to facilitating slow, ordered crystal growth resulting in good quality diffracting crystals (Caffrey, 2003; Kobilka and Schertler, 2008). Depending on the orientation of the individual proteins and the position of their lattice contacts, crystals of varied shape and symmetry can be formed (Nollert, 2005). However, in the case of membrane proteins the process of obtaining crystals is complicated by the presence of detergent(s) (Arora and Tamm, 2001), which often mask the extra-membrane/hydrophilic regions of the protein, preventing the formation of crystal contacts (Mohanty *et al.*, 2003).

1.10.3 Lipidic Cubic Phase

As well as trying to grow crystals of membrane proteins in detergent solution, the protein of interest can also be crystallised in a lipid environment called a lipidic cubic phase (Nollert, 2005). A lipidic cubic phase is a structured 3-dimensional matrix of lipid bilayers linked through a network of interconnected aqueous pores (Landau and Rosenbusch, 1996; Cherezov *et al.*, 2006). This matrix has been used successfully in conjunction with standard crystallization techniques to solve the structure of a range of proteins including a purple

bacterial photosynthetic RC (Katona *et al.*, 2003) as well as the β2-adrenergic (Cherezov *et al.*, 2007) and adenosine 2a receptors (Jaakola *et al.*, 2008). The principle advantage of the cubic phase is that it allows the formation of protein crystals in a native-like membrane environment (Landau and Rosenbusch, 1996) that tend to be of a superior quality than those grown from detergent solution (Wadsten *et al.*, 2006). This is due to the increase in stability of the membrane protein when it is reconstituted into a bilayer environment (Johansson *et al.*, 2009). However, cubic phase can only be used for crystallising membrane proteins with small extra-membrane domains as the aqueous channels are not big enough to accommodate proteins with large hydrophilic regions (Cherezov *et al.*, 2006; Wohri *et al.*, 2008).

1.10.4 Lipidic Sponge Phase

Recently, it has been shown that the cubic phase can be swollen by treatment with solvents such as PEGs (Bender *et al.*, 2008). This 'swollen' lipidic phase is known as the sponge phase and has aqueous channels *ca.* 3x larger than the cubic phase. This means that the sponge phase has the potential for crystallising membrane proteins with large extra-membrane domains (Cherezov *et al.*, 2006; Wadsten *et al.*, 2006). Several proteins have already had their structures solved in the sponge phase (Wohri *et al.*, 2009), and it is an emerging technique which should prove useful for obtaining diffracting crystals of (mammalian) membrane proteins.

1.11 Project context

Currently, the field of mammalian membrane protein structural biology is in its infancy. Existing technologies and experiences have shown that it is possible to obtain the structures of mammalian membrane proteins if sufficient work and thought has been invested. However, there is still an urgent need to develop new methodologies and approaches to improve all aspects (expression,

purification and crystallisation) of mammalian membrane protein structural biology.

1.12 Aim

The main aim of my Ph.D. was to develop new technologies for the overproduction, purification and crystallisation of human membrane proteins, and then test them with a representative member from each of the GPCR (adenosine 2a receptor) and membrane enzyme (sterol isomerase) superfamilies.

1.13 Objectives

i) to create a series of *P. pastoris* expression vectors that give high production levels of recombinant protein which can be readily purified to homogeneity by affinity chromatography

ii) to develop an optimised methodology that gives maximal production of recombinant membrane proteins in fermenter cultures of *P. pastoris*

iii) to develop a protocol for monitoring the aggregation status of recombinant membrane proteins immediately following detergent solubilisation and over the course of the purification process

iv) to develop a method for determining the stability of purified membrane proteins in a range of buffers and salts

v) to create a lipidic-sponge phase screen and compare its usefulness in obtaining mammalian membrane protein crystals compared to conventional crystallisation experiments in detergent solution

2 G-protein coupled receptors and cholesterol biosynthesis

2.1 Human membrane proteins

In Chapter 1, the problems associated with solving the structures of mammalian (including human) membrane proteins were detailed at length. However, as there are many different families of human membrane protein it is pertinent to ask whether or not they are all equally difficult to work with or are some proteins more amenable to study than others. All membrane proteins can be classified into two main groups — those that have conformational flexibility (e.g. receptors, transporters) and those that are essentially rigid (e.g. enzymes, ion channels). This chapter focuses on the human adenosine 2a receptor (a GPCR) and sterol isomerase (an enzyme involved in the late stages of cholesterol biosynthesis) as representative members of the conformationally-dynamic and rigid-body membrane protein superfamilies, respectively.

2.2 Membrane proteins with conformational flexibility

Receptors, which function in information transfer, and transporters, that translocate specific substances across lipid bilayers, have considerable conformational flexibility consistent with their function. However, the inherent flexibility that allows these proteins to exist in multiple conformational states is not readily compatible with growing protein crystals. It is not surprising then that there are so few structures of mammalian receptors and transporters compared to, for example, ion channels (which have rigid structures). Furthermore, the ligand-activated GPCRs that have had their structures solved had all been locked into particular conformations using either ligands (Palczewski *et al.*, 2000; Jaakola *et al.*, 2008) or by site-directed mutagenesis (Warne *et al.*, 2008). The available structures of mammalian membrane proteins

in the PDB suggests that those proteins with high mobility are harder to work with than those that are rigid. I chose to work on the human adenosine 2a receptor, a GPCR, as a model human membrane protein with high conformational flexibility.

2.3 G-protein coupled receptors

G-protein coupled receptors (GPCRs) are a large protein super-family that function in the plasma membrane of mammalian cells (Kobilka, 2007). They are highly conserved throughout evolution with approximately 1% of the human genome encoding them (Sarramegna *et al.*, 2003). GPCRs can be classified into 5 distinct groups (Type 1 to 5) by sequence homology (Deupi and Kobilka, 2007) of which type-1 GPCRs (also known as rhodopsin-like receptors) are the largest and best studied group (Cherezov *et al.*, 2007; Deupi and Kobilka, 2007; Fredholm *et al.*, 2007a). Although GPCRs show poor homology between families (Klabunde and Hessler, 2002), they all have a common 7-transmembrane alpha-helical structure with an extracellular amino terminus and an intracellular carboxyl terminus (Figure 2.1) The hydrophobic transmembrane domains are connected by 3 intra- (IL1-3) and 3 extra-cellular (EL1-3) loops. The size of the loops and terminal tails can vary dramatically between receptors (Klabunde and Hessler, 2002). GPCRs perceive specific chemical messages in the extracellular environment and then convert this information to a signal within the cell (Lefkowitz, 1998; Bockaert and Pin, 1999; Kobilka, 2007). GPCRs are so-called because they activate specific heterotrimeric GTP-binding proteins (G-proteins) via their intracellular loops and C-terminal tail following agonist-binding (Ladds *et al.*, 2003).

GPCRs can be activated by a wide-range of ligands including single photons of light, ions, nucleosides, hormones and large polypeptides (Kobilka, 2007). This enables GPCRs to control a wide range of physiological activities within the body including the perception of light and smell, heart rate as well as mood regulation. The ligand binding site can be located anywhere on the extracellular

side of the GPCR including the N-terminal region, extracellular loops and/or within the transmembrane segments. This said, the type and size of ligand does not necessarily determine where it binds to the receptor (Deupi and Kobilka, 2007; Kobilka, 2007).

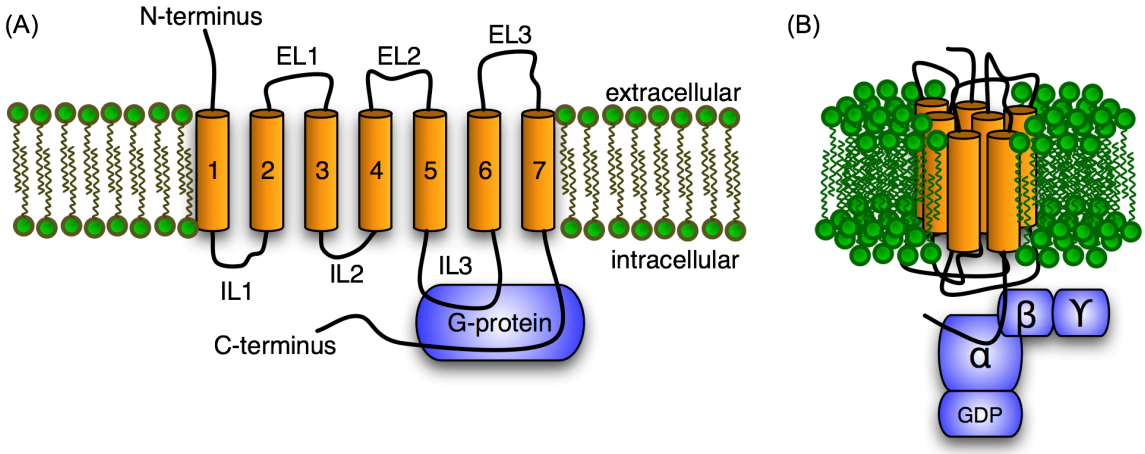


Figure 2.1 Schematic representation of a G-protein coupled receptor

(A) The basic topology of a GPCR consists of 7-transmembrane regions connected by 3 intra- (IL1-3) and 3 extra-cellular (EL1-3) loops. Ligands bind either at the extracellular surface or within the protein's TM domain whereas receptors couple to heterotrimeric G-proteins mainly via their third intracellular loop and C-terminal tail. (B) All GPCRs interact with a G-protein Gα subunit which in the inactive (GDP) state is tightly associated with the Gβγ complex.

Receptors were originally classified according to the endogenous ligand found to act upon it (Figure 2.2) (Cherezov *et al.*, 2007). However, it was quickly discovered that certain endogenous ligands caused a range of physiological effects in different tissues by binding to more than one receptor (Palmer *et al.*, 1995). Upon further investigation, it was found that receptors that bound the same ligand had quite distinct protein sequences. For example, the three β-adrenergic receptors have only 50% amino acid sequence identity in their transmembrane regions (Warne *et al.*, 2008). Therefore, there was a need to sub-classify close family relatives. Receptor sub-types usually have different tissue distributions as well as distinct pharmacological properties (both ligand binding and efficacy of G-protein activation), and in many cases bind different G-proteins (Bockaert and Pin, 1999).

Due to their abundance and cellular location, GPCRs are the major site of action for therapeutic drugs (Kobilka, 2007). However, other than their basic 7-transmembrane topology little is known about the intricacies of their structure or mechanism of action.

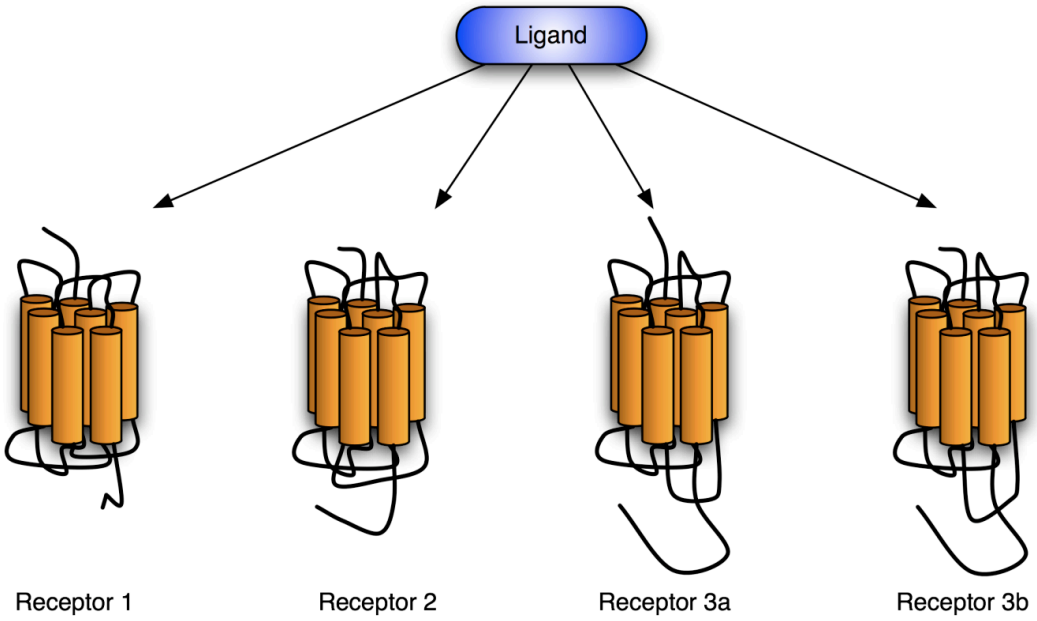


Figure 2.2 Schematic diagram illustrating GPCR sub-classification

Many endogenous ligands bind to several highly-related receptors in the body. Receptor sub-types mainly differ from each other in the length of their N- and C-termini as well as their extra- and intra-cellular loop regions.

2.3.1 GPCRs - signalling mechanisms

Following ligand-binding GPCRs selectively activate signalling cascades within the cell via the activation of specific G-proteins (Bockaert and Pin, 1999; Kobilka, 2007). When a ligand binds to the receptor (Lundström, 2005) a conformational change occurs within the intracellular region of the protein that allows specific G-proteins to bind and become activated (Bockaert and Pin, 1999; Kobilka, 2007). In their inactive state, G-proteins exist as heterotrimers consisting of alpha (α), beta (β) and gamma (γ) subunits that form a high affinity complex (Lambright *et al.*, 1996). Activated GPCRs can initiate the exchange of the Gα subunits bound GDP for GTP, by acting as guanine-nucleotide exchange factors (GEFs) (Ritter and Hall, 2009). This then leads to the dissociation of Gα-

GTP from the Gβγ subunits (Ladds *et al.*, 2003), which can alter the activity of downstream effecter proteins (such as adenylyl cyclase). By either inhibiting or stimulating the production of second messengers (such as cyclic AMP) an intracellular signal is generated (Flower, 1999).

The DRY motif is found at the boundary between the second transmembrane helix and the second intracellular loop of type I GPCRs (Rovati *et al.*, 2007). This motif is essential for G-protein coupling to the ligand-bound receptor and plays an important role in GPCR activation. Other than forming stabilizing interactions within the receptor constraining the inactive form, the role of the DRY motif in effecting G-protein coupling is not yet understood (Yao *et al.* 2006).

2.3.1.1 G-proteins

When G-proteins were originally discovered there was thought to be just one of each kind of sub-unit. Since then, however, 20 α, 6 β and 11 or 12 γ subunits have been identified in humans (Hamm, 1998; Neves *et al.*, 2002) which give rise to a large number of distinct G-protein hetero-trimers each of which generate a specific range of intracellular signals (Figure 2.3) (Ladds *et al.*, 2003). The α subunits can be classified into 4 distinct groups by sequence homology – G_s, G_i, $G_{q/11}$, $G_{12/13}$ (Neves *et al.*, 2002). The first group, G_s, is the best studied and upon GDP-GTP exchange activates adenylyl cyclase increasing the intracellular levels of cyclic AMP (cAMP). G_i acts in direct opposition to G_s inhibiting adenylyl cyclase thereby decreasing cAMP levels (Hamm, 1998). Members of the $G_{q/11}$ sub-family activate phospolipase C which hydrolyses phosphatidylinositol biphosphate to give the second messengers inositol triphosphate (IP_3) and diacylglycerol (DAG) (Neves *et al.*, 2002). G_{12} acts on guanine nucleotide exchange factors (GEFs) to activate GTPase-activating proteins.

In addition, the Gβγ subunits regulate the activity of other proteins/signalling events in the cell. For example, Gβγ can activate inward rectifying K^+ channels (Logothetis *et al.*, 1987; Dascal, 1997) as well as modulate the voltage-dependence of specific calcium channels (Herlitze *et al.*, 1996; Ikeda, 1996; De Waard *et al.*, 1997). Furthermore, Gβγ can regulate Ras which, in turn, controls

the activity of certain mitogen-activated protein kinases (MAPK) (Schmitt and Stork, 2002) (including the extracellular signal-regulated kinases (ERKs) (Koch *et al.*, 1994; Obara *et al.*, 2008)), as well as the c-Jun NH_2-terminal kinases (JNKs) (Coso *et al.*, 1997).

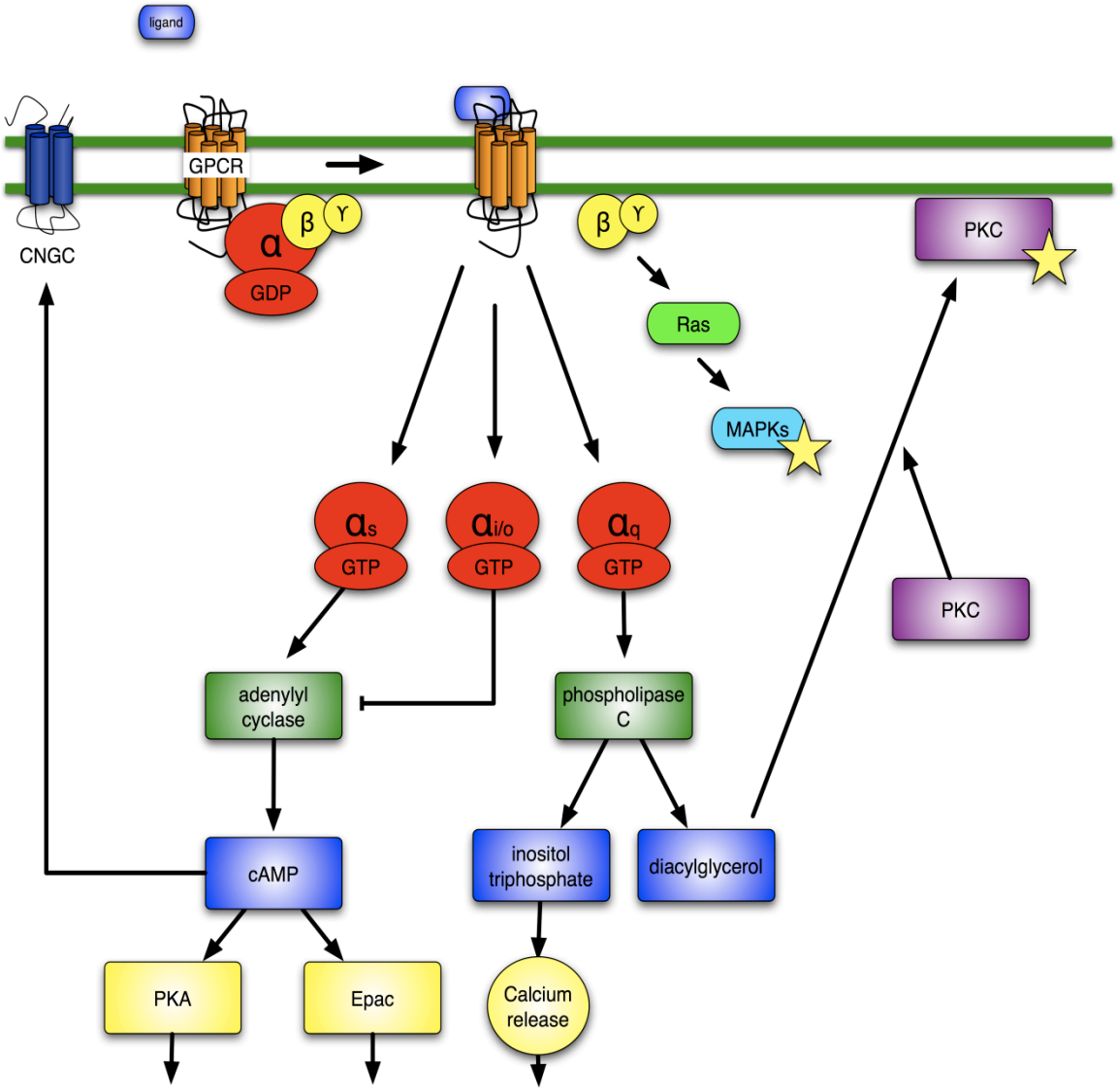


Figure 2.3 Overview of intracellular signals generated by G-proteins following GPCR activation

Agonist binding at the GPCR leads to the selective activation of associated G-proteins by catalysing the exchange of GDP for GTP at the Gα sub-unit. This causes the Gα and Gβγ sub-units to dissociate. Gα-GTP alters the activity of downstream signalling enzymes such as adenylyl cyclase and phospholipase C which, in turn, alters the levels of second messenger molecules cAMP and IP3/DAG, respectively. The Gβγ subunits modulate other signalling pathways. For example, Gβγ has been found to activate mitogen activated protein kinases (MAPKs) via Ras.

The duration of signals propogated by the Gα and Gβγ sub-units is determined by the GTPase activity of the Gα subunit (Cabrera-Vera *et al.*, 2003). Following hydrolysis of the bound GTP to GDP, the Gα and Gβγ sub-units re-associate, returning the G-protein to its resting state (Lambright *et al.*, 1996).

Although the third intracellular loop of GPCRs is known to be a critical determinant of which G-protein complex is activated by each receptor (Kobilka and Schertler, 2008) certain GPCRs have been shown to couple to more than one type of G-protein depending on the particular ligand bound to the receptor as well as the pool of G-proteins available within each cell (Ladds *et al.*, 2003).

2.3.1.2 GPCR desensitisation

To prevent chronic over-stimulation of cells GPCR signalling needs to be switched off (Figure 2.4) (Ferguson, 2001; Ma and Pei, 2007). This is accomplished by phosphorylation of specific residues in the cytoplasmic regions of agonist-bound GPCRs by G-protein coupled receptor kinases (GRKs) (Tobin *et al.*, 2008). In certain cases, phosphorylation is enough to prevent G-protein binding (Maudsley *et al.*, 2005). However, in most cases, GPCR phosphorylation is followed by recruitment of an arrestin protein to the receptor's C-terminal tail that physically blocks G-protein activation (Gurevich and Gurevich, 2006; Ma and Pei, 2007). There are four arrestin proteins (Ferguson, 2001; Rajagopal *et al.*, 2010) that associate with phosphorylated GPCRs (Brzostowski and Kimmel, 2001; Gurevich and Gurevich, 2006; Ma and Pei, 2007). By blocking G-protein activation, arrestins "switch off" the intracellular signalling cascade (Bockaert *et al.*, 2004). The arrestins also target the GPCR for internalisation where they are either dephosphorylated and recycled to the plasma membrane or targeted to the lysosome for degradation (Hall and Lefkowitz, 2002; Bockaert *et al.*, 2004; Gurevich and Gurevich, 2006; Kelly *et al.*, 2008). Receptor desensitisation is an important aspect of GPCR biology not just in switching off the current signalling event but also priming the receptor for receiving a new signal (Lefkowitz, 1998).

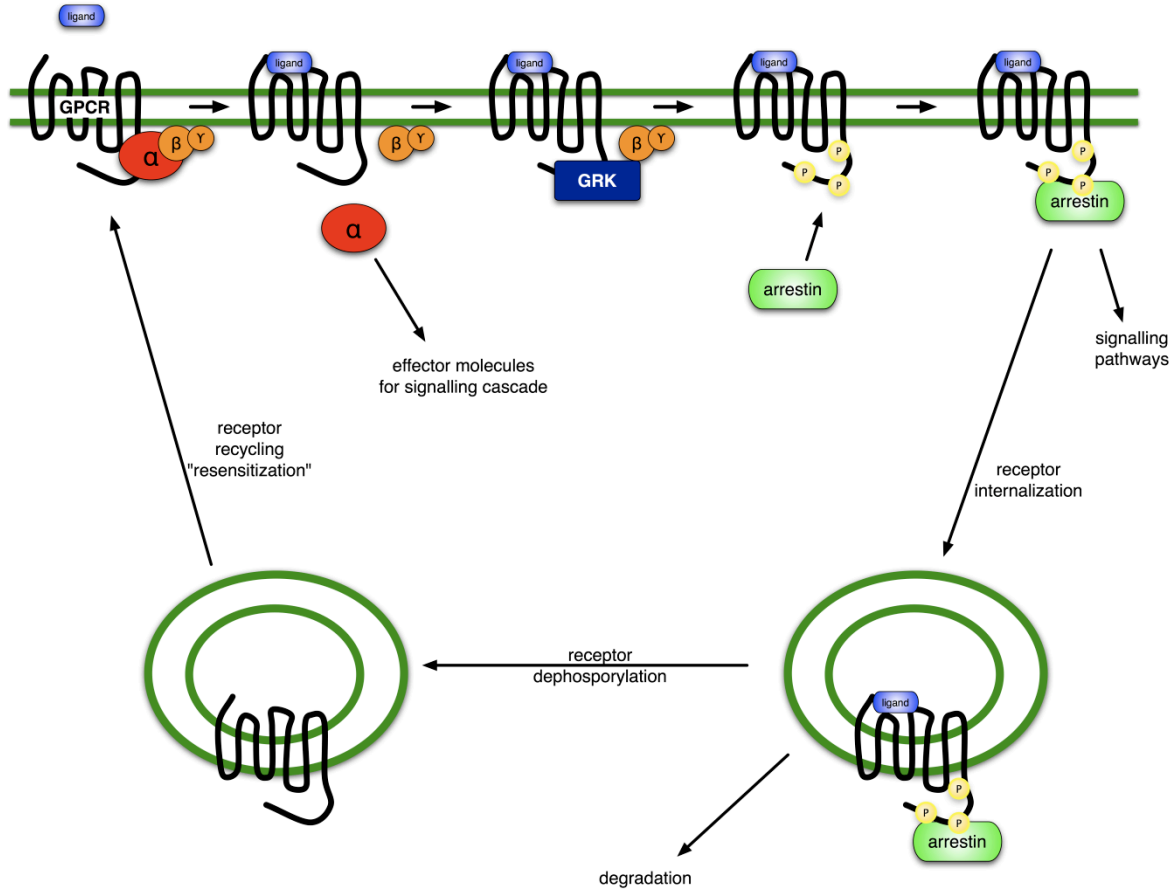


Figure 2.4 GPCR desensitisation via the combined action of G-protein receptor kinases (GRKs) and arrestins

Following agonist binding, GPCRs are able to selectively activate specific G-proteins starting a signalling cascade within the cell. To switch off signalling by the receptor, G-protein receptor kinases (GRKs) phosphorylate specific residues in the cytoplasmic tail of the agonist-bound receptor. This drives the recruitment of arrestin proteins to the receptor and ultimately results in receptor internalization where it is either dephosphorylated and recycled to the cell surface or targeted to the lysosome for degradation. Prior to internalisation, arrestins can stimulate G-protein independent signalling. In certain cases, phosphorylation of the agonist-bound GPCR by the GRK is sufficient to prevent G-protein binding, switching off the signal. This figure is adapted from ((Ma and Pei, 2007)).

2.3.1.3 G-protein independent signalling

Recently, it has been shown that GPCRs can also signal in a G-protein independent fashion (Bockaert and Pin, 1999; Kobilka, 2007) by binding regulatory proteins called GPCR interacting proteins (GIPs) (Bockaert *et al.*, 2004; Gsandtner *et al.*, 2005; Ritter and Hall, 2009). Due to the large number of

proteins that act as GIPs their modes of action can vary considerably. Known effects include altering the pharmacological properties, function and trafficking of GPCRs (Brzostowski and Kimmel 2001; Ritter and Hall 2009).

Two of the best characterised GIPs are β-arrestin 1 and 2 which, following recruitment to a GPCR, act as multi-functional adaptor proteins, scaffolding a diverse group of signaling proteins at the receptor, thereby modulating the downstream activity of a number of signaling networks (reviewed in (DeWire *et al.*, 2007)). For example, β-arrestin 1 recruits an activated non-receptor tyrosine kinase (SRC), to selected receptors causing the downstream activation of extracellular-signal related kinase (ERK) (Luttrell and Gesty-Palmer, 1999; Luttrell *et al.*, 2001). β-arrestins are also known to cause transactivation of the epidermal growth factor receptor (EGFR) in a SRC-dependent manner, as well as being involved in the regulation of receptor trafficking and the activity of small GTPases (Rajagopal *et al.*, 2010).

Other GIPs include (i), the Janus family of kinases (JAKs) and members of the signal transducers and activators of transcription (STATs) (Marrero *et al.*, 1998), (ii) PDZ domain containing proteins such as the Na^+ H^+ exchange regulatory factor (which controls the activity of the Na^+-H^+ exchanger type 3 (NHE3)) (Hall *et al.*, 1998 a,b), and (iii) phosphoinositide-3-kinase (PI3K) sub-unit p85 that influences the activity of the Akt signaling pathway (Bousquet *et al.*, 2006). These examples give just a flavour of the ways that GIPs extend the capacity of GPCRs to regulate cellular processes. In summary, by binding GIPs, GPCRs are able to generate an enormous diversity of intracellular signals in distinct cellular compartments in response to a relatively small number of endogenous ligands.

2.3.1.4 GPCR oligomerisation

Another important aspect of GPCR signalling is that of receptor oligomerisation. In recent years, considerable evidence has been obtained showing that many proteins exist as dimers, and sometimes even as higher order oligomers in both homo- and hetero-oligomeric complexes (Javitch, 2004). It has even been suggested that all GPCRs exist in some form of multimeric complex (Breitwieser,

2004; Fredholm *et al.*, 2007a). Dimerisation was first established for members of the type III class of GPCRs (including the metabotropic glutamate receptor 5 (Romano *et al.*, 1996; Romano *et al.*, 2001) and γ-aminobutyric acid (GABA) receptors (Jones *et al.*, 1998; Kaupmann *et al.*, 1998; White *et al.*, 1998; Kuner *et al.*, 1999; Ng *et al.*, 1999)). Since then, it has become increasing apparent that GPCRs from all families interact with different partner GPCRs on the surface of cells in a variety of complexes. There is considerable evidence that GPCR dimerisation occurs at the level of protein synthesis in the endoplasmic reticulum or during maturation in the Golgi and it is thought that many, if not most GPCRs, are trafficked to the plasma membrane as part of a dimeric complex (Bulenger *et al.*, 2005; Milligan, 2007).

The biological significance of heterodimerisation is not well understood, or exploited, but is predicted to have a major impact on drug discovery in the coming years (George *et al.*, 2002). This said, certain ligands have been shown to have direct and selective effects on the conformation of the individual receptors that make up hetero-dimers. For example, in the case of the GABA(B) receptor, binding of the neurotransmitter in the N-terminal region of the GABA(B)R1 sub-unit results in activation of a G-protein by the GABA(B)R2 protomer (Jones *et al.*, 1998; Breitwieser, 2004; Milligan, 2009). This implies communication between two receptor units with distinct conformational changes in both proteins following ligand binding.

Although there are now numerous reports of GPCR oligomerisation in the literature, not all of the suggested dimers are likely to occur in native tissues. There are two main reasons for this: i) in transient transfection experiments possible GPCR dimers/oligomers that are detected may just be an artifact of receptor overexpression in a non-native background, and ii) even when a number of clear differences in the function and pharmacology of co-expressed pairs of GPCRs have been observed care has to be taken to discriminate between those effects that are due to direct protein-protein interactions and those that are produced indirectly via downstream signaling and feedback control (Milligan, 2007). A physiologically-relevant and therapeutically interesting example of receptor hetero-oligomerisation is that between the adenosine 2a receptor

(A2aR) and the dopamine D2 receptor (D2R). These GPCRs have been shown to exist as heteromeric complexes in both transfected and striatal tissue (Ferre *et al.*, 2001; Fuxe *et al.*, 2007). Furthermore, in brain A2aR antagonists have a strong inhibitory effect on the activity of the D2R reducing tremor and rigidity in patients with Parkinson's disease (Ferre *et al.*, 2001; Fuxe *et al.*, 2007).

2.3.2 GPCRs: conformational complexity

In the past, GPCRs have been described as "bimodal switches" in which they exist in either an active or inactive conformation (Kobilka and Deupi, 2007) (Figure 2.5 (A)). However, this "lock and key" model did not account for the association of G-protein with the receptor and so the simple ternary complex model as a description of GPCR biology was devised (Figure 2.5 (B)). This model was later modified to take account of the high and low affinity states of both receptor and receptor-G protein complexes for ligand (Figure 2.5 (C)). Many receptors have been shown to have constitutive activity (Klinger *et al.*, 2002) allowing them to signal in the absence of an agonist (Vauquelin and Van Liefde, 2005). This finding gave rise to the extended ternary complex model (Figure 2.5 (D)). Common to all of these models is the assumption that the receptor has a unique conformation depending on which partner(s) (eg agonist, G-protein) are associated with it. Although these models may appear complicated the reality is considerably more convoluted with the fact that most GPCRs bind partial (inverse) agonists that stabilise the receptor in a series of conformational intermediates each with their own signalling properties (Ladds *et al.*, 2003; Kobilka and Deupi, 2007). Although both ligands and G-proteins are known to preferentially interact with discrete GPCR states, fluctuations between conformations can occur spontaneously (Vauquelin and Van Liefde, 2005).

The vast range of conformations that GPCRs are known to exist in has proven to be a substantial barrier in efforts to crystallise GPCRs. Well-ordered crystals (required for structure determination) can only be formed from a conformationally homogeneous sample. Therefore, strategies for stabilising specific receptor conformations have to be used with all GPCRs before crystals that diffract to high resolution can be obtained.

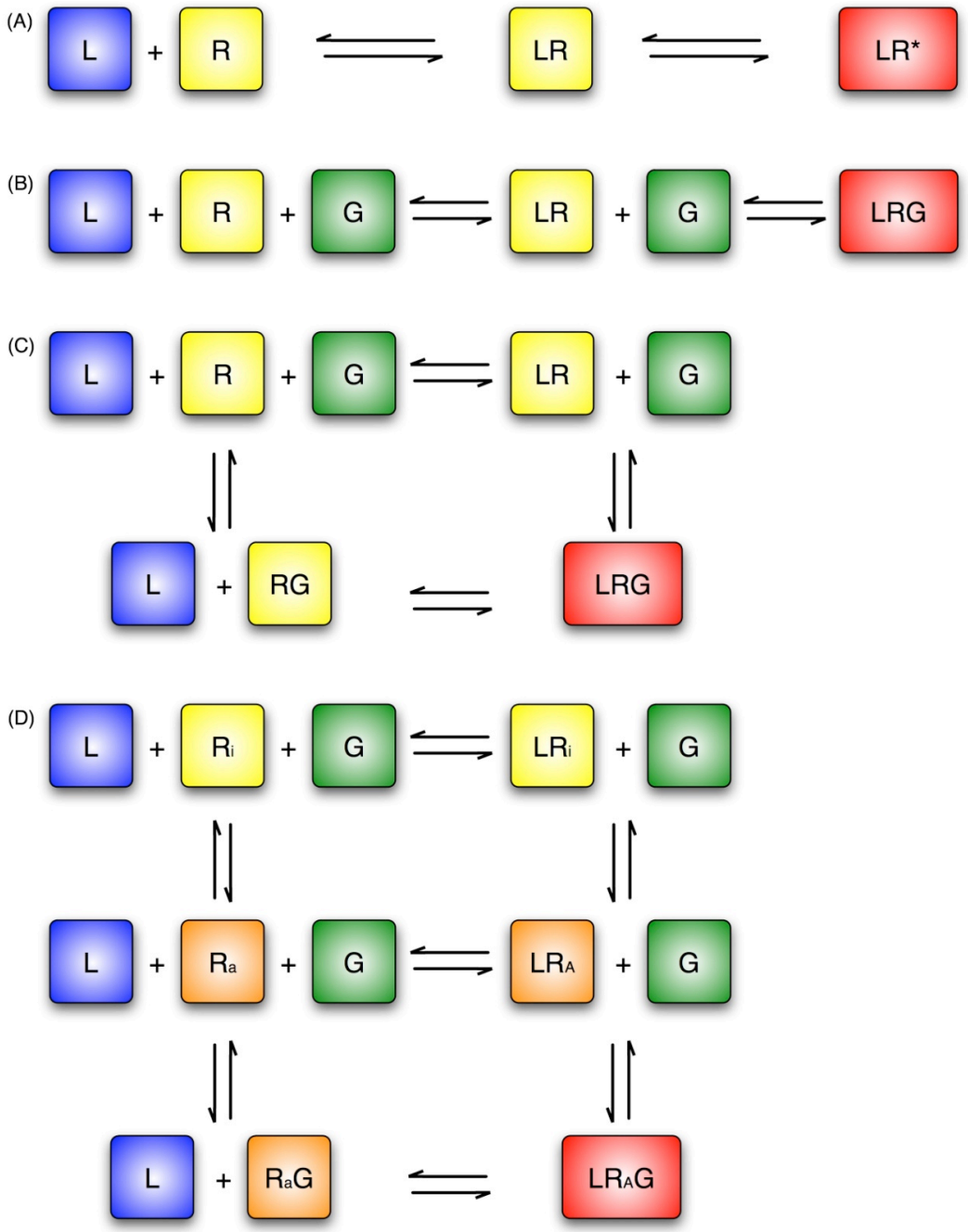


Figure 2.5 Various models describing the possible conformational states of GPCRs

(A) The two-state model in which ligand (L) binding to the receptor (R) results in the formation of a LR complex generating an active receptor conformation (R*) that signals to downstream effectors. (B) The simple ternary complex model (where G is for G-protein). (C) The full ternary complex model which accounts for the high and low affinity states of both R and RG for L. (D) The discovery of constitutively active GPCR mutants resulted in the development of an extended ternary complex model in which the receptor can adopt an active conformation (R_a) in the absence of L or G binding. R_i denotes the inactive conformation of the receptor. (Red indicates a fully active complex, orange is constitutive activity, yellow is inactive receptor.) This figure was modified from (Rajagopal *et al.*, 2010).

2.3.3 Current structural knowledge of GPCRs

To date, only four G-protein coupled receptors have had their structures solved to atomic resolution (figure 2.6) (Palczewski *et al.*, 2000; Cherezov *et al.*, 2007; Rasmussen *et al.*, 2007; Jaakola *et al.*, 2008; Warne *et al.*, 2008) which, considering their biological importance, gives some indication of how difficult it is to obtain a GPCR structure.

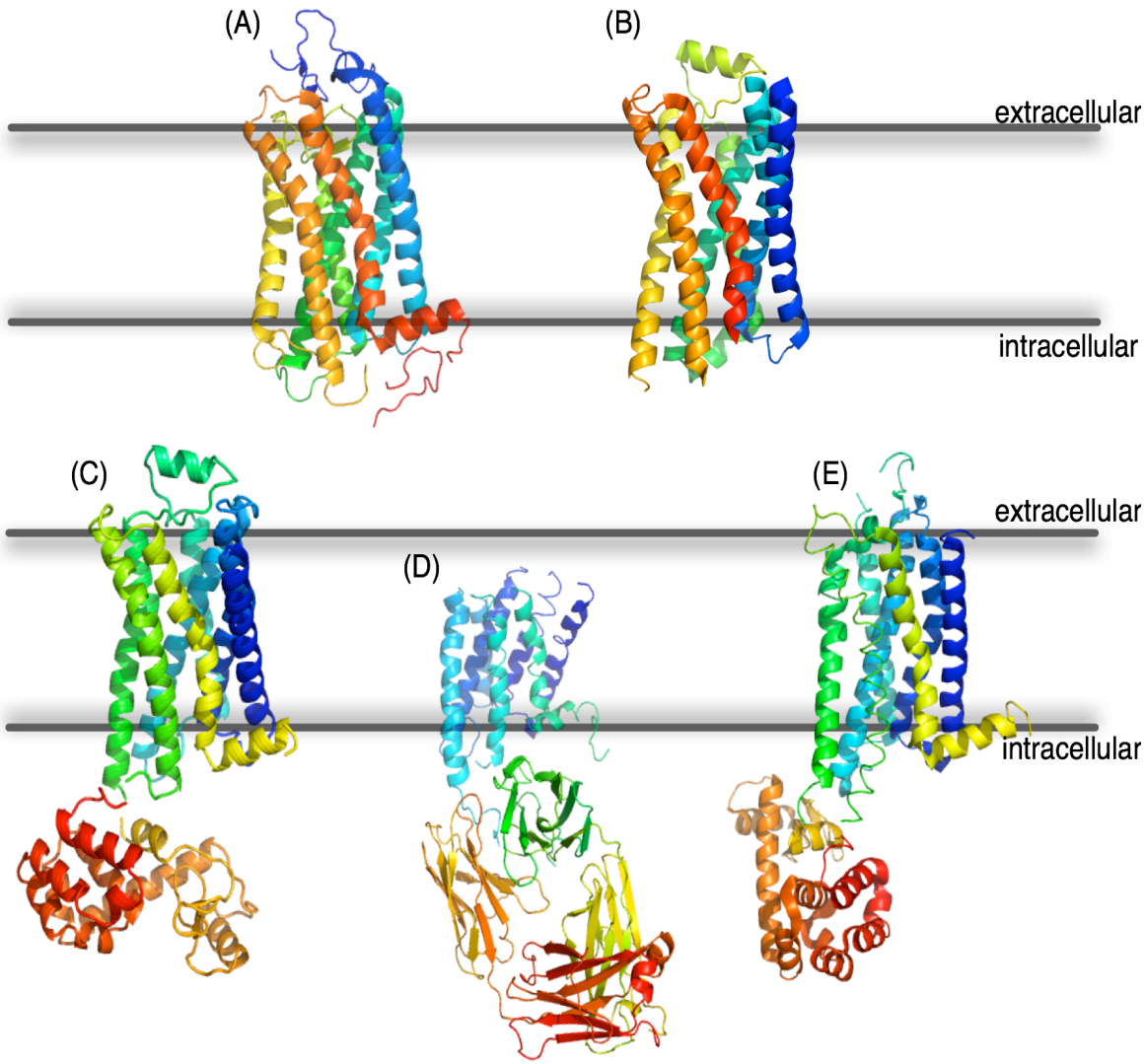


Figure 2.6 GPCR structures in the PDB

Available structures include (A) rhodopsin (1F88), (B) turkey β1-adrenergic receptor (2VT4), (C) human β2- adrenergic receptor with IL3 replaced with T4 lysozyme (2RH1), (D) human β2-adrenergic receptor in complex with a F_{ab} (2R4R), and (E) the human adenosine 2a receptor with T4 lysozyme (3EML), (Palczewski *et al.*, 2000; Cherezov *et al.*, 2007; Rasmussen *et al.*, 2007; Jaakola *et al.*, 2008; Warne *et al.*, 2008).

The first mammalian GPCR to have its structure solved to high resolution was bovine rhodopsin (Palczewski *et al.*, 2000). This is not surprising as rhodopsin has several properties that make it easier to work with compared to other GPCRs. Firstly, as rhodopsin has a high natural abundance it did not have to be overexpressed in a heterologous host before structural studies could begin. Secondly, rhodopsin covalently binds 11-cis-retinal (which acts as an inverse agonist) stabilising the receptor in its inactive conformation (Palczewski *et al.*, 2000). Finally, the bound retinal chromphore gives rhodopsin a red colour which makes it easy to identify those fractions containing receptor throughout the purification process (Teller *et al.*, 2001).

Rhodopsin functions in the perception of light (Palczewski *et al.*, 2000; Teller *et al.*, 2001). On absorbing a photon of light energy the bound retinal changes from an all-*cis* to an all-*trans* conformation that activates rhodopsin which, in turn, activates the G-protein transducin (Sheikh *et al.*, 1996). Thus far, 15 discrete rhodopsin structures have been solved and deposited in the PDB, each giving different insights into rhodopsin's mechanism of activation.

The first rhodopsin structure confirmed the 7TM alpha helical topology that had been predicted for all GPCRs. However, an additional irregular α-helix (helix 8) that is situated along the surface of the cytoplasmic membrane was also observed (Figure 2.6 (A)) (Palczewski *et al.*, 2000). The 7 TM helices vary in length according to the angle that they cross the membrane. Proline residues in TM I, IV, VI and VII creates kinks in the α-helices. The receptor is held together by hydrogen bonds between helices, a disulphide bond (between extracellular loop 2 and TM3) as well as a salt bridge (between the E/DRY motif (Glu134, Arg135 and Tyr136) in TM3 and Glu247 & Thr251 (both TM5)).

Further structural studies have identified several water molecules that occur within the transmembrane region of rhodopsin which are essential for the formation of the hydrogen-bond network between helices (Okada *et al.*, 2002; Li *et al.*, 2004). To understand the changes in protein conformation that accompany rhodopsin's photochemical reactions structures of various intermediate photostates (including inactive (or "dark" state) rhodopsin through

bathorhodopsin, lumirhodopsin, metarhodopsin I and deprotonated metarhodopsin II (the active "light" state) have been determined (Choi *et al.*, 2002; Nakamichi and Okada, 2006 a,b; Salom *et al.*, 2006).

Although heterologous expression of rhodopsin has proved difficult, by introducing a stabilising disulphide bond and removing its only glycosylation site it has been possible to obtain milligram quantities of recombinant rhodopsin from which diffracting crystals and a high-resolution structure have been obtained (Standfuss *et al.*, 2007). This was a landmark achievement as it was the first time that any GPCR had been produced in a heterologous host and then crystallised.

More recently, a crystal structure has been obtained for ligand-free opsin that shows the structural changes that occur within the E/DRY motif and TM5-7 when the receptor is in its active conformation (Park *et al.*, 2008). Furthermore opsin has recently been crystallised in complex with a peptide from the alpha subunit of the G-protein transducin (Scheerer *et al.*, 2008). This structure has given the first insights into the conformational changes that occur on G-protein binding and activation by a GPCR.

As rhodopsin is an atypical GPCR, it was not clear how relevant the structures of rhodopsin would be to the wider GPCR family. Within the past few years, however, structures of the human β2 adrenergic receptor (Figure 2.6 (C, D)), the turkey β1 adrenergic receptor (Figure 2.6 (B)) and the human adenosine 2a receptor (Figure 2.6 (E)) have all been deposited in the PDB. For all of these proteins, diffracting crystals were only obtained using modified receptor constructs in which the highly mobile third intracellular loop (known to be involved in G-protein coupling) had been replaced with T4 lysozyme (Cherezov *et al.*, 2007; Jaakola *et al.*, 2008), shortened (Warne *et al.*, 2008) or bound by a F_{ab} fragment (Rasmussen *et al.*, 2007). Furthermore, diffracting crystals were only obtained using purified receptor that was deglycosylated, had a truncated C-terminal tail and had either an antagonist or inverse agonist bound (which locked the receptor into an "inactive" state).

The β2-adrenoreceptor (β2AR) is probably the best characterised ligand-activated GPCR primarily as it was the first non-rhodopsin GPCR to be cloned (Dixon *et al.*, 1986) but also because it is an important therapeutic drug target in the treatment of cardiovascular and pulmonary disease (Kolb *et al.*, 2009; Overington *et al.*, 2006; Takeda *et al.*, 2002). Structures of carazolol-bound β2AR have been solved in two different ways (Cherezov *et al.*, 2007; Rasmussen *et al.*, 2007). In the first method, the β2AR was co-crystallised with a specific antibody fragment (F_{ab}) bound to the flexible third intracellular loop. Using this approach a structure at a resolution of 3.7 Å was obtained though the electron density around the ligand binding site was poorly resolved (Rasmussen *et al.*, 2007). In the second method IL3 was replaced with T4 lysozyme. Crystals of this receptor construct were grown in a lipidic-cubic phase that diffracted to a resolution of 2.4 Å (Cherezov *et al.*, 2007). Reassuringly, the two methods gave structures where the arrangement of the TM helices and the specific interactions between carazolol and the receptor were essentially the same (Cherezov *et al.*, 2007; Rasmussen *et al.*, 2007; Rosenbaum *et al.*, 2007). Furthermore, the arrangement of transmembrane helices in the β2AR was found to be largely similar to that in rhodopsin. However, an additional short helix in EL2 occurs in the β2AR which differs from that in rhodopsin where there is a flexible loop. The helix in EL2 is thought to prevent the movement (and closing) of the ligand-binding pocket in the β2AR (which is permanently capped in rhodopsin) (Standfuss *et al.*, 2007).

Although carazolol and retinal occupy similar spaces in the β2AR and rhodopsin, the β-ionone ring of retinal extends deeper into the ligand binding pocket than carazolol contacting the highly conserved Trp265 in helix VI of rhodopsin (Cherezov *et al.*, 2007). Changes in the side-chain of Trp 265 are thought to occur upon activation of rhodopsin (as well as other type I GPCRs). The interactions between retinal (a full inverse agonist) and Trp 265 are thought to contribute to rhodopsin's lack of basal activity. In contrast, the β2AR structure shows that carazolol does not interact with this "toggle switch", likely explaining why it is a partial and not a full inverse agonist (Cherezov *et al.*, 2007).

A high resolution structure of the turkey β1-adrenergic receptor was obtained from diffraction quality crystals of a detergent-resistant thermostable mutant (that had been created following extensive mutagenesis of the receptor) (Warne *et al.*, 2008; Warne *et al.*, 2009). By identifying mutations that conferred thermostability on the receptor, as well as locking the protein in an inactive conformation, it was possible to obtain pure, monodisperse protein in short chain detergents (Serrano-Vega *et al.*, 2008) which gave crystals that diffracted to 2.7 Å (Warne *et al.*, 2008).

When the structures of the β1- and β2-ARs are compared, there are no significant differences in the positions of the TM helices. Furthermore, the short α-helix in EL2 identified in the β2AR structures is present in β1AR also. This structural feature may allow GPCRs to bind diffusable ligands rapidly and reversibly (Warne *et al.*, 2008). The β1- structure varies from the β2-AR, however, in that it has an additional helix located within IL2. A hydrogen bond between Tyr149 in the IL2 helix and Asp138 in TM3 may potentially play a role in receptor G-protein specificity but further studies are required to clarify this suggestion (Warne *et al.*, 2008).

The most recent GPCR to have its structure determined was the human adenosine 2a receptor which was solved using the T4-lysozyme insertion approach (Jaakola *et al.*, 2008). The receptor was co-crystallised with the antagonist ZM241385 and a structure to 2.6 Å resolution obtained. Although the hA2aR has the same helical scaffold as both the β1- and β2-ARs, the way in which the antagonist ZM241385 binds to the receptor (perpendicular to the lipid membrane interacting with EL2 and EL3) is completely different (Jaakola *et al.*, 2008). This suggests that further GPCR structures may reveal a diverse range of ligand binding pockets (Jaakola and Ijzerman, 2010). In contrast to the helix that occurs in the EL2 of the β-ARs, the hA2aR has three disulphide bridges (which are essential for high-affinity antagonist binding) instead (Jaakola and Ijzerman, 2010).

While these structures have greatly enhanced our knowledge of GPCR biology at the molecular level, only through obtaining further structures of different GPCRs in a range of conformational states will it be possible to fully understand the molecular details of GPCR activation.

2.3.4 Structure-based drug design against GPCRs

Given that GPCRs are one of the largest protein superfamilies, function at the cell surface and regulate a vast array of physiological processes, it is not surprising that a large number of existing therapeutic molecules act against them. However, many drugs have side-effects arising from non-specific interactions of the therapeutic with other cellular proteins. It is hoped that by solving the structures of ligand-activated GPCRs to atomic resolution it will be possible to design drugs that act specifically against particular receptors, reducing the problem of unwanted side-effects.

2.3.5 Human adenosine receptors

Adenosine is an important chemical involved in the inflammatory response with its extracellular accumulation being the first major signal in response to ischaemia and inflammation (Ohta and Sitkovsky, 2001; Murphree *et al.*, 2002; Sands and Palmer, 2005). The levels of adenosine that accumulate can vary dramatically dependent on the level of stress and the tissue affected (Jacobson and Gao, 2006). Adenosine is formed from AMP by the action of the AMP selective 5' nucleotidase, and its rate of formation is directly proportional to the amount of available AMP (Fredholm *et al.*, 2007b).

In humans, there are four adenosine receptor subtypes - A1, A2a, A2b and A3. Although all of them are activated by adenosine (Gsandtner and Freissmuth, 2006), there is only ~50% sequence homology between the receptor subtypes (Figure 2.7) (Moreau and Huber, 1999). Upon adenosine binding, these receptors activate different G-proteins and signalling pathways within the cell (Palmer *et al.*, 1995; Gsandtner *et al.*, 2005). For example, the A2a receptor activates $G\alpha_s$

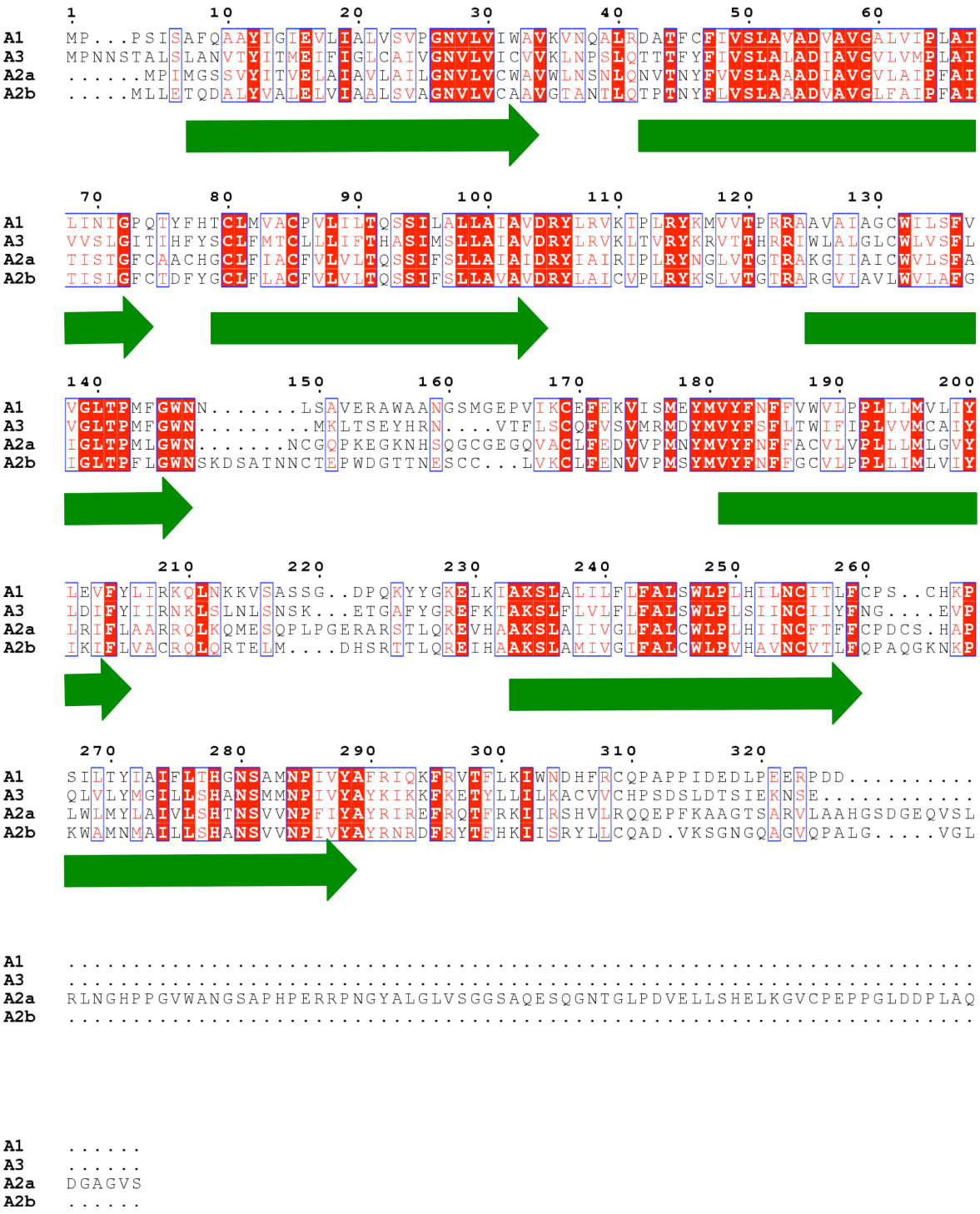


Figure 2.7 Sequence alignment of the four human adenosine receptor subtypes

When the amino acid sequences of the adenosine receptor subtypes were compared it was found that they had *ca.* 37% sequence identity. The alignment also shows that the adenosine 2a receptor subtype is unique in that it has an extremely long C-terminal tail - over 90 residues longer than any of the other adenosine receptor subtypes. Residues highlighted in red have 100% consensus while those boxed in blue have amino acid substitutions maintaining type. The green arrows indicate potential transmembrane domains.

whereas A1 activates $G\alpha_i$. This means that according to the type of adenosine receptor possessed by a cell, the presence of extracellular adenosine may either increase or decrease its cAMP level stimulating different intracellular signalling pathways (Olah, 1997). As well as their distinct signalling properties, the tissue distribution of the different adenosine receptor subtypes varies (Moreau and Huber, 1999) - A1 and A2a are present in both the cardiovascular and central nervous systems, A2b is ubiquitously expressed throughout the body while A3 receptors are highly enriched in the brain.

2.3.6 The human adenosine 2a receptor

The human adenosine 2a receptor (hA2aR) is expressed in high levels in lymphoid tissues such as spleen, leukocytes and blood platelets (Sands *et al.*, 2004) as well as vascular endothelial cells and striatal neuronal cells (Ferre *et al.*, 2001; Gsandtner and Freissmuth, 2006). hA2aR signaling has been extensively studied (Gomez and Sitkovsky, 2003; Sands *et al.*, 2004; Sands and Palmer, 2005) and it has been shown to be a key therapeutic target in the treatment of many inflammatory diseases (Jacobson and Gao 2006). Agonist molecules that interact with the A2aR down-regulate inflammation by inhibiting pro-inflammatory responses within the cell (Sands and Palmer, 2005), and are useful for treating diseases such as atherosclerosis, sepsis and tumour progression. More recently, hA2aR antagonists have become of interest to researchers developing new treatments for Parkinson's disease as the hA2aR dimerises with dopamine D2 receptors (Ciruela *et al.*, 2004) increasing the affinity of the D2 receptor for agonists in the presence of an A2aR antagonist (Ferre *et al.*, 2001).

The hA2aR has a unique C-terminal tail 122 amino acids in length (Figure 2.8). This is considerably longer than what is typical for a family I GPCR whose C-tails are usually in the region of 10-40 amino acids long. The hA2aR C-tail shows no sequence similarity with any other GPCR yet it is highly conserved between species, and is thought to play an important role in either trafficking or G-protein independent signaling but is not involved in G-protein coupling (Olah, 1997; Moreau and Huber, 1999). This region is predicted to be highly disordered, and likely explains why it is able to interact with a diverse range of adaptor

proteins (Fredholm *et al.*, 2007b). For example, the C-tail of the hA2aR has been shown to interact with α-actinin (which mediates receptor internalization) (Burgueno *et al.*, 2003), Usp4 (which is important for maintaining the cell surface position of the hA2aR) (Milojevic *et al.*, 2006) and ARNO (which activates MAP kinase) (Gsandtner *et al.*, 2005).

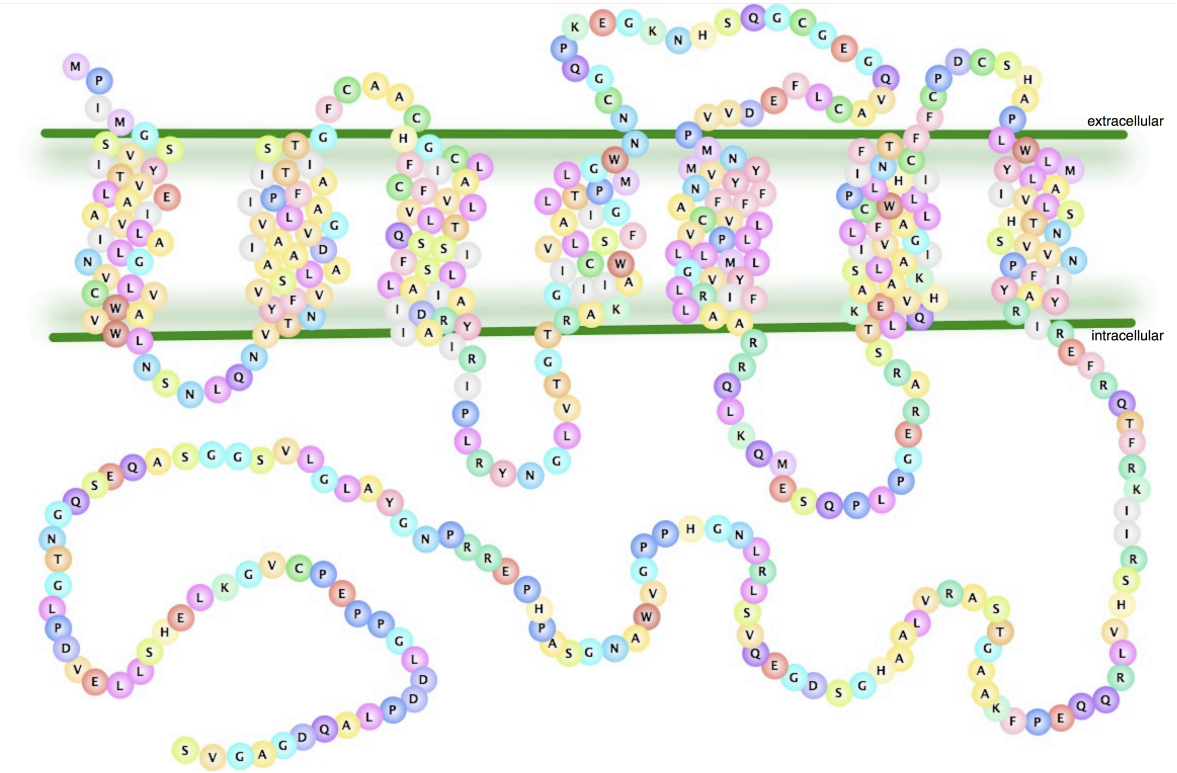


Figure 2.8 Schematic diagram of the human adenosine 2a receptor

Of note are the regions of the hA2aR predicted to be transmembrane as well as the receptor's large intracellular C-terminal tail. Colours are dependent on individual amino acid.

2.4 Rigid-bodied membrane proteins

In contrast to those membrane proteins with high conformational flexibility, rigid bodied membrane proteins are considered relatively easy to work as they have defined structures with very little, if any, regions that are natively disordered. Furthermore, these MPs tend to exist in a limited number of conformational states only. For example, ion channels tend to exist in either an open or closed state. For these reasons a relatively large number of mammalian rigid-bodied MPs have had their 3-dimensional structures solved including water

channels (Murata *et al.*, 2000; Horsefield *et al.*, 2008; Ho *et al.*, 2009), ion channels (Long *et al.*, 2005; Tao *et al.*, 2009), transporters (Aller *et al.*, 2009) and enzymes (Binda *et al.*, 2002; Hernandez-Guzman *et al.*, 2003; Ago *et al.*, 2007; Martinez Molina *et al.*, 2007).

As yet, only three mammalian integral membrane enzymes have had their structures solved (Binda *et al.*, 2002; Hernandez-Guzman *et al.*, 2003; Ago *et al.*, 2007; Martinez Molina *et al.*, 2007). This is rather surprising considering enzymes have highly-ordered structures with fixed active sites where specific substrates bind in a particular orientation in proximity to those amino acid residues that participate in the catalytic reaction (Nelsetuen, 1995; Welch, 1995; Binda *et al.*, 2002). In my Ph.D., I worked on the cholesterol biosynthetic enzyme sterol isomerase as an example of a human rigid-bodied MP.

2.5 Cholesterol

Cholesterol is the principal mammalian sterol. It is an amphipathic molecule consisting of a hydrophilic hydroxyl group, a rigid four-ring structure and a hydrocarbon tail (Figure 2.9). Cholesterol is an essential component of mammalian cell membranes where it regulates the membrane's structure, permeability and fluidity. The hydroxyl group of cholesterol interacts with the polar head groups of both phospho- and sphingo-lipids while the hydrocarbon part restricts the movement of the fatty acid chains of neighbouring lipid molecules (Rietveld and Simons, 1998). Cholesterol rich regions of membranes are known as lipid rafts. These have been shown to play important roles in regulating a diverse range of cellular processes including polarised trafficking events, endo- and exo-cytosis as well as signal transduction (Lucero and Robbins, 2004; Ramstedt and Slotte, 2006). Cholesterol is essential for the full activity of many different mammalian membrane proteins including GPCRs (Gimpl *et al.*, 1995; Klein *et al.*, 1995; Albert *et al.*, 1996; Gimpl *et al.*, 1997; Pang *et al.*, 1999; Lagane *et al.*, 2000; Gimpl and Fahrenholz, 2002), transporters (Shouffani and Kanner, 1990; Rothnie *et al.*, 2001; Tate, 2001; Magnani *et al.*, 2004),

pumps (Ding *et al.*, 1994; Cornelius, 2001) and ion channels (Fernandez-Ballester *et al.*, 1994). Cholesterol can modulate the activity of membrane proteins in two different ways: firstly, cholesterol contributes to the stability and functionality of membrane proteins by exerting lateral pressure on their transmembrane domains (Gimpl *et al.*, 1997); secondly, specific interactions between the protein and cholesterol can be essential for the protein's activity (Gimpl *et al.*, 1997; Tate, 2001; Hanson *et al.*, 2008). In addition to its role in the membrane, cholesterol is also the biosynthetic precursor of a wide-range of biologically active molecules including bile acids (Song *et al.*, 2007; Ferdinandusse *et al.*, 2009), steroid hormones (eg cortisol, aldosterone, progesterone, estrogen and testosterone) (He *et al.*, 2010) as well as several fat-soluble vitamins (eg Vitamin D) (Holick, 1995). In short, cholesterol is a pivotal molecule at the centre of mammalian cell biology.

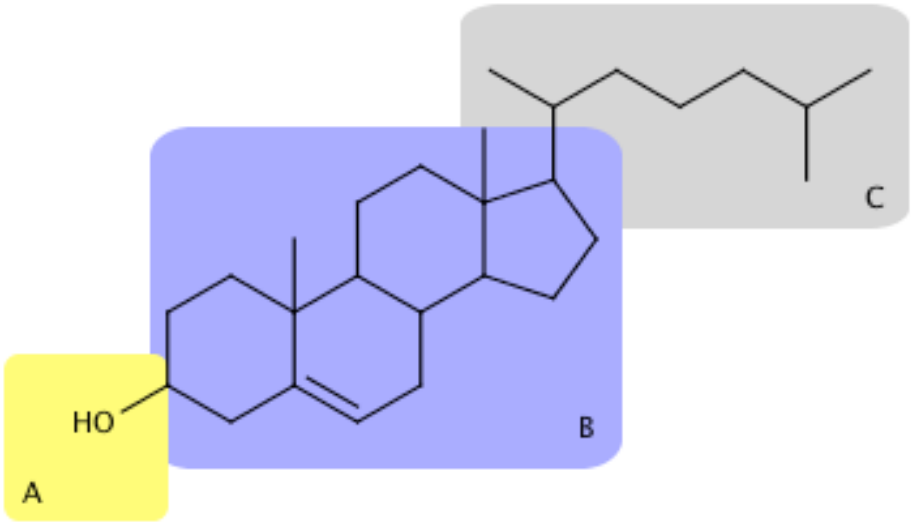


Figure 2.9 Chemical structure of cholesterol

Cholesterol has three distinct regions comprising of (A) a hydroxyl head group, (B) a rigid ring structure and (C) a hydrocarbon tail.

2.5.1 Cholesterol homeostasis

The amount of cholesterol in our bodies is more or less the same from day to day, and in a typical male adult is *ca.* 35g. However, cholesterol is made and turned over on a daily basis. Although each day we make *ca.* 1000 mg of

cholesterol and take-in an additional 200-500 mg from our diet, we lose an equivalent amount of cholesterol via excretion in bodily fluids (Charlton-Menys and Durrington, 2008). Cholesterol is continuously moved between different parts of the body and, due to its hydrophobic nature, is transported in the blood within lipoproteins (Moebius *et al.*, 2000; Ikonen, 2006; Daniels *et al.*, 2009). For example, very low-density lipoproteins (VLDL) and low-density lipoproteins (LDL) transport cholesterol from the liver to peripheral tissues (Simons and Ikonen, 2000). In contrast, high-density lipoproteins (HDL) carry cholesterol back from the extremities of the body to the liver where it is either broken down or converted into bile salts (Schmitz and Grandl, 2009). The cholesterol content of cells is dynamic, and the steady-state levels of cholesterol in individual cells are determined by the relative rates of cholesterol synthesis, uptake and export (Daniels *et al.*, 2009). Moreover, within each cell a cholesterol gradient exists with the highest and lowest levels occurring in the plasma membrane and endoplasmic reticulum, respectively (Schmitz and Grandl, 2009). When excess amounts of cholesterol accumulate in the body, disease always results. For example, elevated blood LDL-cholesterol levels are a major risk factor for atherosclerosis (and from this other cardiovascular diseases such as heart attacks and stroke) (Ikonen, 2006; Daniels *et al.*, 2009).

2.5.2 Cholesterol biosynthesis

In humans, cholesterol is primarily made in the liver and cells of the central nervous system (Charlton-Menys and Durrington, 2008). An overview of cholesterol synthesis is presented in Figure 2.10. Cholesterol synthesis starts with the reaction of acetyl CoA with acetoacetyl-CoA to form 3-hydroxy-3-methylglutaryl CoA (HMG-CoA). Reduction of this molecule to mevalonate is the committed step in cholesterol biosynthesis. This reaction is catalysed by the enzyme HMG-CoA reductase which is inhibited by the statin family of drugs (which are commonly used to lower blood cholesterol levels) (Charlton-Menys and Durrington, 2008). Following this, a series of molecular reactions occurs that results in the formation of squalene which is then cyclicised to form lanosterol (Do *et al.*, 2009). Lanosterol is converted to zymosterol through another series

of chemical modifications. Finally, cholesterol is made from zymosterol via three intermediates - cholesta-7,24-dienol, cholesta-5,7,24-trienol and 7-dehydrocholesterol – requiring the activity of a C8-C7 sterol isomerase, a C5 desaturase, a C24- and a C7-sterol reductase, respectively (Figure 2.10) (Moebius *et al.*, 2000).

2.5.3 C8-C7 sterol isomerase

Sterol isomerase (SI) catalyses the movement of the double bond in the second sterol ring from the C8 to the C7 position (Figure 2.10) (Moebius *et al.*, 1998). SI does not require any redox co-factors or metal ions to perform the isomerisation reaction, and it is unique amongst the cholesterol biosynthetic enzymes in that it is able to catalyse the reverse reaction also (Bae *et al.*, 2001). SI is located in the membrane of the endoplasmic reticulum (Simons and Ikonen, 2000), and is predicted to have four transmembrane domains (Figure 2.11). SI was originally identified as an emopamil-binding protein (EBP) (Silve *et al.*, 1996), and was thought to have an effect on ischemia, immunosuppression and neuroprotection (Moebius *et al.*, 2000). Since then, it has been shown that emopamil does not exert its effects through sterol isomerase activity as originally thought.

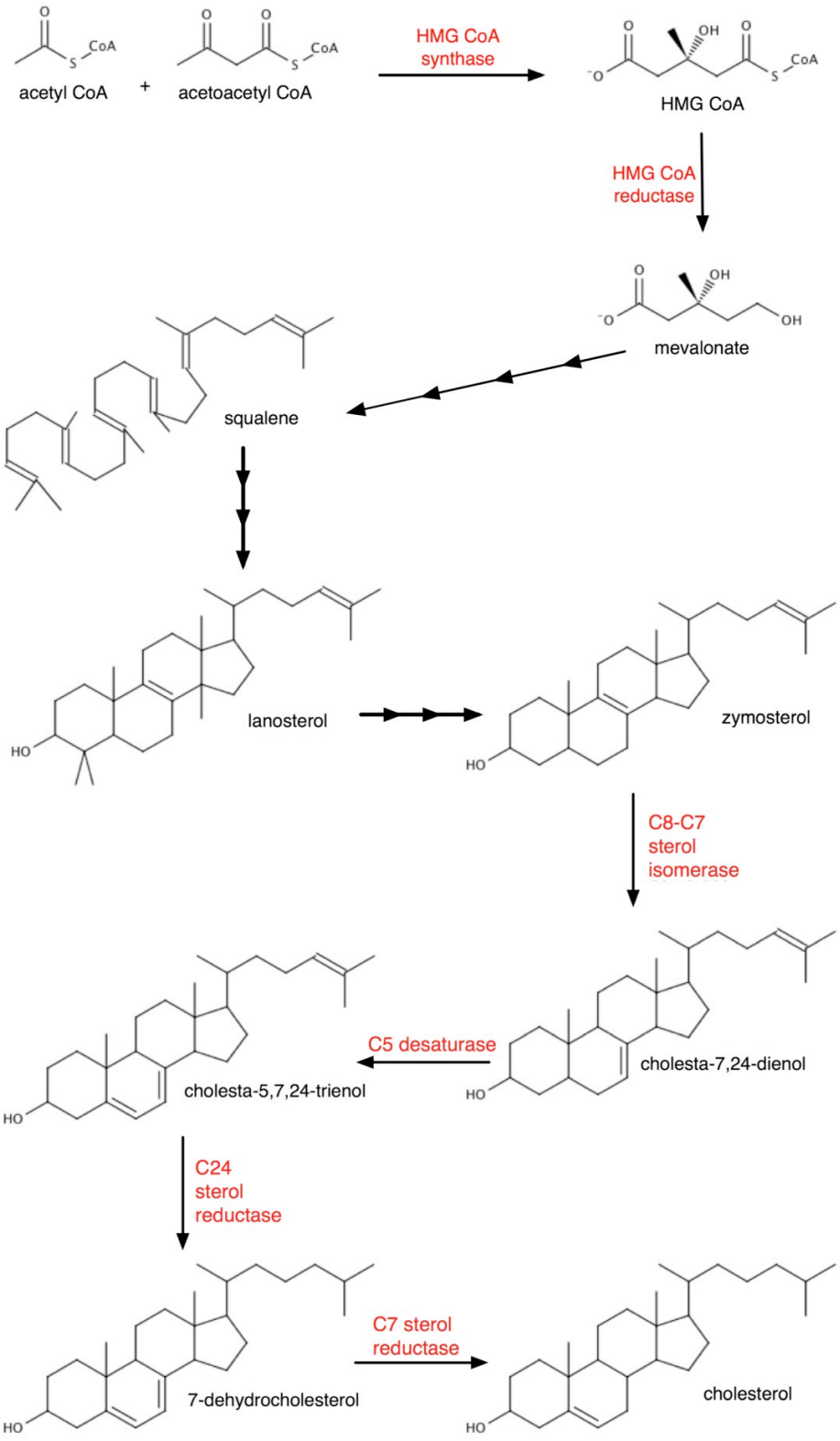


Figure 2.10 Overview of cholesterol biosynthesis

Cholesterol is formed from acetyl CoA and acetoacetyl CoA through a series of molecular reactions. Of note, C8-C7 sterol isomerase catalyses the conversion of zymosterol to cholesta-7,24-dienol.

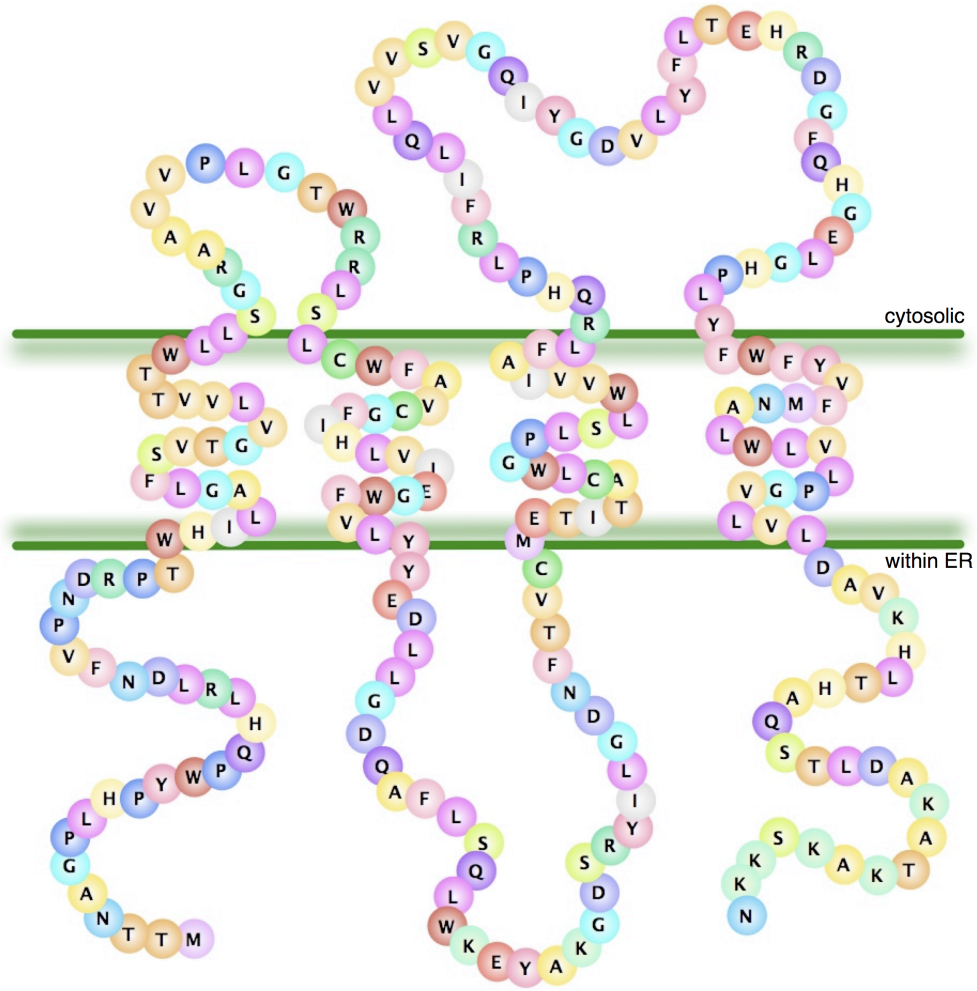


Figure 2.11 Schematic representation of sterol isomerase

Sterol isomerase is predicted to have four transmembrane domains (amino acids coloured individually).

2.5.3.1 Monogenic disease of sterol isomerase

At present, seven inherited disruptions of cholesterol biosynthesis have been described, five of which are post-squalene (Moebius *et al.*, 2000). Conradi-Hünermann-Happle (CHH) syndrome is caused by stable loss-of-function mutations within the gene encoding sterol isomerase (Has *et al.*, 2000). Due to its X-linked dominant inheritance pattern (Moebius *et al.*, 2000), CHH is also called chondrodysplasia punctata type II (CPDX2) (Becker *et al.*, 2001). The disease occurs almost exclusively in females but there has been a recent report of CDPX2 in a male patient (Aughton *et al.*, 2003). The key clinical features of CDPX2 are due to the skeletal dysplasia causing assymeteric limb shortening, short stature and scoliosis however it also affects the skin (linear ichthyosis),

eyes (cataracts) and hair (alopecia) (Figure 2.12). This severe phenotype is discernable from birth, and is caused by an inability to make cholesterol during fetal development. Several loss-of-function mutations (E80K, R110Q, R147G, R147H) in SI have been identified in CDPX2 patients (Braverman *et al.*, 1999; Derry *et al.*, 1999; Has *et al.*, 2000; Becker *et al.*, 2001). Currently, there is no treatment for this debilitating disease.

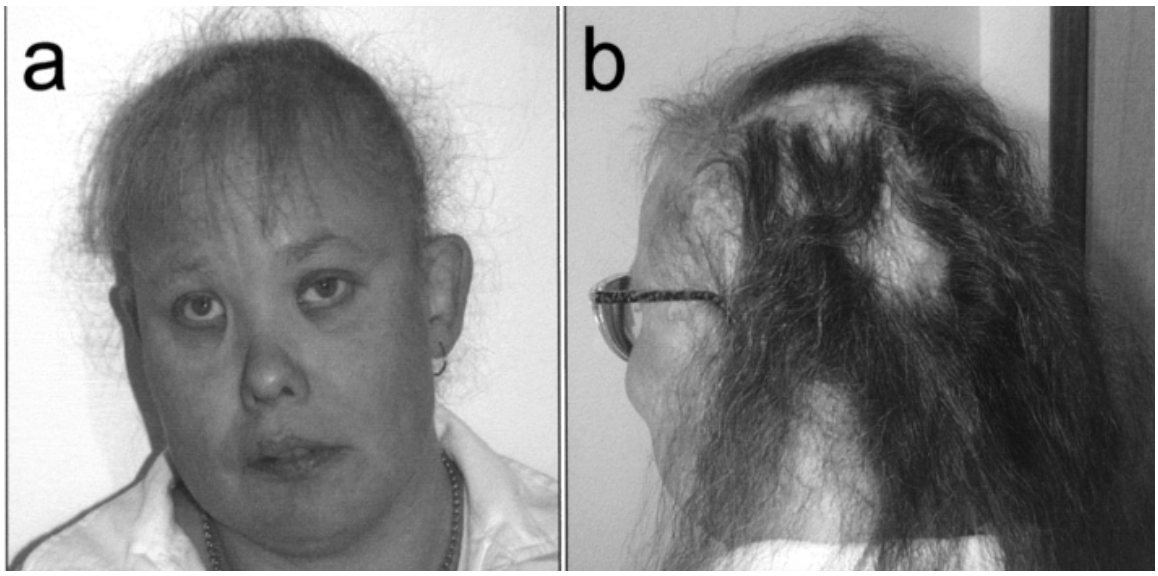


Figure 2.12 The physical abnormalities of a person with CDPX2
(a) Asymmetric facial appearance, showing coarse, sparse hair, undirected growth of eyebrows with sparse lateral growth, a flat nasal bridge, bilateral epicanthus, and strabismus. (b) Cicatricial alopecia and sparse irregular hair growth and premature graying. Figure courtesy of (Whittock *et al.*, 2003).

2.6 Conclusion

By working with the adenosine 2a receptor (a flexible membrane protein) and sterol isomerase (a rigid-bodied membrane protein) together, I wanted to determine whether or not it is equally easy (or difficult) to produce highly-dynamic and rigid human membrane proteins in milligram quantities, purify them to homogeneity without aggregation and grow diffracting crystals. Moreover, over the course of my Ph.D., I wanted to develop new technologies/methodologies that would make it easier to work with my two target proteins (and could be readily adapted for use with other mammalian membrane proteins).

3 Materials and Methods

3.1 Strains

Escherichia coli strain DH5α (Invitrogen) was routinely used for transforming and propagating both the cloning and recombinant expression plasmids. Occasionally, the Dam- *E. coli* strain SCS110 (Stratagene) was used when unmethylated plasmid DNA was required for digestion with a methylation-sensitive restriction enzyme (such as XbaI). All protein production in *E. coli* used the BL21 (DE3) strain (Novagen). Recombinant MP expression was performed in the *Pichia pastoris* strain X33 (Invitrogen).

3.2 Media

E. coli was grown either on LB (1% tryptone, 0.5% yeast extract and 1% NaCl) or LS-LB (1% tryptone, 0.5% yeast extract and 0.5% NaCl) media according to the antibiotic being used for selection. Bacteria containing ampicillin-resistant plasmids were grown on LB, zeocin resistant ones on LS-LB.

P. pastoris strains were routinely cultured using YPD (1% yeast extract, 2% peptone and 2% glucose) medium. For small-scale expression experiments, the yeast transformants were grown on BMGY (1% yeast extract, 2% peptone, 100 mM potassium phosphate pH 6, 1.34% yeast nitrogen base, 4 x 10^{-5}% biotin and 1% glycerol) and then induced using BMMY (same as BMGY except that it contains 0.5% methanol instead of 1% glycerol).

3.3 Expression plasmid construction

The original human adenosine 2a receptor cDNA was obtained from Dr. Timothy Palmer (University of Glasgow). Previously, the N-linked glycosylation site at Asn154 had been removed by mutating this residue to Gln (Fraser, 2006). This

deglycosylated version of the receptor cDNA was denoted dG hA2aR. cDNA of thermostabilized mutants of hA2aR which were conformationally-fixed to preferentially bind either antagonist or agonist, (known as Rant21 and Rag23, respectively) were provided by Dr. Chris Tate (MRC-LMB, Cambridge).

Human sterol isomerase cDNA was purchased from OriGene (US).

hSI and the hA2aR were expressed using modified versions of the basic *Pichia pastoris* expression vectors pPICZA and pPICZαA (both Invitrogen), respectively.

3.3.1 PCR amplification and sub-cloning strategies

Oligonucleotide primers used in all PCR reactions had at least 21 bases complimentary to the template DNA, and had a GC content of *ca.* 40%. Primers were synthesised by either Sigma-Genosys (UK) or Eurofins MWG Operon (Germany). Primers were resuspended in H_2O to a concentration of 50 pmol/μl, and stored at -20°C.

All PCR reactions were performed using KOD HotStart DNA Polymerase (Merck Biosciences) using the appropriate template and primer combinations. Typically, PCR reactions consisted of an initial 2 min Polymerase activation step at 98°C. This was followed by between 25-35 cycles each consisting of a DNA denaturation step (98°C for 20s), primer annealing (55°C for 10s) and DNA extension (70°C for 20s per kb of DNA being amplified). After the last cycling step, the PCR reaction was concluded with a further 10 min at 70°C to ensure all *de novo* DNA synthesis had occurred.

PCR products were run on 1% (w/v) agarose gels alongside a 1kb DNA ladder (Promega). DNA was visualised using SYBR Safe DNA gel stain (Invitrogen). DNA fragments were extracted from the agarose gel using QIAquick Gel Extraction kit (QIAGEN). A-tails were added to all extracted PCR products using Taq polymerase (New England Biolabs) with dATP as substrate prior to ligation into the cloning vector pGEM-T Easy (Promega). The ligation mixture was transformed into DH5α cells by heat shock (42°C for 1 min; 2 min on ice; 1h

shaking at 37°C in LB), plated-out on LB agar plates containing 100 μg/ml ampicillin (Invitrogen) and 100 μg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) (Roche Applied Science), and incubated overnight at 37°C. Colonies containing recombinant vector where white whereas those that contained vector only were blue.

Several putative white colonies for each pGEM construct were grown up in 5 ml of LB containing 100 μg/ml ampicillin overnight at 37°C. Plasmid DNA was isolated from cell cultures using the QIAprep spin miniprep kit (QIAGEN). The identity of recombinants was confirmed first by restriction digestion and then by sequencing (which was carried out at the DNA Sequencing and Services facility, University of Dundee).

At a later stage, DNA fragments were sub-cloned from the pGEM-T Easy vectors into the *P. pastoris* expression plasmids. Ligation reactions were performed using T4 DNA ligase (New England Biolabs). Following transformation into DH5α, the cells were plated-out onto low-salt LB plates containing the antibiotic zeocin (25 μg/ml). pPICZ(α)A plasmid DNA was isolated and charaterised as described for pGEM-T Easy except that the colonies were grown up in LS-LB/Zeocin. Glycerol stocks were prepared for all *E. coli* transformants containing the desired plasmids (both cloning and expression), which were stored at -80°C.

3.3.2 Making a truncated version of dG hA2aR

Truncated dG hA2aR was amplified using the primers GCGCCTGAATTCATGCCCAT CATGGGCTCCTCG and ATCTTAGCGGCCGCACTGGTGCCAGCTGCCTTGAAAGG that contained the restriction sites EcoRI and NotI, respectively (restriction sites are coloured red). The reverse primer truncated the C-terminal tail of the hA2aR at Ser 320. The PCR was carried out as described in Section 3.3.1 using a full-length dG hA2aR cDNA as template (Fraser, 2006). The PCR fragment was sub-cloned into pGEM-T Easy as described in Section 3.3.1.

3.3.3 Construction of the AαHT – TB expression plasmid

A His10-tag followed by a TEV protease recognition site was incorporated into a pPICZαA vector that already contained dG hA2aR immediately upstream of the EcoRI cloning site using the forward primer GAGGCTGAAGCTCATCATCATC ATCATCATCATCATCATCATGAAAATCTGTATTTCCAGAGTGAATTCATGCCCATCATGG GC and reverse primer GCCCATGATGGGCATGAATTCACTCTGGAAATACAGATTTTC ATGATGATGATGATGATGATGATGATGATGAGCTTCAGCCTC (where the hA2aR sequence is in green, the EcoRI site is in red, the TEV site is in purple, the His10 tag is in blue and the pPICZaA sequence is in black).

Furthermore, the biotin acceptor domain from *P. shermanii* was amplified by PCR using the PinpointTM Xa-1 vector (Promega) as template using the primers GCGGCCGCTGAAAATCTGTATTTCCAGAGTGCCGGTAAGGCCGGAGAGGGCGAG and TCTAGATCACCCGATCTTGATGGACCCTGACCGCCCTGCACGGC which contained NotI and XbaI restriction sites, respectively. [Restriction sites are in red, the TEV site is in purple and the biotin acceptor site is in aqua]. The PCR fragment was sub-cloned into pGEM-T Easy (Section 3.3.1) before it was transferred to the AαHT- dG hA2aR expression vector introducing the TEV-biotin acceptor domain between the NotI and XbaI restriction sites.

3.3.4 Creating HT – trunc dG hA2aR – TB

The truncated dG hA2aR cDNA described in 3.3.2 was sub-cloned into the AαHT – TB expression vector detailed in 3.3.3 between the EcoRI and NotI restriction sites.

3.3.5 Creating HT – trunc dG hA2aR – TLB

The HT – trunc dG hA2aR – TB vector was modified by Dr. Niall Fraser introducing SpeI and NdeI restrictions immediately downstream of the TEV protease site and upstream of the biotin acceptor site, respectively.

The iLOV domain was amplified by PCR with the primers GCTTAGACTAGTATAGA GAAGAATTTCGTCATCACTG (forward) and ACGATCCATATGTACATGATCACTTCCAT CGAGCTG (reverse) using a cDNA template gifted to me by Dr. John Christie (University of Glasgow). [Restriction sites are coloured red and ilOV sequence is in pink]. The forward and reverse primers contained SpeI and NdeI restriction sites, respectively. The iLOV PCR product was first sub-cloned into pGEM-T Easy (Section 3.3.1) and then transferred to the modified HT- trunc dG hA2aR – TB expression vector by restriction digestion with SpeI and NdeI.

3.3.6 Creation of GU hA2aR, Rant21 and Rag23 cDNAs

The trunc dG hA2aR cDNA was modified to i) remove a part of IL3 predicted to be disordered, and ii) shorten the truncation so that it was the same as the Rant21/Rag23 constructs. This was accomplished by PCR in several stages using the HT- trunc dG hA2aR -TLB construct as template. PCR reaction 1 was performed using the 5AOX1 primer that binds to the pPICZαA vector, GACTGGTTCCAATTGACAAGC (forward) with the reverse IL3 deletion primer, AGCATGGACCTCCTTCAGCTGTCGTCGCGCCGCCAG). [C-terminal to the IL3 deletion is in orange; N-terminal is blue]. PCR reaction 2 was performed using the forward IL3 deletion primer CGACAGCTGAAGGAGGTCCATGCTGCCAAGTCACT G and reverse reduced C-terminal truncation primer ACGATTTTCAGCGGCCGC TGCCTTGAAAGGTTCTTGCTGCCT). These two PCR products were combined in a third PCR reaction using the forward 5AOX1 and reverse truncation primers to create the Glasgow University (GU) version of the hA2aR.

A Rant21 PCR product was generated in exactly the same way as for the GU construct. Rag23, however, required an additional modification to remove its glycosylation site so it was made by combining three PCR products (instead of two for GU & Rant21). Firstly, the glycosylation site was removed by performing a PCR reaction using the 5AOX forward primer with a reverse primer that overlapped the glycosylation site GGAGTGCGCCTTGCCCTCCTTTGGCTGACCGCA [underline denotes the Asn154Gln change]. PCR reaction 2 was performed using the glycosylation site forward primer GGCAAGGCGCACTCCCAGGGCTGCGGGGAG GGC with the reverse IL3 deletion primer AGCATGGACCTCCTTCGCCTGTCGTCGC

GCCGCCAG. [C-terminal to the IL3 deletion is in orange; N-terminal is blue]. PCR reaction 3 was performed using the forward IL3 deletion primer CGACAGGCGAAGGAGGTCCATGCTGCCAAGTCACTG and the reverse reduced C-terminal truncation primer. These three PCR products were combined in a fourth PCR reaction using the forward 5AOX1 and reverse truncation primers to create the deglycosylated IL3 deleted version of Rag23.

The GU, Rant21 and Rag23 PCR products were first sub-cloned into pGEM-T Easy (Section 3.3.1) and then transferred into the modified HT- TLB expression vector by restriction digestion with EcoRI and NotI.

3.3.7 Creating the -CLBH expression plasmids

The CLBH tag was generated by PCR with the primers GCGGCCGCTGTCCTATTCC AGGGCCCAACTAGTATAGAGAAGAAT (forward) and TTCTAGATCAGTGATGGTGATG GTGATGGTGATGGTGATGCCCGATCTTGATGAGACCCTGACC (reverse). [Key: the C3 protease site is denoted by purple, the His10 tag in blue]. The -CLBH PCR product was first sub-cloned into pGEM-T Easy (Section 3.3.1) and then transferred into the basic pPICZA and pPICZαA vectors by restriction digestion with NotI and XbaI.

GU-, Rant21- and Rag23-CLBH were created by sub-cloning the three hA2aR cDNAs from the HT-TLB vectors into the Aα-CLBH vector by restriction digestion with EcoRI and NotI.

Sterol isomerase was amplified from the cDNA obtained from Origene using the primers GAATTCATGACTACCAACGCGGGCCCCTTG (forward) and GCGGCCGCAGTT CTTCTTGCTCTTGGCTTTTG (reverse). The hSI PCR product was first sub-cloned into pGEM-T Easy (Section 3.2.1) and then transferred into the A-CLBH vector by restriction digestion with EcoRI and NotI.

3.4 Transformation of recombinant expression plasmids into P. pastoris

E. coli that had been previously transformed with the *P. pastoris* expression vectors were grown up overnight from glycerol stocks in 100 ml of LS-LB/Zeocin at 37°C with constant shaking. The cells were pelleted by centrifugation and plasmid DNA isolated using the QIAGEN plasmid Maxi kit. The DNA was then concentrated by precipitating it using ethanol. 10-50 μg of plasmid DNA in a final volume of 20 μl was linearised overnight by digesting with PmeI (New England Biolabs). Linearised plasmid was mixed with 80 μl of competent X33 cells (that had been made as described in the Pichia Easy Select manual (Invitrogen)). After incubating on ice for 5 min, the yeast were transformed by electroporation (2kV, 5 ms). The yeast were allowed to recover for 2h at 30°C in 1 M sorbitol before they were plated out on YPDS plates containing different concentrations of Zeocin (from 0.2 to 2.0 mg/ml). Following a 72 h incubation at 30°C, numerous colonies appeared on the plates. These were re-streaked on fresh YPD plates containing 0.5 mg/ml Zeocin.

3.5 Recombinant MP production in P. pastoris

3.5.1 1 ml cultures in 48-well plates

Clonal selection of yeast that had been transformed with iLOV-containing expression plasmids was performed by measuring whole cell fluorescence from yeast cells that had been grown and induced on a 1 ml scale in a 48-well sterile plate (Section 5.3). Per clonal selection, the expression level was determined for 12 different colonies. For each of the transformants, two 1 ml cultures were established in BMGY. The plate was then shaken for 24 h at 30°C. The cells were pelleted by centrifugation (1500 g, 5 min), and then resuspended in BMMY. The plate was returned to the shaking incubator for a further 24 h at 22°C. After this, cells were harvested by centrifugation and washed twice with distilled water before being resuspended in 1 ml of distilled water. 20 μl of the cells from every culture were transferred to a 96-well black plastic plate (Corning) and

made upto to a final volume of 200 μl with distilled water for determining whole cell fluorescence levels. Glycerol stocks were prepared for the transformant found to have the highest expression level of the target protein by mixing equal volumes of an overnight culture grown in YPD with sterile 50 % (v/v) glycerol in a sterile cryovial. The cultures were snap-frozen using liquid nitrogen and stored at –80°C.

3.5.2 Small-scale (10 ml) expression cultures

Clonal selection of yeast transformed with αHT – trunc dG hA2aR – TB and αHT – full-length dG hA2aR –TB was performed by single-point radioligand binding analysis on membranes isolated from cells grown and induced on a 10 ml scale (Section 4.3.2). Individual transformants were grown up in 10 ml of BMGY in sterile 50 ml falcon tubes at 30°C for 24h. The cells were spun down (1500 g, 5 min) and re-suspended in 10 ml of BMMY, and grown for a further 24 h at 22°C. Cells were harvested by centrifugation (4000g, 5min, 4°C), and membranes made as described in Section 3.6.2.

3.5.3 Medium-scale (100 ml) expression cultures

Yeast transformed with αHT- dG hA2aR – TB and αHT- trunc dG hA2aR – TB were grown at 30°C in 500 ml shake-flasks containing 100 ml of BMGY until the cultures had an OD_{600} of between 2-6. At this point, receptor production was induced by adding 0.5% methanol, reducing the temperature to 22°C and adding 10 mM theophylline. 0.5% methanol was also added 12 and 24 h post-induction. Cells were harvested by centrifugation (5000g, 20 min, 4°C) after 36 h. Membranes were made as described in Section 3.6.2 which were later used in radioligand binding assays for assessing the pharmacological properties of the expressed receptors (Section 4.4).

3.5.4 Recombinant MP production in fermenter culture

All fermentations were carried out using an Applikon® 7L fermentor vessel and ADI 1025 bioreactor system. Dissolved oxygen (dO_2) was maintained above 25% through feedback of agitation (minimum 1000 r.p.m.) and compressed air (0.8 litre min^{-1}). Oxygen (1 bar pressure) was automatically drawn in as necessary to help maintain the dO_2 setpoint during growth at high-cell densities. The pH was maintained at around 5 through addition of 23% NH_4OH. Antifoam A (Fluka) was added drop-wise as required to reduce foaming during culture growth.

4 litres of basal salts media containing 4% glycerol (Table 3.1) was autoclaved within the glass fermentation vessel. To this, 15 ml of PTM_1 trace salts was added (Table 3.1). The media was inoculated with 200 ml of yeast that had been grown up overnight in YPD so that the initial fermenter culture had an OD_{600} of 0.5-1.0. Cultures were grown on basal salts media for 18 h at 30°C until all of the original glycerol had been consumed. Following this, the culture was feed 50% glycerol (v/v) for 11 h at a rate of 1.09 ml/min. The yeast were then starved for 1 h to ensure that any residual glycerol had been consumed before the culture was adapted for growth on methanol (4 h at 0.07 ml/min followed 12 h at 0.09 ml/min methanol). Target MP production was induced by reducing the culture temperature to 22°C. [When hA2aR was being made in fermenter culture, 50 mM theophylline powder was added at this stage]. Methanol feeding continued for a further 12 h at 0.13 ml/min followed by 36 h at 0.33 ml/min. Samples were taken at regular timepoint intervals to assess culture growth by measuring OD_{600} and/or wet cell weight. Recombinant MP production levels were determined either by single-point radioligand binding analysis or measuring iLOV fluorescence. After 80 – 96 h cultivation, the fermentations were stopped and the cells harvested by centrifugation (5000g, 20min, 4°C). The pellets were stored at -20°C, and used as they were required.

	Components	Amount per litre
Basal salts medium	Phosphoric Acid (85%)	(26.7 ml)
	Calcium Sulphate	0.93 g
	Potassium Sulphate	18.2 g
	Magnesium Sulphate – $7H_20$	14.9 g
	Potassium Hydroxide	4.13g
	Glycerol	40.0 g
	Water	to 1 litre
		Sterilised by autoclaving
PTM_1 trace salts	Cupric Sulphate – $5H_2O$	6.0 g
	Sodium Iodide	0.08 g
	Magnesium Sulphate – H_2O	3.0 g
	Sodium Molybdate – $2H_2O$	0.2 g
	Boric Acid	0.02 g
	Cobalt Chloride	0.5 g
	Zinc Chloride	20.0 g
	Ferrous Sulphate – $7H_2O$	65.0 g
	Biotin	0.2 g
	Sulphuric Acid	5.0 ml
	Water	to a final volume of 1 litre
		Sterilised by filtration

Table 3.1 Recipes for fermentation media and trace feed solution

4 litres of basal salts media was prepared and autoclaved immediately prior to commencing a fermentation. PTM_1 trace salts were made and stored in 50 ml aliquots at 4°C, and used as it was required.

3.6 Cell lysis and membrane preparation

3.6.1 Chemical cell lysis

After the iLOV fluorescence intensity had been measured in intact cells, the cells from the 1 ml yeast cultures in the 48 well plates were re-pelleted by centrifugation (1500g, 5 min) and the water discarded. The cells were lysed using 100 μl of Yeast Buster™ Protein Extraction Reagent (Merck) and 10 μl Benzonase nuclease (Merck). To ensure complete rupturing the cells were placed in a shaking incubator at 30°C for 30 min. Following this, 1 ml of distilled water was added to each well before the plate was centrifuged (1500g, 5 min) to pellet cell debris. The supernatants containing the cell contents were removed and their protein concentration determined by BCA assay (Pierce) using BSA as standard.

3.6.2 Small-scale membrane preparation

10 ml cell pellets (Section 3.5.2) were resuspended in Breaking Buffer (4 mg/ml BSA, 5 mM EDTA, 150 mM NaCl in 50 mM Tris HCl pH 7.4). An equal volume of glass beads was added and the cells were broken by vortexing (5 x 1 min). The beads were separated by filtration through a chromatography column and the flow-through containing membranes was subjected to centrifugation. A low speed spin (10,000g, 10 min) was used to pellet cell debris and unbroken cells. The supernatant was retained and centrifuged at high speed (150,000g, 30 min) to pellet membranes. Membrane pellets were resuspended in 1 ml of 150 mM NaCl in 50 mM Tris HCl pH 7.4, and the protein concentration determined by BCA assay (Pierce) using bovine serum albumin (BSA) as standard.

3.6.3 Large-scale membrane preparation

Cell pellets from fermentation cultures were thawed overnight at 4°C and resuspended in Breaking Buffer (5 mM EDTA, 150 mM NaCl in 50 mM Tris HCl pH 7.4). Homogenised cells were passed through a Constant Systems cell disrupter three times at a pressure of 30 kPsi. Following this, 10% glycerol (v/v) and protease inhibitors (phenylmethanesulfonylfluoride (PMSF) and benzamidine) were added to the ruptured cells. Membranes were then isolated by centrifugation as described in Section 3.5.2. Membrane pellets were resuspended in a Membrane Resuspension Buffer consisting of 10% glycerol and 150 mM NaCl in either 50 mM Tris HCl, HEPES or sodium phosphate buffer at a pH of 7.4). The protein concentration of the membranes was determined by BCA assay (Pierce) using BSA as standard. Membranes were diluted with the appropriate Membrane Resuspension Buffer so that they had a final concentration of 20 mg/ml. These were stored at -20°C in 50 ml aliquots.

3.7 Measuring iLOV fluorescence

3.7.1 iLOV excitation and emission spectra

Excitation and emission spectra for pure iLOV (a gift from Dr. John Christie, University of Glasgow) were recorded and compared against those obtained for flavin mononucleotide (FMN) and *P. pastoris* X33 membranes. Measurements were made in quartz cuvettes (pathlength 1 cm) using a Perkin-Elmer L550B spectrophotometer. Fluorescence was measured at right angles to the excitation beam. Excitation spectra were recorded from 300 – 460 nm measuring fluorescence emission at 485 nm. Emission spectra were recorded between 470 and 600 nm, exciting the samples at 450 nm.

3.7.2 Measuring iLOV fluorescence in 96 well plates

The iLOV fluorescence of individual samples (e.g. whole cells or FSEC elution fractions) was determined using a FLUOstar OPTIMA plate-reader. 200 μl of each sample was added to a single well in a black Corning 96-well plate. Samples were excited at 450 nm and iLOV fluorescence emission measured at 520 nm. Results were analysed using the OPTIMA data analysis software. Fluorescence readings from intact yeast cells were used in combination with protein concentration measurements to determine expression levels in clonal selection experiments.

3.8 Radioligand binding analysis

All radioligand binding assays on 'wild-type' hA2aR were carried out using the potent, selective A2a antagonist [^{3}H] ZM241385 (American Radiolabelled Chemicals). Assays on Rag23 were performed using the agonist [^{3}H] NECA (Amersham Biosciences). The same Binding Buffer (5 mM EDTA in 50 mM Tris HCl pH 7.4) was used with both radioligands. All measurements were made in quadruplicate.

3.8.1 [^{3}H] ZM241385 binding

Saturation and single-point binding assays on membrane-bound hA2aR contained between 1-10 μg protein in a final volume of 500μl. 0.1U of adenosine deaminase (Roche Applied Science) was also added to each cocktail in order to degrade any adenosine released from the membranes. Saturation binding was performed by incubating [^{3}H] ZM241385 (final concentration 0.02 – 12 nM) with receptor for 1 h at 30°C. Non-specific binding was determined in the presence of 10 mM theophylline. Following incubation, bound and free radioligand were separated by sucking the suspension onto glass fibre filters (Whatman) pre-soaked in 0.3% polyethyleneimine under vacuum using a Brandel Harvester. The filters were washed three times with 50 mM Tris HCl pH 7.4 before they were

counted using Ultim Gold™ XR scintillation fluid (Perkin Elmer). The counts in dpm were converted to fmol of [^{3}H] ZM241385 bound using the specific activity of the radioligand. Specific binding was determined by subtracting the binding measured in the presence of theophylline (non-specific) away from those measured in buffer only (total). Binding data was analysed using the computer package GraphPad Prism. Curves of total, non-specific and specific binding against [radioligand] (nM) were plotted and fitted using a non-linear regression. The curve fit gave values for the maximum number of binding sites occupied by the ligand (B_{max}) as well as the receptor's equilibration dissociation constant (K_d) for [^{3}H] ZM241385. The K_d is the concentration of radioligand required to occupy half of the available binding sites and, therefore, is a measure of the affinity of the ligand for the receptor. The B_{max} is a reliable estimate of the receptor's expression level.

Single-point binding assays on membrane bound hA2aR had a final volume of 200 μl, contained between 1-10 μg of protein and *ca.* 10 nM [^{3}H] ZM241385. Non-specific binding was determined in the presence of 10 mM theophylline.

3.8.2 [^{3}H] NECA binding

Saturation binding using [^{3}H] NECA was performed on membranes containing Rag23 in the same way as that described for [^{3}H] ZM241385 with the following differences: i) the assay cocktail had a volume of 100 (instead of 500) μl, ii) [^{3}H] NECA was used at considerably higher concentrations (20 – 500 nM), and iii) adenosine deaminase was omitted as it metabolises NECA rendering the radioligand unable to bind to the receptor.

3.9 SDS-PAGE

20 μl of each protein sample to be analysed by SDS-PAGE was mixed with 5 μl of NuPAGE® LDS Sample Buffer (Invitrogen) and heated for 10 min at 60°C. Denatured samples were loaded onto a 12-well 10% NuPAGE® Bis–Tris pre-cast gel (Invitrogen) along with an aliquot of SeeBlue Plus2 pre-stained marker. The

gel was run in MES buffer according to the manufacturer's instructions (200V for 40-50 min). Protein bands were visualised by staining with SimplyBlue™ SafeStain (Invitrogen).

3.10 Detergent screening and fluorescence size-exclusion chromatography (FSEC)

2 ml of yeast membranes (at a concentration of 20 mg/ml) containing either HT-Rag23-TLB or SI-CLBH were solubilised using the detergents DDM, DM, NM and βOG as well as the detergent-lipid mixture DDM/CHS. 0.6 ml of 10% detergent stock solutions were added to the membranes and the mixtures stirred for 1 h at 4°C. Samples were then centrifuged at high speed (50,000g, 10 min) to pellet any unsolubilised material. The supernatant containing the iLOV-tagged MP was filtered before 500 μl of it was injected onto a Superose™ 6 FLPC gel filtration column equilibrated in buffer (either Tris, HEPES or sodium phosphate – all pH 7.4) containing 5% glycerol and an appropriate concentration of the detergent being studied (0.05% DDM, 0.15% DM, 0.3% NM and 0.8% βOG). The column was run at a rate of 0.3 ml/min, and 0.6ml fractions were collected. 200 μl of each fraction was transfered to a black Corning 96-well plate and the intensity of iLOV fluorescence measured using a platereader. Following this, FSEC profiles were plotted.

3.11 Expression and purification of accessory enzymes

3.11.1 Tobacco Etch Virus (TEV) protease

BL21(DE3) cells that had been previously transformed with pET15b-TEV protease (a gift from Dr. Alan Riboldi-Tunnicliffe, University of Glasgow) were grown overnight from a glycerol stock in 5 ml of LB media at 37°C. The next morning, the cells were added to 500 ml of fresh media and grown to an OD_{600} of 0.7. TEV protease expression was induced by adding 1 mM IPTG. The temperature was

reduced to 22°C, and the cells shaken for a further 16 h. Cells were harvested by centrifugation (4000g, 20 min) and then resuspended in lysis buffer (150 mM NaCl, 25 mM imidazole in 50mM Tris HCl pH 8.0). Cells were broken using a Constant Systems cell disruptor (3 passages at 20 kPsi). The broken cells were centrifuged at 10,000g and the supernatant containing TEV protease passed over a 5 ml Nickel column that had been equilibrated in lysis buffer. The column was washed extensively with lysis buffer to remove any contaminant proteins before the TEV protease was released from the column using an imidazole (0 – 800 mM) gradient in 150 mM NaCl, 50 mM Tris HCl pH 8.0. DTT and EDTA were added to the elution fractions (both to a final concentration of 1 mM) to prevent the TEV protease precipitating out of solution. Fractions containing TEV protease were pooled and desalted into 150 mM NaCl in 50 mM Tris HCl pH 7.4 using a PD10 column. TEV protease was aliquoted (1 ml samples with a protein concentration of 5 mg/ml) before being snap-frozen with liquid nitrogen and stored at -20°C.

3.11.2 Human Rhinovirus C3 protease

BL21(DE3) cells that had been previously transformed with pET15b-C3 protease (a gift from Dr. Alan Riboldi-Tunnicliffe, University of Glasgow) were grown overnight from a glycerol stock in 5 ml of LB media at 37°C. Expression and Ni^{2+}-affinity purification of C3 protease was performed as described for TEV protease (Section 3.11.1). An additional gel filtration step (Sepharose C6 column equilibrated in 150 mM NaCl, 50 mM Tris HCl pH 7.4) was then performed to improve enzyme purity.

3.11.3 Biotin ligase

The biotin ligase (BirA) gene was amplified from *E. coli* strain 0157 genomic DNA using the primers TGACCACATATGATGAAGGATAACACCGTGCCACTG and CGTAGAGGATCCTTATTTTTCTGCACTACGCAGGGA. The PCR product was introduced into pGem T-Easy as described in Section 3.3.1 before it was sub-cloned into NdeI/BamHI-digested pET15b vector. Correct recombinant plasmids were identified by restriction analysis and sequencing. The BirA expression plasmid was transformed into BL21 (DE3) cells and expressed and purified as

described in Section 3.11.1. The only exception being that no DTT and EDTA were added to the elution fractions.

3.12 Purification of HT – trunc dG hA2aR – T(L)B

The basic buffer (HT-TLB buffer) used to purify HT – trunc dG hA2aR – T(L)B was 150 mM NaCl, 10% glycerol in 50 mM Tris HCl pH 8, 0.05% DDM, 0.01% CHS.

3.12.1 Membrane solubilisation

Membranes containing either HT – trunc dG hA2aR – TB or HT – trunc dG hA2aR – TLB at a concentration of 20 mg/ml were solubilised by adding a 10:2 DDM:CHS stock solution to the membranes so that the final DDM concentration was 2.5%. The mixture was stirred for 2 h at 4°C. Following this, the sample was centrifuged at 150,000g for 30 min. The supernatant containing the detergent-solubilised receptor was added to an equal volume of HT-TLB buffer prior to the Ni^{2+}-affinity chromatography step.

3.12.2 Ni^{2+}-affinity purification of HT– trunc dG hA2aR –T(L)B

A 5ml Ni^{2+} column was equilibrated in HT-TLB buffer containing 75 mM imidazole. The supernatant containing receptor was then passed over the column three times. The Ni^{2+} column was washed with HT-TLB buffer containing 500 mM (instead of 150) NaCl and 75 mM imidazole. Receptor was eluted from the column using 500 mM imidazole in HT-TLB buffer. Elution fractions were analysed by SDS-PAGE (Section 3.9)

3.12.3 Biotinylation and purification using streptavidin

In vitro biotinylation of Ni^{2+}-affinity purified HT-trunc dG hA2aR – TLB was performed by adding 1 mg of pure BirA to the receptor along with 50 μM biotin, 20 μM ATP and 10 mM Mg-o-acetate. The mixture was stirred continuously

overnight at 4°C. The next morning the receptor protein was doubly desalted into HT-TLB buffer using PD-10 columns (GE Healthcare) to remove free biotin. Streptavidin resin (Pierce) was equilibrated with HT-TLB buffer. Biotinylated receptor was added to the streptavidin slurry and mixed by rotating on a wheel at 22°C for 4 h. The sample was centrifuged (4500 g, 10 min) to pellet the resin. The supernatant containing contaminant proteins was discarded. The beads were washed extensively with HT-TLB buffer to ensure that any proteins that had weakly bound to the beads were removed.

3.12.4 Tag removal using TEV protease

Tag removal from biotinylated HT – trunc dG hA2aR – TLB bound to streptavidin beads was performed by adding 5 mg of TEV protease to the resin in HT-TLB buffer containing 0.5 mM EDTA, 3 mM reduced glutathione and 0.3 mM oxidised glutathione. The slurry was stirred slowly overnight at 4°C. The next morning, the sample was centrifuged (4500g, 10 min), and the supernatant reverse-purified with Ni^{2+} beads to remove the HT tag and TEV protease. The cleaned-up sample was retained for SDS-PAGE analysis.

3.12.5 Gel filtration of enriched HT – trunc dG hA2aR – TB

Ni^{2+}-affinity purified HT – trunc dG hA2aR – TB was concentrated to a volume of *ca.* 500 μl and injected onto a Superose™ 6 FLPC gel filtration column equilibrated in HT-TLB buffer. The column was run at 0.3 ml/min and 1 ml fractions collected. The elution profile (A_{280} v time) was recorded. Peak fractions were retained for analysis by SDS-PAGE.

3.13 Purification of Rag23 in DDM

Membranes containing Rag23-CLBH were resuspended in 100 mM NaCl, 10% glycerol in 10 mM HEPES pH 7.4 (Rag buffer) at a concentration of 20 mg/ml. Membranes were stirred with 100 μM NECA for 30 min at 4°C before DDM was added from a 10% (w/v) stock solution so that the final detergent concentration

was 2.5%. The solubilisation mixture was stirred for a further 2 h before it was centrifuged at 150,000g for 30 min. The supernatant containing Rag23-CLBH was diluted with an equal volume of Rag buffer containing 0.05% DDM. A 5 ml Ni^{2+} charged His-trap column was equilibrated in Rag buffer containing 75 mM imidazole and 0.05% DDM. The detergent-solubilised material was batch-loaded onto the column. This process was repeated three times. The column was washed with Rag buffer containing 75 mM imidazole and 0.05% DDM. Rag23-CLBH was eluted from the column using Rag buffer containing 500 mM imidazole and 0.05% DDM. Aliquots of those fractions that were fluorescent (i.e contained iLOV-tagged receptor) were retained for analysis by SDS-PAGE. The remainder of the material was pooled and desalted into Rag buffer containing 0.05% DDM using PD-10 columns. 20 μl of C3 protease was added per ml of partially-purified Rag23-CLBH along with 1 mM DTT and 1 mM EDTA. The cleavage cocktail was mixed overnight using a rotating wheel at 4°C. The next morning, Ni^{2+} beads equilibrated in Rag buffer containing 0.05% DDM was added to the cleavage mixture. The sample was placed back on the wheel for 30 min before it was centrifuged at 3000g for 5 min. This process was repeated until the supernatant was non-fluorescent. At this point, Rag23 was expected to be largely pure. The supernatant was concentrated to a volume of *ca.* 500 μl and applied to a Superose 6 column on an AKTA prime plus system (GE Healthcare) that had been equilibrated in Rag buffer containing 0.05% DDM. The column was run with a flow rate of 0.3 ml/min and 1 ml fractions were collected (which were later analysed by SDS-PAGE).

3.14 Purification of SI in DDM

SI was purified in DDM as described for Rag23 with the following exceptions. The basic purification buffer contained 150 and not 100 mM NaCl. The membranes were not pre-incubated with NECA prior to detergent solubilisation. The wash buffer for the Ni^{2+} column contained 500 and not 100 mM NaCl. Following SDS-PAGE analysis of pure SI, the major band on the gel was excised for protein identification by tandem mass spectroscopy (MS-MS). This work was performed at the Proteomics Facility at the University of Dundee. The protein

concentration of pure SI was determined by measuring the A_{280} of the sample and dividing this number by 2.8 (which is the theoretical extinction coeffiecient at 280 nm for SI as determined by its amino acid content). This value was obtained using the ProtParam tool on the ExPASy Proteomics Server (http://www.expasy.ch/tools/protparam.html).

3.15 Purification of SI in DM

The purification of SI in DM was the same as that described in Section 3.14 with the following exceptions. The basic purification buffer was 150 mM NaCl, 10% glycerol in 50 mM sodium phosphate pH 7.4. The membranes were solubilised in DM and not DDM. All of the purification buffers contained 0.15% DM instead of 0.05% DDM. The gel-filtration column was equilibrated with and run in 150 mM NaCl, 10% glycerol in 50 mM sodium phosphate pH 6, 0.15% DM.

3.16 Circular dichroism

All CD measurements were made at the Biophysical Characterisation Unit, University of Glasgow by Dr. Sharon Kelly.

3.16.1 CD characterisation of SI

Far- (region 190 - 260 nm) and near-UV (250 – 320 nm) circular dichroism spectra of pure SI were recorded on a Jasco J-810 spectropolarimeter using quartz cuvettes with 0.02 and 0.5 cm pathlengths, respectively. The SI sample had a protein concentration of 0.6 mg/ml. The far-UV CD data was fitted using the CDSSTR alogorithm (Johnson, 1999) to give an estimate of the secondary structure content of SI.

3.16.2 Thermostability analysis of SI by CD

A series of far-UV CD spectra of pure SI were recorded between 5 and 80°C at 5°C intervals. Measurements were made in a cuvette with a pathlength of 0.2 cm

using protein with a concentration of 0.2 mg/ml protein. The extent of SI unfolding with temperature was determined by plotting the the size of the CD signal at wavelengths 209 and 222 nm as a % of the original values prior to heat denaturation.

3.17 CPM assay

A stock solution of CPM reagent (7-Diethylamino-3-(4-maleimidophenyl)-4-methylcoumarin) (Sigma-Aldrich) was made by first dissolving it in dimethyl sulfoxide (DMSO) to a concentration of 25 mg/ml and then further diluting it with H_2O to a final concentration of 4 mg/ml. Care was taken to prevent photobleaching of the CPM dye by storing it in the dark. The CPM thermal denaturation assays were performed in black Corning 96-well plates. For each condition investigated, protein and buffer control measurements were made in triplicate. Each protein well contained 10 μl of pure SI (0.5 mg/ml), 20 μl of a 10x buffer solution and 20 μl of either a 10x additive (such as a salt) or H_2O. Plates were incubated at room temperature for 5 min before 140 μl of 10% SDS was added to each well followed by 10 μl of the stock CPM reagent. Control samples were set-up in a similar fashion except that the protein aliquot was replaced with 10 μl of buffer. The plate was immediately transferred to a fluorescence platereader set at a temperature of 45°C. Fluorescence intensity readings were taken at regular intervals over a period of 3 h. Baseline fluorescence recorded in the buffer controls were subtracted from the values recorded in the protein wells, before the three difference values for each condition were averaged. Curves showing SI denaturation were plotted as an increase in fluorescence over time. To allow comparison of the extent of SI unfolding in different conditions, the maximum amount of unfolding was defined as the fluorescence intensity at t = 8000 s.

3.18 Crystallisation trials with SI

3.18.1 Crystallisation of SI from detergent solution

Pure SI (in both DDM and DM) was concentrated to *ca.* 10 mg/ml. SI purified in HEPES buffer containing DDM was used with the MemStart, MemSys and PegIon crystallisation screens (Table 3.2). SI purified in sodium phosphate buffer containing DM was used with the MemGold screen (Table 3.2). A Cartesian HoneyBee nanoliter dispensing system (Genomic Solutions Ltd) was used to add 500 nl of the protein solution to 500 nl of the reservoir in each of the wells in the 96 well plate. The plates were stored in a Rhombix Crystal Plate Imager (Thermo Scientific) at 20°C and trays monitored automatically over a 3 month period. Possible crystal hits were followed-up manually by microscope. Crystals obtained in the initial screens were tested for X-ray diffraction using the in-house X-ray facility at Glasgow (consisting of a Rigaku MicroMax 007 rotating anode generator coupled with MarResearch 345 image plate detector).

Screen	Manufacturer
MemGold	Molecular Dimensions
MemStart	Molecular Dimensions
MemSys	Molecular Dimensions
Peg/Ion 1 & 2	Hampton Research

Table 3.2 Commercially-available screens used for crystallising SI in detergent solution

3.18.2 Lipidic-sponge phase preparation

Cholesterol doped sponge phases were prepared by dissolving either 0.5 or 5% (w/w) of cholesterol directly into melted mono-olein (MO) (1-mono-oleyl-rac-glycerol) (Nu-Chek Prep). The molten MO/cholesterol mixture was weighed into small glass vials along with the appropriate amounts of either PEG400 or PEG1500 with 100 mM of a buffer solution (with or without 100 mM of a salt). The compositions of all the sponges are listed in Table 6.1. The vials were sealed tight and incubated on a roller for one day at 37°C. In cases where the sample was not homogenous after this treatment, it was heated to 50°C for 2-3 h and then re-incubated at 37°C. Afterwards, the samples were centrifuged (20°C, 5500g, 30 min) and left until a transparent phase was observed.

3.18.3 Crystallisation of SI in the lipidic-sponge phase

Pure SI in DM was concentrated to *ca.* 10 mg/ml and then dialysed for 2h against sodium phosphate buffer containing 0.15% DM using mini-dialysis cassettes (Pierce). Trays were set-up as a series of hanging drops using EasyXtal Tool 24-well crystal trays (QIAGEN). 1 μl of each sponge was pipetted onto the screw-in crystallisation support before 1 μl of the protein solution was added on top. The crystallization supports were screwed into the reservoir chambers which all contained 500 μl of reservoir solution (0.1 g l^{-1} Na_3PO_4, 0.55 M sodium acetate and 0.75 M HEPES pH 6.3). Trays were stored in a 20°C incubator and checked regularly for crystals.

4 Creating a generic system for the over-expression and purification of human membrane proteins

4.1 Introduction

Although there are a few exceptions (e.g. rhodopsin), most mammalian membrane proteins have a low natural abundance. Therefore, the first step towards solving the structure of any mammalian membrane protein usually involves devising a robust methodology for making milligram quantities of the target protein in a heterologous system. The functional over-expression of membrane proteins is a non-trivial issue, and is one of the key "bottlenecks" in determining the structure of a membrane protein.

The methylotrophic yeast *Pichia pastoris* is commonly used for producing soluble proteins of mammalian origin. There are several reasons for this including the yeast's capacity to perform a range of higher-eukaryotic post-translational modifications as well as its ability to grow to high cell densities (*ca.* 10^9 cells/ml) on cheap media in fermenter culture. Recombinant protein production is driven from the strong, methanol-inducible AOX1 promoter (Macauley-Patrick *et al*. 2005; Hartner and Glieder 2006), and can result in the production of tens or even hundreds of milligrams of target protein per litre of culture (Cregg *et al*. 2000; Cereghino *et al*. 2002). It has also been recently shown that a diverse range of mammalian membrane proteins, including G-protein coupled receptors, transporters and ion channels, can be successfully made in *P. pastoris* (Weiss *et al*. 1998; Weiss *et al*. 1998; Newton-Vinson *et al*. 2000; Andre *et al*. 2006; Horsefield *et al*. 2008; Aller *et al*. 2009; Ho *et al*. 2009; Tao *et al*. 2009; Alisio and Mueckler). However, a potential limitation of *P. pastoris* as a host for making mammalian MPs is its inability to make cholesterol. Instead *P. pastoris*

synthesises another type of sterol called ergosterol which, despite being highly similar to cholesterol, is unable to stabilise mammalian MPs in the ways that cholesterol does (Opekarova and Tanner 2003).

Even when a heterologous host as powerful as *P. pastoris* is used, the expression conditions for each target membrane protein must be optimised so that the maximal yield of protein possible is obtained (Singh *et al*. 2008). The amount of functional GPCR obtained in *P. pastoris* can be increased by reducing the induction temperature from 30 to 22°C (Andre *et al*. 2006; Fraser 2006) as well as adding either DMSO (which stimulates membrane synthesis) (Andre *et al*. 2006) or a ligand (e.g. an antagonist) that stabilises the target protein (Weiss *et al*. 1998; Andre *et al*. 2006; Fraser 2006). With *P. pastoris*, a dramatic increase in protein production level occurs when expression is performed in a fermenter as compared to shake flask-culture. Careful optimisation of the methanol feed rate is essential to ensure that there is sufficient methanol present in the fermenter to drive maximal protein production without accumulating excess levels of methanol (which is toxic to the yeast at high (>4%) concentrations) (Cereghino *et al*. 2002).

Once the target MP has been successfully expressed in milligram quantities, it must be purified to homogeneity before crystallisation trials can be started. MPs can be readily purified by affinity chromatography using tags such as His10 (which binds to Ni^{2+} resin) and biotin (which binds to (strept)avidin) (Terpe 2003). Tags can be added to a recombinant protein at either its N- or C-terminus by manipulating the DNA sequence of the expression vector using PCR. When working with MPs, however, it is imperative to determine the expression level of the recombinant protein with and without tags to ensure that addition of the tag has not affected the amount or functionality of the protein produced by the host cell.

Previous work in our laboratory focused on producing the human adenosine 2a receptor (hA2aR) in shake-flask cultures of *P. pastoris* (Fraser 2006). A full-length cDNA encoding the human adenosine 2a receptor was cloned into a modified version of the *P. pastoris* vector pPICZαA. The basic expression plasmid

obtained from Invitrogen had been altered by inserting a His10 tag between the Kex2 protease cleavage site and the EcoRI cloning site (Fraser 2006) (Figure 4.1). hA2aR was made as a fusion with the α-factor leader sequence from *S. cerevisiae* (Figure 4.1). When fused to the N-terminus of GPCRs expressed in *P. pastoris*, the α-factor leader sequence has been shown to increase the yield of functional receptor made by the yeast (Grunewald *et al*. 2004). The α-factor leader sequence is removed from the receptor following cleavage by the yeast protease Kex2 (which resides in the golgi complex) (Julius *et al*. 1984; Julius *et al*. 1984; Redding *et al*. 1991). The modified expression plasmid described in (Fraser 2006) was the starting point of all the work described in this thesis.

Figure 4.1 Schematic diagram of the initial hA2aR expression construct (αH – dG hA2aR)
Previously, a deglycosylated version (dG) of a full-length hA2aR cDNA had been cloned into a modified version of the *P. pastoris* expression vector pPICZαA (Fraser, 2006). The hA2aR is made as a fusion protein with the α-factor leader sequence from *S. cerevisiae*. The leader sequence is later cleaved off endogenously by the protease Kex2 resulting in the yeast making dG hA2aR with a His10 tag at its N-terminus.

4.2 Identifying and removing the disordered C-terminal tail of the hA2aR

Signalling proteins, including receptors, are involved in molecular information transfer. For these proteins to be able to function, they must be able to adopt multiple distinct conformations (Kobilka and Deupi 2007). This is accomplished through the proteins having great inherent flexibility and it is not uncommon for signalling proteins to contain natively-disordered regions (Gsponer and Babu 2009). To identify those regions of the hA2aR that are highly flexible its amino acid sequence was analysed using the disorder prediction program RONN (www.strubi.ox.ac.uk/RONN). The results suggest that the C-terminal region

(αα 347 – 412) of the receptor is largely disordered as well as part of the third intracellular loop (αα 206 – 228) (Figure 4.2).

The disordered regions that were identified are likely to hinder the formation of lattice contacts necessary for crystal formation (Lundstrom 2005; Warne *et al.* 2009). As the receptor's third intracellular loop is involved in G-protein activation (and necessary for function), it was decided at this stage to retain this part of the protein. In contrast, others have shown that the last 100 ααs of the hA2aR can be removed without substantially altering the ability of the receptor to activate G_s (Piersen *et al.* 1994; Olah 1997; Klinger *et al.* 2002). Therefore, a truncated hA2aR mutant (αα 1 – 320) in which the C-terminal region had been largely removed was created by PCR amplification and sub-cloned into pGEM T-Easy (Figure 4.3). The truncated version of the hA2aR (with the C-terminal disordered region removed) was thought more likely to crystallise than the wild-type receptor.

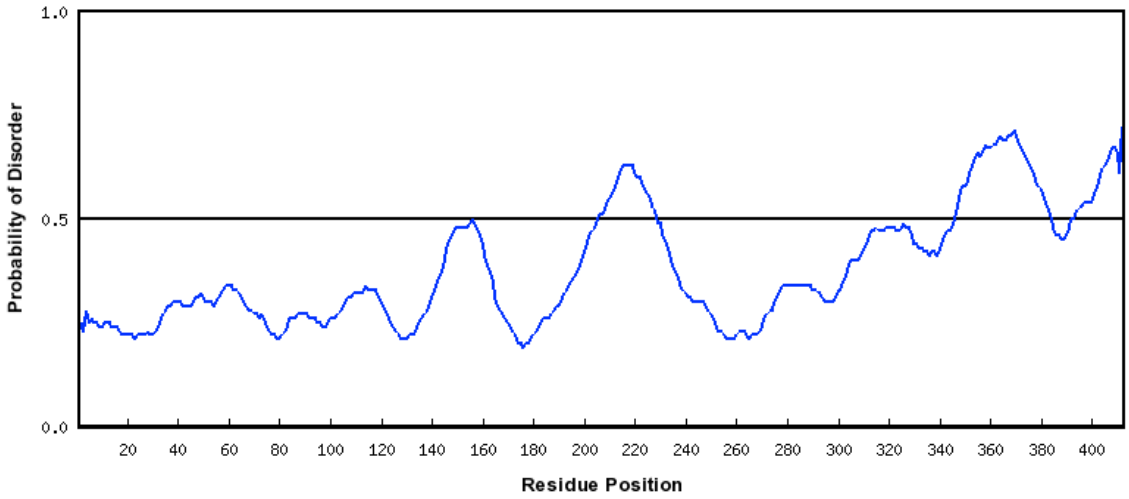


Figure 4.2 RONN disorder prediction plot for the human adenosine 2a receptor

The prediction software analyses the primary sequence of the target protein to try to identify those regions with high flexibility. Residues with a "probability of disorder" value >0.5 (indicated by the solid black line) are predicted to be natively disordered. The plot suggests that the C-terminal region (αα 347 – 412) of the receptor is largely disordered as well as part of the third intracellular loop (αα 206 – 228).

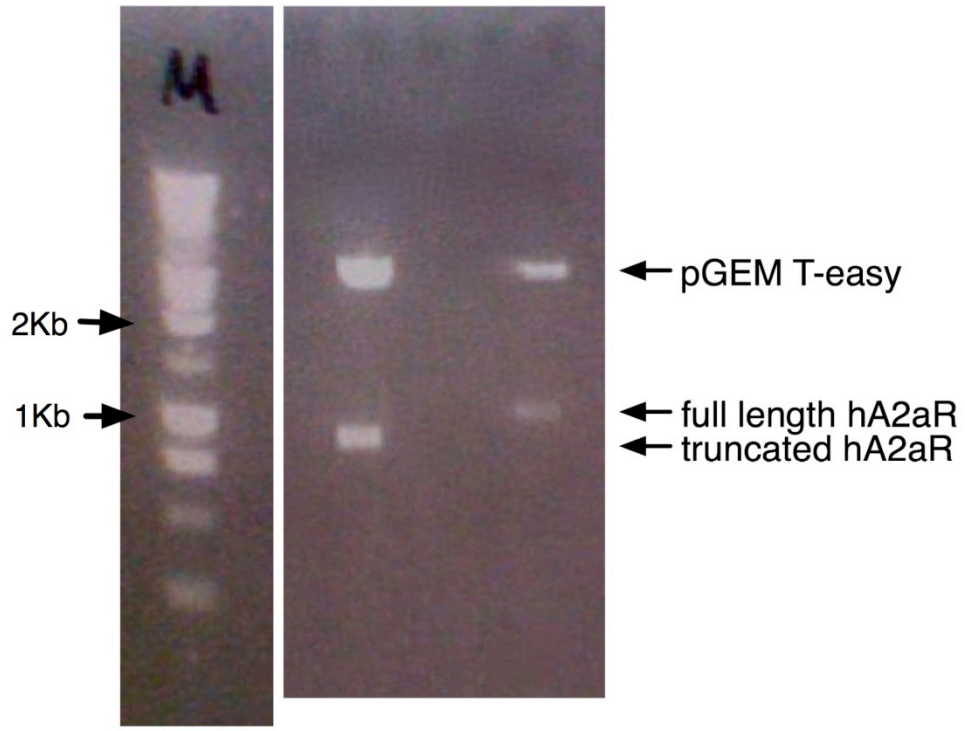


Figure 4.3 Restriction analysis of pGEM T-Easy containing either full-length or truncated deglycosylated hA2aR

Digestion of recombinant plasmid with EcoRI gave a vector band of 3.0 Kb and another band of either 1.2 or 0.9 Kb corresponding to full-length and truncated receptor, respectively.

4.3 Fusion of a biotin acceptor domain to the C-terminus of the hA2aR increases its expression in P. pastoris

4.3.1 Vector construction

Previously, it has been shown that fusion of the biotin acceptor domain from *Proprionibacterium shermanii* to the C-terminus of selected GPCRs improves both the yield and stability of the protein when it is made in *P. pastoris* (Grunewald *et al*. 2004; Lundstrom *et al*. 2006). Here, we wanted to determine whether or not attachment of this domain to the C-terminus of the hA2aR would similarly increase its expression level in *P. pastoris*. Another reason for wanting to fuse this domain to our protein of interest was to create a second affinity tag for purification purposes. This domain can be specifically biotinylated *in vitro* using the enzyme biotin ligase. Following this, the target protein can be isolated

by affinity chromatography using (strept)avidin beads (Chapman-Smith and Cronan 1999). This approach has a high purification factor as the abundance of biotinylated proteins in nature is very rare (Cronan 1990).

The 75αα biotinylation domain from *P. shermanii* was cloned into αH – trunc dG hA2aR (Figure 4.1) downstream of the NotI site according to Materials and Methods Section 3.3.3 (Figure 4.4). At the same time, two Tobacco Etch Virus (TEV) protease recognition sites were incorporated into the vector. The first one was introduced between the N-terminal His10 tag and the EcoRI site used for cloning the target proteins (Figure 4.4). The second is situated between the C-terminal tail of the receptor and the biotin acceptor domain (Figure 4.4). The purpose of adding the TEV sites was 2-fold. Firstly, cleaving off the His tag with TEV opens up the possibility of performing a reverse purification step using Ni^{2+} resin allowing cleaved tag, His-tagged TEV protease as well as contaminant His-rich proteins to be removed (Ford *et al*. 1991; Walker *et al*. 1994). Secondly, cleavage at both TEV sites will generate "tag-free" target protein that can be used in crystallisation experiments (Ford *et al*. 1991). Constructs were made that contained both full-length and truncated versions of the hA2aR (Sections 3.3.3-4) (Figure 4.5).

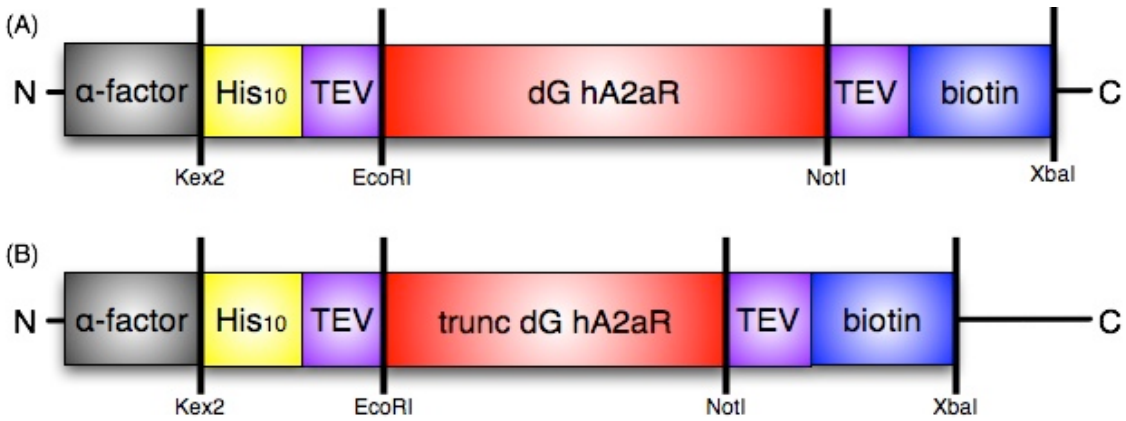


Figure 4.4 Schematic diagram showing the αHT-TB vectors containing either (A) full-length or (B) truncated hA2aR

The original GU *P. pastoris* expression vector (Figure 4.1) was modified by the introduction of a biotin acceptor domain at the receptor's C-terminus. Following biotinylation, this tag allows the protein to be purified by affinity chromatography using (strept)avidin beads. TEV protease recognition sites were inserted immediately up- and down-stream of the receptor coding sequence. Cleavage with TEV allows untagged hA2aR to be isolated.

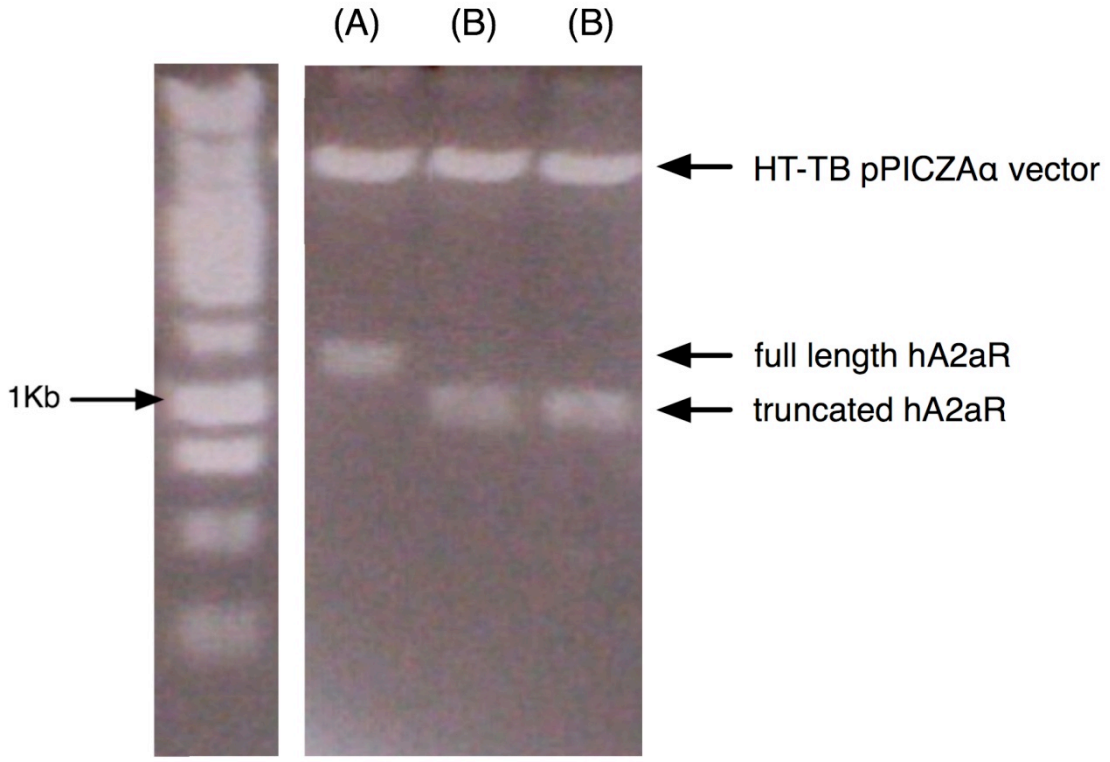


Figure 4.5 Restriction analysis of αHT-TB recombinant plasmids containing either (A) full length or (B) C-terminal truncated hA2aR

Digestion of the vectors with EcoRI/NotI resulted in release of the hA2aR cDNA (1.2 Kb full-length; 0.9 Kb truncated) from the HT-TB vector (3.9 Kb).

4.4 Yeast transformation and clonal selection

There are several possible *P. pastoris* strains that can be used for making recombinant proteins. Two of the most common are the wild-type strain X33 and the protease-deficient strain SMD1163 (his4 pep4 prb1). Following synthesis, recombinant proteins are susceptible to attack by site-specific proteases (Milojevic *et al*. 2006). Heterologously expressed MPs can be cleaved by proteases located in the culture medium as well as those within the cell (Macauley-Patrick *et al*. 2005). Proteins with extensive natively disordered regions tend to show high protease-sensitivity. Proteolysis can be overcome either by removing the susceptible protease site(s) by mutagenesis or by making the protein in a host cell where several of the key proteaseσs have been genetically knocked-out. Previously, the Fraser group used the *P. pastoris* protease-deficient strain SMD1163 for making full-length hA2aR (Fraser 2006).

Here, however, the αHT - (trunc) dG hA2aR – TB constructs were introduced into X33 as this strain gives significantly higher biomass in fermenter culture compared to SMD1163 (Niall Fraser, personal communication).

Both constructs were transformed into X33 according to Section 3.3 where the linearised DNA integrated into the yeast genome. Multi-copy integrants were isolated following selection with the antibiotic zeocin. Previously, it has been shown that the production level of recombinant proteins in *P. pastoris* is influenced by the number of expression cassettes that are incorporated into the genome following transformation (Grunewald *et al*. 2004; Macauley-Patrick *et al*. 2005). As different transformants make different amounts of the target protein it is essential to screen a range of colonies to identify the clone making the highest levels of recombinant protein. Small-scale (10 ml) expression trials were performed with eight transformants for both constructs according to Section 3.5.2. Membranes were isolated for each colony and receptor expression levels were determined by single-point radioligand binding assay as described in Section 3.8 using a tritiated version of the high-affinity hA2aR-specific antagonist ZM241385 (Palmer *et al*. 1995; Poucher *et al*. 1995) (Figure 4.6). For both constructs, an approximate 3-fold difference in receptor expression level was observed between clones that had been transformed with the same plasmid (Figure 4.6). The average expression levels for the full-length and truncated receptors across all the tested transformants were 18.0 +/- 2.5 and 16.4 +/- 2.5 pmol/mg protein, respectively, showing that the yeast made the full-length receptor and truncated mutants to essentially the same level. The highest expressing colony for both constructs was retained for subsequent experiments.

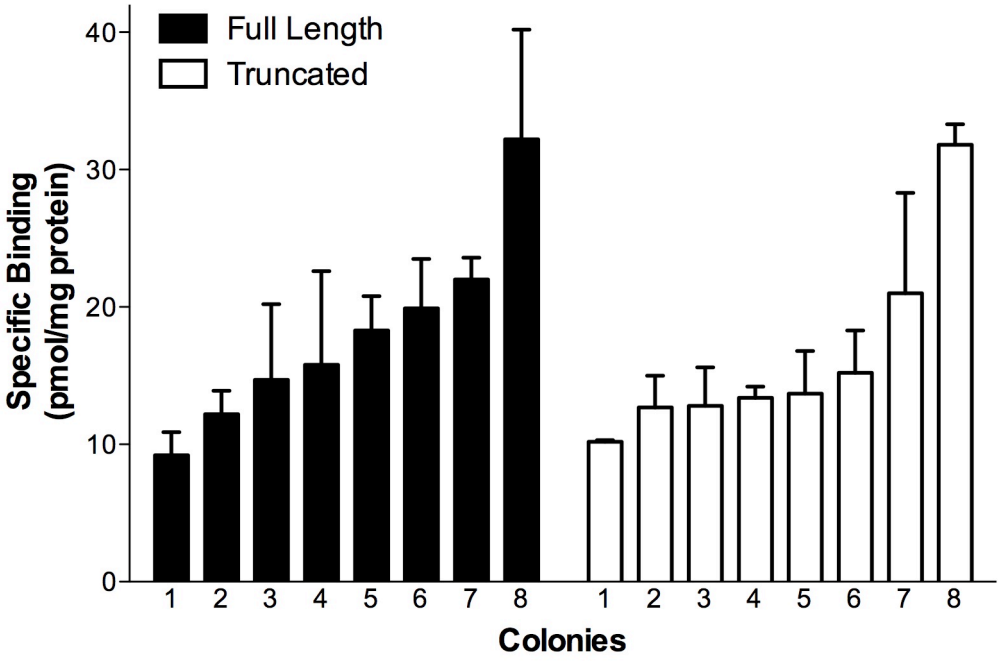


Figure 4.6 Clonal analysis of Pichia pastoris X33 transformants making either full-length or truncated hA2aR

Small-scale expression trials were performed with eight transformants for both αHT-TB expression constructs (Figure 4.4). Membranes were isolated for each colony and receptor expression levels determined by single-point radioligand binding assay using a saturating (*ca.* 12 nM) concentration of [^{3}H] ZM241385. The highest expressing colony for both constructs was retained for subsequent experiments.

4.5 Saturation binding analysis of full-length and truncated hA2aR made in shake flask cultures

The highest expressing colonies isolated from section 4.4 were grown on a medium scale (500 ml) in shake flasks. Following induction, cells were harvested and membranes made according to Section 3.6.3. The pharmacological properties of the expressed receptors were determined by saturation radioligand binding. In this approach, the amount of radioactivity that specifically binds to membrane-bound hA2aR was measured at a series of concentrations of [^{3}H] ZM241385. For both αHT-TB tagged full-length and truncated receptor constructs specific radioligand binding against radioligand concentration was plotted and the dissociation constant (K_d) as well as the total number of binding sites (B_{max})

determined. Representative saturation binding curves are shown in Figure 1.7 with the B_{max} and K_d values for each construct shown in Table 4.1.

Construct	B_{max} (pmol)	K_d (nM)
αH – full length dG hA2aR	8.2 +/- 0.9	1.8 +/- 0.4
	8.5 +/- 0.1	0.4 +/- 0.1
	9.6 +/- 2.2	1.2 +/- 0.6
αHT- full length dG hA2aR – TB	15.8 +/- 6.8	1.5 +/- 0.3
	18.8 +/- 0.3	0.7 +/- 0.1
	24.3 +/- 3.9	4.2 +/- 1.6
αHT- truncated dG hA2aR – TB	20.9 +/- 2.0	2.9 +/- 0.7
	30.1 +/- 1.2	4.4 +/- 0.4
	31.8 +/- 1.5	3.8 +/- 0.8

Table 4.1 Pharamcological data for three different hA2aR constructs grown in shake flask culture

Each construct was subjected to three separate radioligand binding experiments as indicated by the three B_{max} and K_d values.

Comparison of the B_{max} values for the different hA2aR constructs suggests that the fusion of the biotin acceptor domain to the C-terminus of hA2aR increases the receptor expression level by approximately 2-fold. However, the truncation of the receptor's C-terminal tail does not significantly affect its expression level. In native tissues the hA2aR has a K_d for ZM241385 of *ca*. 1.0 nM. In shake flask cultures of *P. pastoris*, K_d values in the range 1-4 nM were observed, indicative of weaker drug-receptor binding than in the physiological background. The most likely explanation for this observation is that oxygen starvation during receptor production prevented complete formation of its disulphide bonds (which are known to be essential for high-affinity ligand binding (Jaakola *et al*. 2008)). It

was hoped that this problem could be overcome by making receptor in fermenter culture where the dissolved oxygen (DO) levels can be readily regulated (Cereghino *et al*. 2002; Jahic *et al*. 2006).

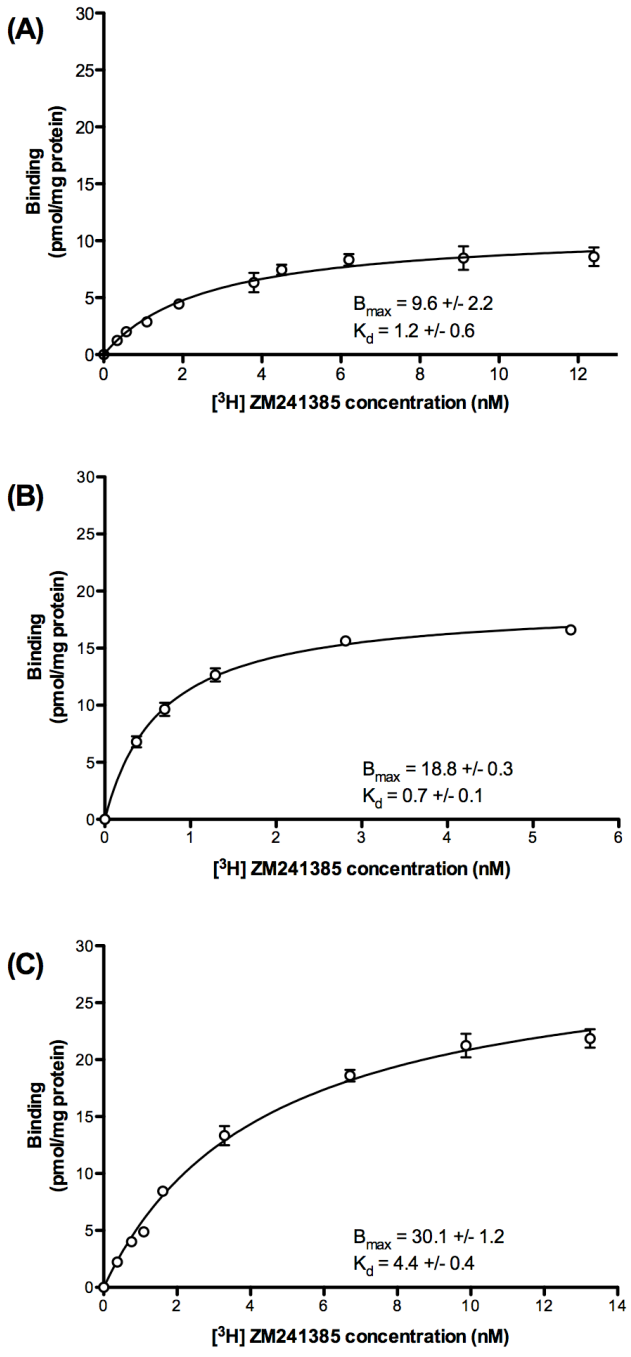


Figure 4.7 Representative saturation binding curves for three different hA2aR constructs made in X33 shake flask cultures: (A) αH– dG hA2aR, (B) αHT– dG hA2aR –TB, and (C) αHT– trunc dG hA2aR –TB

Saturation binding analysis was performed according to Section 3.8.1. The data was fitted using GraphPad Prism software.

4.6 hA2aR production in fermenter culture

One of the main advantages of using *P. pastoris* for recombinant protein production is its ability to grow to high cell densities (*ca.* 10^9 cells/ml) in fermenter culture (Jahic *et al.* 2006). This can be achieved by maintaining the DO content of the culture greater than 20% as well as tightly regulating the methanol feed rate (Cregg *et al.* 2000). The transition from recombinant protein production in low cell-density shake flask cultures to that in high cell-density fermentations requires considerable optimisation to obtain maximal yields of the target protein (Cereghino *et al.* 2002; Singh *et al.* 2008). This is particularly true when the methanol induction (protein expression) phase is performed at low (22 instead of 30°C) temperatures where care has to be taken not to overfeed the culture with methanol. In fermentations of *P. pastoris* the yeast are first grown on glycerol to increase the culture biomass before switching the carbon source to methanol for recombinant protein production.

A series of fermentations were performed on a 5 litre scale using the X33 transformant found to make the highest levels of αHT – trunc dG hA2aR – TB, in which a series of parameters including length of the glycerol feed phase as well as the methanol feed rates were optimised. Maximal cell biomass was obtained by growing the yeast for 18h on basal medium followed by feeding the culture with glycerol for a further 11 h (Figure 4.8). This resulted in a culture OD_{600} of over 300 which is equivalent to *ca.* 200g wet weight of cells per litre of culture. This phase was followed by a 1h starvation period during which time the culture was not fed any carbon source. This step was included to ensure that all of the added glycerol had been completely consumed by the yeast as glycerol represses protein expression from the AOX1 locus even in the presence of methanol (Cereghino *et al.* 2002).

At the start of the methanol-feed phase, the hA2aR antagonist theophylline was added to the culture to enhance the levels of functional receptor made by the yeast (Fraser 2006). In addition, receptor production was carried out at 22°C as previously it has been shown that twice as much functional hA2aR can be

obtained in shake-flask cultures of *P. pastoris* by performing induction at 22 instead of 30°C (Fraser 2006). The induction phase was started on a low methanol feed rate (4 h at 0.09 ml min-1) to allow time for the cells to adapt to growth on methanol as its sole carbon source preventing unnecessary cell death (Cereghino and Cregg 2000; Cereghino *et al*. 2002).

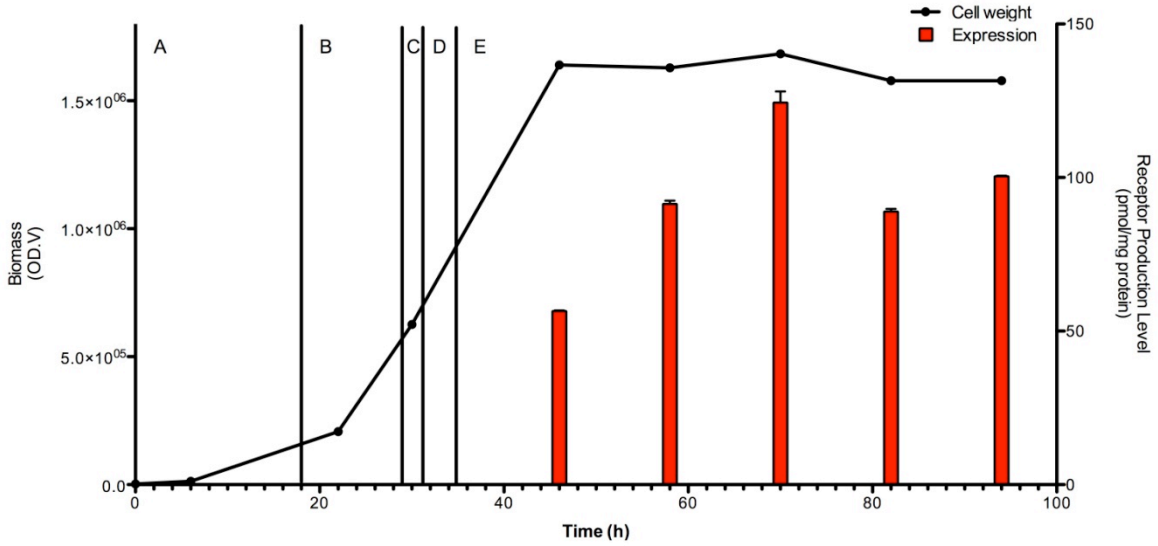


Figure 4.8 A typical fermentation showing how yeast biomass and hA2aR production level varies with time

The yeast were grown on basal media for 18h (A) followed by a glycerol-feed phase lasting a further 11h (B). A starvation period (C) ensured that all of the added glycerol had been completely consumed. A short methanol adaptation phase (D) allowed the cells to accommodate the change of carbon source reducing the possibility of cell death. Finally, receptor production occurred during the methanol-feed phase (E) with maximum expression levels obtained after *ca.* 30 h.

The methanol feed-rate was increased in a step-wise fashion in order to identify the rate that gave the maximum hA2aR expression level without causing cell-toxicity arising from the accumulation of excess methanol. The yeast were fed at a particular rate for 12 h and then sample removed from the culture to allow the receptor expression level to be determined by single-point radioligand binding analysis (results not shown). This allowed the development of an optimised feeding protocol which after the cells had acclimatised to the methanol carbon source consisted of 12 h at 0.13 ml/min followed by 36 h at 0.33 ml/min. Under optimal conditions, the human adenosine 2a receptor could

be routinely made with expression levels greater than 100 pmol receptor per mg protein (eg Figure 4.9).

The αHT-trunc dG hA2aR-TB made in fermenter culture had K_d values for [^{3}H] ZM241385 of 0.7 +/- 0.1, 0.8 +/- 0.1 and 0.9 +/- 0.2 nM. These values are essentially the same as that observed for the hA2aR in native tissues (Uustare *et al.*, 2005), and reflected tighter binding of the radioligand to the receptor when it was made in fermenter as compared to shake-flask culture (Section 4.5). The improvement in the quality of receptor made by the yeast in fermenter culture is most likely due to the consistently high levels of dissolved oxygen that will ensure all of the disulphide bonds in the recombinant protein are fully formed.

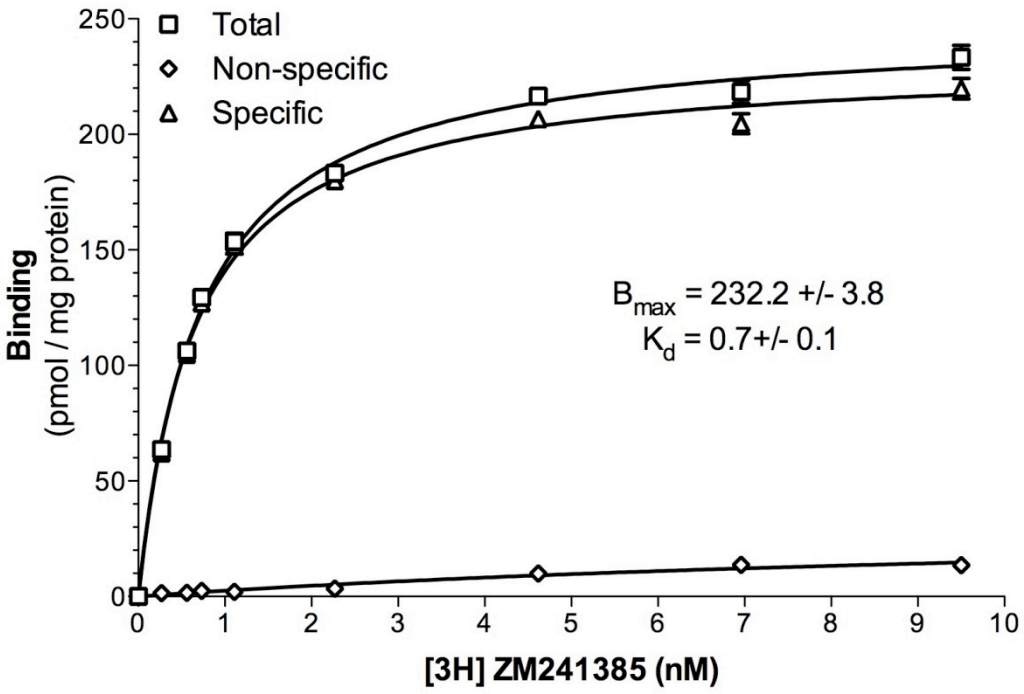


Figure 4.9 A representative saturation binding curve of αHT- trunc dG hA2aR-TB produced in fermenter culture

Saturation binding analysis was performed as described in Section 3.8.1. The data was fitted using Graph Pad Prism. Receptor production levels greater than 100 pmol/mg protein were routinely obtained. The K_d values obtained for hA2aR made in fermenter culture were essentially the same as that observed in native tissue (Uustare *et al.*, 2005), suggesting that the receptor made was of high quality.

4.7 Purification of HT- trunc dG hA2aR - TB

To solve the 3-dimensional structure of any mammalian membrane protein it is first necessary to obtain milligram quantities of pure, monodisperse and stable protein. Mild non-ionic detergents (e.g. DDM) are often used to remove proteins from biological membranes as they are not sufficiently harsh to strip away all of the native-lipids (some of which are often essential for protein function) (Johansson *et al.* 2009). Previously, it has been shown that the hA2aR can be released from yeast membranes in an active state using a combination of DDM with the cholesterol derivative cholesterol hemisuccinate (CHS) (Weiss and Grisshammer 2002; Fraser 2006). Here, αHT – trunc dG hA2aR – TB in X33 membranes was similarly treated with DDM/CHS. Insoluble material was removed by centrifugation and the supernatant containing receptor passed over Ni^{2+} resin. After the receptor had bound to the column, it was washed with buffer containing 75 mM imidazole (to remove contaminant His-rich proteins) and 500 mM NaCl (to remove those proteins that had bound to the column by ionic interactions). Receptor was released from the column using a 0-1 M imidazole gradient and the elution fractions analysed by SDS-PAGE (Figure 4.10).

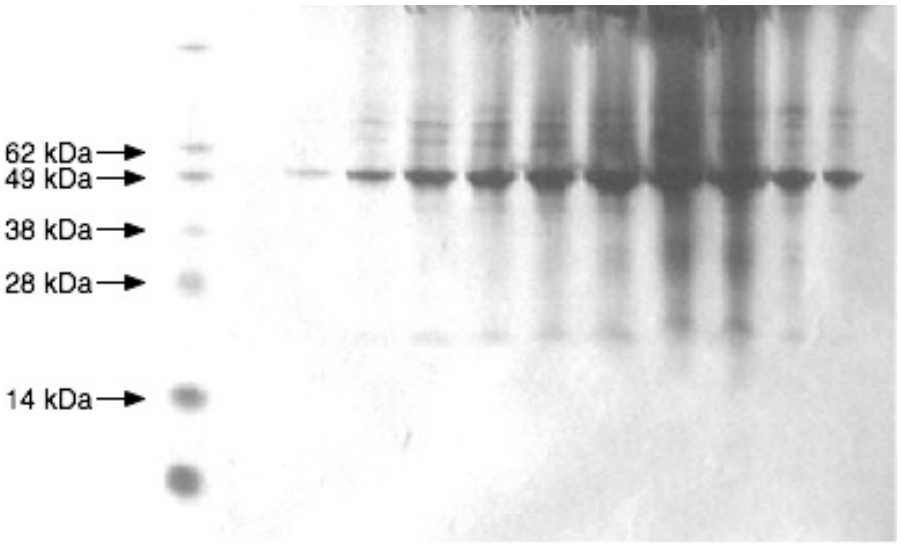
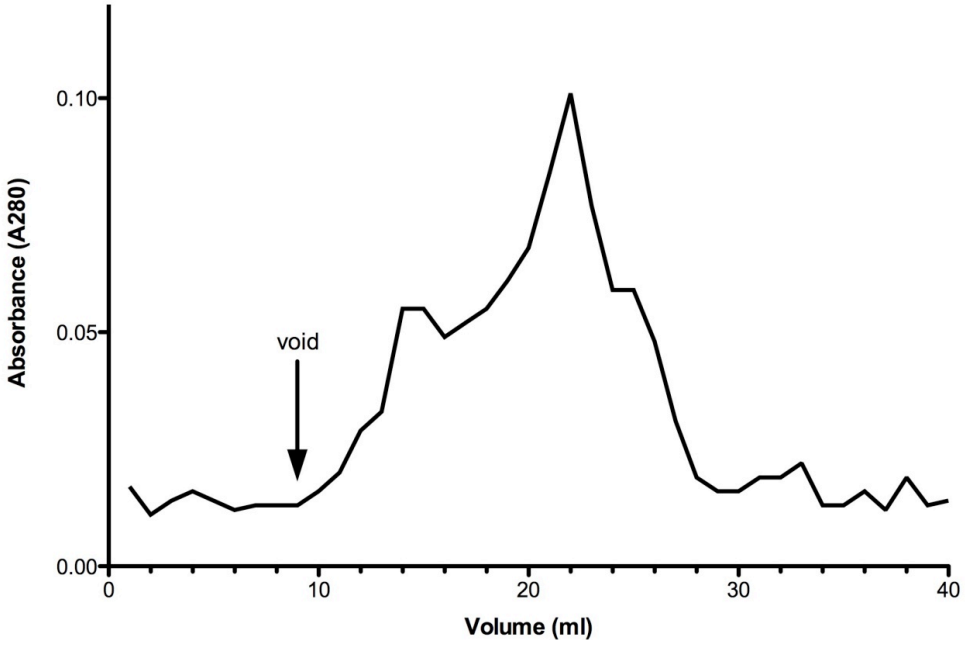


Figure 4.10 SDS-PAGE analysis of partially purified HT – trunc dG hA2aR – TB

Receptor containing membranes were solubilised using DDM/CHS. The recombinant protein was purified by Ni^{2+} affinity chromatography. 20 μl aliquots from each elution fraction were analysed by SDS-PAGE. Following treatment with Simply Blue Safe Stain (Invitrogen), a major band at *ca.* 49 kDa corresponding to the tagged receptor was observed.

As the receptor was relatively pure following the Ni^{2+} column, the hA2aR fractions were pooled, concentrated and loaded onto a Sepharose C6 gel filtration column (which allows proteins to be separated according to their size). In addition to its purification capacity, size-exclusion chromatography (SEC) is also a good indicator of protein quality as the size and shape of the elution peak gives insights into the heterogeneity and aggregation status of the applied protein(s) (Gutmann *et al*. 2007). The ideal result denoting pure, monodisperse protein is a single large symmetrical elution peak while several peaks arising from the one protein suggests that it may be aggregated and as such is unsuitable for structural studies (Kawate and Gouaux 2006). When the highly-enriched hA2aR sample was applied to the column several major peaks were observed in the SEC elution profile (Figure 4.11), suggesting that the receptor population contained aggregate complexes. However, it was not obvious at what stage receptor aggregation had occurred – whether it was during the detergent-solubilisation step, Ni^{2+} purification and/or application to the gel filtration column. This result made it clear that substantially more work was going to be required to develop an optimal purification protocol for the hA2aR. Furthermore, it highlighted the need for an experimental tool that would allow the receptor's dispersity/aggregation status to be followed throughout the purification procedure.

(A)



(B)

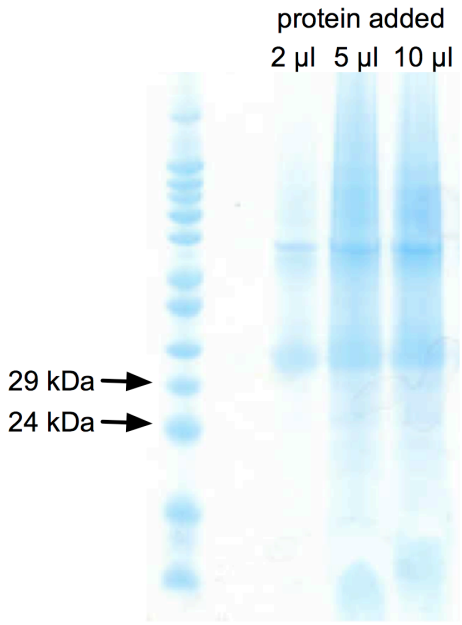


Figure 4.11 Analysis of a highly-enriched hA2aR sample

(A) Elution profile following application to a Sepharose C6 gel filtration column. If the partially-purified receptor was monodisperse, the elution profile would have contained essentially a single, large symmetric peak. However, several heterogeneous peaks were observed, demonstrating that the receptor was aggregated. However, it was not obvious at what stage during the purification process receptor aggregation had occurred. (B) SDS-PAGE analysis showing multiple bands therefore substantiating aggregation had formed.

4.8 Conclusion

Pichia pastoris is an excellent host cell for the overproduction of recombinant proteins including membrane proteins of mammalian origin. However, as illustrated by the work described in this chapter, the commercially available vector is far from what is required to maximise the production levels as well as simplify the detergent extraction and subsequent purification procedures for human membrane proteins.

The addition of the biotinylation domain from *P. shermanii* had previously been shown to increase protein expression (Grunewald *et al*. 2004; Lundstrom *et al*. 2006) and through its addition to the C-termini of the hA2aR expression construct a minimum of a 2-fold expression level increase was seen. However, while the truncation of the hA2aR C-termini did not have signification effects on its expression it was retained to help improve its crystallizability.

An optimised protocol was developed that gave maximal production of our target proteins in fermenter culture at an induction temperature of 22°C. Care was taken to find a methanol feed rate that gave the highest levels of protein production without causing the accumulation of excess methanol in the culture (which is known to be toxic to the yeast). The hA2aR produced in fermenter culture had more native-like pharmacology than when the yeast cells were grown in shake flasks. This is likely due to the incomplete formation of disulphide bonds (which are known to be necessary for high-affinity ligand binding) in the hA2aR when it is made in shake-flask cultures of yeast. This is not surprising as the DO levels in fermenter culture can be maintained >20% whereas the DO levels in shake flasks are essentially zero. This observation will be applicable to other proteins made in *P. pastoris* and suggests that following clonal selection the highest-expressing transformant should be put straight into fermenter culture to assess whether or not the yeast is going to be a suitable host for making milligram quantities of the target protein for structural studies.

In short, by investing time optimising both the expression construct and expression method it is possible to obtain yields of active protein suitable for structural studies.

5 iLOV: a novel fluorescent reporter for membrane protein applications

5.1 Introduction

The over-expression and purification of membrane proteins can be a labour-intensive process. One of the many challenges of working with MPs concerns the detection of recombinant MPs in both expression and purification experiments. As most, though not all, MPs are colourless, target protein identification is usually performed by methods such as radioligand binding and/or Western blotting. However, these techniques can be time-consuming, use a lot of protein and do not give any information on the aggregation state of the protein in detergent solution.

The process of identifying target membrane proteins can be accelerated by fusing a fluorescent reporter to the protein of interest (Sarramegna *et al.*, 2002; Drew *et al.*, 2006). This allows recombinant MPs to be monitored indirectly by measuring fluorescence from the tag (Drew *et al.*, 2006). To be successfully employed as a reporter protein, the fusion partner must be stable in the host cell used for recombinant expression, easily quantifiable and give fluorescence levels that accurately correlate with the amount of recombinant protein present (Drew *et al.*, 2008).

Green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* was the first, and remains the most commonly used, fluorescent reporter (Chalfie *et al.*, 1994). GFP is so-called due to its strong green fluorescence emission upon illumination with visible light. GFP consists of 238 amino acids, and has a molecular weight of *ca.* 25 kDa (Prasher *et al.*, 1992). The GFP chromophore is situated in a central α-helix located within an 11-stranded β-barrel (Figure 5.1) (Ormo *et al.*, 1996), and is formed by the autocatalytic cyclisation of the

tripeptide Ser65 – Tyr66 - Gly67 (Remington, 2006; Huang *et al.*, 2007; Malo *et al.*, 2007). Chromophore maturation consists of four distinct phases: folding of the β-barrel, peptide cyclisation, oxidation and dehydration (Huang *et al.*, 2007). The oxidation reaction is the rate-limiting step in chromophore formation, and requires molecular oxygen. Despite its usefulness as a fluorescence reporter, GFP has a couple of drawbacks: i) it has a tendency to oligomerise, and ii) it cannot be used in conditions that are anaerobic/have low oxygen content (Drepper *et al.*, 2007).

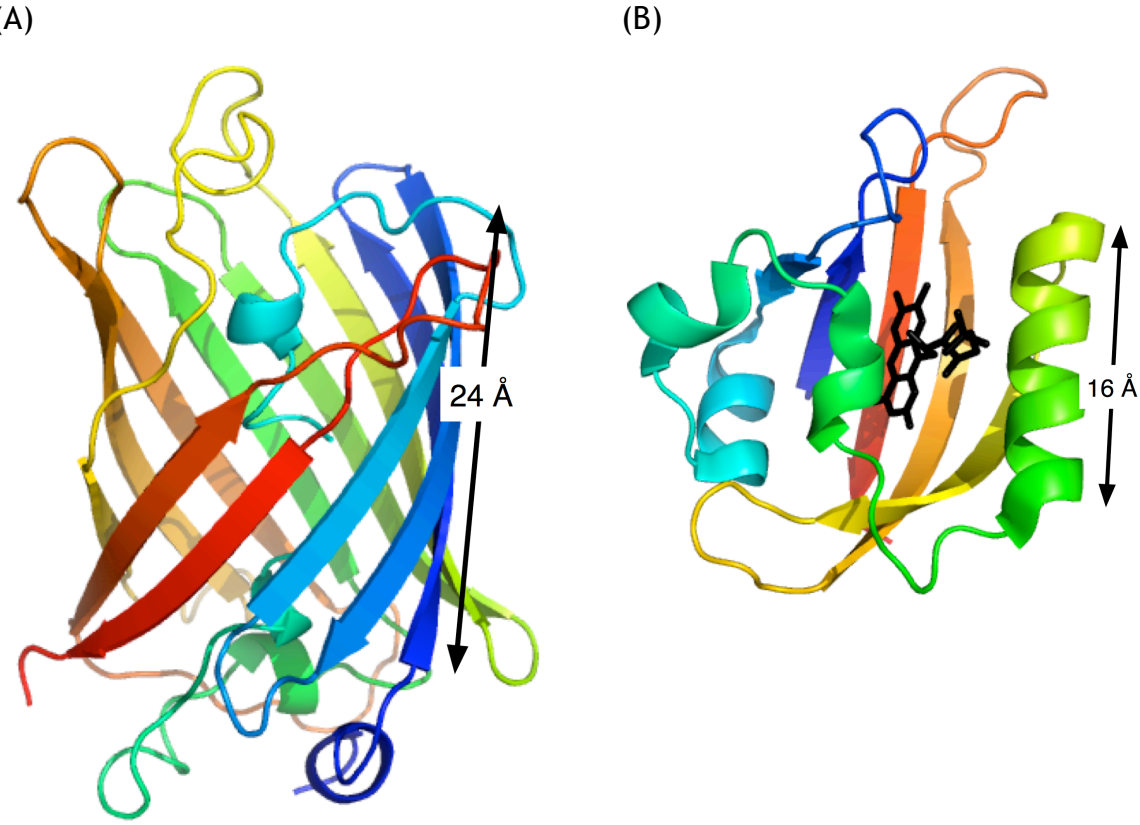


Figure 5.1 The three-dimensional structures of (A) GFP and (B) LOV2
The structure of GFP (accession code 1EMA) was described in (Ormo *et al.*, 1996) and that for LOV2 (accession code 1G28) in (Crosson and Moffat, 2001).

Although GFP has been traditionally used to monitor biological events *in vivo*, it has recently been used as a tool to monitor protein folding *in vitro*. For example, it has been shown that the folding state of soluble recombinant proteins made in *E. coli* can be rapidly determined by fusing GFP to their C-termini as the level of GFP fluorescence observed was found to correspond with

the amount of correctly folded target protein (Waldo *et al.*, 1999). In contrast, fusion protein that had been misfolded ended up in inclusion bodies and was non-fluorescent (as the GFP chromophore did not form). Since this first observation, this methodology has been successfully adapted for use with membrane proteins where C-terminal GFP fusions have been used to identify those membrane proteins that express well in both *E. coli* (Drew *et al.*, 2002; Hamman *et al.*, 2002; Drew *et al.*, 2005; Drew *et al.*, 2006; Link *et al.*, 2008) and *S. cerevisiae* (Newstead *et al.*, 2007; Drew *et al.*, 2008) and those that do not. The GFP tag has also been shown to be useful for identifying those detergents that give monodisperse protein immediately following solubilisation of the membrane via fluorescence size-exclusion chromatography (FSEC) (Kawate and Gouaux, 2006; Newstead *et al.*, 2007; Drew *et al.*, 2008).

Although highly useful for organisms such as *E. coli* and S. *cerevisiae*, GFP is not ideal for use with *P. pastoris* as this yeast grows to high cell densities in culture largely consuming the dissolved oxygen in the media. This is particularly true in shake-flask cultures of *P. pastoris*. The strict requirement for molecular oxygen in the formation of the GFP chromophore suggested to us that a LOV-type of fluorescence reporter (derived from plant phototropins) may be better suited for use in *P. pastoris* than GFP as LOV fluorescence is not affected by oxygen availability (Drepper *et al.*, 2007).

Phototropins are blue-light (region 320-500 nm) receptors that regulate several different physiological processes in plants that serve to maximise the efficiency of photosynthesis. These include phototropism, light-induced stomatal opening, as well as chloroplast movements in response to changes in light intensity (Christie, 2007). Phototropins consist of two distinct regions: a N-terminal photosensory domain with a serine/threonine kinase domain at its C-terminus. Within the photosensory domain are two highly conserved regions that have sequence homology with a family of proteins that are activated by light, oxygen or voltage (LOV) called LOV1 and LOV2 (Christie *et al.*, 1999; Jones and Christie, 2008). Light-induced phototropin kinase activity requires LOV2 (but not LOV1) photoreactivity though the mechanism by which LOV2 regulates the kinase

domain is poorly understood (Christie *et al.*, 2002; Matsuoka and Tokutomi, 2005).

Purification of milligram amounts of individual LOV domains expressed in bacteria has allowed their structure and photochemistry to be investigated in fine-detail. Individual, purified LOV domains are bright yellow and emit a strong green fluorescence when irradiated with UV/blue light. LOV domains non-covalently bind flavin mononucleotide (FMN) in a central cavity through a combination of hydrogen bonds and van der Waals interactions (Figure 5.1) (Christie *et al.*, 1998; Crosson and Moffat, 2001; Christie, 2007). Upon irradiation, a covalent adduct forms between the flavin chromophore and a conserved cysteine within the LOV domain (Salomon *et al.*, 2000). This reaction is fully reversible, and in darkness the LOV domains return to their initial ground state (Salomon *et al.*, 2000).

Recently, Christie and co-workers have created a novel LOV-based fluorescence reporter called iLOV that can be used for both *in vivo* and *in vitro* applications (Chapman *et al.*, 2008). iLOV was derived from the LOV2 domain of the *Arabidopsis thaliana* phototropin 2 by performing several rounds of DNA shuffling. The iLOV protein has increased fluorescence quantum yield and reduced photobleaching compared to wild-type (Chapman *et al.*, 2008). iLOV has certain intrinsic properties that suggested that it may make a useful reporter of recombinant membrane protein production in all systems but particularly in *P. pastoris*. These include: i) iLOV is relatively small (10 as compared to 25 kDa for GFP), ii) fluorescence from the flavin co-factor occurs immediately from the time that the iLOV domain is assembled (in contrast to GFP where formation of the fluorophore takes *ca.* 4 h (Heim *et al.*, 1994)) making it useful for accurate time-course studies, and iii) iLOV fluorescence is unaffected by oxygen availability (unlike GFP where chromophore maturation has an absolute requirement for molecular oxygen) (Drepper *et al.*, 2007). This last property of iLOV is of particular relevance for studying recombinant (membrane) protein production in shake flask cultures of *P. pastoris* as their oxygen content is essentially zero.

Over the course of my Ph.D., I made several different *P. pastoris* expression constructs containing the iLOV fluorescent reporter which are describes in this chapter. In all of the vectors, an iLOV cDNA was inserted into the plasmid C-terminal to the cloning sites for the recombinant protein, and was introduced to simplify the expression and purification of my target MPs. In addition this chapter describes the photochemical properties of iLOV, as well as how this tag can be used to monitor protein expression (clonal selection; time course of MP production in fermenter culture), perform detergent screening and facilitate protein purification.

5.2 Fusion of a fluorescent reporter protein to the receptor's C-terminus

As the vast majority of mammalian membrane proteins are colourless they cannot be followed visually during the purification process. Rather those fractions containing the target protein have to be found either by visualisation techniques (e.g. Western blotting) or by performing a functional assay (e.g. radioligand binding). While these approaches are highly successful in identifying the desired fractions they have many drawbacks not least in the amount of time and protein required to do the assay. However, by fusing the protein of interest to a fluorescent marker it is possible to track the expression and purification of the recombinant protein by eye following excitation of the reporter at an appropriate wavelength (Drew *et al.*, 2006). Although green fluorescent protein (GFP) from *Aequorea victoria* has been successfully used for tracking membrane protein production in *E. coli* and *S. cerevisiae* (Chalfie *et al.*, 1994; Drew *et al.*, 2008), we decided to use a fluorescent protein derived from a plant phototropin Light-Oxygen-Voltage (LOV) domain (Briggs, 2007 a,b). This reporter protein fluoresces from the time it is synthesised (upon binding its co-factor flavin mononucleotide), and has oxygen-independent fluorescence intensity (both unlike GFP) (Christie, 2007). The native LOV domain has been modified by Christie and co-workers to give the iLOV reporter which has improved stability and fluorescence intensity compared to the WT protein (Chapman *et al.*, 2008).

iLOV cDNA was incorporated into the HT-TB expression plasmid between the C-terminal TEV recognition site and the biotinylation domain (Figure 5.2). This was accomplished by introducing two new restriction sites (SpeI and NdeI) between the TEV site and the biotinylation domain, and then sub-cloning the iLOV DNA between the new restriction sites (Figure 5.3). By positioning the iLOV domain downstream of the target protein, it was possible to measure recombinant protein expression level via iLOV fluorescence (Section 5.6). Furthermore, the tag was useful for determining the dispersity of the receptor following detergent solubilisation as well as during the purification process by fluorescence SEC.

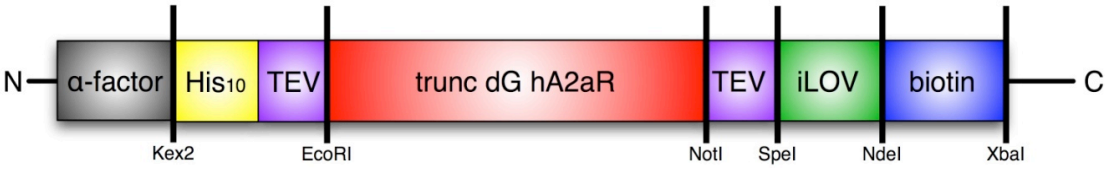


Figure 5.2 Schematic diagram showing the αHT-TLB expression plasmid

The fluorescent reporter protein iLOV was incorporated between the downstream TEV protease recognition site and the biotin acceptor domain. The iLOV tag was useful for determining recombinant membrane protein expression levels as well as their aggregation status following detergent solubilisation.

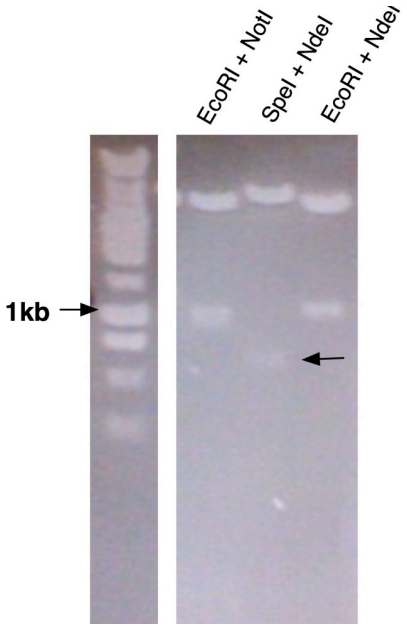


Figure 5.3 Restriction analysis of HT – trunc dG hA2aR – TLB

Digestion of the recombinant plasmid with (a) EcoRI/NotI, (b) SpeI/NdeI and (c) EcoRI/NdeI gave bands at 1.0, 0.3 and 1.2 kb corresponding to trunc hA2aR, iLOV and trunc hA2aR-iLOV fragments of DNA, respectively. (The 0.3kb band is indicated with an arrow due to poor contrast).

5.3 Using conformationally-fixed hA2aR mutants to improve receptor stability

Many signalling proteins including GPCRs have intrinsic flexibility that allows them to adopt multiple conformations (Kobilka and Deupi 2007). While this flexibility is essential for receptor activity it is also an impediment to crystallisation (Warne *et al*. 2009). This problem can, however, be overcome by trapping GPCRs in particular conformations using appropriate ligands. Indeed, the ligand-activated GPCRs that have had their structures solved thus far have all been co-crystallised with either a high-affinity antagonist or inverse agonist bound (e.g. (Palczewski *et al*. 2000; Cherezov *et al*. 2007; Rasmussen *et al*. 2007; Rosenbaum *et al*. 2007; Hanson *et al*. 2008; Jaakola *et al*. 2008; Warne *et al*. 2008). In these examples, the ligands were not just used to fix the receptor's conformation but to increase the thermostability of the protein in detergent solution also.

An alternative approach to solving the twin problem of conformational heterogeneity and poor stability of GPCRs in detergent solution has recently been described by Tate and co-workers who, after performing an extensive mutational analysis of the turkey β1-adrenergic receptor, created a modified version of the receptor called M23 that preferentially binds antagonists and was *ca*. 30°C more thermostable than the WT receptor (Warne *et al*. 2009). This mutant was even stable in C8 detergents including octylthioglucoside in which it was crystallised with the antagonist cyanopindolol bound and structure solved to atomic resolution (Warne *et al*. 2009).

More recently, Tate and co-workers have described two conformationally-fixed thermostable hA2aR mutants called Rant21 (which preferentially binds antagonists) and Rag23 (agonists) (Magnani *et al*. 2008) (Figure 5.4). To determine whether or not these mutants aggregated in the same way that had been found for the truncated hA2aR construct (Section 4.7), Rant21 and Rag23 cDNAs (a kind gift from Dr. Chris Tate) were cloned into the αHT-TLB vector. The Rant21 and Rag23 cDNAs that had been provided to us had been truncated

at αα316. We, therefore, made an equivalent truncation of the WT receptor construct denoted GU. The glycosylation site in all three constructs was removed by site-directed mutagenesis. Finally, to further improve their stability and potential crystallisability the flexible region in intracellular loop 3 (αα 210-227) that had been identified previously using RONN (Section 4.2) was removed from the GU, Rant21 and Rag23 constructs. The integrity of the desired constructs was explored first by restriction analysis (Figure 5.5) and then by sequencing.

All three constructs were transformed into yeast strain X33. Following fermentation, membranes containing iLOV-tagged receptor were isolated and a series of trial solubilisation experiments performed. Using FSEC, it was possible to determine the extent of receptor aggregation immediately after treatment with detergent. These results are presented fully in Section ?. In short, the CHS was found to be the cause of receptor aggregation observed in Section 4.6. My work focussed on Rag23 which could be purified with monodispersity in DDM in the absence of CHS (Section 6.2).

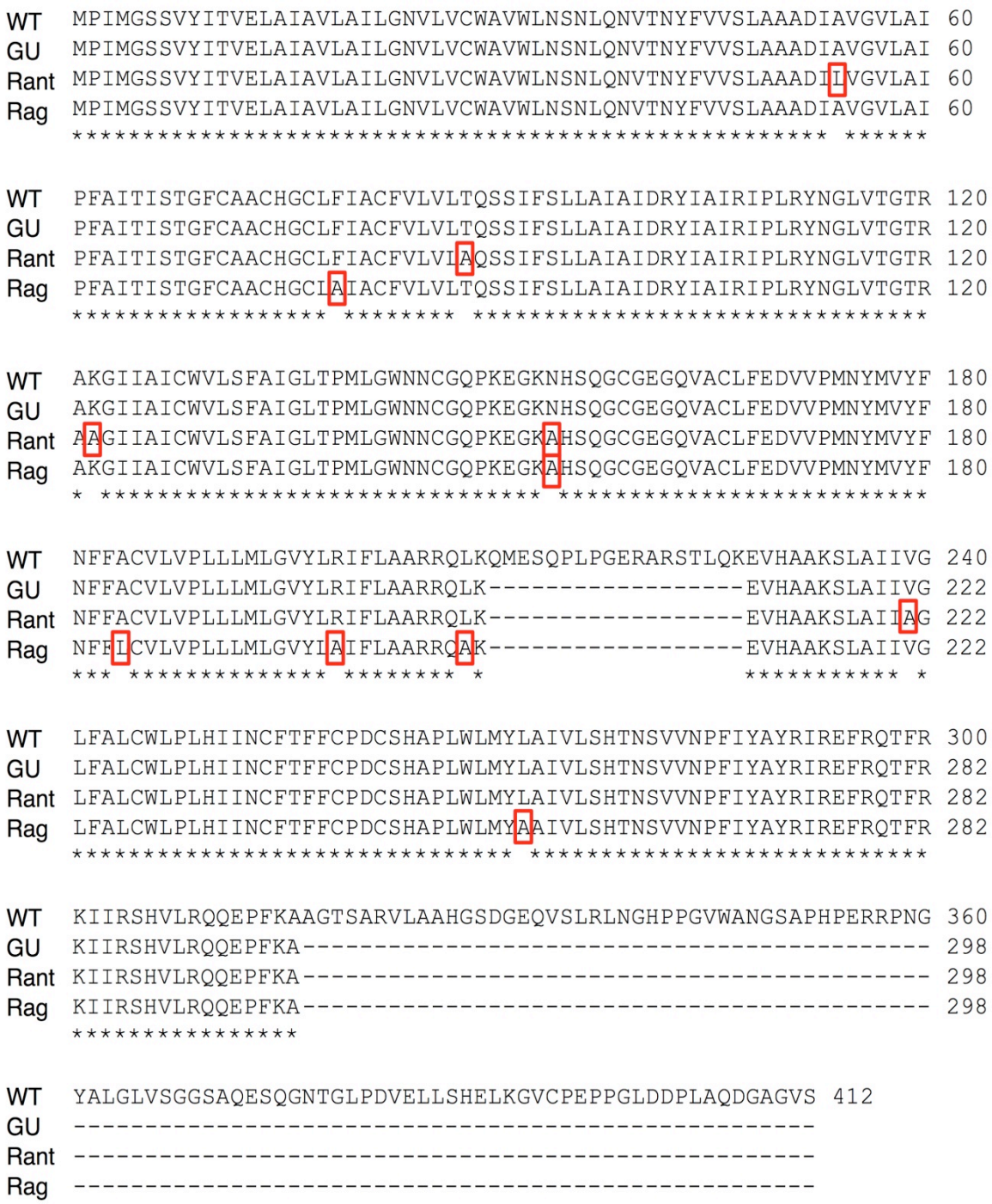


Figure 5.4 Sequence alignment of GU, Rant21 and Rag23 DNA sequences against WT hA2aR
All three modified forms of the hA2aR had their glycosylation sites as well as 18 residues from the third intracellular loop and the C-terminal 96 residues removed. The $\alpha\alpha$ changes in the Rag23 and Rant21 mutants are highlighted in red boxes.

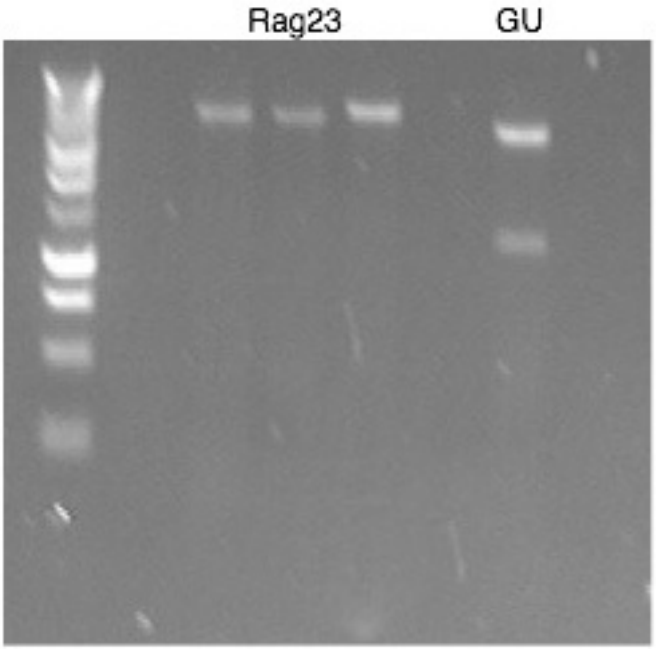


Figure 5.5 Restriction analysis of recombinant αHT-TLB vectors containing either Rag23 or GU hA2aR

Due to the mutations forming the conformationally-fixed Rag23 mutant, the Rag23 cDNA does not have a PvuII restriction site while the GU sequence does. On digestion of the Rag23 construct with PvuII a single band was observed due to linearization of the vector. In contrast, two bands were seen for the GU plasmid due to a single cut each within the insert and vector.

5.4 Optimised tag removal by digestion with C3 protease

The expression construct containing iLOV was re-engineered to simplify the tag removal process. With the HT-TLB vector, complete tag removal required successful cleavage at both TEV sites. Even allowing for the problems encountered with TEV solubility (section 5.10), it was thought that requiring two cleavage events for tag removal was unnecessarily complicated. To simplify the proteolytic procedure, the His10 tag was moved from upstream of the receptor coding sequence to downstream of the biotin-acceptor domain (Figure 5.6), creating a iLOV-biotin-His10 (LBH) tag. The remaining TEV site was replaced with one for human rhinovirus C3 protease (C), creating an αCLBH vector. A version without the α-factor leader sequence was also made -CLBH (Figure 5.6). C3 protease has considerably higher cleavage activity than TEV which means that it gives efficient removal of tags from recombinant proteins even at low (4°C) temperatures (Knott *et al*. 1989; Terpe 2003). Moreover, unlike TEV where

milligram amounts of protease are required for tag removal (Mohanty *et al.* 2003), only microgram quantities of C3 need to be added to the recombinant protein. It was envisaged that the superior enzymatic properties of C3 would overcome the solubility problems that had been encountered with TEV.

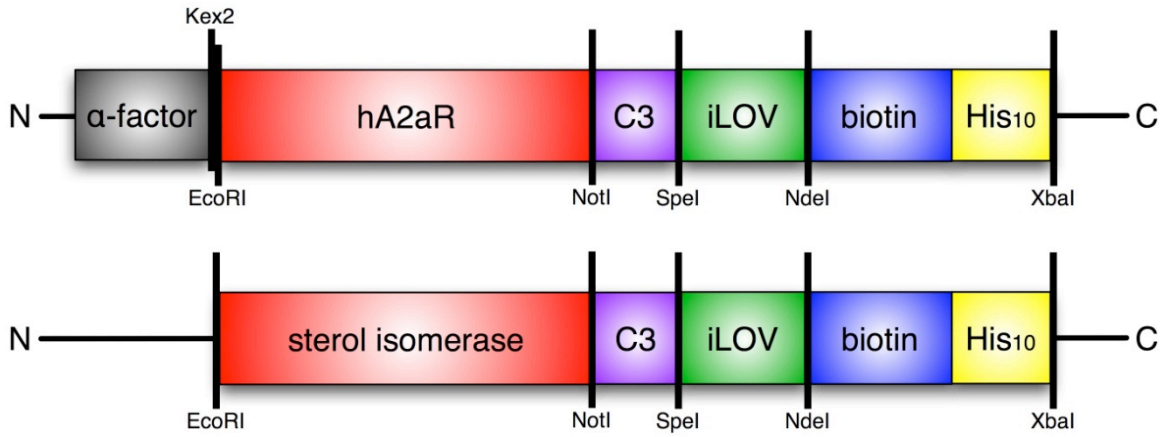


Figure 5.6 Schematic diagram of the CLBH expression constructs with and without α-factor
To simplify the proteolytic processing of the recombinant protein, the His10 tag in the αHT-TLB vector was re-positioned downstream of the biotin acceptor domain. Furthermore, the TEV protease recognition site was replaced with one for human C3 rhinovirus protease. A version of the CLBH vector without α-factor leader sequence was made, and a cDNA for human sterol isomerase cloned into it.

Vector modifications were performed as described in Section 3.3.7. GU, Rant21 and Rag23 versions of hA2aR were cloned into the Aα-CLBH vector and hSI cloned into A-CLBH. To ensure that all of the constructs had been made correctly, they were analysed by restriction digestion (Figure 5.7) prior to sequencing. The CLBH constructs allowed for the rapid purification of both Rag23 and SI on a milligram scale. This work is described in Sections 6.2, 6.3 and 6.7.

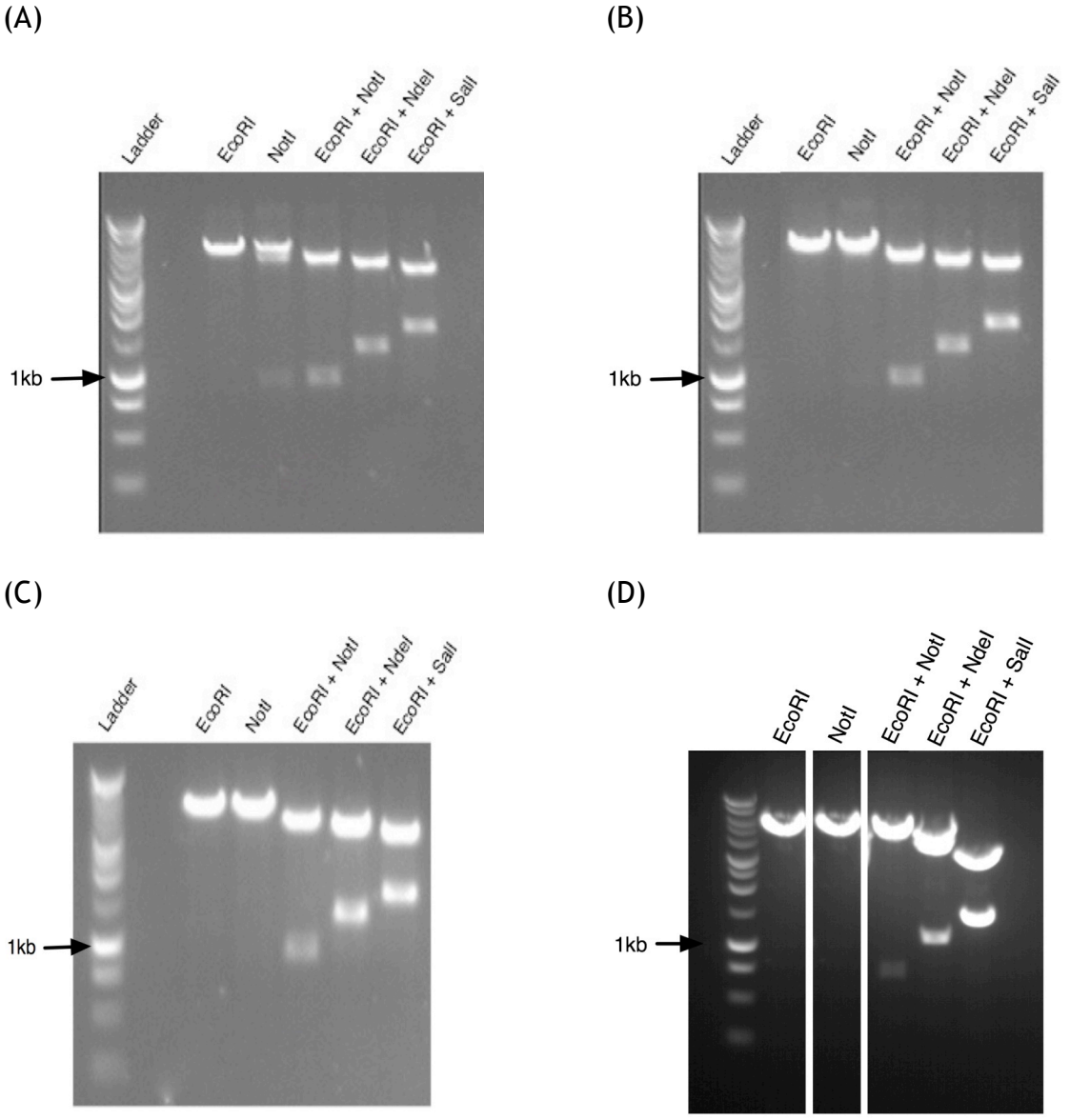


Figure 5.7 Restriction digest analysis of (A) αGU-, (B) αRant21-, (C) αRag23- and (D) SI-CLBH constructs

All of the recombinant plasmids could be linearised using either EcoRI or NotI. Treatment with both of these enzymes resulted in release of either the hA2aR (0.9 Kb) or SI (0.7 Kb) cDNA from the CLBH vector (α factor – 4.2 Kb; no α factor – 3.9 Kb). Treatment with EcoRI and NdeI or SalI instead of NotI resulted in the release of larger inserts containing either iLOV (0.3 Kb) or LB (0.6 Kb), respectively, in addition to the target protein cDNA.

5.5 The spectral properties of iLOV

Before iLOV could be used as a fluorescence reporter, it was necessary to characterise its spectral properties so that optimal excitation and emission

wavelengths could be determined (Figure 5.8). As the flavin co-factor is itself fluorescent, it was necessary to measure (and compare) the excitation and emission profiles for the same concentration of flavin (with that for iLOV). Lastly, excitation and emission spectra were recorded for yeast membranes to assess the extent of background fluorescence from endogenous flavoproteins.

The iLOV emission spectrum has a distinctive peak at 485 nm with a pronounced shoulder at 525 nm. This is appreciably different from flavin which has a single broad peak from 510 - 540 nm (Figure 5.8 (A)). The fluorescence emission spectrum of the iLOV tag was the same when it was fused to the hA2aR (results not shown). As the spectral properties of iLOV are unchanged when it is attached to another protein iLOV can be used as a fluorescence reporter. Negligible fluorescence from the yeast membranes was observed (Figure 5.8 (A)), indicating that any fluorescence subsequently observed in membranes containing recombinant protein had to come from the iLOV tag. The optimal wavelength for measuring emission from iLOV was 485 nm. However, the shape of the emission spectrum suggested that iLOV fluorescence could be effectively measured in the range 485 – 530 nm.

iLOV has a strong, broad absorption in the region 320-460 nm which gives rise to peaks at 350, 370 and 450 nm in its excitation spectrum (Figure 5.8 (B)). The equivalent spectrum for flavin is subtly but significantly different with just peaks at 355 and 390 nm. These results show that it is possible to selectively excite iLOV over flavin by using light with wavelength around 450 nm.

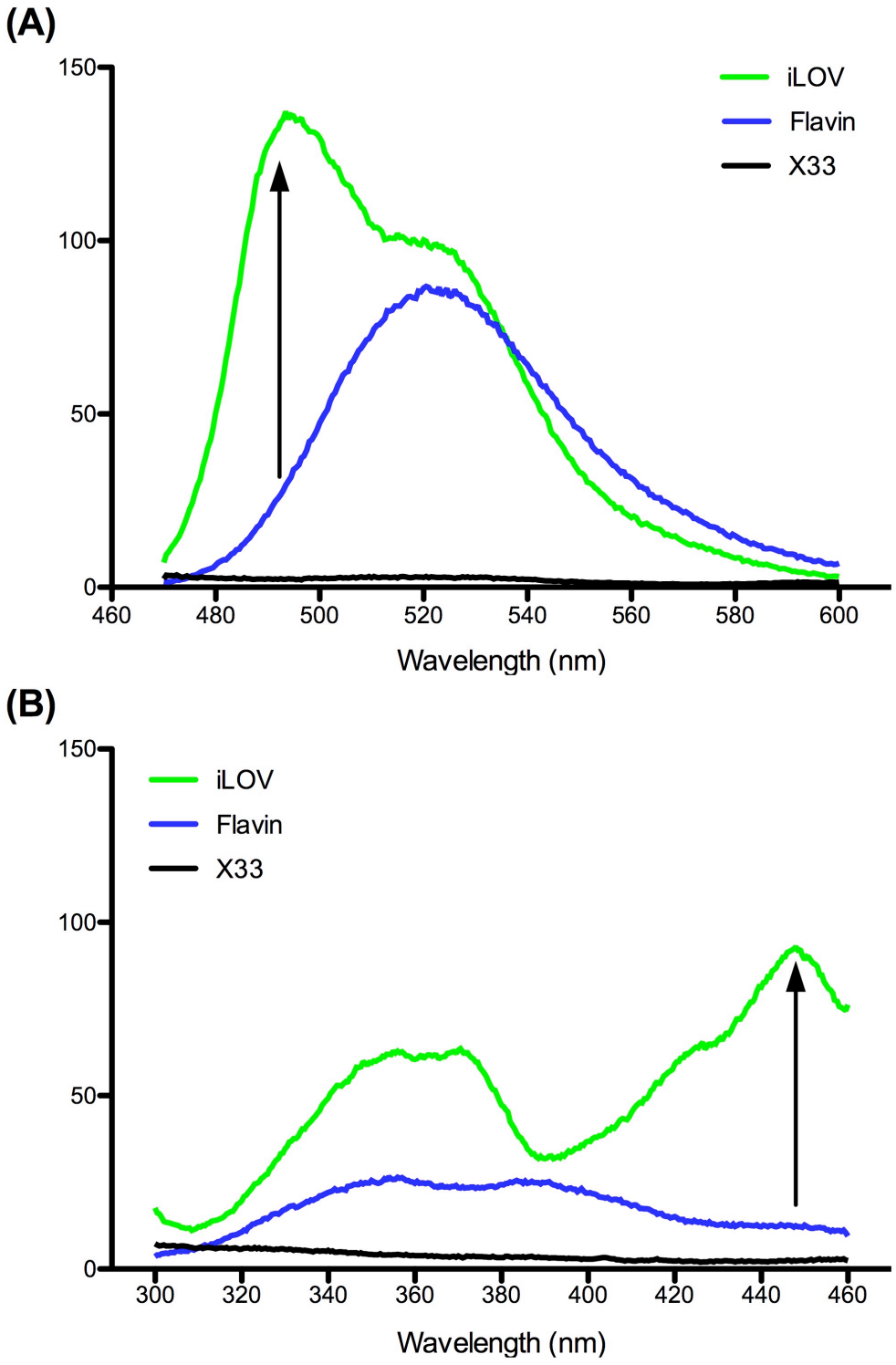


Figure 5.8 The fluorescence properties of equal concentrations of iLOV and flavin compared to native yeast membranes

(A) Emission spectra were recorded from 470 – 600 nm following chromophore excitation at 450 nm. (B) Excitation spectra were recorded from 300 – 460 nm measuring fluorescence emission at 485 nm. The optimal wavelength for exciting iLOV is 450 nm, whereas emission can be effectively measured in the range 485 – 530 nm.

5.6 Clonal selection using iLOV

Once a recombinant expression plasmid has been transformed into *P. pastoris*, it is essential to screen a range of colonies to identify the clone making the highest levels of recombinant protein (as different transformants make different levels of the target protein). Screening is usually done by radioligand binding (where one is available) or by Western blotting. In Chapter 4, X33 transformants expressing HT-trunc dG hA2aR-TB were screened using the radioligand [^{3}H] ZM241385. To accomplish this, small-scale (10 ml) expression trials were performed with eight transformants, membranes isolated for each colony and receptor expression levels determined by single-point radioligand binding assay. Although this process allowed the highest expressing transformant to be identified, this approach is cumbersome and consumed valuable time and resources. Here, clonal selection was performed on X33 colonies that had been transformed with GU-, Rant21- and Rag23-CLBH by measuring iLOV fluorescence in whole cells following induction.

Twelve transformants for each construct were grown up overnight at 30°C in 1 ml cultures of BMGY in a 48-well plate. To induce receptor production, the BMGY was replaced with BMMY and the yeast incubated at 22°C for 24 h. The test expression experiments for each colony were performed in duplicate. The cells were harvested by centrifugation, washed twice and then resuspended in 1 ml of water. The cells were transferred (in duplicate) to a 96 well plate, where iLOV fluorescence was measured ($\lambda_{excitation}$ = 450, $\lambda_{emission}$ = 520 nm). Expression levels are always in the form amount of target protein per total amount of protein. To convert iLOV fluorescence into an expression level, it was necessary to have an accurate measure of the number of cells that contributed to the measured fluorescence. Initially, the optical density (OD_{600}) of each culture was determined in a 96-well plate. However, it proved difficult to obtain reproducible data largely caused by cell sinkage as well as high turbidity (results not shown). Instead, cell biomass was quantitated by chemically permeabilising the cells and then measuring the protein concentration of the lysate. The main benefit of this approach was the improved reproducibility of the cell biomass

data which allowed the expression levels of different constructs to be compared. The clonal selection results for GU-, Rant21- and Rag23-CLBH are presented in Figure 5.9. For all three constructs, the vast majority of the transformants differed in their expression level by a factor of 2-3, similar to that which had been observed in the clonal analysis of HT- trunc dG hA2aR – TB (where receptor amount had been determined by radioligand binding analysis).

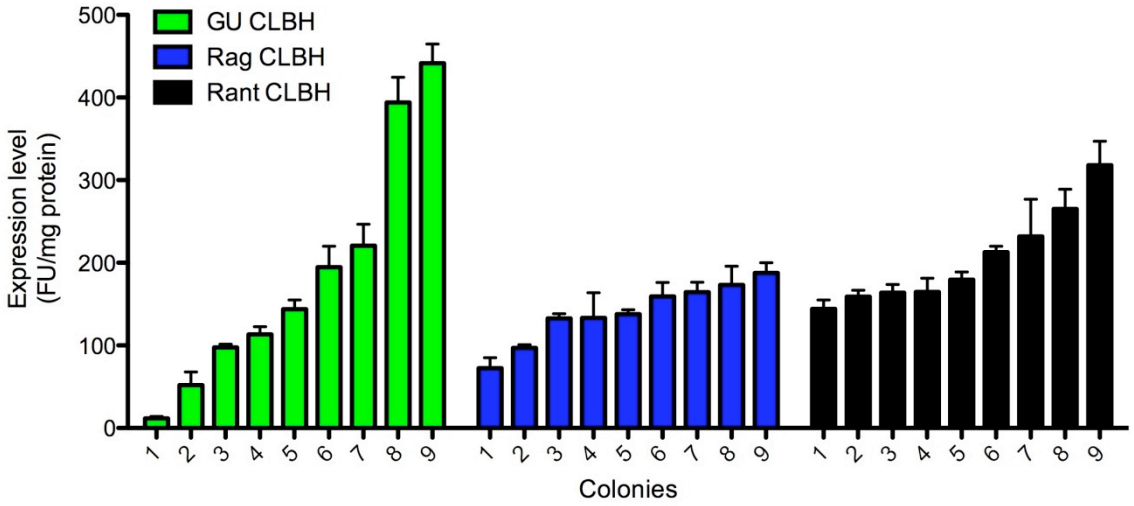


Figure 5.9 Clonal selection using iLOV fluorescence

GU-, Rant21- and Rag23-CLBH transformants were grown up overnight at 30°C in 1 ml cultures of BMGY. Receptor production was induced by replacing the media with BMMY and then incubating the yeast at 22°C for 24 h. Receptor levels were determined by measuring iLOV fluorescence from whole cells. Cell biomass was quantitated by chemically permeabilising the cells and then measuring the protein concentration of the lysate. For all five constructs, the majority of the transformants differed in their expression level by a factor of 2-3.

To address whether or not higher expressing clones could be isolated on plates with high levels of antibiotic, X33 cells that had been transformed with GU-CLBH were grown on YPD plates containing different (0.2, 0.5, 1 and 2 mg/ml) concentration of zeocin. Twelve colonies from each plate were grown in 1 ml cultures, and receptor production induced as described in Section 3.5.1. The expression level was determined by measuring iLOV fluorescence per amount of total cell protein (Figure 5.9).

Although colonies isolated from the plates with higher zeocin concentrations did show the highest expression levels, they did not vary significantly from colonies isolated on plates with lower levels of antibiotic (Figure 5.10). As the expression level profile was essentially the same for all four plates, it suggested that there was no significant benefit in performing clonal selection at high (2 mg/ml) concentrations of zeocin. Rather, the results suggested that the best way to identify a highly expressing transformant is to screen a large number of colonies.

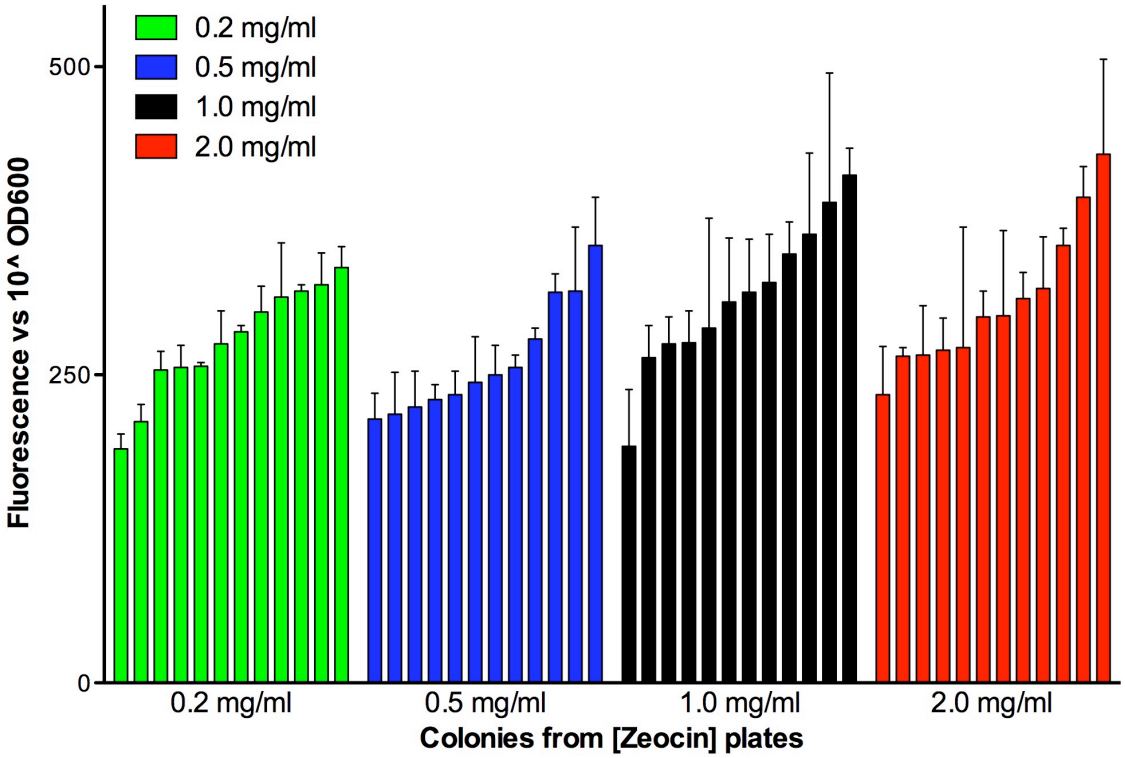


Figure 5.10 The expression level of GU-CLBH in different transformants selected on plates with increasing concentrations of the antibiotic Zeocin

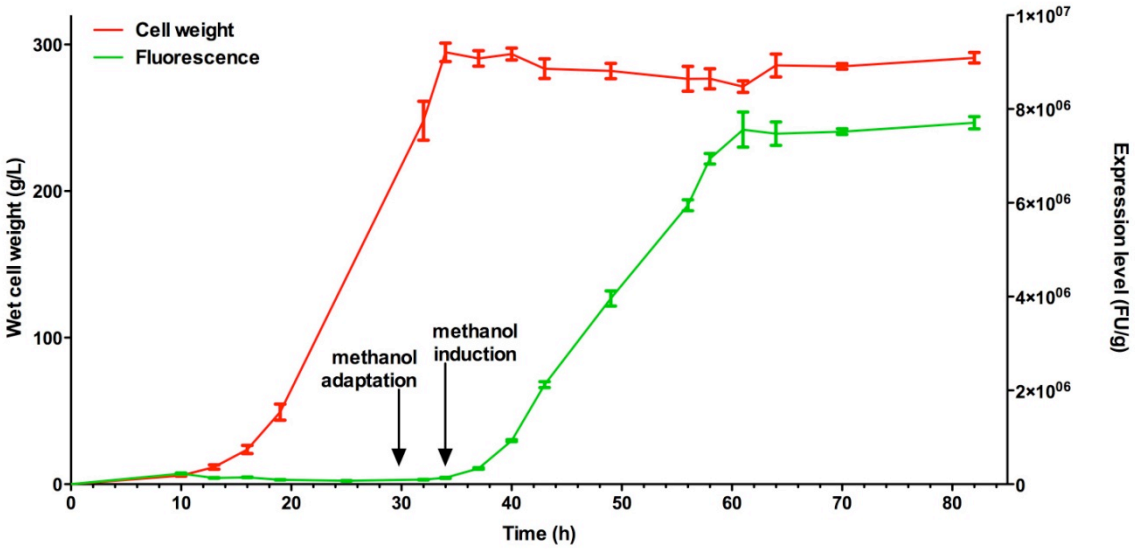
The X33 *P. pastoris* strain was transformed with an expression plasmid containing GU-CLBH. The cells were grown on YPD plates containing different (0.2, 0.5, 1 and 2 mg/ml) concentrations of zeocin. Twelve colonies from each plate were grown in 1 ml cultures, and receptor production induced as described in Section 3.5.1. The expression level was determined by measuring iLOV fluorescence per amount of total cell protein. The expression level profile was essentially the same for all four plates, suggesting that there was no significant benefit of performing clonal selection at high (2 mg/ml) concentrations of zeocin.

5.7 Using iLOV to optimise MP production in fermenter culture

The main benefit of using *P. pastoris* for recombinant MP production is its ability to grow to high cell densities in fermenter culture (where dissolved oxygen content, pH, temperature and feed rates can all be tightly controlled (Cereghino *et al.*, 2002; Jahic *et al.*, 2006)). For each different MP, the basic fermentation protocol needs to be optimised to ensure that maximal levels of the target protein are obtained. If recombinant protein expression levels have to be determined by either radioligand binding or Western Blotting, the optimisation process can be unnecessarily long.

Previously, HT- trunc dG hA2aR – TB production in fermenter culture was optimised by determining the levels of membrane bound receptor at various time points by radioligand binding (Section 4.3.2). Receptor levels were only determined once the fermentation had been completed as it was simplest to make membranes for all of the time-points together before performing radioligand binding. Here, Rag23- and Rant21-CLBH production in fermenter culture was followed by measuring iLOV fluorescence from intact cells at regular intervals (Figure 5.11). Culture biomass was determined by measuring the wet cell weight per ml of culture at each of the time-points. From these two measurements, the receptor expression level could be quantified promptly (generally within 30 minutes of taking the sample). This speed of analysis is beneficial as it allows time-courses of recombinant MP production to be followed in real-time, allowing decisions about when is the best time to stop the fermentation during each run. The iLOV marker is ideal for time-course applications as, unlike GFP, there is no delay period between iLOV assembly and fluorescence from the chromophore. For both constructs, growth on glycerol gave a maximal cell density of *ca.* 300 g wet weight of cells per litre of culture. Maximal receptor expression was achieved by feeding the yeast on methanol for *ca.* 40 h.

(A)



(B)

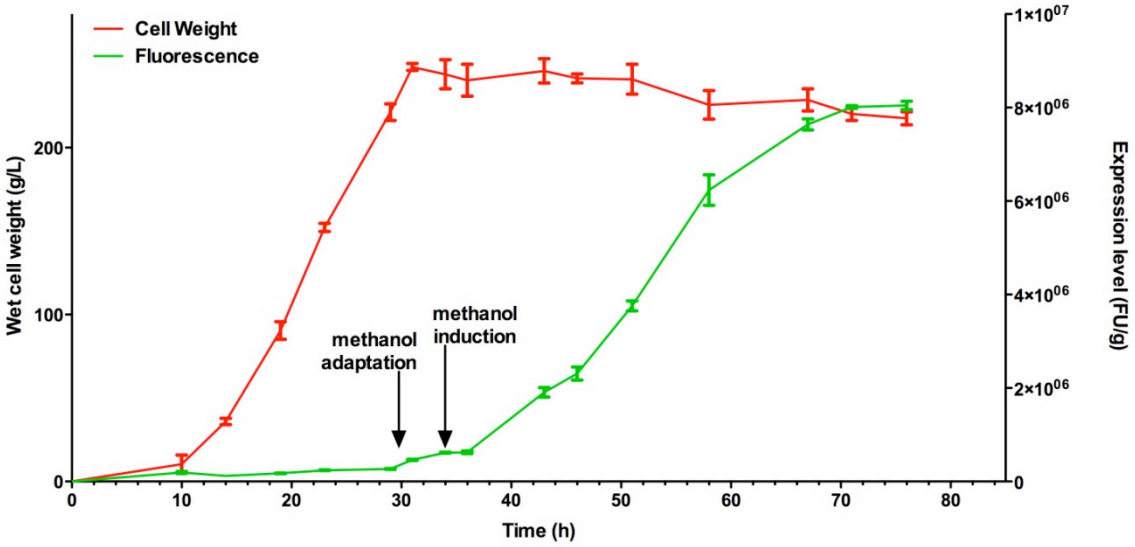


Figure 5.11 Time-course of (A) Rag- and (B) Rant-CLBH production in fermenter culture

Samples were taken at regular intervals throughout the fermentation. For each time point, 1 ml of cells was centrifuged and the pellet weight measured. The cell pellet was washed twice with (to remove free flavin present in the media) and then resuspended in water. iLOV fluorescence was measured using a platereader. Rag23- and Rant21-CLBH production levels were expressed as amount of iLOV fluorescence per amount of wet weight of cells. For both constructs, growth on glycerol gave a maximal cell density of *ca.* 300 g wet weight of cells per litre of culture. Maximal receptor expression was achieved by feeding the yeast on methanol for *ca.* 40 h.

5.8 Monitoring MP aggregation status using iLOV

Although it is relatively easy to assess a protein's purity (e.g. by SDS-PAGE) it is rather more difficult to determine its aggregation status. Often, it is not realised that the target protein has aggregated until it has been purified to homogeneity, rendering it useless for structural biology purposes (Postis *et al.*, 2008). Aggregation is a common problem encountered when working with membrane proteins, and is usually caused by protein instability in the detergent that has been chosen for membrane solubilisation and purification purposes. Recently, the approach of fluorescence size-exclusion chromatography (FSEC) has been developed as a method for monitoring the aggregation status of MPs throughout the purification process following membrane solubilisation. When a pure monodisperse protein is applied to a SEC column, it elutes as a single peak with Gaussian symmetry (Gutmann *et al.*, 2007). By attaching a fluorescent tag to the MP of interest, it is possible to assess whether or not the target protein is monodisperse at any stage of the solubilisation/purification procedure by applying it to a SEC column and then measuring fluorescence (target-protein specific) rather than A_{280} (all proteins) in the elution fractions. The elegance of FSEC is that it allows the aggregation status of recombinant MPs to be monitored even in the presence of other proteins (Kawate and Gouaux, 2006). A single, symmetric fluorescent SEC peak throughout the purification indicates that the target protein is stable. If the SEC elution profile has multiple fluorescent peaks with one of them located in the void volume it suggests that the target protein has become aggregated, and is unsuitable for further purification and crystallisation trials. GFP is the tag most commonly used for FSEC experiments (Kawate and Gouaux, 2006). Here, the usefulness of iLOV to monitor the aggregation status of HT – trunc dG hA2aR – TLB as well as Rag23- and SI-CLBH following detergent solubilisation is described.

Previously, it has been shown that functional solubilisation of the hA2aR can be accomplished using a combination of DDM and cholesterol hemisuccinate (CHS) (Weiss and Grisshammer, 2002; Fraser, 2006). However, CHS is bulky, hydrophobic and has a tendency to drive MP aggregation (Chris Tate, personal

communication). Here, I wanted to investigate whether or not i) HT – trunc dG hA2aR – TLB was monodisperse in detergent solution following solubilisation with DDM/CHS, and ii) the DDM:CHS ratio affected the hA2aR aggregation status. Membranes were solubilised with DDM (final concentration of 2.5%) containing different amounts of CHS (CHS:DDM ratios of 1:5, 1:10, 1:25, 1:50, no CHS). Following centrifugation, the supernatant was applied to a Sepharose C6 gel filtration column in buffer containing 0.05% DDM with CHS added in the appropriate ratio. iLOV fluorescence was measured in all of the elution fractions. The FSEC profiles showed that only a DDM:CHS ratio of 50:1 gave monodisperse receptor following solubilisation (Figure 5.12). At both higher and lower DDM:CHS ratios, the receptor was found to aggregate (Figure 5.12). Receptor that had been solubilised with a 50:1 DDM:CHS ratio was able to bind [^{3}H] ZM241385 in solution (results not shown). These results showed that if enough CHS was included in the solubilisation mixture monodisperse receptor that retained its activity in detergent solution could be obtained. If, however, too much CHS was added, the additive caused the receptor to aggregate. In all of the profiles, a small broad peak with an elution volume around 25 ml was observed. On measuring the fluorescence emission spectrum of this peak, it was found to be flavin. The other, higher MW peaks all had the characteristic LOV emission profile consistent with them being oligomers/aggregates of LOV-tagged receptor. The free flavin comes from denaturation of the iLOV domain on treatment with detergent. As this peak did not obscure monitoring of the receptor's aggregation status, it was concluded that iLOV could be used as an alternative reporter to GFP for FSEC experiments.

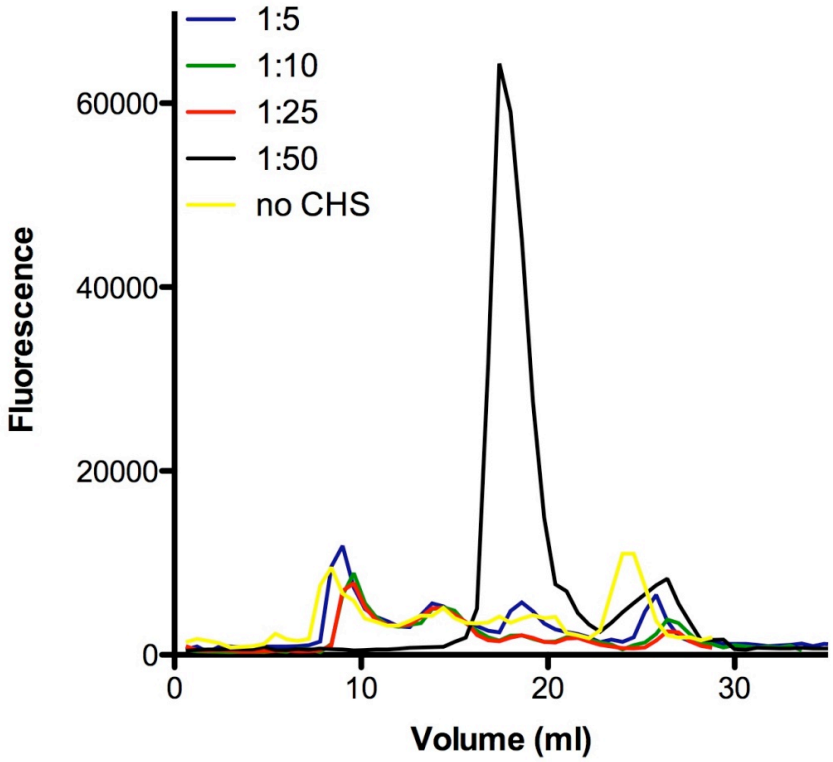


Figure 5.12 The effect of different DDM:CHS ratios on the aggregation status of HT – trunc dG hA2aR - TLB

Membranes containing HT – trunc dG hA2aR – TLB were solubilised with DDM (final concentration of 2.5%) containing different amounts of CHS. Following solubilisation, the samples were centrifuged and the resulting supernatants analysed by FSEC. iLOV fluorescence in each of the elution fractions was measured using a platereader.

Having identified conditions that gave monodisperse receptor following solubilisation, a Ni^{2+} purification step was performed. (This is detailed in Section). An aliquot of the partially purified receptor was applied to the gel filtration column, and the fluorescence content of the elution fractions measured (Figure 5.13). Disappointingly, a single broad peak that eluted at *ca.* 12 ml was observed (that was appreciably different from the sharp peak around 20 ml that had been seen in the post-solubilisation sample), showing that purification with the Ni^{2+} column had caused the receptor to aggregate. This effect may have of been caused by one of the chemicals contained within the purification buffers used with the Ni^{2+} column. To explore this possibility, a series of test solubilisation experiments were performed in the presence of either NaCl, imidazole or histidine. The detergent-solubilised material was then applied to the gel-filtration column equilibrated in a buffer containing the

chemical under investigation. Fluorescent elution profiles are shown in Figure 5.14. The results showed that none of the reagents tested, in and of themselves, caused the receptor to aggregate. Rather it looked like the receptor became aggregated during the time it was bound to the column perhaps through a local enrichment in the concentration of CHS. Despite considerable effort to overcome this problem of aggregation, none could be found. In all subsequent hA2aR solubilisation and purification experiments conformationally-fixed thermostable mutants were used (Magnani *et al.*, 2008).

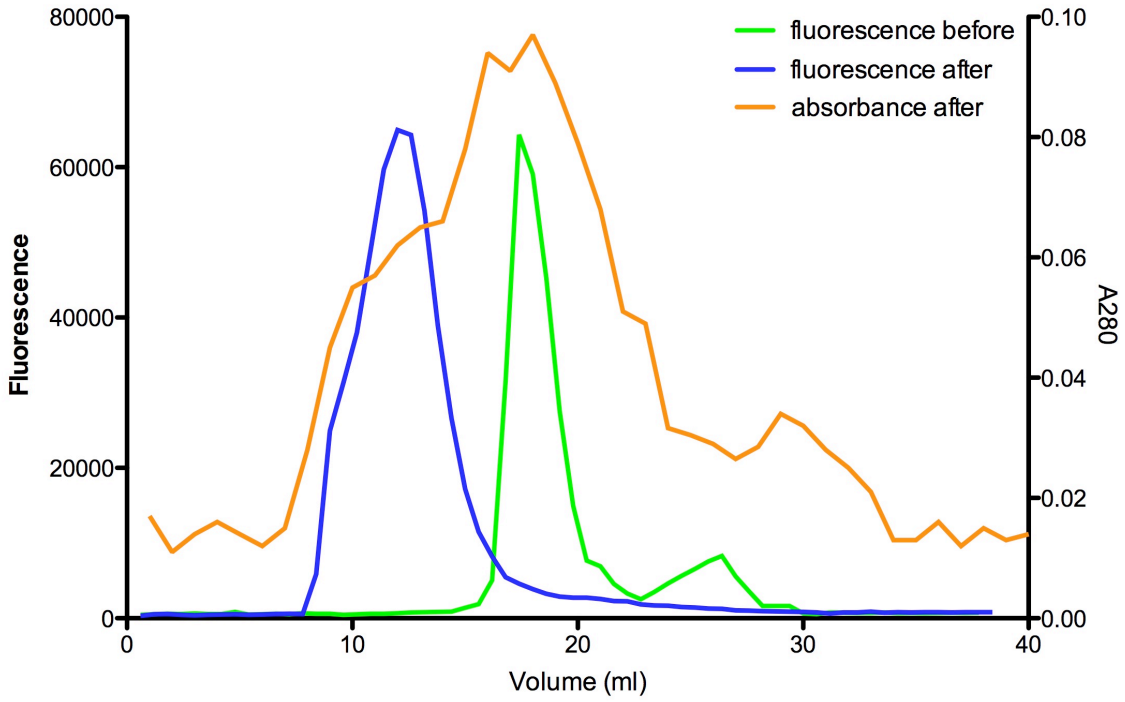


Figure 5.13 Analysis of HT - trunc dG hA2aR - TLB before and after Ni^{2+}-affinity purification by FSEC and SEC

The hA2aR was solubilised from yeast membranes using a combination of DDM (final concentration 2.5%) and CHS (ratio 50:1). The receptor was purified by Ni^{2+} chromatography as described in Section 3.12. Aliquots containing the hA2aR before and after purification were applied to a Sepharose C6 gel-filtration column, and the fluorescence elution profile measured for both samples (before – green; after blue). A single sharp peak at *ca.* 20 ml was observed for the post-solubilisation sample. However, after Ni^{2+} purification, the peak was shifted to an earlier elution time (12 ml) and was considerably broader suggesting that during the purification step the receptor had become aggregated. This was confirmed by the absorbance profile following purification.

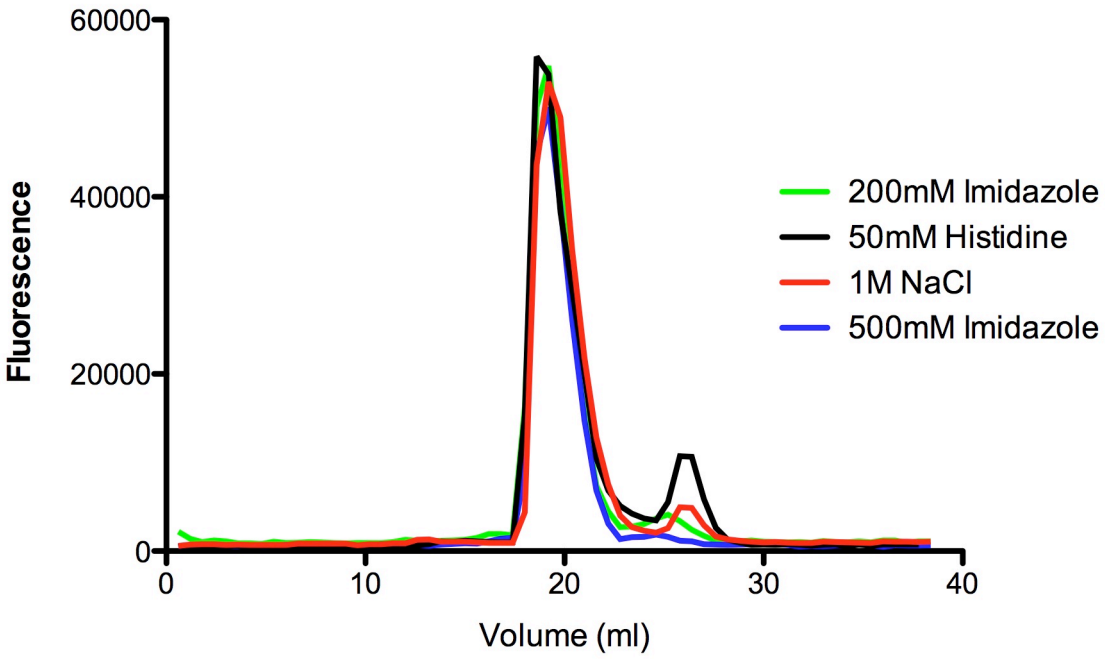


Figure 5.14 The effects of a range of buffer components on the aggregation status of HT-trunc dG hA2aR – TLB in detergent solution

A series of test solubilisation experiments were performed using a combination of DDM (final concentration of 2.5%) and CHS (50:1) in the presence of either NaCl, imidazole or histidine. The detergent-solubilised material was then applied to the gel-filtration column equilibrated in a buffer containing the chemical under investigation. The elution profiles showed that none of the reagents tested, in and of themselves, caused the receptor to aggregate.

The hA2aR thermostable mutants Rant21 and Rag23 can bind [^{3}H] ZM241385 and [^{3}H] NECA in short chain detergents, respectively (Magnani *et al.*, 2008). Here, I wanted to explore whether or not these mutants were monodisperse in a range of different detergents following solubilisation. As the hA2aR has already had its structure solved with ZM241385 bound (Jaakola *et al.*, 2008), I focussed on the hA2aR agonist mutant Rag23. Membranes containing HT-Rag23-TLB were incubated with a saturating (100 μM) concentration of NECA and then solubilised with either DDM (alkyl chain C=12), DM (C=10), NM (C=9) or βOG (C=8). After centrifugation, the supernatants were applied to the gel filtration column pre-equilibrated in buffer containing NECA as well as the appropriate detergent. Rag23 was found to be monodisperse in all of the non-ionic detergents tested even in βOG (Figure 5.15). Furthermore, monodisperse peaks were also obtained (for all of the detergents tested) when the receptor containing supernatants

were applied to the column 48 h after solubilisation, demonstrating that Rag23 has prolonged stability in detergent solution (results not shown). Unfortunately, due to time constraints, it was not possible to characterise the Rant21 mutant by FSEC.

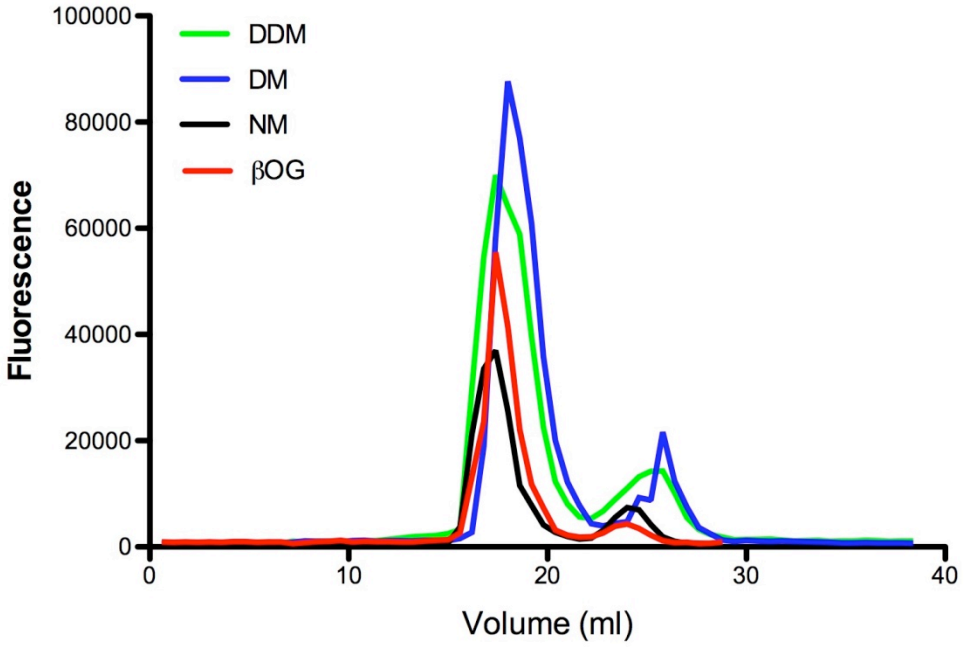


Figure 5.15 FSEC analysis of HT-Rag23-TLB

Membranes containing HT-Rag23-TLB were incubated with a saturating concentration (100 µM) of NECA before they were solubilised using detergents with different alkyl chain lengths – DDM (C = 12), DM (10), NM (9) and βOG (8). Even with the harshest detergent βOG, monodisperse receptor was obtained. This experiment suggested that Rag23 is an excellent candidate for structural studies as it is remarkably stable even in short chain detergents.

In order to determine the stability of SI in detergent solution membranes containing SI-CLBH in sodium phosphate buffer were solubilised using either DDM, DM, NM or βOG. After centrifugation, the supernatants were applied to the gel filtration column pre-equilibrated in sodium phosphate buffer containing the appropriate detergent. SI was found to be monodisperse in both DDM and DM (Figure 5.15). In NM, although a single fluorescent peak was observed it was appreciably broader than that which had been observed in either DDM or DM suggesting that the enzyme was starting to lose its stability in NM (Figure 5.15). When SI was solubilised with βOG, the protein was completely aggregated

(Figure 5.15). Remarkably, for a wild-type protein, SI displays good stability in the short-chain detergent DM (and to a lesser extent NM) suggesting that this protein may be an excellent candidate for structural studies.

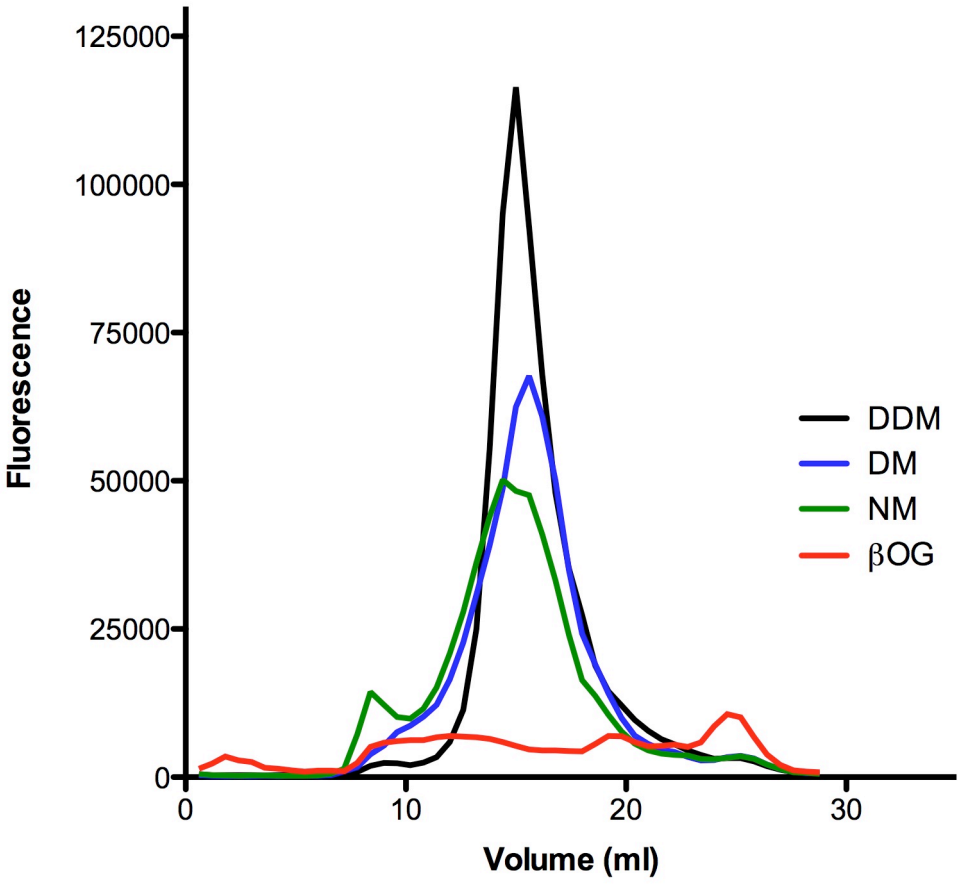


Figure 5.16 FSEC analysis of SI-CLBH

Membranes containing SI-CLBH in sodium phosphate buffer were solubilised with either DDM, DM, NM or βOG. Remarkably for a native protein, SI was monodisperse in both DDM and DM. By NM, however, the fluorescent peak was starting to broaden suggesting that the enzyme was starting to lose its stability in this detergent. When SI was solubilised with βOG, the protein was completely aggregated.

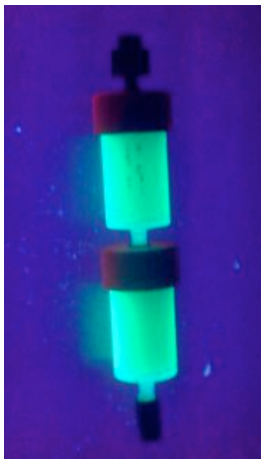
5.9 iLOV as a purification aid

All protein purification steps involve binding of the target protein to a resin, washing the column or beads and then selectively eluting the protein of interest.

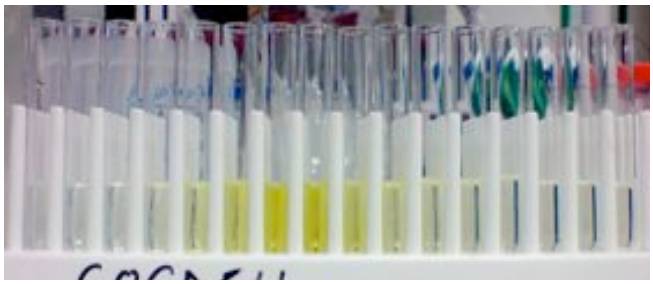
When performing any purification step, it is necessary to know how much of the target protein has bound to the column as well as which elution fractions contain the protein of interest. Such basic information can be hard to obtain if there is not a simple assay to measure the activity/presence of the protein of interest. Here, I describe how the iLOV tag can be used to overcome these purification difficulties.

Yeast membranes containing HT – trunc dG hA2aR – TLB were solubilised with 2.5%:0.05% DDM:CHS according to Section 3.103.12.1. Following centrifugation, the soluble fraction containing receptor was passed over a Ni^{2+} column. Receptor binding was monitored by comparing the fluorescence of the supernatant before and after passage over the column. Receptor binding to the Ni^{2+} column was readily visualised by illuminating it with UV light (Figure 5.17 (A)). The resin was washed with buffer containing 500 mM NaCl and 75 mM imidazole according to Section 3.13. This was to remove any contaminant proteins that had non-specifically-bound to the column. The receptor was released from the column by passing buffer over it containing 300 mM imidazole, and fractions collected. Those containing iLOV-tagged receptor were readily recognised, as they were yellow in the light and fluorescent on exposure to UV (Figure 5.17 (B-C)). The fractions containing receptor were pooled and analysed by FSEC which showed that the hA2aR had become aggregated during the Ni^{2+} purification step (Section ?).

(A)



(B)



(C)

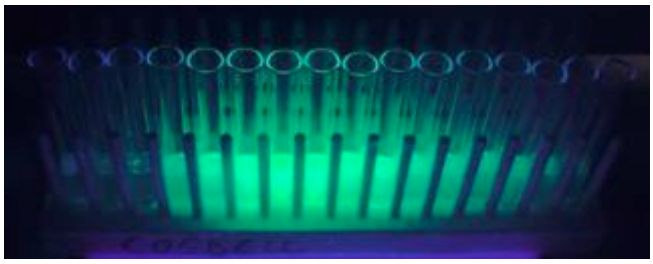


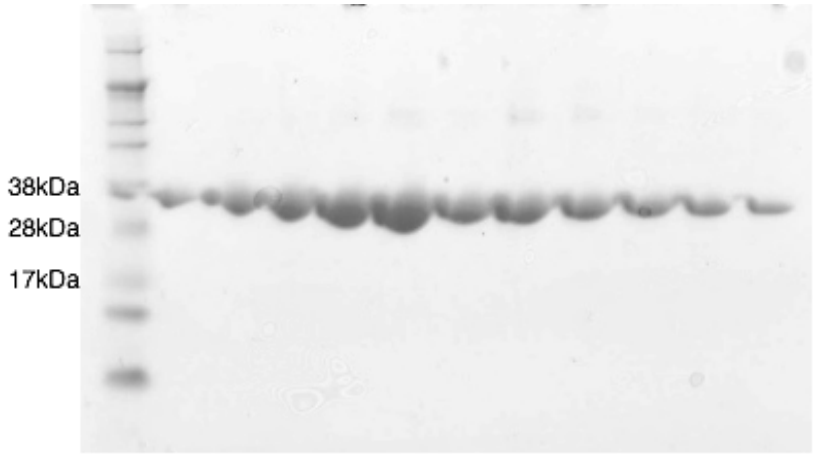
Figure 5.17 Visualisation of iLOV-tagged hA2aR during Ni^{2+}-affinity purification
(A) iLOV-tagged hA2aR bound to a Ni^{2+} column has a fluorescent green colour under UV-light. The receptor was eluted from the column using 300 mM imidazole. Fractions containing HT-trunc dG hA2aR -TLB are (B) yellow in colour and (C) fluorescent green following UV-irradiation.

5.10 Immobilisation of fluorescently-tagged hA2aR using streptavidin beads

The biotinylation domain was primarily included in the in-house expression vector as a second affinity tag for protein purification purposes. Biotinylated proteins are rare in nature with only one known in *E. coli* (Fall 1979), five in *S. cerevisiae* (Lim *et al*. 1987) and four in mammalian cells (Robinson *et al*. 1983; Chandler and Ballard 1988). As the levels of endogenous biotinylated proteins in *P. pastoris* are very low, it was envisaged that purification using this tag would result in a substantial enrichment of the target protein. Although it has been previously shown that the biotin acceptor domain from the *P. shermanii* transcarboxylase can be biotinylated endogenously by *S. cerevisiae* (Cronan, 1990) to ensure that all of the fusion domains were biotinylated Ni^{2+} enriched HT- GU hA2aR-TLB was incubated at 4°C overnight with recombinant *E. coli*

biotin ligase *in vitro* (Figure 5.18 (A)). After biotinylation was complete, the fluorescently tagged receptor was desalted before application to streptavidin beads. All of the fluorescently-labelled receptor bound to the beads (Figure 5.18 (B)), showing that all of the fusion domains had been biotinylated *in vitro*. As the interaction between biotin and streptavidin is incredibly tight it was not possible to elute bound receptor from the beads using free biotin. Although the biotin-streptavidin bond can be broken at low pH (<3) these conditions were far too harsh to expose the receptor to. Rather, it was thought that the best way to remove the receptor from the beads was through tag cleavage. Initial studies using TEV protease were unable to get satisfactory cleavage, at either end of the recombinant protein, as the TEV that I made in the lab had poor solubility (results not shown). However, the usefulness of this immobilization technique has been confirmed using constructs containing the C3 protease recognition site (data not shown).

(A)



(B)

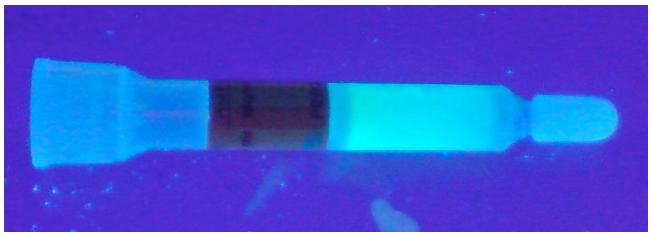


Figure 5.18 *In vitro* treatment of HT – GU hA2aR – TLB with biotin ligase allows the receptor to be immobilised on streptavidin beads

(A) SDS-PAGE of purified recombinant *E. coli* biotin ligase, (B) Following *in vitro* biotinylation of Ni^{2+}-enriched hA2aR the receptor could be readily immobilised on streptavidin beads. This was visualised by looking at fluorescence from the iLOV tag under a UV light source.

5.11 Conclusion

Having made a series of related expression constructs, optimal (CLBH) vectors were created that gave high level production of two different human membrane proteins (one N_{in} – hSI; one N_{out} – hA2aR) that could be rapidly purified to homogeneity using a generic protocol. The biotin acceptor site could be fully biotinylated *in vitro* using recombinantly expressed BirA. Due to the high expression levels achieved with both the hA2aR and hSI in fermenter culture, this tag was not required for their successful purification. However, the ability to immobilise any target protein by its C-terminus is an incredibly useful feature of the expression vector and should be compatible with drug discovery approaches such as Biacore and target-immobilised NMR spectroscopy. The inclusion of a fluorescent reporter C-terminal to the protein of interest helped greatly to assess the expression, detergent-extraction and purification processes. Finally, C3 protease was found to be a considerably better enzyme for removing the expression/purification tags from the recombinant protein than TEV as the TEV made in our lab had poor solubility. For the sake of simplicity, the final optimised expression vector had the iLOV-biotin-His10 (LBH) tags fused to the C-terminus of the recombinant protein. This allowed removal of all the tags in a single cleavage event with C3. The major disadvantage of having all of the purification tags at one end of the recombinant protein as compared to tags at both ends is that there is no guarantee of purifying protein that is completely intact. When all of the purification tags are at one end of the protein, it is possible to purify a proteolytic fragment of the target protein. However, when there are tags at either end of the target protein, by performing a two-step purification procedure it is possible to isolate protein that has both its N- and C-termini intact.

The iLOV fluorescent reporter protein is a useful tool for optimising processes such as clonal selection, protein production in fermenter culture, detergent and construct screening as well as protein purification. The primary benefits of using iLOV are that it saves time, labour and cost as well as giving important insights into the stability of the target MP immediately after solubilisation as well as

during the purification process. This second advantage is considerable given the propensity of MPs to aggregate. An unexpected but incredibly useful property of the iLOV tag was discovered during the FSEC experiments where all of the elution profiles had a "late" peak consisting of flavin (that had been released on breakdown of the iLOV tag). The size of the flavin peak gave a direct measure of the amount of target protein denaturation that had occurred on treating the membranes with detergent. The presence of a large flavin peak in a FSEC experiment (following either detergent-solubilisation or -exchange) suggests that the target protein is unstable in the chosen detergent whereas the absence of a flavin peak in the elution profile suggests that the iLOV-tagged MP is stable. In short, by monitoring the free-flavin content of iLOV-tagged MPs in detergent solution it should be easy to identify those detergents in which the MP is unstable/has become denatured.

One disadvantage of iLOV compared to GFP is that it is not compatible with in-gel fluorescence experiments. As fluorescence from the iLOV chromophore comes from the flavin co-factor, when iLOV tagged receptor is treated with SDS, the iLOV domain denatures rendering it non-fluorescent. This means that iLOV-tagged proteins on a SDS-PAGE cannot be visualised by placing it under a UV source. This is different from GFP-tagged proteins which retain their structure and fluorescence even after treatment with SDS (Drew *et al.*, 2006).

This Chapter also highlights the usefulness of conformationally-fixed thermostable mutants of GPCRs for structural biology purposes. Rag23 was stable in DDM without any CHS present. This is important as CHS has a tendency to drive MP aggregation. Furthermore, as Rag23 was monodisperse in a series of short-chain detergents it suggested that this mutant will make an excellent candidate for structural trials. Although time-consuming to make, thermostable mutants will likely allow structures of numerous proteins to be solved in the future that otherwise would likely to have proved to be intractable.

6 Structural studies on Rag23 and sterol isomerase

6.1 Introduction

Having created an optimal expression vector for making fluorescently-tagged mammalian membrane proteins in *P. pastoris*, I now wanted to use it to obtain milligram quantities of pure Rag23 and SI for structural studies. Although I could have worked on any of the three hA2aR constructs that I had made (Section 5.4) I chose to focus on Rag23 as it is the one that I considered most biologically interesting. Currently, the only structures of ligand-activated GPCRs in the PDB are in the inactive state. The next major challenge for structural biologists is to solve the structure of a receptor in an active conformation. Although Rag23 is a thermostable mutant, it is considerably less stable than its antagonist equivalent Rant21 (Magnani *et al.*, 2008). I was, therefore, interested to know if Rag23 could be purified to homogeneity in a mild-detergent (such as DDM) without the receptor becoming aggregated. As SI is an integral membrane enzyme, it was thought that this protein would be significantly more stable than the hA2aR. To test this hypothesis, I wanted to determine whether or not wild-type SI could be purified to homogeneity while retaining its monodispersity in short- (e.g. DM) as well as long-chain detergents (DDM). This Chapter describes my efforts to purify and biophysically characterise both Rag23 and SI. In addition, crystallisation trials (detergent and lipidic sponge phase) were performed with SI, and the results presented.

6.2 Purification of Rag23 in DDM

Aα-Rag23-CLBH was produced in fermenter cultures of *P. pastoris* as described in Section 3.5.4. Membranes were isolated according to Section 3.6.3, and saturation radioligand binding analysis performed using [^{3}H] NECA (Figure 6.1). A

B_{max} value of *ca.* 26 pmol/mg protein was obtained with a K_d value of *ca.* 3 μM. Accurate B_{max} or K_d values for [^{3}H] NECA binding to my Rag23 construct could not be obtained as the experimental data suggested that complete receptor saturation would require the addition of between 10-50 μM radioligand – this was neither practical nor affordable. This result was both disappointing and unexpected. The K_d of the original Rag23 construct for [^{3}H] NECA is *ca.* 21 nM (Magnani *et al.*, 2008), displaying two-orders of magnitude higher-affinity for NECA than the version I had made. The most likely explanation for the difference in pharmacological properties between constructs is that my version of Rag23 has a substantial part of IL3 removed, trapping the receptor in a low-affinity state for NECA. To test this hypothesis, the original Rag23 construct (which has its third intracellular loop intact) needs to be cloned into the Aα-CLBH expression vector and, following transformation, the receptor's pharmacological properties in yeast membranes determined. Unfortunately, due to the lack of time, it was not possible to do this.

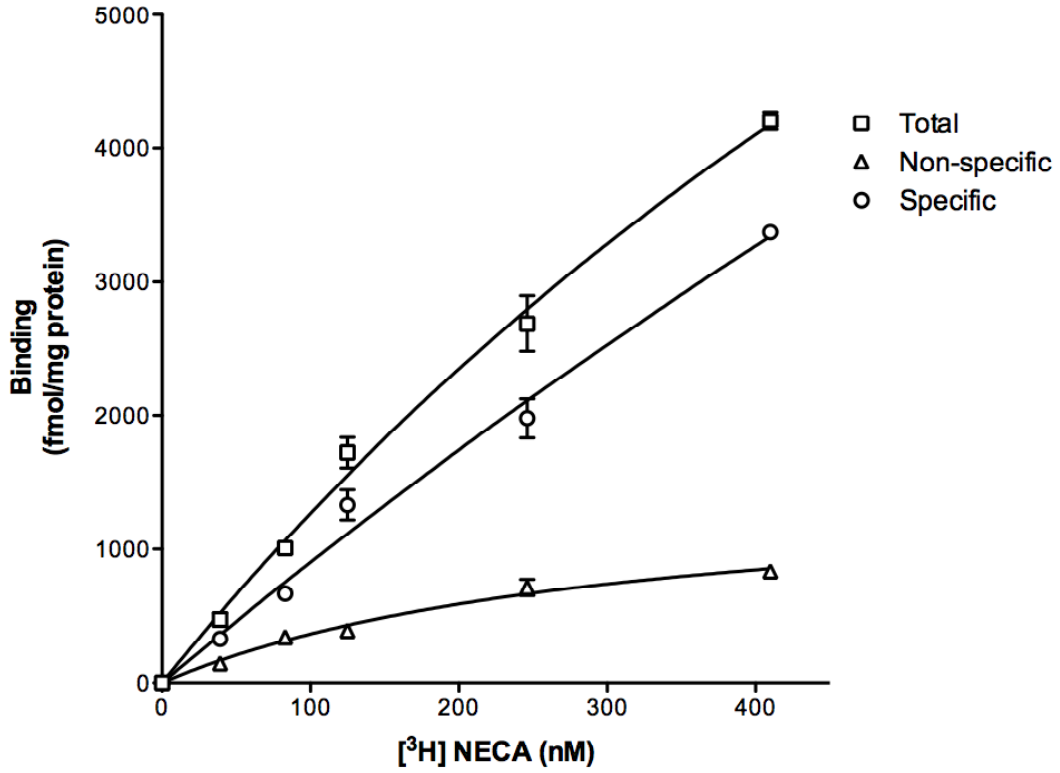


Figure 6.1 Saturation binding of [^{3}H] NECA to Rag23-CLBH in *P. pastoris* membranes
Saturation binding analysis was performed according to Section 3.8.2. The data was fitted using GraphPad Prism Software. A B_{max} value of *ca.* 26 pmol/mg protein was obtained with a K_d value of *ca.* 3 μM.

Even though the Rag23 construct I had made did not have the expected pharmacology, I still went onto to purify it. Prior to solubilisation, yeast membranes containing Rag23-CLBH were stirred with 100 µM NECA for 1h at 4°C. This was to ensure that all of the receptor binding sites were occupied with agonist prior to detergent treatment (as it has been previously shown that NECA confers an extra 16°C of thermostability on Rag23 in detergent solution (Magnani *et al.*, 2008)). The receptor was released from the membranes by treatment with DDM. Following centrifugation, the supernatant containing receptor was applied to a Ni^{2+} column. The material that bound was washed with HEPES buffer containing 100 mM NaCl and 75 mM imidazole (to remove His rich proteins). It was not possible to wash the column with buffer containing a high (e.g. 500 mM) salt concentration as NaCl abrogates NECA binding to Rag23 (results not shown). Finally, the protein was eluted using 500 mM imidazole. Those fractions that were fluorescent (i.e. those containing iLOV tagged receptor) were analysed by SDS-PAGE (Figure 6.2). At this stage, Rag23 was only partially pure and it was not readily obvious which band corresponded to Rag23-CLBH. In contrast, after HT- trunc dG hA2aR - TB had been Ni^{2+}-purified the receptor was largely pure with a single band clearly visible at *ca.* 49 kDa (Figure 4.10). The difference in purity of the Rag23 and trunc dG hA2aR constructs following Ni^{2+}-affinity purification reflects their relative abundance in yeast membranes (compare Figure 6.1 with Figure 4.7). HT – trunc dG hA2aR- TB had an expression level 4-8 times greater than that of Rag23-CLBH explaining why this receptor construct was significantly purer than Rag23 following a single Ni^{2+}-affinity chromatography step. An aliquot of partially-purified Rag23-CLBH was analysed by FSEC, and a single sharp symmetric peak observed (Figure 6.3). This showed that my version of Rag23 was sufficiently stable in the presence of NECA to be purified using DDM as detergent.

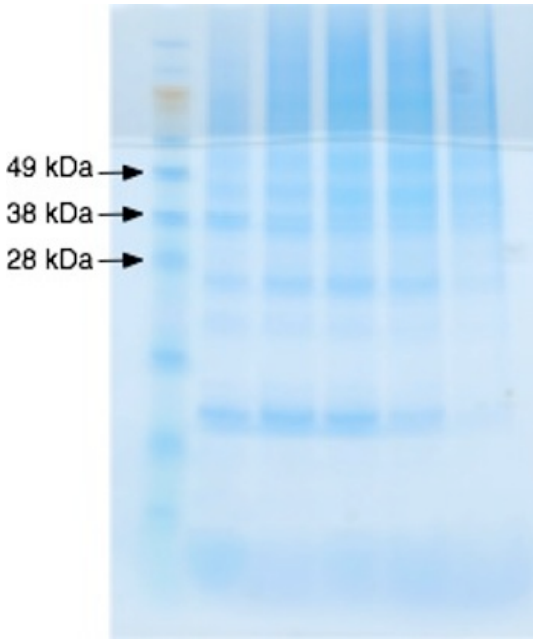


Figure 6.2 SDS-PAGE of Ni^{2+}-affinity purified Rag23-CLBH

Elution fractions from the Ni^{2+} column that were fluorescent (i.e. those that contained iLOV-tagged Rag23) were analysed by SDS-PAGE. At this stage in the purification process, it was not readily obvious which band corresponed to Rag23-CLBH.

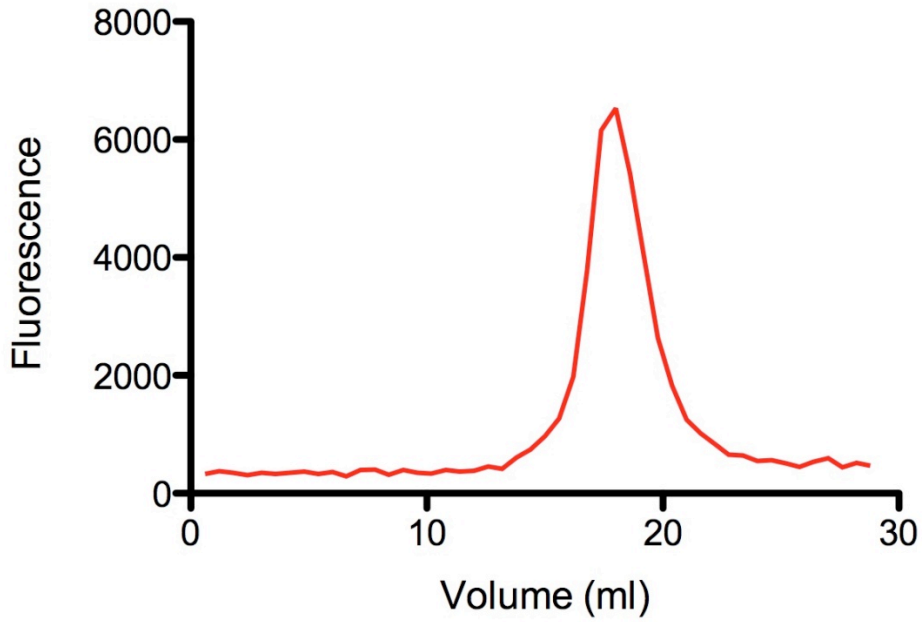


Figure 6.3 FSEC elution profile following application of Ni^{2+}-affinity purified Rag23-CLBH to a Sepharose C6 gel filtration column

A single, symmetric fluorescent peak was observed demonstrating that monodisperse Rag23 could be purified in the presence of NECA using DDM as detergent.

Pooled Rag23-CLBH was desalted to remove the imidazole before it was cleaved overnight at 4°C with C3 protease. The following morning, a reverse Ni^{2+} purification step was performed. It was envisaged that when the cleavage mixture was applied to the Ni^{2+} beads the cleaved CLBH tag, the His-tagged C3 protease as well as all of the contaminant proteins that bound to the original Ni^{2+} column would bind to the beads, leaving largely Rag23 in solution. The reverse-purified sample was concentrated and applied to a Sepharose C6 gel-filtration column. The peak fraction from the gel filtration column was run on a gel and a single faint band at *ca.* 26 kDa was observed (Figure 6.4), consistent with the receptor's MW (32.8 kDa). If time had allowed, the band thought to contain Rag23 would have been cut out from the gel and sent for analysis by tandem mass spectrometry (MS-MS) to unequivocally identify the purified protein.

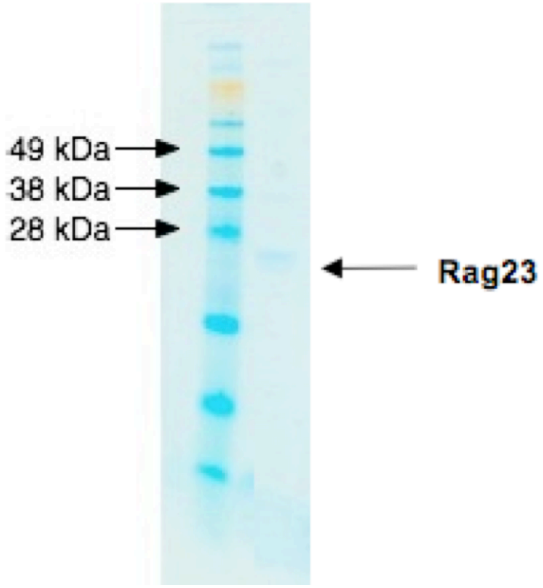


Figure 6.4 SDS-PAGE of pure Rag23

Following treatment of partially-purified Rag23-CLBH (Figure 6.2) with C3 protease, the sample was reverse-purified using Ni^{2+} beads. The supernatant containing Rag23 was concentrated and applied to a Sepharose C6 gel-filtration column. The peak elution fraction was analysed by SDS-PAGE and a single faint band at *ca.* 26 kDa corresponding to Rag23 was observed.

Here, it has been shown that conformationally-fixed GPCR mutants that preferentially binds agonists (such as Rag23) can be purified to homogeneity with monodispersity as long as they are sufficiently thermostable to prevent

aggregation in detergent solution. If time had allowed, I would have liked to have repeated this purification using the shorter chain detergent DM. If monodisperse Rag23 could be purified in DM, it would significantly increase the exposed polar surface area of the receptor available for making lattice contacts, potentially allowing the formation of ordered protein crystals from detergent solution. However, the results in this Section also highlight the need to perform rigorous functional analysis of any mutants that are made for structural studies.

6.3 Large-scale production of SI

The A-SI-CLBH vector (Section ?) was introduced into the *P. pastoris* expression strain X33 by electroporation. Transformants were selected by growing the yeast on plates containing zeocin. Twelve colonies were grown up in 1 ml cultures of BMGY (basal media) overnight at 30°C. The following morning, the BMGY was replaced with 1ml of BMMY (induction media), and the cultures left for 24 h at 22°C. The production level of SI for each clone was determined by measuring the amount of LOV fluorescence (FU) (emitted by whole cells) per milligram of protein (following chemical cell lysis) (Figure 6.5). Of the twelve colonies tested, one did not grow. In the eleven which did grow, a 2.5 fold difference in the SI production level was observed between transformants (Figure 6.5). As the expression levels seen for individual SI clones were considerably higher than those seen for native hA2aR (an increase in fluorescence of 1.2 to 8-fold for the same amount of protein measured) (compare Figure 6.5 to Figure 5.9), it suggested that wild-type SI was more stable than the receptor, and may make a good target for structural studies. The transformant making the highest levels of SI was retained for subsequent experiments.

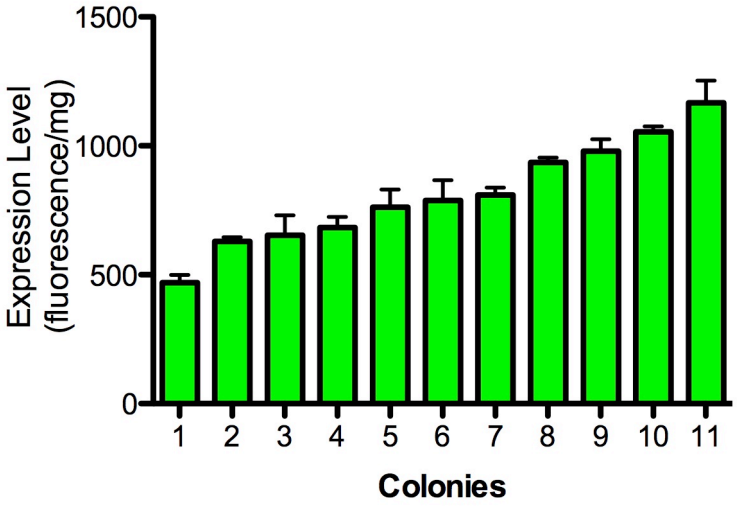


Figure 6.5 Production of sterol isomerase in different *P. pastoris* transformants
Twelve transformants were grown for 24h in BMGY at 30°C followed by another 24h in BMMY at 22°C. SI expression levels for the eleven colonies that grew were determined by measuring the amount of LOV fluorescence (FU) (emitted by whole cells) per milligram of protein (following chemical cell lysis). A 2.5 fold difference in the SI production level was observed between transformants.

Next, the SI clone that had been selected was grown in fermenter culture according to the protocol that had been optimised for the hA2aR. Samples were taken at selected time-points throughout the fermentation to determine both the culture's biomass (by measuring wet cell weight) and the SI expression level (by measuring iLOV fluorescence from whole cells) (Figure 6.6). Biomass levels of *ca.* 300 g wet weight of cells per litre of culture were obtained, similar to that observed in hA2aR fermentations. Maximal SI expression levels were in the range 1 x 10^7 fluorescence units per milligram of cellular protein (Figure 6.6). This compares with 8 x 10^6 FU/mg observed for both Rant21 and Rag23 (Figure 5.11).

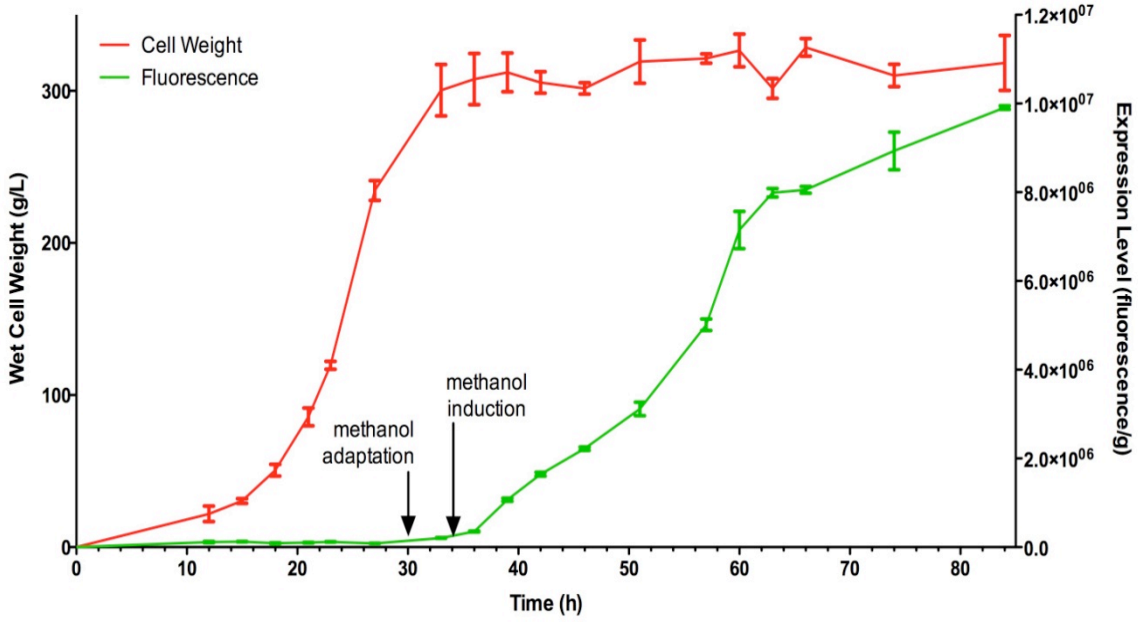


Figure 6.6 Time-course of sterol isomerase expression in fermenter culture

SI-CLBH was made in fermenter culture using the optimised protocol developed for hA2aR production. Wet cell weight and SI expression levels were determined at regular intervals. Typically, biomass levels of *ca.* 300g wet weight of cells per litre of culture were obtained with SI expression levels in the region of 1 x 10^7 fluorescence units per milligram of protein.

6.4 Purification of SI in DDM using HEPES buffer

Membranes containing SI-CLBH were prepared from yeast cells grown in fermenter culture (Section 3.6.3). The membranes were re-suspended in HEPES-buffer pH 7.4 and then solubilised with DDM. Following centrifugation, the supernatant fraction (containing SI) was applied to a Ni^{2+} column where the SI bound to it via its His10 tag. Contaminant proteins were removed from the column by washing it with HEPES buffer containing 500 mM NaCl (to remove those contaminant proteins that bound to the Ni^{2+} resin via ionic interactions) and 75 mM imidazole. SI-CLBH was eluted from the column using an imidazole gradient (Figure 6.7). Those fractions containing SI-CLBH (as determined by measuring iLOV fluorescence) were analysed by SDS-PAGE to determine how pure the sample was after this purification step (Figure 6.8). The gel showed that the SI-CLBH was largely pure at this stage with a major band clearly visible

at *ca*. 44 kDa (SI-CLBH has a predicted MW of 45 kDa). The SI-CLBH expression level was so high that it essentially prevented any contaminant host proteins from binding to the Ni^{2+} column, explaining why it was largely pure following a single chromatographic step.

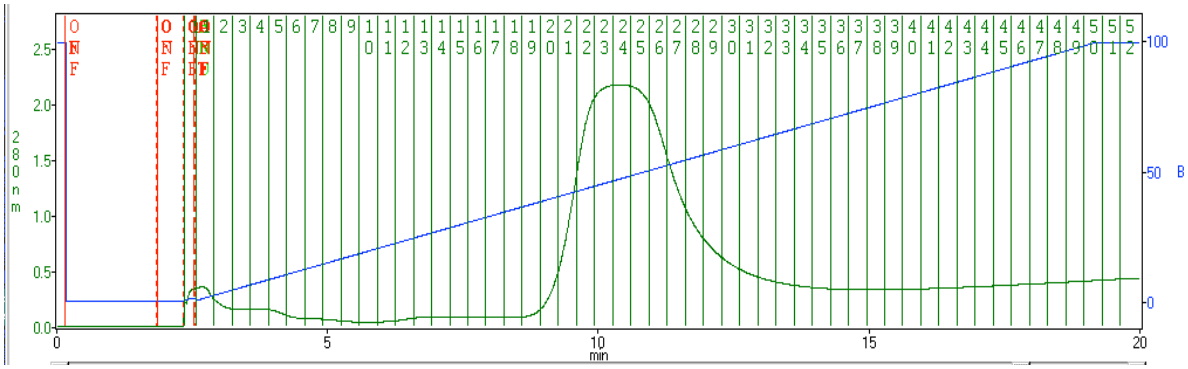


Figure 6.7 Elution of SI-CLBH from a Ni2+ column using an imidazole gradient

When detergent-solubilised membranes were passed over a Ni^{2+} column the recombinant SI-CLBH bound to the Ni^{2+} resin via its His10 tag. After the column had been washed, SI-CLBH was eluted from the column using a 0-1M imidazole gradient. A single, fluorescent peak corresponding to SI-CLBH was released by *ca.* 500 mM imidazole.

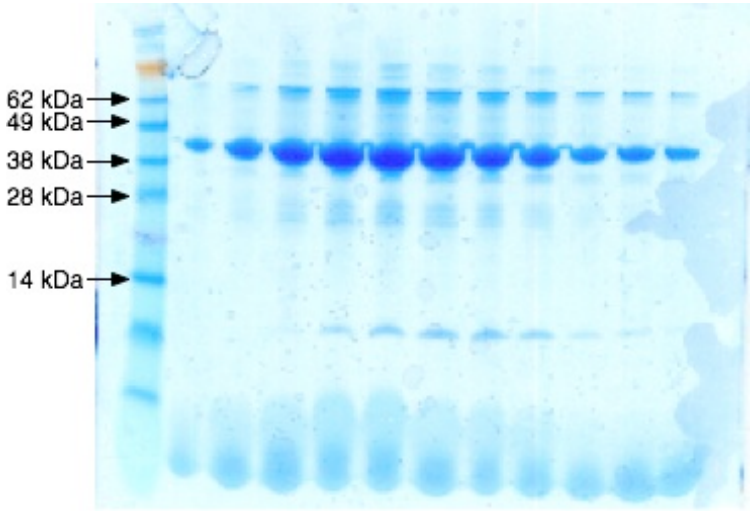


Figure 6.8 SDS-PAGE of partially purified SI-CLBH

Peak fractions eluted from the Ni^{2+} column (Figure 6.7) were analysed by SDS-PAGE to determine the purity of the SI-CLBH. A major band at *ca.* 44 kDa (corresponding to SI-CLBH) can be clearly observed in all of the fractions. The gel showed that SI-CLBH was largely pure following application to the Ni^{2+} column.

The enriched SI-CLBH sample was desalted before it was cleaved overnight with C3 protease. The sample was reverse purified the following day using Ni^{2+} beads,

concentrated to a volume of 0.5 ml and then applied to a Sepharose C6 gel filtration column. The elution profile only contained a single large symmetrical peak (Figure 6.9), demonstrating that SI was both pure and monodisperse in DDM making it useful for structural studies. Fractions across the elution peak were analysed by SDS-PAGE (Figure 6.10). The gel showed there was a single band with a molecular weight of *ca.* 22 kDa in all of the fractions tested. As SI has a predicted MW of 26.2 kDa, it was likely that the band contained SI. To confirm this possibility the protein band was excised and analysed by MS-MS. Peptide MTTNAGPLHPYWPQHLRL corresponding to αα 1-18 of SI was detected, confirming the identity of the purified protein.

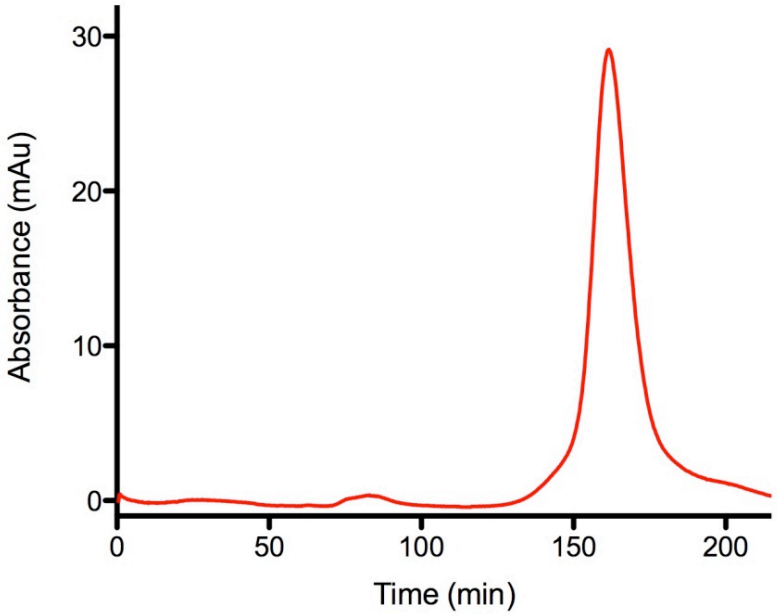


Figure 6.9 Elution profile of pure SI in DDM from a Sepharose C6 gel filtration column
Following reverse-purification with Ni^{2+} beads, the SI-containing sample was concentrated and applied to the gel-filtration column. The buffer was 10 mM HEPES pH 7.4 containing 10% glycerol and 0.05% DDM. A single symmetric peak was observed showing that the purified SI was monodisperse. Fractions across the peak were analysed by SDS-PAGE (Figure 6.10)

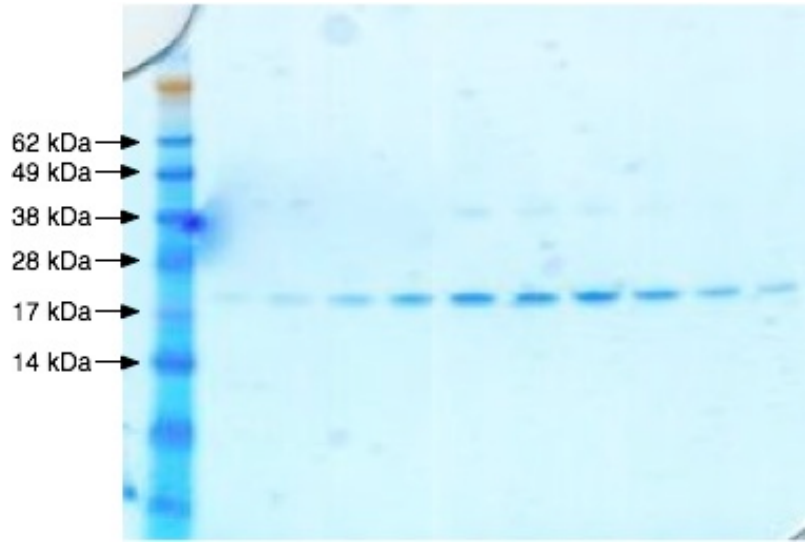


Figure 6.10 SDS-PAGE analysis of SI-containing gel-filtration fractions

Aliquots of the fractions from the gel-filtraton peak were analysed by SDS-PAGE. In all of the lanes, a strong band of *ca.* 22 kDa was observed suggesting that it was SI (actual MW 26.2 kDa). This band was cut-out and sent for MS-MS analysis. A peptide of MTTNAGPLHPYWPQHLRL was identified corresponding to $\alpha\alpha$ 1 - 18 of SI confirming the identify of the purified protein. A further faint band could be seen at *ca.* 40 kDa which appears to be an SDS-induced dimer, as only one protein peak was observed by SEC.

6.5 Secondary structure determination by circular dichroism

Circular dichroism (CD) refers to the differential absorption of left- and right-handed polarised light by chiral molecules (Kelly *et al.*, 2005). CD signals only occur at those wavelengths where the asymmetric molecule absorbs. All proteins have several chromophores that absorb in the ultra-violet (UV) region of the electromagnetic spectrum including peptide bonds (which absorb between 180 and 240 nm), aromatic amino acids (260 to 320 nm) and disulphide bonds (which have weak broad absorption bands around 260 nm).

The different kinds of secondary structure (e.g. α-helix, β-sheet) that occur in proteins have characteristic CD spectra in the far UV (which arise from the chiral arrangement of the peptide bonds in each structure). A number of algorithms (e.g. SELCON (Sreerama and Woody, 1993), VARSLC (Manavalan and Johnson, 1987), CDSSTR (Johnson, 1999) and CONTIN (Provencher and Glockner, 1981))

have been developed that provide an estimation of a protein's secondary structure composition from its far-UV CD spectra.

Proteins also have measurable CD in the near-UV (260–320 nm) arising from its constituent aromatic amino acids (Trp, Tyr, Phe). The actual shape and magnitude of the near-UV CD spectrum for each protein depends on the number of each type of aromatic amino acid present, their mobility, their environment (H-bonding, polar groups and polarisability) as well as their position within the protein. The near UV-CD spectrum of a protein provides a valuable fingerprint of its tertiary structure, and can be used to compare, for example, the structural similarity of wild-type and mutant forms of the protein.

Here, the far- and near-UV CD spectra of SI are presented.

6.5.1 CD analysis of SI

The near-UV CD spectrum (region 250-320 nm) of purified SI was recorded and found to consist of a series of peaks and troughs with maxima at 256, 265, 277 & 291 nm and minima at 262, 270, 282 & 296, respectively (Figure 6.11 (A)). The secondary structure composition of SI was determined by recording a far-UV CD spectrum (region 190–260 nm) (Figure 6.11 (B)), and then fitting it using the CDSSTR alogorithm (Johnson, 1999). From this, the secondary structure content of SI was estimated to be 31 % α-helix, 26 % β-sheet, 23 % turn with 20 % of the protein disordered. This prediction suggests that SI is largely ordered (consistent with its catalytic function) and probably has transmembrane α-helices.

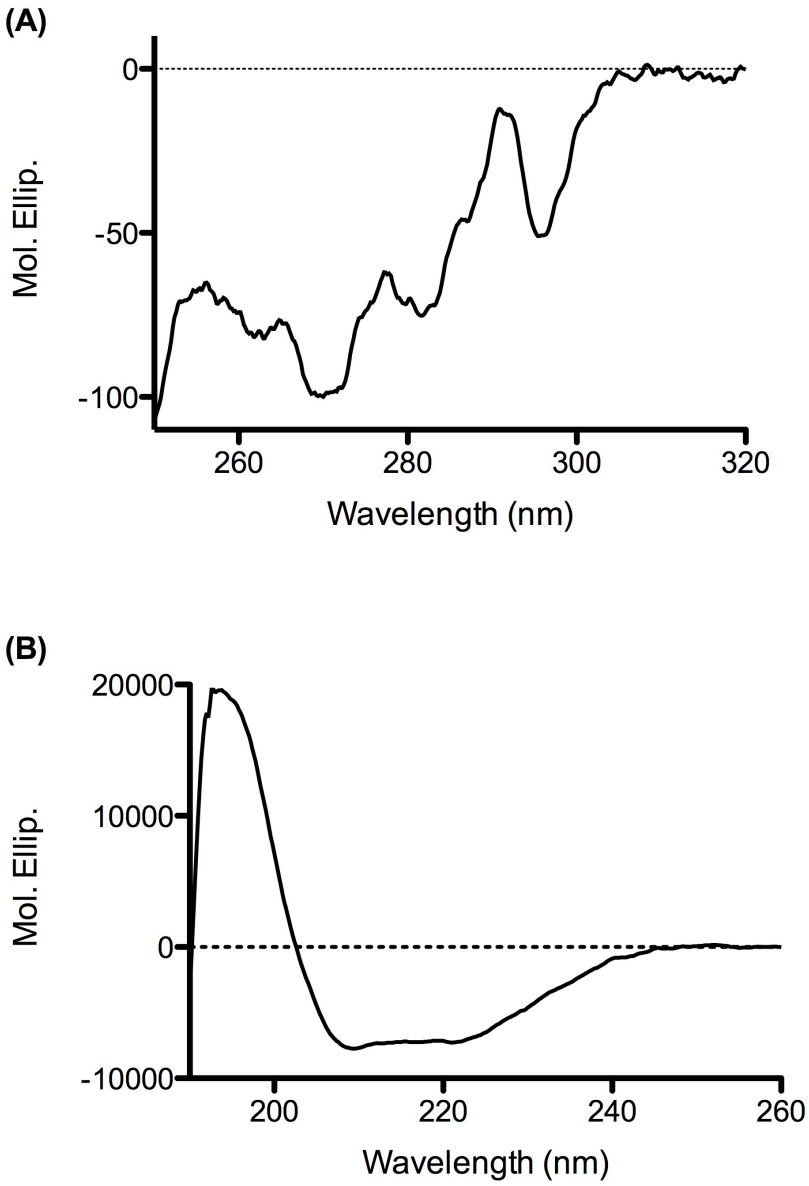


Figure 6.11 CD spectra of SI in the (A) near- and (B) far-UV

The near-UV CD spectrum consists of a series of peaks and troughs with maxima at 256, 265, 277 & 291 nm and minima at 262, 270, 282 & 296, respectively. On fitting the far-UV CD spectrum with the CDSSTR alogorithm (Johnson, 1999) the secondary structure content of SI was estimated to be 31 % α-helix, 26 % β-sheet, 23 % turn with 20 % of the protein disordered.

6.6 Assessing and maximising the thermostability of SI in detergent solution

Removal of membrane proteins from their native lipid environment with detergent confers, to a greater or lesser extent, instability on them. Furthermore, as it can take weeks or even months for protein crystals to grow it is essential that the membrane protein has long-term stability in detergent solution. The stability of all membrane proteins in aqueous solution are affected by a range of factors including the buffer type as well as its pH, the ionic composition and strength of the buffer as well as the type of detergent used. Therefore, considerable effort needs to be invested for each protein of interest to find the conditions in which it is most stable. Through accomplishing this, it may be possible to purify the target protein in a short chain detergent increasing the probability of getting diffracting protein crystals. In Chapter 5, the effects of choice of detergent on hA2aR and hSI stability immediately following membrane solubilisation were investigated by FSEC. Although buffer and salt effects on MP stability can be assessed by FSEC too, this method is incredibly time consuming. An alternative approach involves purifying the target MP to homogeneity in a mild-detergent (such as DDM), and then measuring the thermostability of the protein in a range of buffer/salt conditions. The stability of any protein can be thought of in terms of its resistance to denaturation. Those proteins that are unstable denature easily whereas those that are stable require harsher conditions (e.g. high temperatures) before they start to unfold. Here, two methods (CD, CPM assay) for studying protein unfolding were explored with a view to finding those conditions that gave maximum SI thermostability in detergent solution.

6.6.1 Investigating SI thermostability by circular dichroism

CD can be used to study how a protein's secondary structure changes with temperature giving insights into its thermostability. Here, a thermal denaturation of SI in 10 mM HEPES pH 7.5, 150 mM NaCl, 0.05% DDM was performed. The far-UV CD spectrum of SI was recorded at regular intervals

between 5 and 80°C (Figure 6.12 (A)). As the temperature increased, SI became partially denatured with a decrease in the minima peak intensities at both 209 and 222 nm. This said, SI largely retained its near-UV CD profile even after heating to 80°C showing that even at this high temperature only *ca.* 30% of the total secondary structure had become lost (Figure 6.12 (B-C)). As it was not possible to remove all of the protein's secondary structure by thermal denaturation, this assay was considered to be too insensitive for screening the effects of different buffers and salts on SI stability.

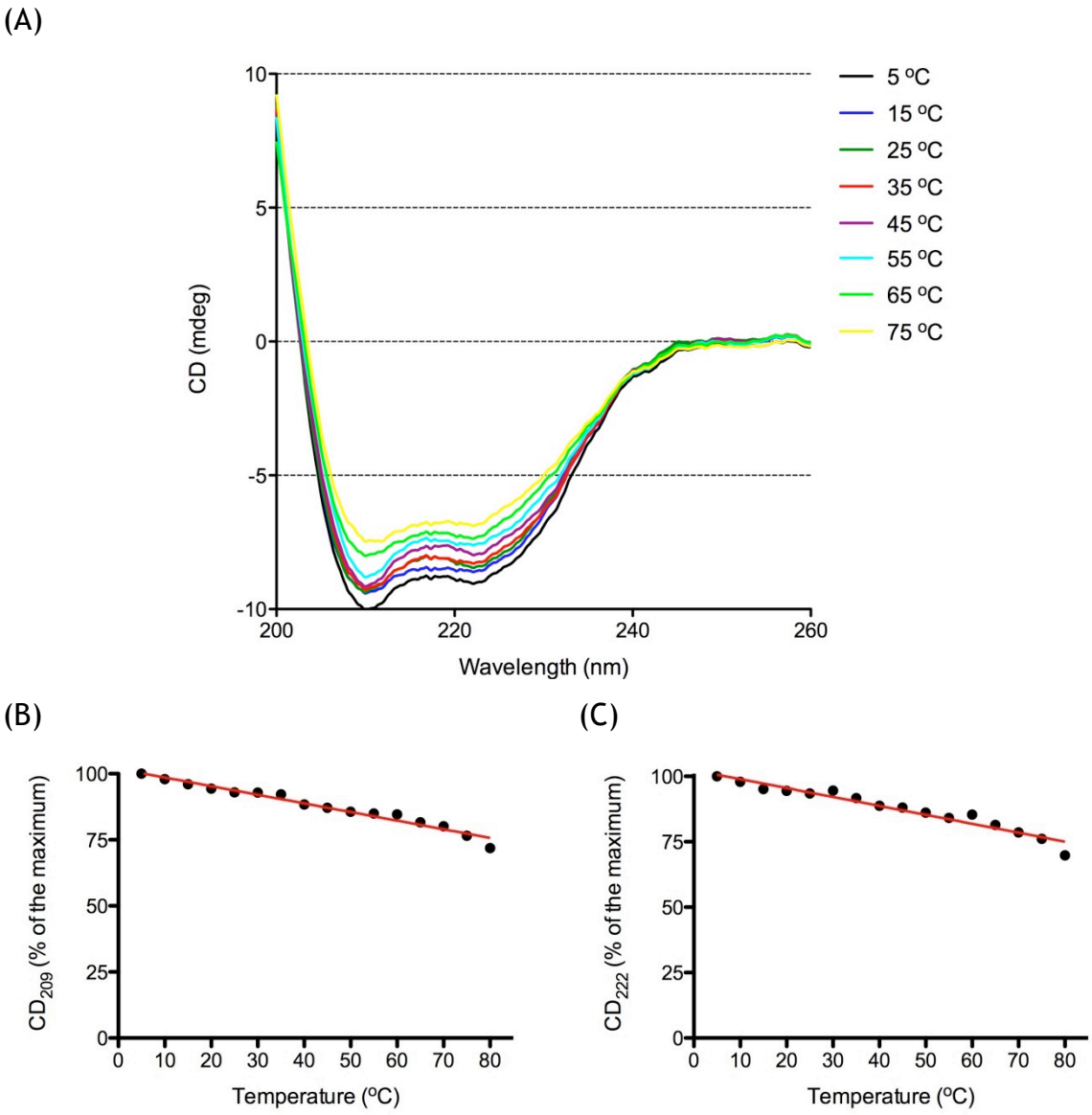


Figure 6.12 Thermal denaturation of SI assessed by circular dichroism

(A) Change in the far-UV CD spectrum of SI over the temperature range 5 - 80°C. The extent of SI secondary structure remaining following heat denaturation was determined by measuring the size of the CD signal at (B) 209 and (C) 222 nm.

6.6.2 Using the CPM fluorophore to assess SI protein stability

A sensitive fluorescence-based (CPM) thermal stability assay that can be used with MPs has recently been described by Ray Stevens and co-workers (Alexandrov *et al.*, 2008). In this assay, reaction of a protein's solvent-exposed cysteine residues with the thiol-specific fluorochrome N-[4-(7-diethylamino-4-methyl-3-coumarinyl)phenyl]maleimide (CPM) results in the generation of a fluorescent signal. When a protein is thermally denatured in the presence of CPM as the cysteine residues located in the protein interior became solvent-exposed an increase in CPM fluorescence intensity occurs over time. By altering the assay conditions (e.g. different buffers - types and pH; salts - types and concentrations; addition of specific ligands) in which the thermal denaturations are carried out it is possible to identify those reagents that confer thermostability on the target protein. This assay can be readily-adapted for use with 96-well plates allowing numerous conditions to be rapidly assessed. Another advantage of this assay is that it consumes relatively small amounts (1-10 µg) of protein as compared to CD measurements (which require several 100 µgs). Here, the CPM assay was used to identify the optimal buffer and salt conditions that conferred maximal stability on SI.

The approach of Alexandrov *et al.* (Alexandrov *et al.*, 2008) was modified so that SI was denatured in the presence of SDS (to enhance the rate of unfolding) at a constant temperature of 45°C. Initial experiments were performed to determine the amount of SI that needed to be added to each assay to give measurable amounts of fluorescence (5 µg) as well as the length of time necessary for maximal unfolding to occur (2.5 h) (results not shown). Initial experiments were performed denaturing SI in the presence of different (50, 100, 150 and 200 mM) concentrations of NaCl (Figure 6.13). The amount of fluorescence due to the unfolding of SI was determined by subtracting the background fluorescence in a buffer control from that measured in the presence of protein. For each condition, the protein and buffer measurements were made in quadruplicate. For all four NaCl conditions, the amount of specific CPM fluorescence increased with time (Figure 6.13). The amount of fluorescence at t = 8000 s for each

condition was plotted in a bar chart (Figure 6.14A). This value is a measure of how much SI had become unfolded during the course of the assay. The higher the fluorescence, the more SI had become unfolded. The results showed that the amount of denatured SI progressively increased with the concentration of NaCl. From this, it was concluded that SI becomes less stable with increasing concentrations of NaCl.

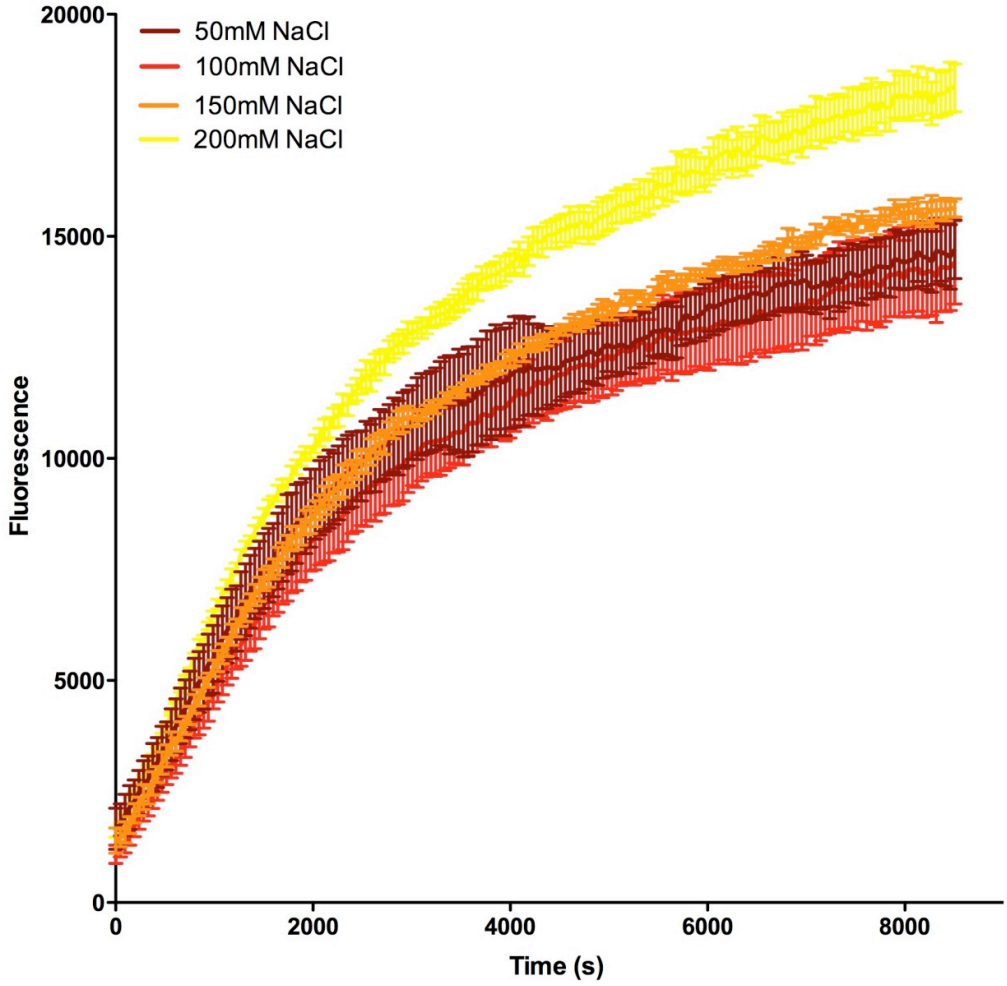


Figure 6.13 The unfolding of SI in the presence of different concentrations of NaCl

The extent of SI denaturation was determined by reaction with the CPM reagent (which labeled solvent-exposed cysteine residues). As the protein unfolded, cysteines from the inner core became accessible to the fluorochrome resulting in an increase in fluorescence with time.

This approach was used with a range of different salts and buffers (Figure 6.14-16). To compare the effects of Na^+, K^+, Mg^{2+}, Ca^{2+} and NH_4^+ ions on SI stability, CPM denaturation assays were performed in the presence of four different concentrations (50, 100, 150, 200 mM) of their chloride salt (Figure 6.14 A-C).

The trends obtained suggested that in terms of their effect on SI stability NH_4^+ > Na^+ > K^+ and that Mg^{2+} > Ca^{2+}. The effects of anions on SI stability were investigated in a similar way (Figure 6.14 C-D). The results showed that Cl^-, SO_4^{2-} and NO_3^- have similar effects on SI stability. However, PO_4^{3-} appeared to potently stabilise SI (Figure 6.14 D). NH_4^+ and PO_4^{3-} are known to promote hydrophobic interactions within proteins, and it is likely that these ions stabilise SI by strengthening existing hydrophobic contacts (most likely in the transmembrane region of the protein) within SI.

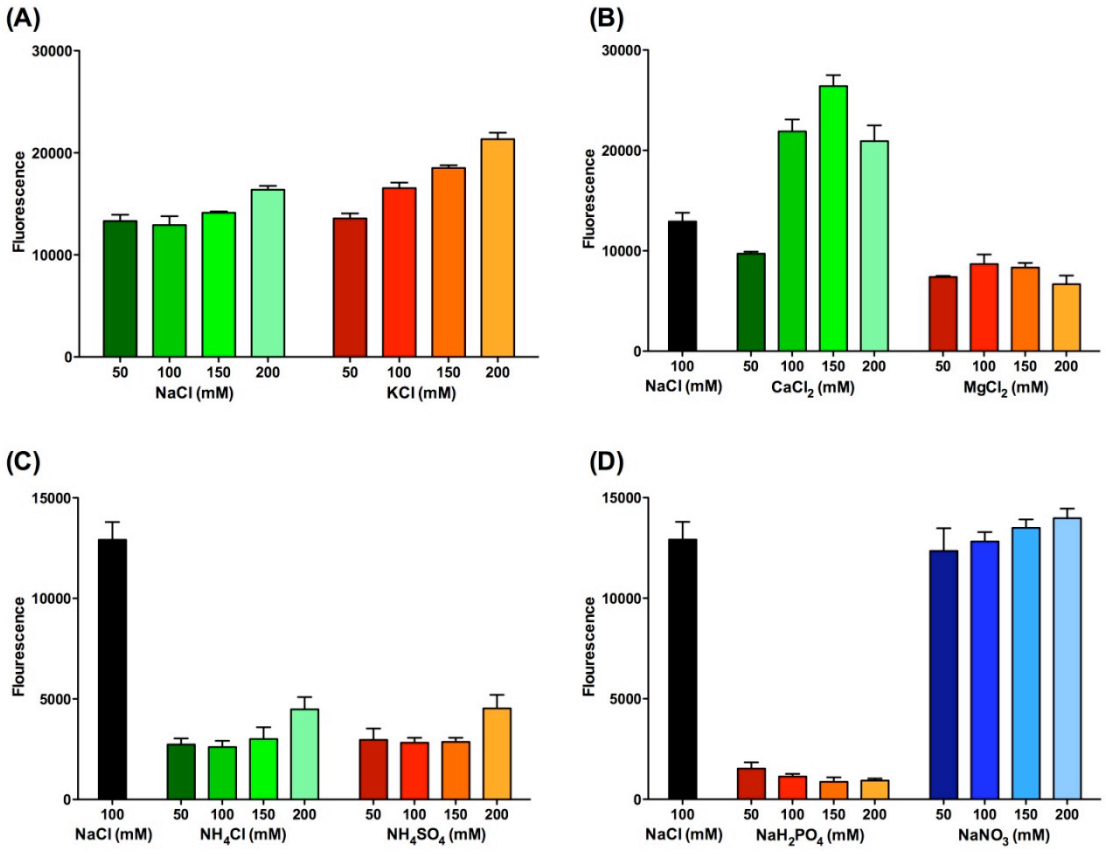


Figure 6.14 The effect of a range of salts on SI thermostability

SI was denatured in the presence of the CPM reagent. The ability of different ions to stabilise SI was assessed by measuring the specific increase in CPM fluorescence at the end of the thermal unfold. Those ions that gave the smallest rises in fluorescence (e.g. NH_4^+ and PO_4^{3-}) over the course of the assay had the greatest ability to stabilise SI in detergent solution.

To determine whether phosphate itself was stabilising SI or was having an indirect pH effect, further SI thermostability assays were performed in a series of phosphate buffers that had different pH (range 4 – 9) (Figure 6.15). A clear

trend was observed with greatest SI thermostability observed at pH 4. As the pH of the phosphate buffer increased more SI denaturation occurred. The effect of pH on SI stability was further investigated by performing a series of denaturation assays in four different buffers (MES, MOPS, HEPES, Tris) over the pH range 5.5 to 8 (Figure 6.16 (A)). Similar to that observed in phosphate buffer, SI had greater thermostability at acidic than alkali pH. Lastly, the effect of selected transition metal ions on SI stability was investigated (Figure 6.16 (B)). Zn^{2+}, Cu^{2+}, Ni^{2+} and Co^{2+} were all found to stabilise SI.

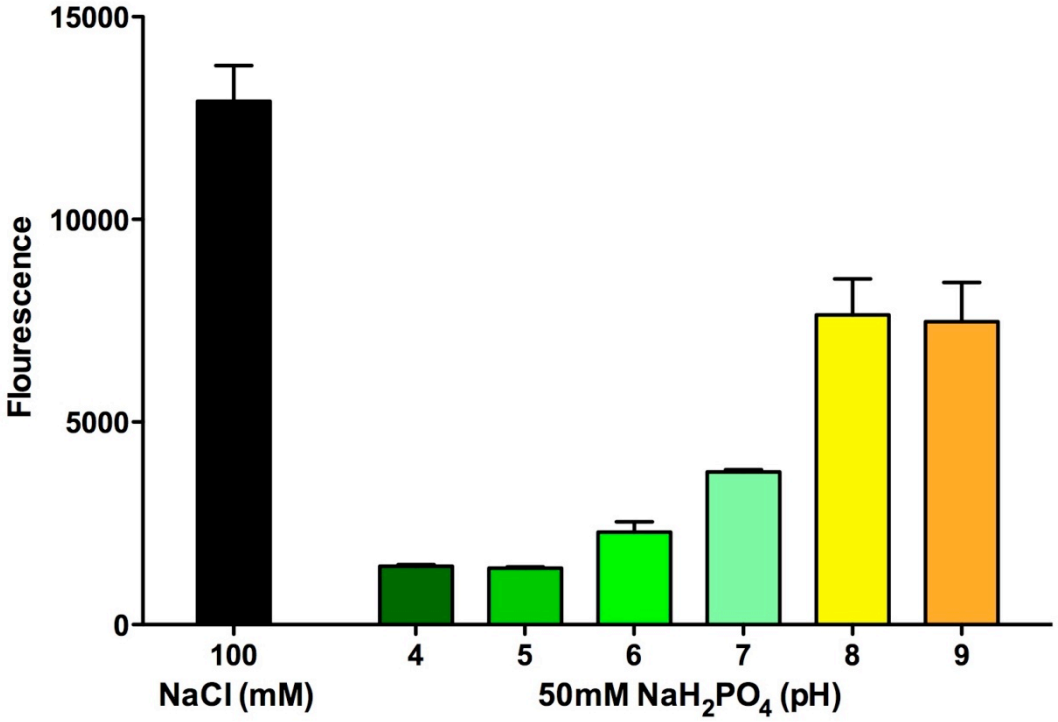


Figure 6.15 Effect of the pH of phosphate buffer on SI thermostability

The thermostability of SI was pH-dependent. SI was most stable at pH 4. As the pH of the phosphate buffer increased more SI became denatured.

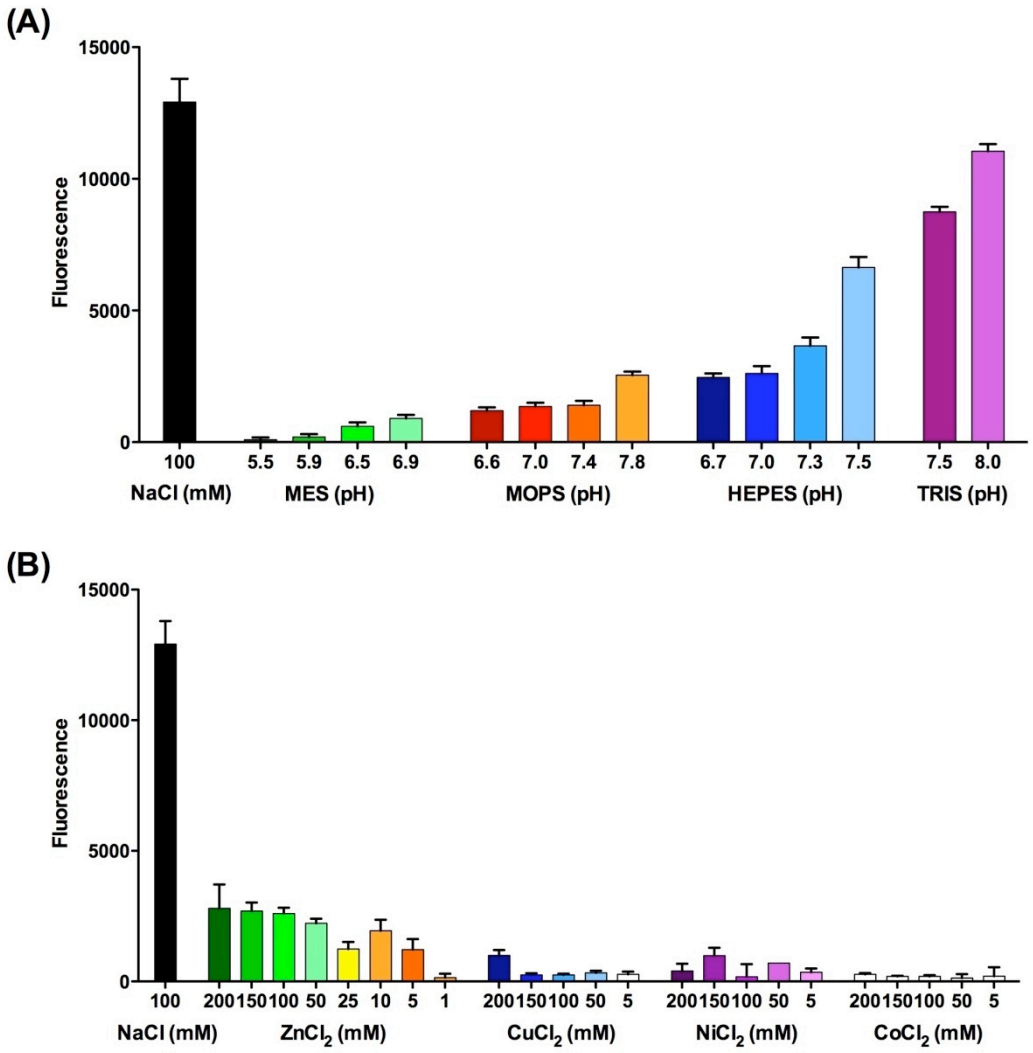


Figure 6.16 The effect of (A) pH and (B) the addition of transition metals on SI thermostability

Similar to that observed in phosphate buffer (Figure 6.15), SI had greater thermostability at acidic than alkali pH. The transition metal ions Zn^{2+}, Cu^{2+}, Ni^{2+} and Co^{2+} were all found to stabilise SI.

6.7 Purification of SI in DM using a phosphate-based buffer

SI stability in detergent solution is clearly dependent on the choice of buffer used as well as its ionic content (Section 6.6). Using the information from the CPM denaturation assays, I set-out to develop a purification protocol that would give monodisperse SI in the short-chain detergent DM. SI is more stable at acidic than alkali pH. However, the efficiency of SI release from the membrane on

detergent solubilisation decreased dramatically at pH <7 (results not shown). Furthermore, His tag binding to a Ni^{2+} column is pH-dependent with the strength of interaction between the transition metal and the tag increasing as the buffer becomes more alkali. Using all of this information, a protocol based on a sodium phosphate buffer was devised in which the Ni^{2+} affinity, C3 cleavage and reverse-purification steps were performed at pH 7.5, and the gel-filtration step performed at pH 6. Under these conditions, pure, monodisperse SI in DM could be obtained (Figure 6.17). It was hoped that this protein would prove amenable to crystallisation.

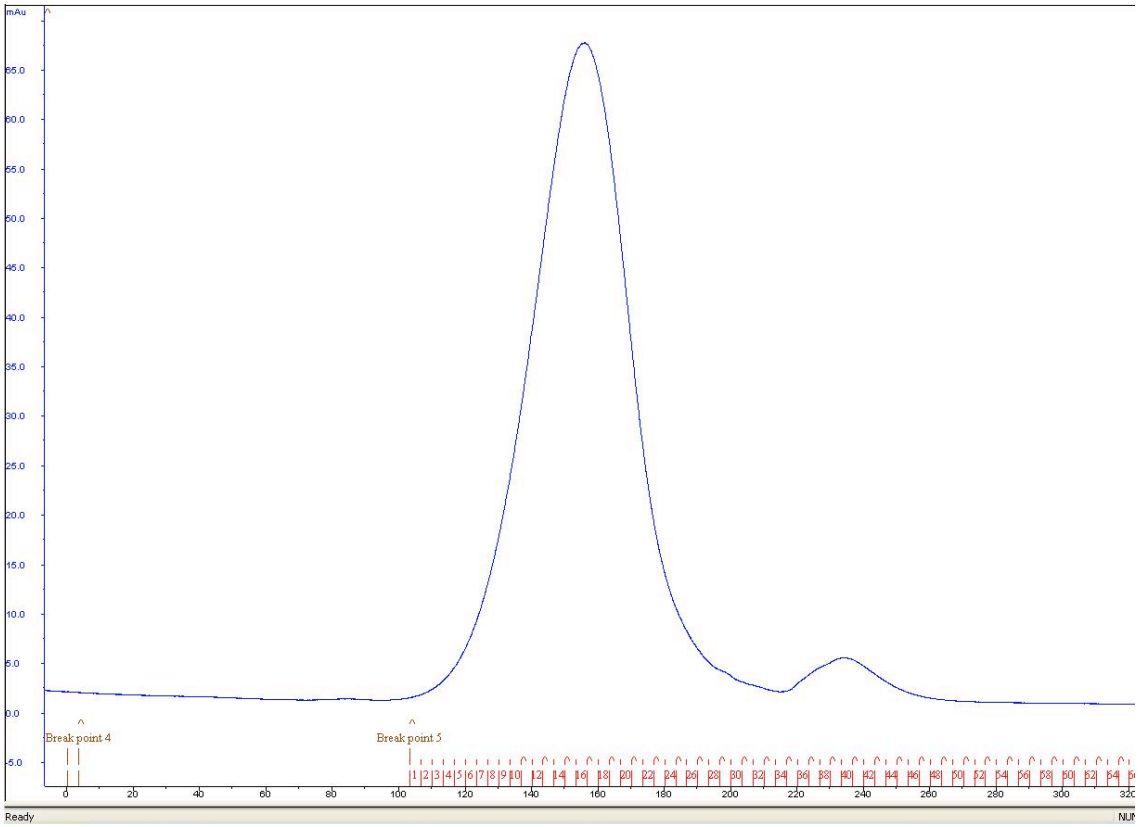


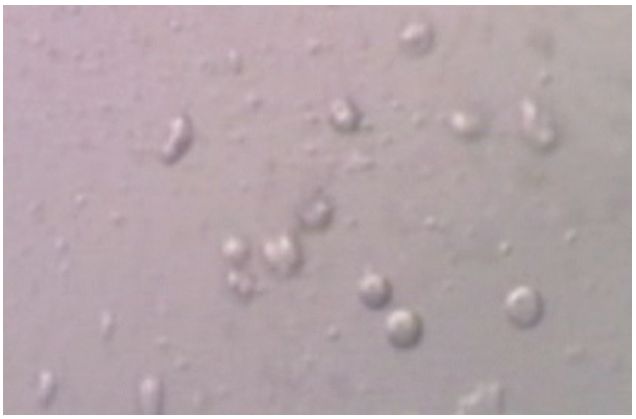
Figure 6.17 Elution of pure SI from the Sepharose C6 gel-filtration column equilibrated in sodium phosphate/DM buffer

A protocol for purifying SI using a sodium phosphate buffer was devised in which the Ni^{2+} affinity, C3 cleavage and reverse-purification steps were performed at pH 7.5, and the gel-filtration step performed at pH 6. Under these conditions, pure, monodisperse SI in DM could be obtained.

6.8 Crystallisation of SI from detergent solution

As SI could be purified to homogeneity with monodispersity (in both DDM and DM), it was thought that this enzyme would make a good candidate for crystallisation trials using conventional techniques. Initial screening with SI was performed using protein that had been purified in HEPES buffer pH 7.4 containing DDM (Section 6.4). 500 nl of 20 mg/ml protein was pipetted into each of the drops in a 96 well plate using a Cartesian robot along with an equal volume of the reservoir solution from the Peg/Ion and MemSys/MemStart sparse matrix crystallisation screens. The plates were incubated in a Rhombix imager at 20°C for 3 months with regular photographs taken (daily for the first week, weekly thereafter) to record crystal growth or precipitation formation. In several conditions, detergent crystals formed (Figure 6.18 (A)). Disappointingly, no protein crystals were observed. A possible explanation for this is that the DDM micelle is so big that not enough (and perhaps none) of the polar surface area of SI is exposed to form lattice contacts. To address this problem, I set-up a crystal tray (MemGold) with SI that had been purified in sodium phosphate buffer (pH 6) containing DM (Section 6.7). To reduce the number of detergent crystals that formed, a micro-dialysis step was introduced immediately prior to setting up trays in order to reduce the detergent's concentration to just above its CMC. Although this helped to substantially reduce the number of detergent crystals formed (results not shown), no wells were found to contain diffracting protein crystals. Several conditions contained phosphate salt crystals (Figure 6.18 (B)). Many wells contained light protein precipitate (Figure 6.18 (C)) which, given sufficient time and protein for optimisation, may in the future lead to the formation of small protein crystals.

(A)



(B)

(C)

Figure 6.18 Efforts to crystallise SI in detergent solution

(A) Initial attempts to crystallise SI in HEPES buffer containing DDM resulted in the formation of numerous detergent crystals. This problem was overcome by performing a micro-dialysis step immediately prior to setting up trays that ensured that the detergent concentration in the SI sample was just above its CMC. At a later stage, SI that had been purified in sodium phosphate buffer containing DM was used with the MemGold screen but no diffracting protein crystals were obtained. Several conditions did contain phosphate crystals (B). Many wells contained light protein precipitate (C) which, given sufficient time and protein for optimisation, may in the future lead to the formation of protein crystals.

6.9 Development of a cholesterol-doped sparse matrix lipidic-sponge phase crystallisation screen

The recent structures of the human β2-adrenergic and adenosine 2a receptors were solved using crystals grown from cholesterol-containing lipidic cubic phases (Cherezov *et al.*, 2007; Jaakola *et al.*, 2008). However, due to difficulties in handling the cubic phase as well as restrictions on the range of MPs that can be used with the cubic phase (due to the relatively small water pores in the lipid matrix) the more fluid 'sponge' phase has become of interest to the membrane protein structural community. While work with lipidic sponge phases is still in its infancy they have already been used to obtain diffracting crystals of the purple bacterial photosynthetic reaction centre (Wohri *et al.*, 2009). At the outset of my Ph.D. there was not a commercially-available lipidic sponge phase crystallisation screen. Recently, however, the group of Richard Neutze (Chalmers University, Sweden) have devised such a screen (in which I played a small part in helping to develop) (Wohri *et al.*, 2008). I have shown that up to 10% cholesterol can be incorporated into lipidic sponge phases (results not shown). Here, I describe my own version of the the Neutze screen consisting of a series of cholesterol-doped sponges.

The conditions used to make the cholesterol containing sponge phase screen were largely the same as that reported in (Wohri *et al.*, 2008), and all contained the lipid mono-olein (MO) (Figure 6.19). Cholesterol-containing versions of the jeffamine sponges detailed in (Wohri *et al.*, 2008) were not made. Similar to the Neutze sponge phase screen, three buffers (HEPES, MES and Tris) and three salts (NaCl, $MgCl_2$, and Li_2SO_4) were used all at a concentration of 100 mM. The pH of the buffers ranged between 6.5 and 9. It was not possible to use buffers with strongly acidic pH as the MO ester bond is extrememly sensitive to hydrolysis (Engstrom *et al.*, 2007). Cholesterol was mixed with the MO prior to sponge phase formation to ensure that the sterol was completely incorporated within the lipid matrix. Sponge phases were created as described in (Wohri *et al.*, 2008) using cholesterol-doped MO for all conditions. Sponge phases were made that had two different concentrations of cholesterol (0.5 and 5 % (w/w)) to increase

the diversity of the screen. The complete cholesterol-doped sponge phase screen is listed in Table 6.1. PEG 400 and PEG 1500 were chosen to form the sponge-phases as they both have cryo-protectant properties which means that any protein crystals formed in any of the conditions can be directly transfered from the drop to the X-ray beam. Similar to that for detergent-based crystallisation screens, any crystals obtained using the sponge-phase screen will likely need optimised to improve their packing before crystals that diffract to high-resolution are obtained.

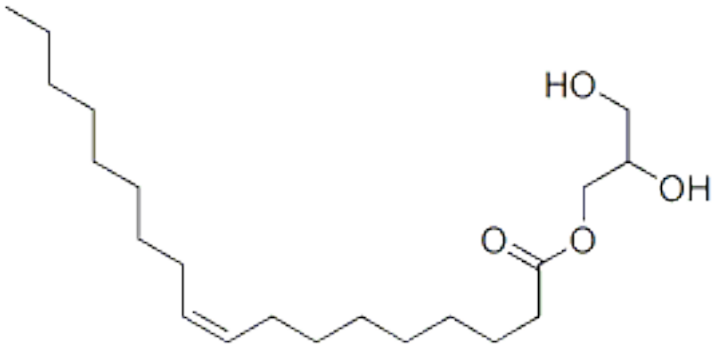


Figure 6.19 Chemical structure of mono-olein

6.10 Crystallisation of SI in the lipidic-sponge phase

As the cholesterol-doped sponge phases described in Section 6.9 closely mimic mammalian membranes they should prove useful for crystallising a wide-range of human MPs. Here, I used the lipidic-sponge screen with SI to try to identify conditions that might yield diffracting crystals.

SI that had been purified in sodium phosphate buffer containing DM was dialysed to remove excess detergent and then concentrated to 10 mg/ml. Hanging drop crystallisation trays were set-up in which 1 μl of the protein solution was directly added to 1 μl of one of the cholesterol-containing sponges. In addition, a buffer control was set-up for each sponge to give confidence that any crystals observed in the protein drops did indeed come from the protein and not from any of the chemicals present in either the buffer solution or sponge itself. The

Condition	**PEG**	%	% MO	% cholesterol	**pH**	**Buffer (100 mM)**	**Salt 1 (100 mM)**	**Salt 2 (100 mM)**
1	400	40	30	5	6.5	MES		
2	400	40	30	5	6.5	MES	NaCl	Li_2SO_4
3	400	40	30	5	6.5	MES	NaCl	$MgCl_2$
4	400	40	30	5	7.5	HEPES	NaCl	Li_2SO_4
5	400	40	30	5	7.5	HEPES		
6	400	40	30	5	7.5	HEPES	NaCl	$MgCl_2$
7	400	40	30	5	8.5	HEPES		
8	400	40	30	5	8.5	HEPES	NaCl	$MgCl_2$
9	400	40	30	5	8.5	HEPES	NaCl	Li_2SO_4
10	400	40	30	5	9.0	TRIS		
11	400	40	30	5	9.0	TRIS	NaCl	$MgCl_2$
12	400	40	30	5	9.0	TRIS	NaCl	Li_2SO_4
13	400	40	30	0.5	6.5	MES		
14	400	40	30	0.5	6.5	MES	NaCl	Li_2SO_4
15	400	30	40	0.5	7.5	HEPES		
16	400	30	40	0.5	7.5	HEPES	NaCl	Li_2SO_4
17	400	40	30	0.5	7.5	HEPES		
18	400	40	30	0.5	7.5	HEPES	NaCl	$MgCl_2$
19	400	40	30	0.5	8.5	HEPES		
20	400	40	30	0.5	8.5	HEPES	NaCl	$MgCl_2$
21	400	40	30	0.5	8.5	HEPES	NaCl	Li_2SO_4
22	400	40	30	0.5	9.0	TRIS		
23	400	40	30	0.5	9.0	TRIS	NaCl	$MgCl_2$
24	400	40	30	0.5	9.0	TRIS	NaCl	Li_2SO_4
25	1500	40	30	5	6.5	MES		
26	1500	40	30	5	6.5	MES	NaCl	Li_2SO_4
27	1500	40	30	5	6.5	MES	NaCl	$MgCl_2$
28	1500	40	30	5	7.5	HEPES		
29	1500	40	30	5	7.5	HEPES	NaCl	Li_2SO_4
30	1500	40	30	5	7.5	HEPES	NaCl	$MgCl_2$
31	1500	40	30	5	8.5	HEPES		
32	1500	40	30	5	8.5	HEPES	NaCl	$MgCl_2$
33	1500	40	30	5	7.0	TRIS		
34	1500	40	30	5	8.0	TRIS		
35	1500	40	30	5	7.5	TRIS		
36	1500	40	30	5	9.0	TRIS		
37	1500	40	30	0.5	6.5	MES		
38	1500	40	30	0.5	6.5	MES	NaCl	Li_2SO_4
39	1500	30	30	0.5	6.5	MES	NaCl	$MgCl_2$
40	1500	30	30	0.5	7.5	HEPES		
41	1500	40	30	0.5	7.5	HEPES	NaCl	Li_2SO_4
42	1500	40	30	0.5	7.5	HEPES	NaCl	$MgCl_2$
43	1500	40	30	0.5	8.5	HEPES		
44	1500	40	30	0.5	8.5	HEPES	NaCl	$MgCl_2$
45	1500	40	30	0.5	7.0	TRIS		
46	1500	40	30	0.5	8.0	TRIS		
47	1500	40	30	0.5	7.5	TRIS		
48	1500	40	30	0.5	9.0	TRIS		

Table 6.1 A cholesterol-containing sponge phase crystallisation screen for use with mammalian membrane proteins

trays were sealed and stored at 20°C for one month. Due to the size of the crystallisation trays and the drops themselves, all visualisation events were performed manually using a microscope. Those conditions thought to be interesting were photographed. After one month most wells showed some protein precipitation over the sponge background (Figure 6.20 (A)). In certain conditions small crystalline material was observed (Figure 6.20 (B)). Due to time limitations, however, it was not possible to test the crystalline matter in our in-house X-ray facility. The preliminary results presented here do suggest that the cholesterol-doped sponge-phase screen will prove to be a useful addition to the range of existing crystallisation screens currently used with mammalian membrane proteins.

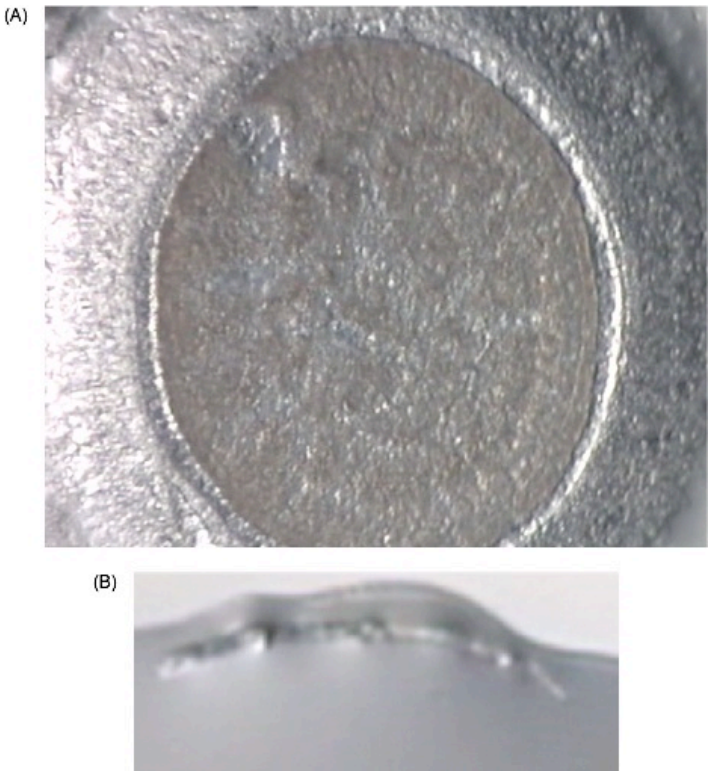


Figure 6.20 Efforts to crystallise SI using cholesterol-doped lipidic sponges

(A) Image of sponge matrix containing protein. (B) Crystalline material in condition 5 (40% PEG400, 30% mono-olein, 5% cholesterol, HEPES pH 7.5) of the sponge-phase screen was observed along the perimeter of the sponge drop.

6.11 Conclusion

This chapter shows that CLBH-tagged recombinant mammalian MPs (both N_{in} and N_{out}) can be purified to homogeneity in four simple steps (Ni^{2+}-affinity purification; cleavage with C3 protease; desalt and reverse-purification with Ni^{2+} beads; gel-filtration). This vector system should be compatible for use with other mammalian MPs including type I GPCRs, membrane enzymes, transporters and ion channels.

My work on Rag23 highlights the need to perform thorough functional assays on any mutants made for structural studies to ensure that the mutations which have been made do not affect the activity of the target protein. The IL3 deletion that I made in Rag23 was introduced to reduce the amount of disorder within the receptor. However, in making this deletion, I made a mutant that had low-affinity for agonist most likely by trapping the modified receptor in a closed state. This was unfortunate, and prevented me from doing large-scale protein preparations for biophysical characterisation and structural studies.

Sterol isomerase is an incredibly stable enzyme for a wild-type mammalian MP. Using information obtained from denaturation assays using the CPM reagent it was possible to devise a purification protocol in which monodisperse SI was obtained in the short-chain detergent DM using sodium phosphate as buffer. The power of the CPM assay is the detailed insights it is able to give on the buffer (both type and pH) and salt preferences of MPs in detergent solution. The approach I used to identify the conditions that gave maximum SI thermostability in detergent solution will be useful with other mammalian MPs. Whether or not it is possible to get diffracting crystals of SI in sodium phosphate/DM buffer remains to be seen.

Finally, a cholesterol-doped lipidic-sponge phase sparse matrix crystallisation screen was created that was tested with SI. While some crystalline material was observed, time limitations prevented me from doing any X-ray analysis and crystal optimisation. However, the preliminary results obtained suggest that the screen may be useful for growing diffracting crystals of mammalian MPs.

7 Discussion and future perspectives

Solving the structure of a human membrane protein is a massive challenge as evidenced by the small (<20) number of structures that have been deposited in the PDB to date. Given the biological and pharmaceutical importance of many human MPs, there is currently a substantial effort within the structural biology community to find technologies that will overcome the existing bottlenecks in solving the structures of mammalian MPs. Unlike their bacterial equivalents (of which there are several hundred unique structures in the PDB), most human membrane proteins are intrinsically unstable in detergent solution. This means that there is an urgent need to develop new technologies: i) that allow the most stabilising detergent for each human MP to be readily identified, ii) for optimising buffer conditions to ensure maximal thermostability of each MP in detergent solution, iii) such as novel detergents and lipids that effectively stabilise mammalian MPs once they have been removed from the membrane, and iv) that allow mammalian MPs to be crystallised in a lipid environment. Furthermore, simple generic systems that allow the overproduction and purification of milligram amounts of a range of human MPs for crystallisation trials are required. Over the course of my Ph.D., I have developed a series of novel technologies that have sought to address some of these bottleneck issues.

The CLBH vector system developed here allowed the successful over-expression and purification of both a conformationally-flexible (hA2aR) and a rigid-bodied (hSI) human membrane protein. Two versions of the CLBH expression vector were made – one with an upstream α-factor (that should be used for recombinant MPs with a N_{out} topology (such as the hA2aR)) and one without (that should be used with MPs that have a N_{in} orientation (such as hSI)). The CLBH vectors allowed both target proteins to be purified to homogeneity in four simple steps (Ni^{2+}-affinity purification; cleavage with C3 protease; desalt and reverse purification with Ni^{2+} beads; gel-filtration). Any recombinant protein made using the CLBH system can be purified in the same way. In fact, the vectors have already been used in our laboratory to obtain milligram amounts of

the plant photoreceptor UVR8, adrenodoxin and the cytochrome p450 enzyme aldosterone synthase (Niall Fraser, personal communication).

The iLOV fluorescent reporter protein was shown to be a useful tool for optimising laborious processes such as yeast clonal selection, protein production in fermenter culture, detergent and construct screening as well as for tracking protein purification, saving both time and energy. iLOV could also be used to assess the monodispersity of MPs in detergent solution by FSEC. Using this approach, it was shown that the cholesterol derivative CHS (which confers radioligand binding capacity on the hA2aR in detergent solution) promotes receptor aggregation during purification. Furthermore, iLOV was used to show that monodisperse Rag23 can be obtained in the short-chain detergent βOG immediately post-solubilisation. Such insights were only realised by measuring fluorescence from the iLOV tag. In contrast to GFP, when iLOV becomes denatured the co-factor flavin is released. This appears as a "late" peak on certain of the FSEC profiles. The size of the flavin peak gave a direct measure of the amount of target protein denaturation that had occurred on treating the membranes with detergent. The presence of a large flavin peak in a FSEC experiment (following either detergent-solubilisation or -exchange) shows that the target protein is unstable in the chosen detergent whereas the absence of a flavin peak suggests that the iLOV-tagged MP is stable. By monitoring the free-flavin content of iLOV-tagged MPs in detergent solution by FSEC it is possible to readily identify those detergents in which the MP is unstable. One disadvantage of iLOV compared to GFP is that it is not compatible with in-gel fluorescence experiments. As fluorescence from the iLOV chromophore comes from the flavin co-factor, when iLOV tagged receptor is treated with SDS, the iLOV domain denatures rendering it non-fluorescent. This means that iLOV-tagged proteins on a SDS-PAGE cannot be visualised by placing it under a UV-source. This is different from GFP-tagged proteins which retain their structure and fluorescence even after treatment with SDS (Drew *et al.*, 2006).

Attachment of the biotin acceptor domain to the C-terminus of the hA2aR increased the receptor's expression level by *ca.* 2-4 fold. Similar effects have been observed previously when this domain was fused to other type I GPCRs

made in *P. pastoris* (Grunewald *et al.*, 2004; Lundström *et al.*, 2006; Andre *et al.*, 2006). It appears likely, therefore, that attachment of this domain to the C-terminus of other MPs will increase their production levels in *P. pastoris* also. Increased expression helps to reduce the detergent costs associated with making large quantities of protein through reducing the volume of membranes required for solubilisation. Moreover, fusion of the biotin acceptor domain to MPs that express poorly by themselves in *P. pastoris* may help to make certain projects viable that otherwise would remain impossible. The biotin acceptor domain can also be biotinylated *in vitro* by treatment with recombinant biotin ligase. This then allows the target MP to be immobilised using (strept)avidin resin. This is an elegant feature of the expression vector, and is compatible not only with purifying low-abundance MPs but also with drug design techniques such as surface Plasmon resonance (SPR) or target immobilized NMR screening (TINS) (Vanwetswinkel *et al.*, 2005) where the protein of interest has to be immobilised prior to use with molecular and/or fragment libraries.

A considerable amount of time was spent developing an optimised *P. pastoris* fermentation protocol for make recombinant MPs at an induction temperature of 22°C. The developed methodology allowed both the hA2aR and hSI to be made with a production level >10 mg protein per litre of culture (which is more than sufficient for structure-determination experiments by way of X-ray crystallography). However, the protein produced could equally be used in nuclear magnetic resonance (NMR) experiments. NMR is an excellent structural approach that gives information regarding protein motion at the molecular level complementary to that obtained from X-ray crystallography (which gives highly-detailed information on the three-dimensional structures of proteins trapped in particular conformations). Several MPs have had their structures solved by NMR most of which have been of bacterial origin (e.g. Arora *et al.*, 2001; Yu *et al.*, 2005; van Horn *et al.*, 2009). Notably exceptions include archael sensory rhodopsin II (Gautier *et al.*, 2010) and human VDAC-1 (Hiller *et al.*, 2008; Bayrhuber *et al.*, 2008). One of the major bottlenecks that has prevented the wide-spread use of NMR with mammalian MPs lies in the difficulties of obtaining labelled, functionally expressed protein for analysis. Traditionally, proteins for

NMR experiments are made in *E. coli* or yeast as these organisms both grow on minimal media into which heavy atom labels can be introduced. At present, however, most academic groups are unable to produce 'active' mammalian MPs in either bacteria or yeast. The main advantage of the optimised fermentation protocol described in this thesis is that it can be readily adapted for labelling mammalian MPs with heavy atoms including ^{15}N, ^{13}C and ^{2}H. As the yeast are grown on a minimal media, ^{15}N labelling can be achieved by using ^{15}N ammonium hydroxide (as this is the sole nitrogen source used during fermentations), ^{13}C labelling through feeding with ^{13}C methanol and ^{2}H labelling through making up the basal media in D_2O (not H_20). The quantities of hA2aR and hSI that can be made per fermentation are more than sufficient for NMR analysis, and this is an area of research that I would like to have pursued if there had been sufficient time.

As demonstrated in this thesis, yeasts are powerful hosts for the overproduction of mammalian MPs. They are, however, far from being ideal cells for making human MPs as they differ from higher eukaryotes both in the type of glycosylation that they perform and sterol that they make. Although fungi and mammals share the initial steps of protein glycosylation (making the same core structure Man_8-$GlcNAc_2$), the final glycans attached to specific asparagine residues in glycoproteins are significantly different (Gemmill and Trimble, 1999). In mammals, regular, complex N-glycans of defined size are attached to the protein. In yeast, however, numerous mannose sugars are added to the core structure as the glycoprotein passes through the Golgi giving hypermannosylated N-glycans (with upto 100 mannose residues sometimes attached). As a result, glycoproteins made by yeast are highly-heterogenous in their composition and are incompatible with protein crystallisation. This problem can be overcome either by using glycosylation-deficient mutants (in which the asparagines that are glycosylated are mutated to glutamine) or by enzymatic treatment of purified glycoproteins with glycosidases *in vitro*. In recent years, however, strains of *P. pastoris* have been engineered in such a way that the yeast performs uniform complex N-glycosylation adding defined glycans that are terminally sialylated (Choi *et al.*, 2003; Hamilton *et al.*, 2003; Wildt and

Gerngross, 2005; Hamilton *et al.*, 2006; Hamilton and Gerngross, 2007). Such strains have been used to make recombinant erythropoietin with biological activity (Hamilton *et al.*, 2006). As yet, these genetically-modified versions of *P. pastoris* have not been used for the recombinant production of human MPs. They would, however, make excellent hosts for those human MPs where glycosylation is essential for the protein's activity.

A further drawback of using yeast for making mammalian MPs is that their membranes do not contain any cholesterol – instead they contain ergosterol. Although cholesterol and ergosterol have similar chemical structures, biologically they are non-equivalent (Opekarova and Tanner, 2003). Moreover, it is known that cholesterol is essential for maintaining the stability and function of many mammalian membrane proteins (e.g. Lagane *et al.*, 2000; Tate, 2001; Opekarova and Tanner, 2003). It is likely, therefore, that a strain of yeast that synthesises cholesterol instead of ergosterol will be a superior host for making human MPs compared to wild-type. Currently, work is underway within the Fraser group to make cholesterol biosynthetic versions of both *S. cerevisiae* and *P. pastoris* with a view to determining their capacity to make mammalian MPs (personal communication, Dr. Niall Fraser).

Although it is widely recognised that mammalian MPs are considerably less stable in detergent solution than ones of bacterial origin, it is poorly understood why. There is an emerging body of evidence that strongly suggests that cholesterol is a critical determinant of mammalian MP stability in the membrane, detergent solution as well as in lipidic crystallisation matrices. It is possible to 'dope' both cubic and sponge phases with cholesterol. Expression hosts can be chosen that have cholesterol-containing membranes. However, the issue of stabilising mammalian MPs in detergent solution remains. One possible approach is to create mutants such as Rag23 (Magnani *et al.*, 2008) which have sufficient thermostability in detergent solution that they do not have any need for cholesterol. However, the most direct approach to solving the problem of protein instability in detergent solution is to devise novel reagents that mimic mammalian membranes more closely than existing commercially-available ones. Of note, three recently described detergents (the maltose-neopentyl glycol

(MNG) amphiphiles (Chae *et al.*, 2010) and two cholesterol based detergents - facial amphiphiles (Zhang *et al.*, 2007) and ChoBiMalt (Howell *et al.*, 2010)) - have been shown to be useful tools for stabilising and crystallising mammalian MPs in and from detergent solution (Chae *et al.* 2010; Annalora *et al.*, 2010; Gay *et al.*, 2010; He *et al.*, 2010; Howell *et al.*, 2010). I would be particularly interested to find out if the cholesterol-detergent ChoBiMalt (cholesterol – bimaltoside) could stabilise the hA2aR in the same way as CHS does without driving receptor aggregation during protein purification (as was observed in my work).

The CPM assay was an excellent tool for understanding the buffer (type and pH) and salt preferences of SI in detergent solution. Using the information obtained from the denaturation assays, it was possible to purify SI in the short-chain detergent DM without it becoming aggregated using sodium phosphate as buffer. Since I finished in the lab, the Fraser group have since gone onto purify monodisperse SI in NM using ammonium phosphate buffer (Dr. Niall Fraser, personal communication). Such thermostability for a wild-type protein is rather unusual but suggests that hSI will make a good construct for structural studies (both by X-ray crystallography and NMR). When trying to crystallise any protein (membrane or otherwise), the first step usually involves using a range of commercially available sparse matrix crystallisation screens to try to identify conditions where the protein of interest comes out of solution to form regular crystals. Once “hits” have been identified, the crystallisation conditions are optimised in an attempt to obtain crystals that diffract to high-resolution. However, the results from the SI thermostability assays suggests that, at least for this enzyme, the probability of getting crystals from a sparse-matrix will be low as the protein is only stable in a limited range of conditions (ammonium or sodium phosphate, acidic pH). This suggestion was borne out from the limited number of crystallisation trays that I set-up. Rather, it seems sensible to try to find crystallisation conditions in which SI has known stability around the enzyme’s isoelectric point (pI) – pH 6.3. If time had allowed, I would have used the robotic systems in the lab to make a series of screens using either sodium or ammomium phosphate as buffer varing the pH between 5.5 and 7.5 in

combination with different types and concentrations of PEG. Hopefully, this approach will yield some SI crystal hits.

Although the CPM assay can be used to devise a purification scheme that yields monodisperse mammalian MPs in short-detergents, there are no guarantees that it will be possible to obtain protein crystals in the limited range of conditions in which the protein of interest is stable. Therefore, it is likely that the best way for getting crystals of mammalian MPs will be in either the lipidic cubic- or sponge-phase. It is no surprise then that all but one of the ligand-activated GPCR structures that have been solved to date (the human β2 adrenoceptor (Cherezov *et al.*, 2007; Hanson *et al.*, 2008); the human adenosine 2a receptor (Jaakola *et al.*, 2008); the human CXCR4 chemokine receptor (Wu *et al.*, 2010) and the dopamine D3 receptor (Chien *et al.*, 2010)) used diffraction data collected from crystals grown in the lipidic cubic phase. Recognising the importance of cubic- and sponge-phases, I have established a cholesterol-doped version of a sparse-matrix lipidic-sponge phase crystallisation screen. Although there was only limited usefullness for this screen with SI, it is likely that cholesterol-sponges will become increasing used within the Fraser lab as their attention turns from making to trying to crystallise human MPs.

Lastly, although it was possible to purify a thermostable mutant of the hA2aR and hSI to homogeneity, the membrane-enzyme was considerably easier to work with. This is not surprising as SI is likely to have a rigid-structure where as the hA2aR is likely to have considerable conformational flexibility. This raises the issue of target-selection when setting out trying to solve the structure of a mammalian MP. Proteins that are similar to SI are likely to be considerably more tractable to work than GPCRs, and maybe amenable to a high-throughput structural pipeline approach. In contrast, GPCRs should only be worked on one at a time by any given person due to the extra difficulties involved in getting sufficient stable protein for structural studies.

Over the next 5-10 years, I expect that there is going to be a rapid increase in the number of human membrane proteins that have had their structures solved. This will in no small part be due to the technologies that have recently been and

are currently being developed (such as the ones presented here) that allow milligram quantities of stable, functional and monodisperse membrane proteins to be produced. Furthermore, the use of lipidic crystallisation approaches (including both the cubic- and sponge-phase) for obtaining MP crystals that diffract to high-resolution are likely to become considerably more prevalent than they are today.

Bibliography

Abrahams, J. P., *et al.* (1994). "Structure at 2.8 Å resolution of F1-ATPase from bovine heart mitochondria." Nature **370**(6491): 621-628.

Abramson, J., *et al.* (2003a). "Structure and mechanism of the lactose permease of *Escherichia coli*." Science **301**(5633): 610-615.

Abramson, J., *et al.* (2003b). "The lactose permease of *Escherichia coli*: overall structure, the sugar-binding site and the alternating access model for transport." FEBS Letters **555**(1): 96-101.

Ago, H., *et al.* (2007). "Crystal structure of a human membrane protein involved in cysteinyl leukotriene biosynthesis." Nature **448**(7153): 609-612.

Albert, A. D., *et al.* (1996). "Effect of cholesterol on rhodopsin stability in disk membranes." Biochimica et Biophysica Acta (BBA) - Biomembranes **1297**(1): 77-82.

Alexandrov, A. I., *et al.* (2008). "Microscale fluorescent thermal stability assay for membrane proteins." Structure **16**(3): 351-359.

Alisio, A. and M. Mueckler (2010). "Purification and characterization of mammalian glucose transporters expressed in *Pichia pastoris*." Protein Expression and Purification **70**(1): 81-87.

Aller, S. G., *et al.* (2009). "Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding." Science **323**(5922): 1718-1722.

Amunts, A., *et al.* (2007). "The structure of a plant photosystem I supercomplex at 3.4 Å resolution." Nature **447**(7140): 58-63.

Andre, N., *et al.* (2006). "Enhancing functional production of G protein-coupled receptors in *Pichia pastoris* to levels required for structural studies via a single expression screen." Protein Science **15**(5): 1115-1126.

Annalora, A.J. *et al.* (2010). "Crystal structure of CYP24A1, a mitochondrial cytochrome P450 involved in vitamin D metabolism." Journal of Molecular Biology **396**(2):441-451.

Arora, A. and L. K. Tamm (2001). "Biophysical approaches to membrane protein structure determination." Current Opinion in Structural Biology **11**(5): 540-547.

Arora, A. *et al.* (2001). "Structure of outer membrane protein A transmembrane domain by NMR spectroscopy." Nature Structural Biology **8**(4):334-338.

Aughton, D. J., *et al.* (2003). "X-linked dominant chondrodysplasia punctata (CDPX2) caused by single gene mosaicism in a male." American Journal of Medical Genetics A **116A**(3): 255-260.

Bae, S., *et al.* (2001). "Cholesterol biosynthesis from lanosterol: molecular cloning, chromosomal localization, functional expression and liver-specific gene regulation of rat sterol Δ8-isomerase, a cholesterogenic enzyme with multiple functions." Biochemical Journal **353**(Pt 3): 689-699.

Bass, R. B., *et al.* (2002). "Crystal structure of *Escherichia coli* MscS, a voltage-modulated and mechanosensitive channel." Science **298**(5598): 1582-1587.

Bayrhuber, M., *et al.* (2008). "Structure of the human voltage-dependent anion channel." Proceedings of the National Academy of Sciences of the United States of America **105**(40): 15370-15375.

Becher, A. and R. A. McIlhinney (2005). "Consequences of lipid raft association on G-protein-coupled receptor function." Biochemical Society Symposium(72): 151-164.

Becker, K., *et al.* (2001). "Identification of a novel mutation in 3β-hydroxysteroid-Δ8-Δ7-isomerase in a case of Conradi-Hunermann-Happle syndrome." Experimental Dermatology **10**(4): 286-289.

Ben-Shem, A., *et al.* (2003). "Crystal structure of plant photosystem I." Nature **426**(6967): 630-635.

Bender, J., *et al.* (2008). "Structure and dynamics of a sponge phase in the methyl delta-aminolevulinate/monoolein/water/propylene glycol system." Journal of Colloid and Interface Science **317**(2): 577-584.

Binda, C., *et al.* (2002). "Structure of human monoamine oxidase B, a drug target for the treatment of neurological disorders." Nature Structural Biology **9**(1): 22-26.

Blois, T. M. and J. U. Bowie (2009). "G-protein-coupled receptor structures were not built in a day." Protein Science **18**(7): 1335-1342.

Bockaert, J. and J. P. Pin (1999). "Molecular tinkering of G protein-coupled receptors: an evolutionary success." EMBO Journal **18**(7): 1723-1729.

Bockaert, J., *et al.* (2004). "GPCR interacting proteins (GIP)." Pharmacology and Therapeutics **103**(3): 203-221.

Bousquet, C., *et al.* (2006). "Direct binding of p85 to sst2 somatostatin receptor reveals a novel mechanism for inhibiting PI3K pathway." EMBO Journal **25**(17): 3943-3954.

Braverman, N., *et al.* (1999). "Mutations in the gene encoding 3β-hydroxysteroid-Δ8-Δ7-isomerase cause X-linked dominant Conradi-Hunermann syndrome." Nature Genetics **22**(3): 291-294.

Breitwieser, G. E. (2004). "G protein-coupled receptor oligomerization: Implications for G protein activation and cell signaling." Circulation Research **94**(1): 17-27.

Briggs, W. R. (2007a). "The LOV domain: a chromophore module servicing multiple photoreceptors." Journal of Biomedical Science **14**(4): 499-504.

Briggs, W. R., *et al.* (2007b). "Phototropins and their LOV domains: Versatile plant blue-light receptors." Journal of Integrative Plant Biology **49**(1): 4-10.

Brown, D. A. and E. London (1998). "Functions of lipid rafts in biological membranes." Annual Review of Cell and Developmental Biology **14**: 111-136.

Brown, R. E. (1998). "Sphingolipid organization in biomembranes: what physical studies of model membranes reveal." Journal of Cell Science **111**: 1-9.

Brzostowski, J. A. and A. R. Kimmel (2001). "Signaling at zero G: G-protein-independent functions for 7-TM receptors." Trends in Biochemical Sciences **26**(5): 291-297.

Buchanan, S. K. (1999). "β-barrel proteins from bacterial outer membranes: structure, function and refolding." Current Opinion in Structural Biology **9**(4): 455-461.

Bulenger, S., *et al.* (2005). "Emerging role of homo- and heterodimerization in G-protein-coupled receptor biosynthesis and maturation." Trends in Pharmacological Sciences **26**(3): 131-137.

Burgueno, J., *et al.* (2003). "The adenosine A2A receptor interacts with the actin-binding protein alpha-actinin." Journal of Biological Chemistry **278**(39): 37545-37552.

Cabrera-Vera, T. M., *et al.* (2003). "Insights into G protein structure, function, and regulation." Endocrine Reviews **24**(6): 765-781.

Caffrey, M. (2003). "Membrane protein crystallization." Journal of Structural Biology **142**(1): 108-132.

Cereghino, J. L. and J. M. Cregg (2000). "Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*." FEMS Microbiology Reviews **24**(1): 45-66.

Cereghino, G. P., *et al.* (2002). "Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*." Current Opinion in Structural Biology **13**(4): 329-332.

Chae, P. S., *et al.* (2010). "Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins." Nature Methods **7**(12): 1003-1008.

Chalfie, M., *et al.* (1994). "Green fluorescent protein as a marker for gene expression." Science **263**(5148): 802-805.

Chandler, C. S. and F. J. Ballard (1988). "Regulation of the breakdown rates of biotin-containing proteins in Swiss 3T3-L1 cells." Biochemical Journal **251**(3): 749-755.

Chang, G., *et al.* (1998). "Structure of the MscL homolog from *Mycobacterium tuberculosis*: a gated mechanosensitive ion channel." Science **282**(5397): 2220-2226.

Chapman, S., *et al.* (2008). "The photoreversible fluorescent protein iLOV outperforms GFP as a reporter of plant virus infection." Proceedings of the National Academy of Sciences of the United States of America **105**(50): 20038-20043.

Chapman-Smith, A. and J. E. Cronan, Jr. (1999). "*In vivo* enzymatic protein biotinylation." Biomolecular Engineering **16**(1-4): 119-125.

Charlton-Menys, V. and P. Durrington (2008). "Human cholesterol metabolism and therapeutic molecules." Experimental Physiology **93**(1):27-42.

Chen, Y. J., *et al.* (2007). "X-ray structure of EmrE supports dual topology model." Proceedings of the National Academy of Sciences of the United States of America **104**(48): 18999-19004.

Cherezov, V., *et al.* (2006). "Room to move: crystallizing membrane proteins in swollen lipidic mesophases." Journal of Molecular Biology **357**(5): 1605-1618.

Cherezov, V., *et al.* (2007). "High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor." Science **318**(5854): 1258-1265.

Chien, E. Y., *et al.* (2010). "Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist." Science **330**(6007): 1091-1095.

Chini, B. and M. Parenti (2009). "G-protein coupled receptors, cholesterol and palmitoylation: facts about fats." Journal of Molecular Endocrinology **42**: 371-379.

Choi, G., *et al.* (2002). "Structural studies of metarhodopsin II, the activated form of the G-protein coupled receptor, rhodopsin." Biochemistry **41**(23): 7318-7324.

Choi, B.K., *et al.* (2003). "Use of combinatorial genetic libraries to humanise N-linked glycosylation in the yeast *Pichia pastoris*." Proceedings of the National Academy of Sciences USA **100**(9):5022-5027.

Christie, J. M., *et al.* (1998). "Arabidopsis NPH1: a flavoprotein with the properties of a photoreceptor for phototropism." Science **282**(5394): 1698-1701.

Christie, J. M., *et al.* (1999). "LOV (light, oxygen, or voltage) domains of the blue-light photoreceptor phototropin (nph1): binding sites for the chromophore flavin mononucleotide." Proceedings of the National Academy of Sciences of the United States of America **96**(15): 8779-8783.

Christie, J. M., *et al.* (2002). "Phototropin LOV domains exhibit distinct roles in regulating photoreceptor function." Plant Journal **32**(2): 205-219.

Christie, J. M. (2007). "Phototropin blue-light receptors." Annual Review of Plant Biology **58**: 21-45.

Ciruela, F., *et al.* (2004). "Combining mass spectrometry and pull-down techniques for the study of receptor heteromerization. Direct epitope-epitope electrostatic interactions between adenosine A2A and dopamine D2 receptors." Analytical Chemistry **76**(18): 5354-5363.

Cornelius, F. (2001). "Modulation of Na^+,K^+-ATPase and Na^+-ATPase activity by phospholipids and cholesterol. Steady-state kinetics." Biochemistry **40**(30): 8842-8851.

Coso, O. A., *et al.* (1997). "Signaling from G protein-coupled receptors to the c-jun promoter involves the MEF2 transcription factor. Evidence for a novel c-jun amino-terminal kinase-independent pathway." Journal of Biological Chemistry **272**(33): 20691-20697.

Cowan, S. W., *et al.* (1992). "Crystal structures explain functional properties of two *E. coli* porins." Nature **358**(6389): 727-733.

Cregg, J. M., *et al.* (2000). "Recombinant protein expression in *Pichia pastoris*." Molecular Biotechnology **16**(1): 23-52.

Cronan, J. E., Jr. (1990). "Biotination of proteins *in vivo*. A post-translational modification to label, purify, and study proteins." Journal of Biological Chemistry **265**(18): 10327-10333.

Crosson, S. and K. Moffat (2001). "Structure of a flavin-binding plant photoreceptor domain: insights into light-mediated signal transduction." Proceedings of the National Academy of Sciences of the United States of America **98**(6): 2995-3000.

Daniels, T. F., *et al.* (2009). "Lipoproteins, cholesterol homeostasis and cardiac health." International Journal of Biological Sciences **5**(5): 474-488.

Dascal, N. (1997). "Signalling via the G protein-activated K^+ channels." Cellular Signalling **9**(8): 551-573.

Dawson, R. J. and K. P. Locher (2006). "Structure of a bacterial multidrug ABC transporter." Nature **443**(7108): 180-185.

De Waard, M., *et al.* (1997). "Direct binding of G-protein $\beta\gamma$ complex to voltage-dependent calcium channels." Nature **385**(6615): 446-450.

Deisenhofer, J., *et al.* (1984). "X-ray structure analysis of a membrane protein complex. Electron density map at 3 Å resolution and a model of the chromophores of the photosynthetic reaction center from Rhodopseudomonas viridis." Journal of Molecular Biology **180**(2): 385-398.

Deisenhofer, J., *et al.* (1985). "Structure of the protein sub-units in the photosynthetic reaction centre of *Rhodopseudomonas viridis* at 3 Å resolution." Nature 318:618-624.

Deisenhofer, J. and H. Michel (1989). "The photosynthetic reaction center from the purple bacterium *Rhodopseudomonas viridis*." Science **245**(4925): 1463-1473.

Derewenda, Z. S. (2010). "Application of protein engineering to enhance crystallizability and improve crystal properties." Acta Crystallography Section D Biological Crystallography **66**: 604-615.

Derry, J. M., *et al*. (1999). "Mutations in a Δ8-Δ7 sterol isomerase in the tattered mouse and X-linked dominant chondrodysplasia punctata." Nature Genetics **22**(3): 286-290.

Deupi, X. and B. Kobilka (2007). "Activation of G protein-coupled receptors." Advances in Protein Chemistry **74**: 137-166.

DeWire, S. M., *et al*. (2007). "β-arrestins and cell signaling." Annual Review of Physiology **69**: 483-510.

Ding, J., *et al*. (1994). "Binding sites for cholesterol on Ca^{2+}-ATPase studied by using a cholesterol-containing phospholipid." Biochemistry **33**(16): 4974-4979.

Dixon, R. A., *et al*. (1986). "Cloning of the gene and cDNA for mammalian β-adrenergic receptor and homology with rhodopsin." Nature **321**(6065): 75-79.

Do, R., *et al*. (2009). "Squalene synthase: a critical enzyme in the cholesterol biosynthesis pathway." Clinical Genetics **75**(1): 19-29.

Doyle, D. A., *et al*. (1998). "The structure of the potassium channel: molecular basis of K^+ conduction and selectivity." Science **280**(5360): 69-77.

Drepper, T., *et al*. (2007). "Reporter proteins for *in vivo* fluorescence without oxygen." Nature Biotechnology **25**(4): 443-445.

Drew, D., *et al*. (2002). "Rapid topology mapping of *Escherichia coli* inner-membrane proteins by prediction and PhoA/GFP fusion analysis." Proceedings of the National Academy of Sciences of the United States of America **99**(5): 2690-2695.

Drew, D., *et al*. (2005). "A scalable, GFP-based pipeline for membrane protein overexpression screening and purification." Protein Science **14**(8): 2011-2017.

Drew, D., *et al.* (2006). "Optimization of membrane protein overexpression and purification using GFP fusions." Nature Methods **3**(4): 303-313.

Drew, D., *et al.* (2008). "GFP-based optimization scheme for the overexpression and purification of eukaryotic membrane proteins in *Saccharomyces cerevisiae*." Nature Protocols **3**(5): 784-798.

Duke, E. M. H. and L. N. Johnson (2010). "Macromolecular crystallography at synchrotron radiation sources: current status and future developments." Proceedings of the Royal Society **466**(2124): 3421-3452.

Dutzler, R., *et al.* (2002). "X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity." Nature **415**(6869): 287-294.

Dutzler, R., *et al.* (2003). "Gating the selectivity filter in ClC chloride channels." Science **300**(5616): 108-112.

Egli, M. (2010). "Diffraction techniques in structural biology." Current Protocols in Nucleic Acid Chemistry **41**: 7.13.1-7.13.35.

Engstrom, S., *et al.* (2007). "Cubic, sponge, and lamellar phases in the glyceryl monooleyl ether-propylene glycol-water system." Langmuir **23**(20): 10020-10025.

Epand, R. M. (2006). "Cholesterol and the interaction of proteins with membrane domains." Progress in Lipid Research **45**(4): 279-294.

Fall, R. R. (1979). "Analysis of microbial biotin proteins." Methods in Enzymology **62**: 390-398.

Fantoni, A., *et al.* (2007). "Improved yields of full-length functional human FGF1 can be achieved using the methylotrophic yeast *Pichia pastoris*." Protein Expression and Purification **52**(1): 31-39.

Ferdinandusse, S., *et al.* (2009). "Toxicity of peroxisomal C27-bile acid intermediates." Molecular Genetics and Metabolism **96**(3): 121-128.

Ferguson, S. S. (2001). "Evolving concepts in G protein-coupled receptor endocytosis: the role in receptor desensitization and signaling." Pharmacological Reviews **53**(1): 1-24.

Fernandez-Ballester, G., *et al.* (1994). "A role for cholesterol as a structural effector of the nicotinic acetylcholine receptor." Biochemistry **33**(13): 4065-4071.

Ferre, S., *et al.* (2001). "Adenosine/dopamine interaction: implications for the treatment of Parkinson's disease." Parkinsonism Related Disorders **7**(3): 235-241.

Ferreira, K. N., *et al.* (2004). "Architecture of the photosynthetic oxygen-evolving center." Science **303**(5665): 1831-1838.

Flower, D. R. (1999). "Modelling G-protein-coupled receptors for drug design." Biochimica et Biophysica Acta (BBA) - Biomembranes **1422**(3): 207-234.

Ford, C. F., *et al.* (1991). "Fusion tails for the recovery and purification of recombinant proteins." Protein Expression and Purification **2**(2-3): 95-107.

Fraser, N. J. (2006). "Expression and functional purification of a glycosylation deficient version of the human adenosine 2a receptor for structural studies." Protein Expression and Purification **49**(1): 129-137.

Fredholm, B. B., *et al.* (2007a). "G-protein-coupled receptors: an update." Acta Physiologica **190**(1): 3-7.

Fredholm, B. B., *et al.* (2007b). "Aspects of the general biology of adenosine A2A signaling." Progress in Neurobiology **83**(5): 263-276.

Fu, D., *et al.* (2000). "Structure of a glycerol-conducting channel and the basis for its selectivity." Science **290**(5491): 481-486.

Fuxe, K., *et al.* (2007). "Adenosine A(2A) receptors, dopamine D(2) receptors and their interactions in Parkinson's disease." Movement Disorders **22**(14): 1990-2017.

Garavito, R. M., *et al.* (1996). "Strategies for crystallizing membrane proteins." Journal of Bioenergetics and Biomembranes **28**(1): 13-27.

Gautier, A. *et al.* (2010). "Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy." Nature Structural and Molecular Biology **17**(6):768-774.

Gay, S.C. *et al.* (2010). "Structures of cytochrome P450 2B4 complexed with the anti-platelet drugs ticlopidine and clopidogrel." Biochemistry **49**(40):8709-8720.

Gemmill, T.R. and R.B. Trimble (1999). "Overview of N- and O-linked oligosaccharide structures found in various yeast strains." Biochimica et Biophysica Acta **1426**(2):227-237.

George, S. R., *et al.* (2002). "G-protein-coupled receptor oligomerization and its potential for drug discovery." Nature Reviews Drug Discovery **1**(10): 808-820.

Gimpl, G., *et al.* (1995). "Expression of the human oxytocin receptor in baculovirus-infected insect cells: high-affinity binding is induced by a cholesterol-cyclodextrin complex." Biochemistry **34**(42): 13794-13801.

Gimpl, G., *et al.* (1997). "Cholesterol as modulator of receptor function." Biochemistry **36**(36): 10959-10974.

Gimpl, G. and F. Fahrenholz (2002). "Cholesterol as stabilizer of the oxytocin receptor." Biochimica et Biophysica Acta (BBA) - Biomembranes **1564**(2): 384-392.

Gomez, G. and M. V. Sitkovsky (2003). "Targeting G protein-coupled A2a adenosine receptors to engineer inflammation *in vivo*." International Journal of Biochemistry and Cell Biology **35**(4): 410-414.

Gorzelle, B. M., *et al.* (1999). "Reconstitutive refolding of diacylglycerol kinase, an integral membrane protein." Biochemistry **38**(49): 16373-16382.

Grunewald, S., *et al.* (2004). "Production of the human D2S receptor in the methylotrophic yeast *P. pastoris*." Receptors and channels **10**(1): 37-50.

Gsandtner, I., *et al.* (2005). "Heterotrimeric G protein-independent signaling of a G protein-coupled receptor. Direct binding of ARNO/cytohesin-2 to the carboxyl terminus of the A2A adenosine receptor is necessary for sustained activation of the ERK/MAP kinase pathway." Journal of Biological Chemistry **280**(36): 31898-31905.

Gsandtner, I. and M. Freissmuth (2006). "A tail of two signals: the C terminus of the A(2A)-adenosine receptor recruits alternative signaling pathways." Molecular Pharmacology **70**(2): 447-449.

Gsponer, J. and M. M. Babu (2009). "The rules of disorder or why disorder rules." Progress in Biophysics and Molecular Biology **99**(2-3): 94-103.

Gurevich, E. V. and V. V. Gurevich (2006). "Arrestins: ubiquitous regulators of cellular signaling pathways." Genome Biology **7**(9): 236.

Gutmann, D. A., *et al.* (2007). "A high-throughput method for membrane protein solubility screening: the ultracentrifugation dispersity sedimentation assay." Protein Science **16**(7): 1422-1428.

Hall, R. A. and R. J. Lefkowitz (2002). "Regulation of G protein-coupled receptor signaling by scaffold proteins." Circulation Research **91**(8): 672-680.

Hall, R. A., *et al.* (1998a). "A C-terminal motif found in the β2-adrenergic receptor, P2Y1 receptor and cystic fibrosis transmembrane conductance regulator determines binding to the Na^+/H^+ exchanger regulatory factor family of PDZ proteins." Proceedings of the National Academy of Sciences of the United States of America **95**(15): 8496-8501.

Hall, R. A., *et al.* (1998b). "The β2-adrenergic receptor interacts with the Na^+/H^+-exchanger regulatory factor to control Na^+/H^+ exchange." Nature **392**(6676): 626-630.

Hamilton, S.R., *et al.* (2003). "Production of complex human glycoproteins in yeast." Science **301**(5637): 1244-1246.

Hamilton, S.R. *et al.* (2006). "Humanisation of yeast to produce complex terminally sialylated glycoproteins." Science **313**(5792):1441-1443.

Hamilton, S. R. and T. U. Gerngross (2007). "Glycosylation engineering in yeast: the advent of fully humanized yeast." Current Opinion in Structural Biology **18**(5): 387-392.

Hamm, H. E. (1998). "The many faces of G protein signaling." Journal of Biological Chemistry **273**(2): 669-672.

Hamman, B. D., *et al.* (2002). "Binding of a Pleckstrin homology domain protein to phosphoinositide in membranes: a miniaturized FRET-based assay for drug screening." Journal of Biomolecular Screening **7**(1): 45-55.

Hanson, M. A., *et al.* (2008). "A specific cholesterol binding site is established by the 2.8 Å structure of the human beta2-adrenergic receptor." Structure **16**(6): 897-905.

Hartner, F. S. and A. Glieder (2006). "Regulation of methanol utilisation pathway genes in yeasts." Microbial Cell Factories **5**:39.

Has, C., *et al.* (2000). "The Conradi-Hunermann-Happle syndrome (CDPX2) and emopamil binding protein: novel mutations, and somatic and gonadal mosaicism." Human Molecular Genetics **9**(13): 1951-1955.

He, J., *et al.* (2010). "Minireview: Nuclear receptor-controlled steroid hormone synthesis and metabolism." Molecular Endocrinology **24**(1): 11-21.

He, X.A. *et al.* (2010). "Structure of a cation-bound multidrug and toxic compound extrusions transporter." Nature **467**(7318):991-994.

Heim, R., *et al.* (1994). "Wavelength mutations and posttranslational autoxidation of green fluorescent protein." Proceedings of the National Academy of Sciences of the United States of America **91**(26): 12501-12504.

Henderson, R. and P. N. Unwin (1975). "Three-dimensional model of purple membrane obtained by electron microscopy." Nature **257**(5521): 28-32.

Herlitze, S., *et al.* (1996). "Modulation of Ca^{2+} channels by G-protein $\beta\gamma$ subunits." Nature **380**(6571): 258-262.

Hernandez-Guzman, F. G., *et al.* (2003). "Structure of human estrone sulfatase suggests functional roles of membrane association." Journal of Biological Chemistry **278**(25): 22989-22997.

Higgins, C. F. (1992). "ABC transporters: from microorganisms to man." Annual Review of Cellular Biology **8**: 67-113.

Hiller, S. *et al.* (2008). "Solution structure of the integral human membrane protein VDAC-1 in detergent micelles." Science **321**(5893):1206-1210.

Ho, J. D., *et al.* (2009). "Crystal structure of human aquaporin 4 at 1.8 Å and its mechanism of conductance." Proceedings of the National Academy of Sciences of the United States of America **106**(18): 7437-7442.

Holick, M. F. (1995). "Defects in the synthesis and metabolism of vitamin D." Experimental and Clinical Endocrinology & Diabetes **103**(4): 219-227.

Holick, M. F. (1995). "Environmental factors that influence the cutaneous production of vitamin D." American Journal of Clinical Nutrition **61**(3 Suppl): 638S-645S.

Hollenstein, K., *et al.* (2007). "Structure of an ABC transporter in complex with its binding protein." Nature **446**(7132): 213-216.

Horsefield, R., *et al.* (2008). "High-resolution X-ray structure of human aquaporin 5." Proceedings of the National Academy of Sciences of the United States of America **105**(36): 13327-13332.

Howell, S. C., *et al.* (2010). "CHOBIMALT: A Cholesterol-Based Detergent." Biochemistry. **49**(44):9572-9583.

Huang, J. R., *et al.* (2007). "Stable intermediate states and high energy barriers in the unfolding of GFP." Journal of Molecular Biology **370**(2): 356-371.

Huang, Y., *et al.* (2003). "Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*." Science **301**(5633): 616-620.

Hunte, C. and H. Michel (2002). "Crystallisation of membrane proteins mediated by antibody fragments." Current Opinion in Structural Biology **12**(4): 503-508.

Hunte, C., *et al.* (2005). "Structure of a Na^+/H^+ antiporter and insights into mechanism of action and regulation by pH." Nature **435**(7046): 1197-1202.

Hunte, C., *et al.* (2010). "Functional modules and structural basis of conformational coupling in mitochondrial complex I." Science **329**(5990): 448-451.

Ikeda, S. R. (1996). "Voltage-dependent modulation of N-type calcium channels by G-protein $\beta\gamma$ subunits." Nature **380**(6571): 255-258.

Ikonen, E. (2006). "Mechanisms for cellular cholesterol transport: defects and human disease." Physiological Reviews **86**(4): 1237-1261.

Iwata, S., *et al.* (1995). "Structure at 2.8 Å resolution of cytochrome c oxidase from *Paracoccus denitrificans*." Nature **376**(6542): 660-669.

Iwata, S. (1998a). "Structure and function of bacterial cytochrome c oxidase." Journal of Biochemistry **123**(3): 369-375.

Iwata, S., *et al.* (1998b). "Complete structure of the 11-subunit bovine mitochondrial cytochrome bc1 complex." Science **281**(5373): 64-71.

Jaakola, V. P., *et al.* (2008). "The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist." Science **322**(5905): 1211-1217.

Jaakola, V. P. and A. P. Ijzerman (2010). "The crystallographic structure of the human adenosine A(2A) receptor in a high-affinity antagonist-bound state: implications for GPCR drug screening and design." Current Opinion in Structural Biology.

Jacobson, K. A. and Z. G. Gao (2006). "Adenosine receptors as therapeutic targets." Nature Reviews Drug Discovery **5**(3): 247-264.

Jahic, M., *et al*. (2006). "Process technology for production and recovery of heterologous proteins with *Pichia pastoris*." Biotechnology Progress **22**(6): 1465-1473.

Jasti, J., *et al*. (2007). "Structure of acid-sensing ion channel 1 at 1.9 Å resolution and low pH." Nature **449**(7160): 316-323.

Javitch, J. A. (2004). "The ants go marching two by two: oligomeric structure of G-protein-coupled receptors." Molecular Pharmacology **66**(5): 1077-1082.

Jiang, Y., *et al*. (2002a). "Crystal structure and mechanism of a calcium-gated potassium channel." Nature **417**(6888): 515-522.

Jiang, Y., *et al*. (2002b). "The open pore conformation of potassium channels." Nature **417**(6888): 523-526.

Johansson, L. C., *et al*. (2009). "Membrane protein crystallization from lipidic phases." Current Opinion in Structural Biology.

Johnson, W. C. (1999). "Analyzing protein circular dichroism spectra for accurate secondary structures." Proteins **35**(3): 307-312.

Jones, K. A., *et al*. (1998). "GABA(B) receptors function as a heteromeric assembly of the subunits GABA(B)R1 and GABA(B)R2." Nature **396**(6712): 674-679.

Jones, M. A. and J. M. Christie (2008). "Phototropin receptor kinase activation by blue light." Plant Signaling and Behavior **3**(1): 44-46.

Julius, D., *et al*. (1984a). "Glycosylation and processing of prepro-α-factor through the yeast secretory pathway." Cell **36**(2): 309-318.

Julius, D., *et al*. (1984b). "Isolation of the putative structural gene for the lysine-arginine-cleaving endopeptidase required for processing of yeast prepro-alpha-factor." Cell **37**(3): 1075-1089.

Kabsch, W. and C. Sander (1983). "Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features." Biopolymers **22**(12): 2577-2637.

Kamiya, N. and J. R. Shen (2003). "Crystal structure of oxygen-evolving photosystem II from *Thermosynechococcus vulcanus* at 3.7 Å resolution." Proceedings of the National Academy of Sciences of the United States of America **100**(1): 98-103.

Katona, G., *et al.* (2003). "Lipidic cubic phase crystal structure of the photosynthetic reaction centre from *Rhodobacter sphaeroides* at 2.35 Å resolution." Journal of Molecular Biology **331**(3): 681-692.

Kaupmann, K., *et al.* (1998). "GABA(B)-receptor subtypes assemble into functional heteromeric complexes." Nature **396**(6712): 683-687.

Kawate, T. and E. Gouaux (2006). "Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins." Structure **14**(4): 673-681.

Kawate, T., *et al.* (2009). "Crystal structure of the ATP-gated P2X(4) ion channel in the closed state." Nature **460**(7255): 592-598.

Kelly, E., *et al.* (2008). "Agonist-selective mechanisms of GPCR desensitization." British Journal of Pharmacology **153 Suppl 1**: S379-388.

Kelly, S. M., *et al.* (2005). "How to study proteins by circular dichroism." Biochimica et Biophysica Acta (BBA) - Biomembranes **1751**(2): 119-139.

Kennedy, S. J. (1978). "Structures of membrane proteins." Journal of Membrane Biology **42**(3): 265-279.

Klabunde, T. and G. Hessler (2002). "Drug design strategies for targeting G-protein-coupled receptors." ChemBioChem **3**(10): 928-944.

Klein, U., *et al.* (1995). "Alteration of the myometrial plasma membrane cholesterol content with beta-cyclodextrin modulates the binding affinity of the oxytocin receptor." Biochemistry **34**(42): 13784-13793.

Klinger, M., *et al.* (2002). "Removal of the carboxy terminus of the A2A-adenosine receptor blunts constitutive activity: differential effect on cAMP accumulation and MAP kinase stimulation." Naunyn-Schmiedeberg's Archives of Pharmacology **366**(4): 287-298.

Knott, J. A., *et al.* (1989). "The expression and purification of human rhinovirus protease 3C." European Journal of Biochemistry **182**(3): 547-555.

Kobilka, B. and G. F. Schertler (2008). "New G-protein-coupled receptor crystal structures: insights and limitations." Trends in Pharmacological Sciences.

Kobilka, B. K. (2007). "G protein coupled receptor structure and activation." Biochimica et Biophysica Acta **1768**(4): 794-807.

Kobilka, B. K. and X. Deupi (2007). "Conformational complexity of G-protein-coupled receptors." Trends in Pharmacological Sciences **28**(8): 397-406.

Koch, W. J., *et al*. (1994). "Direct evidence that G_i-coupled receptor stimulation of mitogen-activated protein kinase is mediated by G βγ activation of p21ras." Proceedings of the National Academy of Sciences of the United States of America **91**(26): 12706-12710.

Koepke, J., *et al*. (1996). "The crystal structure of the light-harvesting complex II (B800-850) from *Rhodospirillum molischianum*." Structure **4**(5): 581-597.

Kolb, P., *et al*. (2009). "Structure-based discovery of β2-adrenergic receptor ligands." Proceedings of the National Academy of Sciences of the United States of America **106**(16): 6843-6848.

Krauss, N., *et al*. (1996). "Photosystem I at 4 angstrom resolution represents the first structural model of a joint photosynthetic reaction centre and core antenna system." Nature Structural Biology **3**(10):965-973.

Krueger-Koplin, R. D., *et al*. (2004). "An evaluation of detergents for NMR structural studies of membrane proteins." Journal of Biomolecular NMR **28**(1): 43-57.

Kuner, R., *et al*. (1999). "Role of heteromer formation in GABA(B) receptor function." Science **283**(5398): 74-77.

Kuo, A., *et al*. (2003). "Crystal structure of the potassium channel KirBac1.1 in the closed state." Science **300**(5627): 1922-1926.

Lacapere, J.-J., *et al*. (2007). "Determining membrane protein structures: still a challenge!" Trends in Biochemical Sciences **32**(6): 259-270.

Ladds, G., *et al*. (2003). "Modified yeast cells to investigate the coupling of G protein-coupled receptors to specific G proteins." Molecular Microbiology **47**(3): 781-792.

Lagane, B., *et al*. (2000). "Role of sterols in modulating the human μ-opioid receptor function in *Saccharomyces cerevisiae*." Journal of Biological Chemistry **275**(43): 33197-33200.

Lambright, D. G., *et al*. (1996). "The 2.0 Å crystal structure of a heterotrimeric G protein." Nature **379**(6563): 311-319.

Landau, E. M. and J. P. Rosenbusch (1996). "Lipidic cubic phases: a novel concept for the crystallization of membrane proteins." Proceedings of the National Academy of Sciences of the United States of America **93**(25): 14532-14535.

Law, C. J., *et al*. (2008). "Ins and outs of major facilitator superfamily antiporters." Annual Review of Microbiology **62**: 289-305.

Lefkowitz, R. J. (1998). "G protein-coupled receptors. New roles for receptor kinases and beta-arrestins in receptor signaling and desensitization." Journal of Biological Chemistry **273**(30): 18677-18680.

Li, J., *et al*. (2004). "Structure of bovine rhodopsin in a trigonal crystal form." Journal of Molecular Biology **343**(5): 1409-1438.

Lim, F., *et al*. (1987). "Pyruvate carboxylase in the yeast pyc mutant." Archives of Biochemistry and Biophysics **258**(1): 259-264.

Link, A. J., *et al*. (2008). "Efficient production of membrane-integrated and detergent-soluble G protein-coupled receptors in *Escherichia coli*." Protein Science **17**(10): 1857-1863.

Liu, Z., *et al*. (2004). "Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution." Nature **428**(6980): 287-292.

Locher, K. P., *et al*. (2002). "The *E. coli* BtuCD structure: a framework for ABC transporter architecture and mechanism." Science **296**(5570): 1091-1098.

Logothetis, D. E., *et al*. (1987). "The $\beta\gamma$ subunits of GTP-binding proteins activate the muscarinic K^+ channel in heart." Nature **325**(6102): 321-326.

Long, S. B., *et al*. (2005). "Crystal structure of a mammalian voltage-dependent *Shaker* family K^+ channel." Science **309**(5736): 897-903.

Lu, M. and D. Fu (2007). "Structure of the zinc transporter YiiP." Science **317**(5845): 1746-1748.

Lucero, H. A. and P. W. Robbins (2004). "Lipid raft-protein association and the regulation of protein activity." Archives of Biochemistry and Biophysics **426**(2): 208-224.

Lundström, K. (2005). "Structural genomics of GPCRs." Trends in Biotechnology **23**(2): 103-108.

Lundström, K., *et al*. (2006). "Structural genomics on membrane proteins: comparison of more than 100 GPCRs in 3 expression systems." Journal of Structural and Functional Genomics **7**(2): 77-91.

Luttrell, L. M., *et al*. (2001). "Activation and targeting of extracellular signal-regulated kinases by β-arrestin scaffolds." Proceedings of the National Academy of Sciences of the United States of America **98**(5): 2449-2454.

Luttrell, L. M. and D. Gesty-Palmer (2010). "Beyond desensitization: physiological relevance of arrestin-dependent signaling." Pharmacological Reviews **62**: 305-330.

Ma, L. and G. Pei (2007). "β-arrestin signaling and regulation of transcription." Journal of Cell Science **120**(Pt 2): 213-218.

Macauley-Patrick, S., *et al.* (2005). "Heterologous protein production using the *Pichia pastoris* expression system." Yeast **22**(4): 249-270.

Magnani, F., *et al.* (2004). "Partitioning of the serotonin transporter into lipid microdomains modulates transport of serotonin." Journal of Biological Chemistry **279**(37): 38770-38778.

Magnani, F., *et al.* (2008). "Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor." Proceedings of the National Academy of Sciences of the United States of America **105**(31): 10744-10749.

Malo, G. D., *et al.* (2007). "X-ray structure of Cerulean GFP: a tryptophan-based chromophore useful for fluorescence lifetime imaging." Biochemistry **46**(35): 9865-9873.

Manavalan, P. and W. C. Johnson, Jr. (1987). "Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra." Analytical Biochemistry **167**(1): 76-85.

Marrero, M. B., *et al.* (1998). "Regulation of angiotensin II-induced JAK2 tyrosine phosphorylation: roles of SHP-1 and SHP-2." American Journal of Physiology **275**(5 Pt 1): C1216-1223.

Martinez Molina, D., *et al.* (2007). "Structural basis for synthesis of inflammatory mediators by human leukotriene C4 synthase." Nature **448**(7153): 613-616.

Matsuoka, D. and S. Tokutomi (2005). "Blue light-regulated molecular switch of Ser/Thr kinase in phototropin." Proceedings of the National Academy of Sciences of the United States of America **102**(37): 13337-13342.

Maudsley, S., *et al.* (2005). "The origins of diversity and specificity in G protein-coupled receptor signaling." Journal of Pharmacology and Experimental Therapeutics **314**(2): 485-494.

McDermott, G., *et al.* (1995). "Crystal structure of an integral membrane light-harvesting complex from photosynthetic bacteria." Nature **374**: 517-521.

McLuskey, K., *et al.* (2001). "The crystallographic structure of the B800-820 LH3 light-harvesting complex from the purple bacteria *Rhodopseudomonas acidophila* strain 7050." Biochemistry **40**(30): 8783-8789.

McLuskey, K., *et al.* (2010). "Crystal structures of all-alpha type membrane proteins." European Biophysical Journal **39**(5): 723-755.

Meinild, A. K., *et al.* (1998). "Bidirectional water fluxes and specificity for small hydrophilic molecules in aquaporins 0-5." Journal of Biological Chemistry **273**(49): 32446-32451.

Michel, H. (1982). "Three-dimensional crystals of a membrane protein complex. The photosynthetic reaction centre from *Rhodopseudomonas viridis*." Journal of Molecular Biology **158**(3): 567-572.

Milligan, G. (2007). "G protein-coupled receptor dimerisation: molecular basis and relevance to function." Biochimica et Biophysica Acta (BBA) - Biomembranes **1768**(4): 825-835.

Milligan, G. (2009). "G protein-coupled receptor hetero-dimerization: contribution to pharmacology and function." British Journal of Pharmacology. **158**: 5-14.

Milojevic, T., *et al.* (2006). "The ubiquitin-specific protease Usp4 regulates the cell surface level of the A(2A) receptor." Molecular Pharmacology **69**(4): 1083-1094.

Misquitta, Y. and M. Caffrey (2003). "Detergents destabilize the cubic phase of monoolein: implications for membrane protein crystallization." Biophysical Journal **85**(5): 3084-3096.

Mize, G. J., *et al.* (2008). "Regulated expression of active biotinylated G-protein coupled receptors in mammalian cells." Protein Expression and Purification **57**(2): 280-289.

Moebius, F. F., *et al.* (1998). "Pharmacological analysis of sterol Δ8-Δ7 isomerase proteins with [3H]ifenprodil." Molecular Pharmacology **54**(3): 591-598.

Moebius, F. F., *et al.* (2000). "Genetic defects in postsqualene cholesterol biosynthesis." Trends in Endocrinology and Metabolism **11**(3): 106-114.

Mohanty, A. K., *et al*. (2003). "Inhibition of tobacco etch virus protease activity by detergents." Protein Expression and Purification **27**(1): 109-114.

Moreau, J. L. and G. Huber (1999). "Central adenosine A(2A) receptors: an overview." Brain Research **31**(1): 65-82.

Murata, K., *et al*. (2000). "Structural determinants of water permeation through aquaporin-1." Nature **407**(6804): 599-605.

Murphree, L. J., *et al*. (2002). "Human A(2A) adenosine receptors: high-affinity agonist binding to receptor-G protein complexes containing Gβ(4)." Molecular Pharmacology **61**(2): 455-462.

Nakamichi, H. and T. Okada (2006a). "Crystallographic analysis of primary visual photochemistry." Angewandte Chemie International Edition **45**(26): 4270-4273.

Nakamichi, H. and T. Okada (2006b). "Local peptide movement in the photoreaction intermediate of rhodopsin." Proceedings of the National Academy of Sciences of the United States of America **103**(34): 12729-12734.

Nelsetuen, G. L. (1995). "How enzymes work." Principles of Medical Biology. **4:** 25-44.

Neves, S. R., *et al*. (2002). "G protein pathways." Science **296**(5573): 1636-1639.

Newstead, S., *et al*. (2007). "High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in *Saccharomyces cerevisiae*." Proceedings of the National Academy of Sciences of the United States of America.

Newton-Vinson, P., *et al*. (2000). "High-level expression of human liver monoamine oxidase B in *Pichia pastoris*." Protein Expression and Purification **20**(2): 334-345.

Ng, G. Y., *et al*. (1999). "Identification of a GABA(B) receptor subunit, gb2, required for functional GABA(B) receptor activity." Journal of Biological Chemistry **274**(12): 7607-7610.

Nollert, P. (2005). "Membrane protein crystallization in amphiphile phases: practical and theoretical considerations." Progress in Biophysics and Molecular Biology **88**(3): 339-357.

Nyholm, T. K., *et al.* (2007). "How protein transmembrane segments sense the lipid environment." Biochemistry **46**(6): 1457-1465.

Obara, Y., *et al.* (2008). "Betagamma subunits of $G_{(i/o)}$ suppress EGF-induced ERK5 phosphorylation, whereas ERK1/2 phosphorylation is enhanced." Cellular Signalling **20**(7): 1275-1283.

Ohta, A. and M. Sitkovsky (2001). "Role of G-protein-coupled adenosine receptors in downregulation of inflammation and protection from tissue damage." Nature **414**(6866): 916-920.

Okada, T., *et al.* (2002). "Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography." Proceedings of the National Academy of Sciences of the United States of America **99**(9): 5982-5987.

Olah, M. E. (1997). "Identification of A2a adenosine receptor domains involved in selective coupling to G_s. Analysis of chimeric A1/A2a adenosine receptors." Journal of Biological Chemistry **272**(1): 337-344.

Oldham, M. L., *et al.* (2007). "Crystal structure of a catalytic intermediate of the maltose transporter." Nature **450**(7169): 515-521.

Opekarova, M. and W. Tanner (2003). "Specific lipid requirements of membrane proteins -a putative bottleneck in heterologous expression." Biochimica et Biophysica Acta (BBA) - Biomembranes **1610**(1): 11-22.

Ormo, M., *et al.* (1996). "Crystal structure of the *Aequorea victoria* green fluorescent protein." Science **273**(5280): 1392-1395.

Overington, J. P., *et al.* (2006). "How many drug targets are there?" Nature Reviews Drug Discovery **5**(12): 993-996.

Palczewski, K., *et al.* (2000). "Crystal structure of rhodopsin: A G protein-coupled receptor." Science **289**(5480): 739-745.

Palmer, T. M., *et al.* (1995). "125I-4-(2-[7-amino-2-[2-furyl][1,2,4]triazolo[2,3-a][1,3,5] triazin-5-yl-amino]ethyl)phenol, a high affinity antagonist radioligand selective for the A2a adenosine receptor." Molecular Pharmacology **48**(6): 970-974.

Pang, L., *et al.* (1999). "Membrane cholesterol modulates galanin-GalR2 interaction." Biochemistry **38**(37): 12003-12011.

Park, J. H., *et al.* (2008). "Crystal structure of the ligand-free G-protein-coupled receptor opsin." Nature **454**(7201): 183-187.

Pauling, L. and R. B. Corey (1951a). "Configuration of polypeptide chains." Nature **168**(4274): 550-551.

Pauling, L. and R. B. Corey (1951b). "The pleated sheet, a new layer configuration of polypeptide chains." Proceedings of the National Academy of Sciences of the United States of America **37**(5): 251-256.

Pauling, L., *et al.* (1951). "The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain." Proceedings of the National Academy of Sciences of the United States of America **37**(4): 205-211.

Pautsch, A. and G. E. Schulz (1998). "Structure of the outer membrane protein A transmembrane domain." Nature Structural Biology **5**(11): 1013-1017.

Piersen, C. E., *et al.* (1994). "A carboxyl-terminally truncated mutant and nonglycosylated A2a adenosine receptors retain ligand binding." Molecular Pharmacology **45**(5): 861-870.

Pinkett, H. W., *et al.* (2007). "An inward-facing conformation of a putative metal-chelate-type ABC transporter." Science **315**(5810): 373-377.

Postis, V. L., *et al.* (2008). "A high-throughput assay of membrane protein stability." Molecular Membrane Biology **25**(8): 617-624.

Poucher, S. M., *et al.* (1995). "The in vitro pharmacology of ZM 241385, a potent, non-xanthine A2a selective adenosine receptor antagonist." British Journal of Pharmacology **115**(6): 1096-1102.

Prasher, D. C., *et al.* (1992). "Primary structure of the *Aequorea victoria* green-fluorescent protein." Gene **111**(2): 229-233.

Prive, G. G. (2007). "Detergents for the stabilization and crystallization of membrane proteins." Methods **41**(4): 388-397.

Provencher, S. W. and J. Glockner (1981). "Estimation of globular protein secondary structure from circular dichroism." Biochemistry **20**(1): 33-37.

Raines, K. S., *et al.* (2010). "Three-dimensional structure determination from a single view." Nature **463**(7278): 214-217.

Rajagopal, S., *et al.* (2010). "Teaching old receptors new tricks: biasing seven-transmembrane receptors." Nature Reviews Drug Discovery **9**(5): 373-386.

Ramstedt, B. and J. P. Slotte (2006). "Sphingolipids and the formation of sterol-enriched ordered membrane domains." Biochimica et Biophysica Acta (BBA) - Biomembranes **1758**(12): 1945-1956.

Rasmussen, S. G., *et al.* (2007). "Crystal structure of the human β2 adrenergic G-protein-coupled receptor." Nature **450**(7168): 383-387.

Redding, K., *et al.* (1991). "Immunolocalization of Kex2 protease identifies a putative late Golgi compartment in the yeast *Saccharomyces cerevisiae*." Journal of Cell Biology **113**(3): 527-538.

Rees, D. C., *et al.* (2009). "ABC transporters: the power to change." Nature Reviews Molecular Cell Biology **10**(3): 218-227.

Reilander, H. and H. M. Weiss (1998). "Production of G-protein-coupled receptors in yeast." Current Opinion in Biotechnology **9**(5): 510-517.

Remington, S. J. (2006). "Fluorescent proteins: maturation, photochemistry and photophysics." Current Opinion in Structural Biology **16**(6): 714-721.

Rietveld, A. and K. Simons (1998). "The differential miscibility of lipids as the basis for the formation of functional membrane rafts." Biochimica et Biophysica Acta (BBA) - Biomembranes **1376**(3): 467-479.

Ritter, S. L. and R. A. Hall (2009). "Fine-tuning of GPCR activity by receptor-interacting proteins." Nature Reviews Molecular Cell Biology **10**(12): 819-830.

Robinson, B. H., *et al.* (1983). "[^{3}H]biotin-labeled proteins in cultured human skin fibroblasts from patients with pyruvate carboxylase deficiency." Journal of Biological Chemistry **258**(10): 6660-6664.

Romano, C., *et al.* (1996). "Metabotropic glutamate receptor 5 is a disulfide-linked dimer." Journal of Biological Chemistry **271**(45): 28612-28616.

Romano, C., *et al.* (2001). "Covalent and noncovalent interactions mediate metabotropic glutamate receptor mGlu5 dimerization." Molecular Pharmacology **59**(1): 46-53.

Rosenbaum, D. M., *et al.* (2007). "GPCR engineering yields high-resolution structural insights into β2-adrenergic receptor function." Science **318**(5854): 1266-1273.

Roszak, A. W., *et al.* (2003). "Crystal structure of the RC-LH1 core complex from *Rhodopseudomonas palustris*." Science **302**(5652): 1969-1972.

Rothnie, A., *et al.* (2001). "The importance of cholesterol in maintenance of P-glycoprotein activity and its membrane perturbing influence." European Biophysical Journal **30**(6): 430-442.

Rovati, G. E., *et al.* (2007). "The highly conserved DRY motif of class A G protein-coupled receptors: beyond the ground state." Molecular Pharmacology **71**(4): 959-964.

Saier, M. H., Jr., *et al.* (1999). "The major facilitator superfamily." Journal of Molecular Microbiology and Biotechnology **1**(2): 257-279.

Salom, D., *et al.* (2006). "Crystal structure of a photoactivated deprotonated intermediate of rhodopsin." Proceedings of the National Academy of Sciences of the United States of America **103**(44): 16123-16128.

Salomon, M., *et al.* (2000). "Photochemical and mutational analysis of the FMN-binding domains of the plant blue light receptor, phototropin." Biochemistry **39**(31): 9401-9410.

Sanders, C. R. and K. Oxenoid (2000). "Customizing model membranes and samples for NMR spectroscopic studies of complex membrane proteins." Biochimica et Biophysica Acta (BBA) - Biomembranes **1508**(1-2): 129-145.

Sands, W. A., *et al.* (2004). "Specific inhibition of nuclear factor-κB-dependent inflammatory responses by cell type-specific mechanisms upon A(2A) adenosine receptor gene transfer." Molecular Pharmacology **66**(5): 1147-1159.

Sands, W. A. and T. M. Palmer (2005). "Adenosine receptors and the control of endothelial cell function in inflammatory disease." Immunology Letters **101**(1): 1-11.

Sarramegna, V., *et al.* (2002). "Optimizing functional versus total expression of the human μ-opioid receptor in *Pichia pastoris*." Protein Expression and Purification **24**(2): 212-220.

Sarramegna, V., *et al.* (2003). "Heterologous expression of G-protein-coupled receptors: comparison of expression systems fron the standpoint of large-scale production and purification." Cellular and Molecular Life Sciences **60**(8): 1529-1546.

Savage, D. F., *et al.* (2003). "Architecture and selectivity in aquaporins: 2.5 Å X-ray structure of aquaporin Z." PLoS Biol **1**(3): E72.

Scheerer, P., *et al.* (2008). "Crystal structure of opsin in its G-protein-interacting conformation." Nature **455**(7212): 497-502.

Schmidt, M., *et al.* (2010). "Five-dimensional crystallography." Acta Crystallographica A **66**: 198-206.

Schmitt, J. M. and P. J. Stork (2002). "Gα and Gβγ require distinct Src-dependent pathways to activate Rap1 and Ras." Journal of Biological Chemistry **277**(45): 43024-43032.

Schmitz, G. and M. Grandl (2009). "The molecular mechanisms of HDL and associated vesicular trafficking mechanisms to mediate cellular lipid homeostasis." Arteriosclerosis, Thrombosis and Vascular Biology **29**(11): 1718-1722.

Schubert, W. D., *et al.* (1997). "Photosystem I of *Synechococcus elongatus* at 4 Å resolution: comprehensive structure analysis." Journal of Molecular Biology **272**(5): 741-769.

Schulz, G. E. (2000). "β-barrel membrane proteins." Current Opinion in Structural Biology **10**(4): 443-447.

Seddon, A. M., *et al.* (2004). "Membrane proteins, lipids and detergents: not just a soap opera." Biochimica et Biophysica Acta **1666**(1-2): 105-117.

Serrano-Vega, M. J., *et al.* (2008). "Conformational thermostabilization of the β1-adrenergic receptor in a detergent-resistant form." Proceedings of the National Academy of Sciences of the United States of America **105**(3): 877-882.

Sheikh, S. P., *et al.* (1996). "Rhodopsin activation blocked by metal-ion-binding sites linking transmembrane helices C and F." Nature **383**(6598): 347-350.

Shi, N., *et al.* (2006). "Atomic structure of a Na^+- and K^+-conducting channel." Nature **440**(7083): 570-574.

Shouffani, A. and B. I. Kanner (1990). "Cholesterol is required for the reconstruction of the sodium- and chloride-coupled, γ-aminobutyric acid transporter from rat brain." Journal of Biological Chemistry **265**(11): 6002-6008.

Shukla, A. K., *et al.* (2007). "Heterologous expression and characterization of the recombinant bradykinin B2 receptor using the methylotrophic yeast *Pichia pastoris*." Protein Expression and Purification **55**(1): 1-8.

Silve, S., *et al.* (1996). "Emopamil-binding protein, a mammalian protein that binds a series of structurally diverse neuroprotective agents, exhibits Δ8-Δ7 sterol isomerase activity in yeast." Journal of Biological Chemistry **271**(37): 22434-22440.

Simons, K. and E. Ikonen (2000). "How cells handle cholesterol." Science **290**(5497): 1721-1726.

Singer, S. J. and G. L. Nicolson (1972). "The fluid mosaic model of the structure of cell membranes." Science **175**(23): 720-731.

Singh, S., *et al.* (2008). "Large-scale functional expression of WT and truncated human adenosine A2A receptor in *Pichia pastoris* bioreactor cultures." Microbial Cell Factories **7**: 28.

Sobolevsky, A. I., *et al.* (2009). "X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor." Nature **462**(7274): 745-756.

Song, K. H., *et al.* (2007). "Hepatocyte growth factor signaling pathway inhibits cholesterol 7α-hydroxylase and bile acid synthesis in human hepatocytes." Hepatology **46**(6): 1993-2002.

Sreerama, N. and R. W. Woody (1993). "A self-consistent method for the analysis of protein secondary structure from circular dichroism." Analytical Biochemistry **209**(1): 32-44.

Standfuss, J., *et al.* (2005). "Mechanisms of photoprotection and non-photochemical quenching in pea light-harvesting complex at 2.5 Å resolution." EMBO journal **24**(5): 919-928.

Standfuss, J., *et al.* (2007). "Crystal structure of a thermally stable rhodopsin mutant." Journal of Molecular Biology **372**(5): 1179-1188.

Stock, D., *et al.* (1999). "Molecular architecture of the rotary motor in ATP synthase." Science **286**(5445): 1700-1705.

Stowell, M. H., *et al.* (1997). "Light-induced structural changes in photosynthetic reaction center: implications for mechanism of electron-proton transfer." Science **276**(5313): 812-816.

Takeda, S., *et al.* (2002). "Identification of G protein-coupled receptor genes from the human genome sequence." FEBS Letters **520**(1-3): 97-101.

Tao, X., *et al.* (2009). "Crystal structure of the eukaryotic strong inward-rectifier K^+ channel Kir2.2 at 3.1 Å resolution." Science **326**(5960): 1668-1674.

Tate, C. G. (2001). "Overexpression of mammalian integral membrane proteins for structural studies." FEBS Letters **504**(3): 94-98.

Tate, C. G., *et al.* (2003). "Comparison of seven different heterologous protein expression systems for the production of the serotonin transporter." Biochimica et Biophysica Acta **1610**(1): 141-153.

Teller, D. C., *et al.* (2001). "Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G-protein-coupled receptors (GPCRs)." Biochemistry **40**(26): 7761-7772.

Terpe, K. (2003). "Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems." Applied Microbiology and Biotechnology **60**(5): 523-533.

Tobin, A. B., *et al.* (2008). "Location, location, location...site-specific GPCR phosphorylation offers a mechanism for cell-type-specific signalling." Trends in Pharmacological Sciences **29**(8): 413-420.

Tornroth-Horsefield, S., *et al.* (2006). "Structural mechanism of plant aquaporin gating." Nature **439**(7077): 688-694.

Tsukihara, T., *et al.* (1996). "The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 Å." Science **272**(5265): 1136-1144.

Tucker, J. and R. Grisshammer (1996). "Purification of a rat neurotensin receptor expressed in *Escherichia coli*." Biochemical Journal **317**: 891-899.

Ujwal, R., *et al.* (2008). "The crystal structure of mouse VDAC1 at 2.3 Å resolution reveals mechanistic insights into metabolite gating." Proceedings of the National Academy of Sciences of the United States of America **105**(46): 17742-17747.

Uustare, A., *et al.* (2005). "Kinetic and functional properties of ^{3}H ZM241385, a high affinity antagonist for adenosine 2a receptors". Life Sciences **76**(13): 1513-1526.

van Horn, W.D. *et al.* (2009). "Solution nuclear magnetic resonance structure of membrane integral diacylglycerol kinase." Science **324**(5935):1726-1729.

Vanwetswinkel, S., *et al.* (2005). "TINS, target immobilized NMR screening: an efficient and sensitive method for ligand discovery." Chemistry and Biology **12**(2): 207-216.

Vauquelin, G. and I. Van Liefde (2005). "G protein-coupled receptors: a count of 1001 conformations." Fundamental and Clinical Pharmacology **19**(1): 45-56.

Vogt, J. and G. E. Schulz (1999). "The structure of the outer membrane protein OmpX from *Escherichia coli* reveals possible mechanisms of virulence." Structure **7**(10): 1301-1309.

Wadsten, P., *et al.* (2006). "Lipidic sponge phase crystallization of membrane proteins." Journal of Molecular Biology **364**(1): 44-53.

Waldo, G. S., *et al.* (1999). "Rapid protein-folding assay using green fluorescent protein." Nature Biotechnology **17**(7): 691-695.

Walker, P. A., *et al.* (1994). "Efficient and rapid affinity purification of proteins using recombinant fusion proteases." Biotechnology **12**(6): 601-605.

Ward, A., *et al.* (2007). "Flexibility in the ABC transporter MsbA: Alternating access with a twist." Proceedings of the National Academy of Sciences of the United States of America **104**(48): 19005-19010.

Warne, T., *et al.* (2003). "Expression and purification of truncated, non-glycosylated turkey β-adrenergic receptors for crystallization." Biochimica et Biophysica Acta **1610**(1): 133-140.

Warne, T., *et al.* (2008). "Structure of a β1-adrenergic G-protein-coupled receptor." Nature **454**(7203): 486-491.

Warne, T., *et al.* (2009). "Development and crystallization of a minimal thermostabilised G protein-coupled receptor." Protein Expression and Purification **65**(2): 204-213.

Weiss, H.M., *et al.* (1995). "Expression of functional mouse 5HT(5A) serotonin receptor in the methylotrophic yeast *Pichia pastoris* : pharmacological characterization and localization." FEBS Letters **377**(3):451-456.

Weiss, H. M., *et al.* (1998). "Comparative biochemical and pharmacological characterization of the mouse 5HT5A 5-hydroxytryptamine receptor and the human beta2-adrenergic receptor produced in the methylotrophic yeast *Pichia pastoris*." Biochemical Journal **330**: 1137-1147.

Weiss, H. M. and R. Grisshammer (2002). "Purification and characterization of the human adenosine A(2a) receptor functionally expressed in *Escherichia coli*." European Journal of Biochemistry **269**(1): 82-92.

Weiss, M. S. and G. E. Schulz (1992). "Structure of porin refined at 1.8 Å resolution." Journal of Molecular Biology **227**(2): 493-509.

Welch, G. R. (1995). "The organization of metabolic pathways *in vivo*." Principles of Medical Biology. **4:** 77-92.

Westenhoff, S., *et al.* (2010). "Time-resolved structural studies of protein reaction dynamics: a smorgasbord of X-ray approaches." Acta Crystallographica A **66**: 207-219.

White, J. H., *et al.* (1998). "Heterodimerization is required for the formation of a functional GABA(B) receptor." Nature **396**(6712): 679-682.

White, S. H. (2009). "Biophysical dissection of membrane proteins." Nature **459**(7245): 344-346.

Whittock, N. V., *et al.* (2003). "Novel mutations in X-linked dominant chondrodysplasia punctata (CDPX2)." Journal of Investigative Dermatology **121**(4): 939-942.

Wildt, S. and T. U. Gerngross (2005). "The humanization of N-glycosylation pathways in yeast." Nature Reviews Microbiology **3**(2): 119-128.

Wohri, A. B., *et al.* (2008). "A lipidic-sponge phase screen for membrane protein crystallization." Structure **16**(7): 1003-1009.

Wohri, A. B., *et al.* (2009). "Lipidic sponge phase crystal structure of a photosynthetic reaction center reveals lipids on the protein surface." Biochemistry **48**(41): 9831-9838.

Wu, B., *et al.* (2010). "Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists." Science **330**(6007): 1066-1071.

Wu, S. and G. J. Letchworth (2004). "High efficiency transformation by electroporation of *Pichia pastoris* pretreated with lithium acetate and dithiothreitol." Biotechniques **36**(1): 152-154.

Yamashita, A., *et al.* (2005). "Crystal structure of a bacterial homologue of Na^+/Cl^--dependent neurotransmitter transporters." Nature **437**(7056): 215-223.

Yang, Z. R., *et al.* (2005). "RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins." Bioinformatics **21**(16): 3369-3376.

Yankovskaya, V., *et al*. (2003). "Architecture of succinate dehydrogenase and reactive oxygen species generation." Science **299**(5607): 700-704.

Yao, X., *et al*. (2006). "Coupling ligand structure to specific conformational switches in the β2-adrenoceptor." Nature Chemical Biology **2**(8): 417-422.

Yeates, T. O., *et al*. (1987). "Structure of the reaction center from *Rhodobacter sphaeroides* R-26: membrane-protein interactions." Proceedings of the National Academy of Sciences of the United States of America **84**(18): 6438-6442.

Yernool, D., *et al*. (2004). "Structure of a glutamate transporter homologue from *Pyrococcus horikoshii*." Nature **431**(7010): 811-818.

Yin, Y., *et al*. (2006). "Structure of the multidrug transporter EmrD from *Escherichia coli*." Science **312**(5774): 741-744.

Yu, L. *et al*. (2005). "Nuclear magnetic resonance structural studies of a potassium channel-charybdotoxin complex." Biochemistry **44**(48):15834-15841.

Yuan, P., *et al*. (2010). "Structure of the human BK channel Ca^{2+}-activation apparatus at 3.0 Å resolution." Science **329**(5988): 182-186.

Zhang, Q.H. *et al*. (2007). "Designing facial amphiphiles for the stabilisation of integral membrane proteins." Angewandte Chemie - International Edition **46**:7023-7025.

Zewail, A. H. (2010). "The new age of structural dynamics." Acta Crystallographica A **66**: 135-136.

Zollner, A., *et al*. (2008). "Purification and functional characterization of human 11beta hydroxylase expressed in *Escherichia coli*." FEBS Journal **275**(4): 799-810.

Zouni, A., *et al*. (2001). "Crystal structure of photosystem II from *Synechococcus elongatus* at 3.8 Å resolution." Nature **409**(6821): 739-743.