Investigating the molecular signatures of β₁-adrenergic receptor activation

Andras Sandor Solt

Department of Biochemistry

Wolfson College

University of Cambridge

Submitted: 2018 September

This dissertation is submitted for the degree of Doctor of Philosophy

[BLANK PAGE]

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. (Please see additional note under Declaration in Chapter 3.)

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

It does not exceed the prescribed word limit for the relevant Degree Committee.

Summary

Name: Mr. Andras Sandor Solt

Thesis title: Investigating the molecular signatures of β_1-adrenergic receptor activation

In this thesis I have investigated the molecular signatures of receptor activation, using the β_1-adrenergic receptor (β_1AR) as a prototypical class A G-protein coupled receptor (GPCR). I have used a minimally thermostabilised turkey β_1AR and expressed it functionally in insect cells using a baculovirus system. The work described here established the labelling, expression, purification and sample preparation of the receptor in LMNG detergent micelles for use in nuclear magnetic resonance (NMR) spectroscopy. GPCRs are highly dynamic molecular systems, and the use of solution NMR is highly suited to the study of fine structural changes that take place within the receptor as a consequence of receptor activation. To this end the receptor was selectively labelled with ^{13}C at the methyl position of methionine residues. The labelling was carried out during insect cell expression, by supplementing methionine deficient media with the labelled amino acid. In this way the methionines throughout the receptor served as reporters. The radio frequency signals emitted by the nuclei of these labelled residues were monitored. NMR experiments were recorded on the receptor in the presence of ligands of various efficacies together with and without G-protein mimetic nanobodies, and changes in the signal were recorded. This allowed for a pattern of molecular signatures to be established, reporting on the effect ligands and G-protein mimetics have on the receptor. This identified two conformational equilibria, between an inactive and a ligand bound-pre-activated state and between a more and a less active ternary state when bound to a G-protein mimetic. Furthermore, it was also observed that ligand binding to the G-protein mimetic saturated basal active state elicits further changes on the receptor cytoplasmic side, demonstrating that ligand efficacy modulates the nature of receptor interaction with G-proteins, which may underpin partial agonism. It was also observed that ligand binding affects the dynamics and rigidity of the receptor, with a full agonist bound receptor exhibiting extensive μs to ms timescale dynamics, compared to a more rigid nanobody bound state. The increased dynamics suggest that full agonist binding primes the receptor for interaction with various downstream signalling partners. Once this coupling takes place, ligand efficacy determines the quality of interaction in this rigidified system.

In addition to activation by ligands, certain proteins, such as antibodies can cause receptor agonism in the absence of a small molecule agonist. An example of this takes place in chronic Chagas' heart disease, where anti-*Trypanosoma cruzi* antibodies inappropriately cross-react to β_1AR, leading to chronic cardiac overstimulation and heart failure. In this thesis, the production of a published monoclonal antibody fragment was explored, in order to generate a tool for the study of this activation mechanism.

Dedication and Acknowledgements

I would like to dedicate this thesis to all those in my life, whether they be teachers or friends, who encouraged me to "go and find out" and to have fun while doing so.

Table of contents

List of Figures

List of Tables

List of abbreviations

µOR	µ-Opioid Receptor
3CP	Prescission Protease
^{3}H-DHA	^{3}H-Dihidroaplrenolol
3Q	Triple Quantum
7mC	7-Methylcyanopindolol
7TM	Seven Trans-Membrane
$A_{2A}R$	A_{2a} Adenosine Receptor
ADR	Adrenaline
Ag	Agonist
AIM	Autioinduction Media
BLT2	Leukotriene B4 Receptor
BRIL	Cytochrome B Fusion
BTFA	3-Bromo-1,1,1-Trifluoroacetone
BTFMA	2-Bromo-4-Trifluoromethylacetanilide
cAMP	Cyclic Adenosine Monophosphate
C_H, C_L	Constant Heavy, Constant Light Immunoglobulin Domain
CHS	Cholesterol Hemisuccinate
CMC	Critical Micellar Concentration
CVD	Carvedilol
CYA	Cyanopindolol
Cα	Backbone α-Carbon
DDM	Dodecyl Maltoside
ECL	Extracellular Loop
ECL (e.g. "ECL2")	Extracellular Loop (2)
EGFR	Epidermal Growth Factor Receptor
EXSY	Exchange Spectroscopy
Fab	Fragment Antigen Binding
FCS	Foetal Calf Serum
GAP	GTPase Activating Protein
GDP	Guanosine Nucleoside Diphosphate
GEF	Guanosine Exchange Factor
GRK	GPCR Kinase
GTP	Guanosine Nucleoside Triphosphate
HDX-MS	Hydrogen Deuterium Exchange Mass Spectrometry
HMQC	Heteronuclear Multiple-Quantum Correlation
HSQC	Heteronuclear Single-Quantum Correlation
HT	HisTrap
htG-protein	Heterotrimeric G-Protein
htG-protein	Heterotrimeric G-Protein
HTS	High Throughput Screening
hβ1AR	Human β1-Adrenergic Receptor
IBP	Intracellular Binding Partner
ICL	Intracellular Loop
ICL	Intracellular Loop
ICL (e.g. "ICL2")	Intracellular Loop (2)
ISO	Isoprenaline
LDS	Lithium Dodecyl Sulphate
LMNG	Lauryl Maltose Neopentyl Glycol

LPC	Lipid Cubic Phase
m23	β_1AR-β44-m23 Construct (See Appendix A)
mAb	Monoclonal Antibody
MBP	Maltose Binding Protein
MoA	Mode of Action
MS	Mass Spectrometry
NMR	Nuclear Magnetic Resonance
NTS1	Neurotensin Receptor 1
PBS	Phosphate Buffered Saline
PGS	*Pyrococcus Abysii* Glycogen Synthase
P_i	Inorganic Phosphate
PKA	Protein Kinase A
RGS	Regulator Of G-Protein Signalling
ScFv	Single Chain Fragment Crystallisable
Scint	Scintillation Counting
SDS	Sodium Dodecyl Sulphate
SLB	Salbutamol
SMD	Small Molecule Drug
STD	Saturation Transfer Difference
T2R	Taste 2 Receptor
TBS	Tris Buffered Saline
TET	2,2,2-Trifluoroethanethiol
TM	Trans-Membrane
T_m	Melting Temperature
TROSY	Transverse Relaxation Optimized Spectroscopy
tβ1AR	Turkey β_1-Adrenergic Receptor
UV 280 nm	280 nm Absorbance
V_H, V_L	Variable Heavy, Variable Light Immunoglobulin Domain
XAM	Xamoterol
β_1AR	β_1-Adrenergic Receptor
β_2AR	β_2-Adrenergic Receptor

Chapter 1 – Introduction

1.1 – G-protein Coupled Receptors

G-protein coupled receptors are 7-trans-membrane alpha helical proteins. With over 800 human genes encoding GPCRs, they represent one of the largest families of integral membrane proteins in our genome (*1*). GPCRs are implicated in a wide array of signalling pathways, affecting a diverse palette of physiological functions such as smooth muscle and myocardial regulation, platelet activation, endocrine and metabolic regulation, immune system activation, synaptic transmission, sight and olfactory regulation, development and growth and cellular proliferation (*2*). It is therefore not surprising that GPCR dysregulation is associated with conditions spanning multiple disease areas, making GPCRs the single biggest family of drug targets. It is estimated that about 30% of current drugs modulate GPCR function (*3*).

1.1.1 – GPCR Classification

Several classifications of GPCRs exist and the methods used have evolved over time. The earliest classification grouped GPCRs on the basis of sequence homology and this was the basis of the well-known A-F Class system (*4*). With the publication of the human genome, the dataset of GPCRs was expanded and a system of classification emerged based on phylogenetic analysis (*5*). This groups GPCRs into the families of Glutamate, Rhodopsin, Adhesion, Frizzled/Taste2 and Secretin, and is hence known as the GRAFS system (*6*, *7*). The two classifications overlap: the Rhodopsin family was known as Class A, the Adhesion and Secretin families both as Class B, Glutamate as Class C. Class D and E were separate groups of fungal mating pheromone receptors and cyclic AMP receptors respectively, and Frizzled receptors were initially designated Class O (other 7TM) and later Class F. Neither Class D nor E receptors were found in vertebrates.

Although considerable variation exists in terms of primary and secondary structure between different classes of GPCRs, they all share the common hallmarks of seven stretches of 25-35 highly hydrophobic amino acids – which form the seven transmembrane helices – as well as the ability to interact with heterotrimeric G-proteins (htG-proteins) (*1*).

The Rhodopsin family (class A) is the largest of the five and is the best structurally characterised. The family is divided into four subfamilies, these are α, β, γ and δ. Class A receptors are variable in primary sequence, but certain motifs such as the DRY motif and the NPxxY motif are highly conserved. The N terminal region is also relatively short and sequence-diverse. Interesting cases of Rhodopsin family GPCRs have been documented, where the N terminus plays a role in receptor activation. This can take

place in the form of N-terminal cleavage, as in the case of the human thrombin receptor (PAR1/ F2R), where cleavage at an N-terminal thrombin binding site reveals a tethered ligand, which in turn activates the receptor. PAR2 receptor is pre-activated by trypsin in a similar manner (*8*, *9*). The N-termini of some receptors, such as the LGR7, LGR8 (relaxin-binding receptors), FSHR (follicle-stimulating-hormone receptor), LHR (leuteinizing hormone receptor) and TSHR (thyrotropin receptor) contain leucine rich segments, which take part in the interactions between the receptor and their glycoprotein ligands (*10*, *11*). However these are unusual cases and most Rhodopsin family GPCRs bind their ligands through pockets formed by the extracellular loops and TM helices (*12*).

The Rhodopsin subfamilies group receptors with a huge variety of ligands, but on the basis of phylogeny only. Thus, receptors belonging to different subfamilies share preference for a given ligand type. The α subfamily contains for example histamine receptors, dopamine receptors, serotonin receptors, adrenoreceptors, muscarinic acetylcholine receptors, muscarinic receptors, cannabinoid and prostanoid receptors amongst others. The β subfamily contains *mostly* peptide binding receptors, such as endothelin receptors, gonadotrophin-releasing hormone receptors, oxytocin receptors and neuropeptide receptors. The γ subfamily groups opioid receptors, somatostatin receptors amongst others. An interesting member of the group is chemokine receptor 5 (CCR5), a target of the drug maraviroc, which blocks the interaction between the HIV virus and CCR5, a crucial interaction in virus entry (*13*). The δ-group contains the aforementioned glycoprotein receptors, FSHR, LHR, TSHR, the protease-activated receptors as well as P2RY. This group also includes the largest number of mammalian olfactory receptors in the Rhodopsin family.

The Secretin family derives its name from the first Class B receptor discovered (*14*), and it contains five subfamilies. These are the calcitonin receptors, corticotropin-releasing factor receptors, glucagon receptors, parathyroid hormone receptors, vasoactive intestinal peptide receptors *and* pituitary adenylate cyclase activating receptors (*7*). Typical of this family are a group of three highly conserved cysteine bridges, which have been shown to be important for ligand binding through stabilisation of the N-terminus (*15*, *16*).

Also previously grouped under Class B are the Adhesion family of receptors; the second largest family in the human genome. Like Secretin family members, Adhesion receptors possess a long N-terminal domain. A difference between these two groups is that Adhesion receptors contain a GPCR proteolytic (GPS) domain. Furthermore, these two families display differences in terms of the nature of the bound ligand, Secretin GPCRs bind peptide hormones and Adhesion receptors – for which the ligand is known – have so far been extracellular matrix components (*7*).

Glutamate family receptors (abbreviated GRMs) (Class C) contain four main groups plus Class C orphans receptors: receptors for which the endogenous ligand has not been identified. These are calcium sensing receptors, GABA$_B$ (gamma-aminobutyric acid) receptors, metabotropic glutamate receptors and taste 1 receptors. Members of this family commonly have extended N-terminal domains, which forms the orthosteric ligand binding site (*17*). In the case of mGluR5 this cavity is formed by two lobes of the N-terminal domain, and ligand binding has been likened to a Venus fly trap mechanism (*18*). However many allosteric sites have been identified formed by extracellular loops. The TM regions of GRMs show a high degree of residue conservation, and highly conserved residues include non-polar hydrophobic amino acids, as well as a number of polar, charged and aromatic residues, which may be important for interhelical contacts. Conserved motifs have also been identified, such as the Rhodopsin NPxxY equivalent PKxY motif in TM7 and the WL motif in TM6 (*1*).

The Frizzled/ Taste2 family groups together frizzled receptors, the smoothened receptor and a number of Taste2 receptors. Frizzled receptors bind the wnt family of small glycoproteins as their ligand, whilst the smoothened receptor is understood to function in a ligand independent manner. Frizzled receptors share a large extracellular domain ranging in length from 200 to 320 amino acids and contain nine conserved cysteines (*12*). Taste2 receptors (T2Rs) have been identified to recognise thousands of bitter taste compounds, a feat which may be explained by the wide spectrum of sequence conservation between different subgroups of T2Rs (*19*, *20*).

1.1.2 – GPCR signalling

GPCRs serve to translate extracellular stimuli through the cell membrane. This stimulus can take form in a wide range of shapes, such as proteins, peptides, nucleosides, amines, ions and light. Ligand binding induces structural changes within the receptor that allow the association of intracellular binding partners and thereby the relay of extracellular signals to within the intracellular environment (*21*). A typical feature of GPCRs is their basal activity (although exceptions exist, such as Rhodopsin), where a low-level downstream signalling activity takes place even in the absence of extracellular stimulus (*22*). In this regard, ligands can be categorised on the basis of how they modulate basal receptor signalling. Ligands which activate receptor signalling through downstream second messengers are called agonists and depending on the magnitude of the elicited signalling output, they may be full, partial, or neutral agonists **(Figure 1.1)**. The latter in the case when ligand binding takes place but no increase above basal activity is observed. Ligands may also be able to decrease basal activity and these are known as inverse agonists. These too may be partial, or full in terms of the output of attenuation. Antagonists are ligands which modulate receptor activity by modifying the way or extent to which endogenous ligands are able to bind to the receptor. Hence antagonists may be

3

inverse agonists, partial agonists or neutral in effect (*23*). Ligands may bind to an orthosteric or "main" ligand binding site, or they may bind away from there to an allosteric or "alternative" site, where ligand binding may modulate binding to the orthosteric site (*24*, *25*).

Pharmacological response to ligands
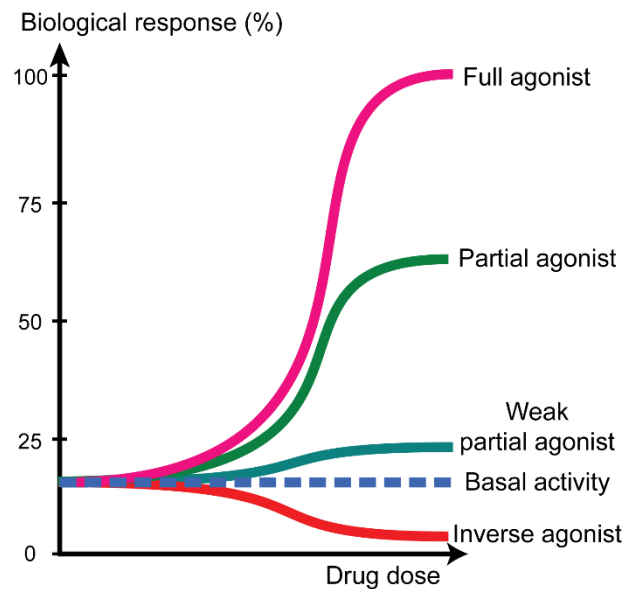
Figure 1.1 Schematic representation of ligand efficacy

A comparison of ligands according to their maximum biological response attained allows their categorisation into full, partial or weak agonists. Ligands that abrogate any basal response observed in the apo state are called inverse agonists. Antagonists are ligands which modulate the effect of an agonised receptor in relation to another ligand.

The traditional view of GPCR activation was understood in terms of a switch-like transition from the inactive R state to the active/signalling R* state. This has evolved to the current understanding of GPCRs as an ensemble of highly dynamic states, stabilised to different extents by various binding events. This theme will be explored further in this thesis.

"Canonical" GPCR signalling takes place through interactions with heterotrimeric G-proteins (htG-proteins). htG-proteins consist of an α and a non-dissociable β and γ subunits. Nucleotide binding and GTPase activity are intrinsic features of the α subunit. Signal transduction via htG-proteins takes place through a cycle of activation. In the basal state the α and βγ subunits form a GDP-associated heterotrimer. GPCR binding causes the exchange of GDP to GTP and the dissociation of the α subunit from the βγ complex. This allows the subunits to bind and regulate downstream signalling partners such as enzymes or ion channels. Subunit α GTPase activity will give rise to an eventual hydrolysis of

bound GTP to GDP, thereby terminating the signalling process. The GTPase activity has been shown to be increased by certain effectors known as GTPase activating proteins (GAPs) or regulators of G-protein mediated signalling (RGSs). Following nucleotide hydrolysis, the α and βγ subunits reassociate, in preparation for another cycle (*21*). **Figure 1.2** illustrates the cycle of htG-protein activation-inactivation.
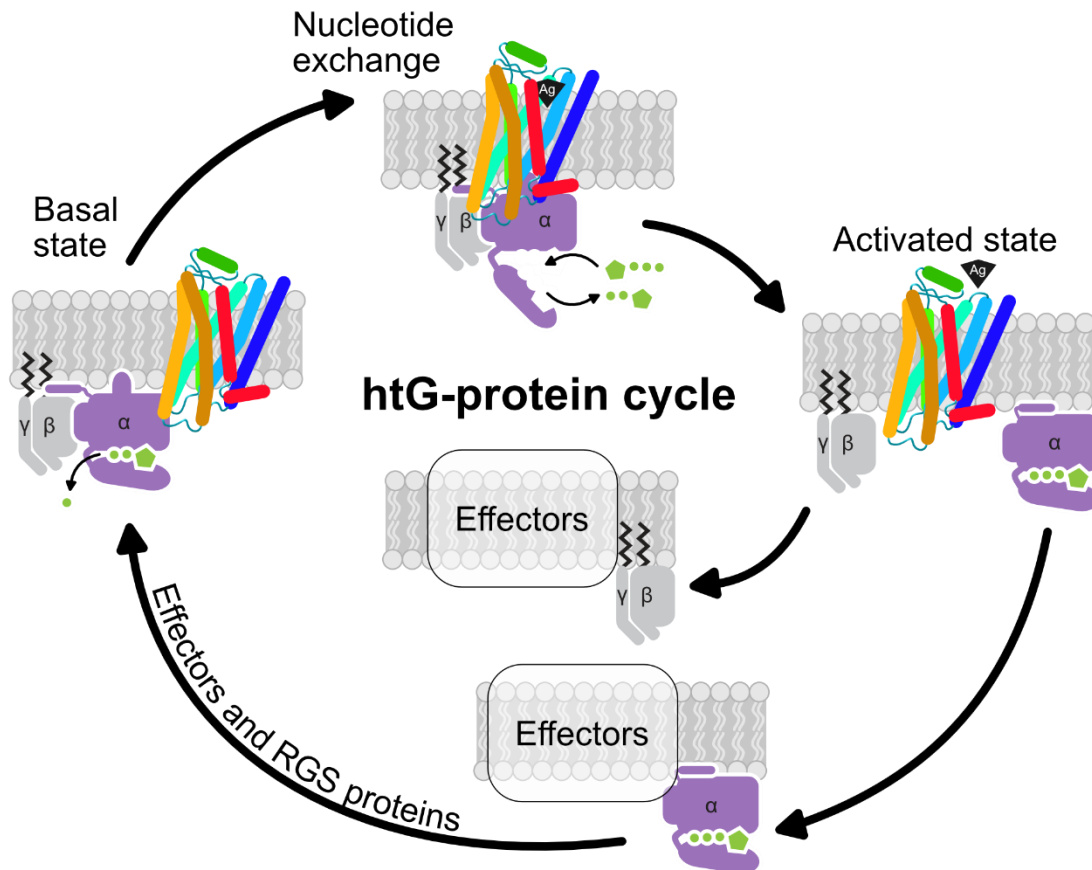
Figure 1.2 Cycle of heterotrimeric G-protein activation

Cycle of heterotrimeric G-protein activation. Binding of GDP associated G-protein to GPCR causes nucleotide exchange and the dissociation of α subunit from βγ complex, which interact with downstream signalling partners (effectors). Subunit α GTPase activity causes GTP hydrolysis to GDP, terminating the signalling process. Abbreviations: RGS, regulator of G-protein mediated signalling. Figure from (*21*).

Heterotrimeric G-protein α subunits can be classed into four broad categories, $G\alpha_s$, $G\alpha_i$/$G\alpha_o$, $G\alpha_q$/$G\alpha_{11}$, $G\alpha_{12}$/$G\alpha_{13}$. The mammalian htG-proteins are further assembled from a panel of five β subunits and 12 γ subunits. The type of G-proteins activated by a given receptor has a defined consequence regarding the downstream signalling output due to the specific identity of effectors which that given G-protein will modulate. Which G-protein a certain GPCR will activate is not *absolutely* defined, as some receptors interact with multiple Gα subunits, but neither is it completely random as a strong

preference exists for a certain class of Gα in many GPCRs. For example the β1-adrenergic receptor couples exclusively to $Gα_s$, whilst the β2-adrenergic receptor (β2AR) has been shown to signal through $Gα_s$ and $Gα_i$ as well (*26*).

G-protein mediated signalling can be oversimplified to the Gα regulation of adenylyl cyclase, in turn modulating the levels of cyclic AMP (cAMP), activating PKA, which in turn affects a wide range of physiological functions. However it is important to consider that the Gα regulation of adenylyl cyclase is ubiquitous to GPCRs in general and depending on which type of Gα they interact with, the effect on the enzyme's activity will be different. For example $Gα_s$ increases the production of cAMP, whilst $Gα_q$ causes the downregulation of adenylyl cyclase activity. Therefore to understand the complete significance of G-protein mediated signalling, the process has to be viewed in relation to the signalling activity of other receptors. This concept is illustrated in **Figure 1.3** in the context of heart function regulation. Here, sympathetic cardiac activation is mediated by the primarily $Gα_s$ coupled βARs, whilst opposing parasympathetic regulation by the muscarinic acetylcholine receptor M_2 via $Gα_i$ and $Gα_{i/o}$ causes adenylyl cyclase downregulation and the activation of the inward rectifier potassium channel and the regulation of the L-type calcium channel (21).
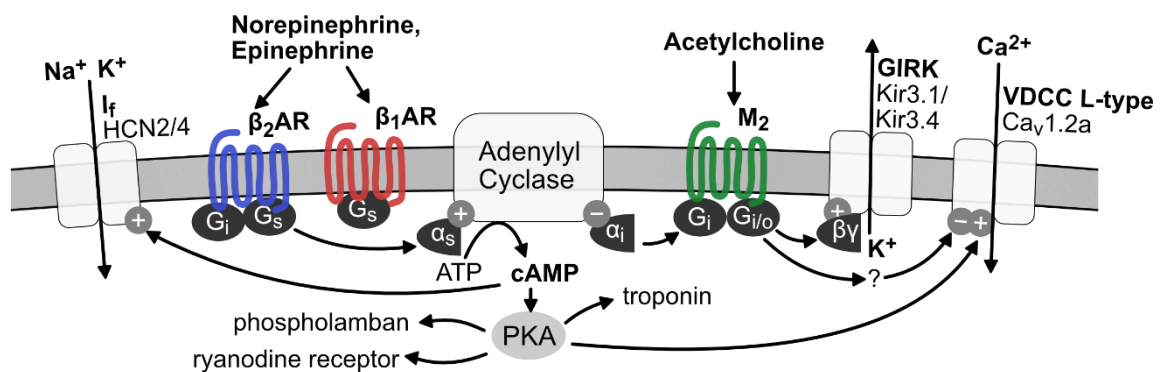
Figure 1.3 G-protein mediated signalling in heart regulation

G-protein mediated signalling concert in heart regulation, illustrating that G-protein mediated GPCR signalling is complex and integrative. Abbreviations: $β_1$: $β_1AR$, $β2$: $β_2AR$, M_2: muscarinic acetylcholine receptor M_2, I_f: pacemaker channel, GIRK: G-protein mediated inward rectifier potassium channel, VDCC: voltage-dependent calcium channel, PKA: protein kinase A.

GPCRs interact not only with htG-proteins but also with β-arrestin and GPCR kinases (GRKs).

Arrestins were first discovered in relation to rhodopsin and hence the first arrestin was later termed visual arrestin (arrestin 1). A homologous protein and a binding partner of the β2AR has soon after been identified and was called β-arrestin. Two types of β-arrestin have been found, β-arrestin 1 and 2 (or arrestins 2 and 3) (*27*).

Our understanding of the role of arrestins have evolved over time, but can be broadly ascribed three functions; as terminators of G-protein mediated signalling, as mediators of receptor internalisation and trafficking and as alternative non-G-protein mediated signalling partners in their own right.

Arrestins were first described to be able to terminate the G-protein mediated signalling output of rhodopsin, β_2AR and other GPCRs. The association of arrestin with the receptor prevents the coupling of G-protein and hence signalling through htG-proteins. Arrestin binding requires prior receptor phosphorylation by GPCR kinases (GRKs), of which seven have been identified to date (*28*). GRKs have been shown to phosphorylate serines and threonines on both the intracellular C-terminus as well as on other intracellular loops, mainly ICL2 (*29*, *30*). Significantly, receptor phosphorylation takes place mostly in the active, agonist bound receptor state, hence arrestin recruitment serves to desensitise the active receptor.

In 1996, the connection between receptor internalisation – a major alternative form of receptor desensitisation to arrestin desensitisation – and arrestin binding have been made, establishing the role of arrestin in receptor internalisation (*31*, *32*). Arrestins interact with clathrins to promote vesicular internalisation and from here GPCRs can be dephosphorylated and recycled in what is called re-sensitisation, or can be targeted for degradation. The most recent view on arrestin mediated receptor trafficking is that arrestins form the foundations of a scaffold of proteins, which determine the fate of receptors in the process of trafficking (*33*, *34*).

The third and most recently identified function of arrestins is the ability to activate non-classical signalling pathways as a consequence of receptor binding. This presents a whole new plethora of pathways through which GPCRs are able to modulate cellular function. The activation of three main signalling pathways have been linked to arrestins and these are: non-receptor tyrosine kinases such as cSrc, Hck, Fgr and Yes, secondly MAP kinase cascade components, including both the ERK and JNK pathways and thirdly signalling through PI3K and PKB/AKT (*27*). Receptor internalisation has shown a close association with the regulation of these signalling pathways both in terms of signalling activation and duration.

GRKs have also been shown to be modulators of alternative GPCR signalling in their own right in addition to simply carrying out a function of receptor modification enabling arrestin binding. GRKs have been shown to phosphorylate multiple non-GPCR targets, thereby fine tuning the activity of downstream non-G-protein mediated signalling pathway components (*28*). Furthermore the recruitment of ubiquitin E3-ligase for the ubiquitination and degradation of β_2AR has also shown GRK involvement (*35*). This gives GRKs a dual role in GPCR desensitisation, both in a short term manner via arrestin binding or in a long term fashion via receptor degradation.

7

The discovery of alternative GRK/arrestin mediated non-G-portein signalling pathways has introduced the concept of biased agonism. Biased agonists are GPCR ligands which are able to trigger the activation of these alternative signalling pathways. Biased ligands may be able to activate both G-protein mediated and alternative pathways to a different extent, hence the nature of their bias might be called perfect or partial (*23*).

1.2 – The β-adrenergic receptors

1.2.1 – $β_1$AR Signalling

The β1-adrenergic receptor is a Class A, Rhodopsin family GPCR which is activated by the endogenous catecholamine ligands; adrenaline and noradrenaline. Three subtypes of the β-adrenergic subfamily have been identified in the human genome; the $β_1$, $β_2$ and $β_3$ receptors. The existence of a fourth subtype, the $β_4$AR has been proposed in turkey erythrocytes (*36*), however the pharmacological evidence supporting this proposal has been questioned and a $β_4$AR is not considered to exist (*37*, *38*).

Agonist binding triggers a conformational change in the receptor, which causes the dissociation of the heterotrimeric G-protein binding partner (htG-protein) to $Gα_s$ and and Gβγ subunits. $Gα_s$ activates the enzyme adenylyl cyclase, resulting in elevated levels of the cytosolic second messenger cyclic AMP (cAMP). cAMP activates the regulatory kinase PKA, which in turn causes the phosphorylation of calcium-channel-regulatory proteins, leading to an efflux of Ca^{2+} ions from the sarcoplasmic reticulum into the cytosol. Increased Ca^{2+} levels in turn lead to myocardial contraction (*39*). Chronic $Gα_s$ mediated βAR signalling has been associated with heart failure and drugs designed to treat myocardial failure through the downregulation of such signalling are aptly called beta-blockers (*40*).

$β_1$AR has been shown to be able to activate β-arrestin mediated alternative signalling pathways described above. Recent studies of $β_1$AR alternative signalling have highlighted its physiological significance in the context of cardioprotection. Here, the transactivation of the EGFR signalling cascade by $β_1$AR through β-arrestin 1 and 2 has been shown to correlate with increased tolerance to catecholamine induced cardiomyocyte toxicity (*41*). The proposed signalling cascade includes receptor phosphorylation by GRK5/6 leading to arrestin recruitment of Src to the receptor. Src causes the MMP activated release of HG-EGF, a ligand of EGFR. The involvement of such non-G-protein mediated, multi-member signalling cascades has the potential to significantly increase the number of pharmaceutical targets for the treatment of heart failure.

1.2.2 – Crystallographic studies of the β-adrenergic receptors

Until 2007, the only GPCR structure available was that of bovine rhodopsin (*42*). The crystallisation of the β_2-adrenergic receptor in 2007 (*43*) and of other GPCRs since have been made possible by a number of technological advances as well as innovations in protein engineering (*44*). The major challenges were threefold, first to make available necessary quantities of receptor for structural studies, secondly to achieve crystal formation of GPCRs and thirdly to be able to collect high resolution diffraction data on crystals obtained. The high natural abundance of bovine rhodopsin in the rod outer segments of the retina, the ease of its isolation through non-chromatographic methods, and its high stability in membrane mimetic detergents was a luxury not available for other GPCRs. Hence considerable work has been done to find suitable heterologous expression systems, to maximise yields and to improve innate stability through the introduction of thermostabilising mutations, stabilising fusion partners (such as T4L, PGS, Rubredoxin, BRIL) or the engineering of stabilising antibody fragments (*45*).

To aid crystal formations, novel detergents, alongside alternative crystallisation methods to the conventional vapour diffusion, such as lipid cubic phase (LPC), were developed (*46–49*). To increase functional homogeneity ligand affinity purification resins were created, which were able to purify GPCRs to a very high degree of functional purity (*50*). GPCR crystals have been much smaller in dimension than conventional non-lipid crystals and their manipulation for cryoprotection and high quality data collection is difficult. Methods to deal with microcrystals which are usually 10-20 μm in their longest dimension have been advanced. These include rastering, where the presence of small crystals can be automatically detected within a larger volume of lipid matrix which is not possible optically (*51*). The use of microfocus beamlines and more recently X-ray-free electron lasers have yielded high resolution crystal structures of a number of GPCRs (*52, 53*). Crystal structures for 52 unique GPCRs have been defined to date, in various states of activation and with different ligands (Source: GPCRdb, retrieved: 25/Sept/2018). This breaks down to 45 Class A (Rhodopsin) receptors, 4 Secretin class receptors, 2 Glutamate class and 1 Frizzled Class receptor. Several receptors have had their structure determined in complex with a G-protein.

The majority of GPCR structures represent inactive states. Active state conformations have been obtained for rhodopsin ($G\alpha_t$ peptide stabilised), β_2AR (nanobody and $G\alpha_s$ bound) and the M_2 muscarinic acetylcholine receptor (nanobody stabilised) (*54–56*). G-protein bound states have been determined for bovine rhodopsin (G_t bound), $A_{2A}R$ (bound to G_s and $MiniG_s$), GLP-1 (G_s bound), 5-HT$_{1B}$ (G_o bound), A_1R (G_{i2} bound), μ-opioid receptor (G_i bound). These have been determined by X-ray crystallography or cryo-EM (*57–68*). Additionally structures are available for the β_1AR, α_2-adenosine

receptor ($A_{2A}R$), neurotensin type 1 receptor (NTSR1) and P2Y purinergic receptor 12 ($P2Y_{12}R$), which show an active-like, partially activated conformation, but without the presence of any cytoplasmic binding partners (*69–73*). Furthermore the structure of human rhodopsin bound to mouse visual arrestin has also been determined (*74*).

Of particular significance and relevance to this project are the eleven crystal structures of β_2AR. Significantly the β_2AR in complex with $G\alpha_s$ (PDB ID: 3SN6), was the first full Gα/GPCR complex and the first model of an active state adrenergic receptor. In this regard this β_2AR structure and others can provide valuable insight into the activation process of the β_1AR. The following table lists β_2AR structures to date.

Table 1.1 List of β_2AR structures currently available

PDB ID	Construct	Binding partner/ ligand	Reference
2R4R	48 AA C-terminal truncation: β2AR365	Fab5	(*75*)
2RH1	β_2AR-T4L in ICL3	Carazolol (inverse agonist)	(*76*)
3D4S	β_2AR(E122W)-T4L (Mutation for stability)	Timolol (beta blocker, partial inverse agonist), cholesterol	(*77*)
3KJ6	^{13}C-methyl-β2AR	Fab5	(*78*)
3NY8		ICI118,551 (selective antagonist)	
3NY9	β_2AR-ICL3-T4L	Compund 2 (inverse agonist)	(*79*)
3NYA		Alprenolol (inverse agonist)	
3P0G	β_2AR-T4L	Nb80 G-protein mimic, BI-167107 (full agonist)	(*80*)
3PDS	β_2AR(H93C)-T4L (Mutation for covalent link)	Covalently bound FAUC50 agonist	(*81*)
3SN6	β_2AR-Nterm-T4L	$G\alpha_s$, Nb35, BI-167107	(*82*)
4GBR	T4L-β_2AR-Δ-ICL3	Carazolol (inverse agonist)	(*83*)
4LDL	β2AR365+Nb6B9	Hydroxybenzylisoproterenol (high-affinity cathecolamine agonist)	(*84*)
4LDE	β2AR365+Nb6B9	BI-167107 (ultra-high-affinity agonist)	(*84*)
4LDO	β2AR365+Nb6B9	Adrenaline (low affinity endogenous cathecolamine agonist)	(*84*)
4QKX	β_2AR(H93C)+Nb6B9	Covalent (nor)adrenaline derived agonist	(*85*)
5D5A	β_2AR	Carazolol (crystallography method paper for LCP method)	(*86*)
5JQH	β_2AR+Nb60	Allosteric nanobody	(*87*)
5D6L	β_2AR	Sample preparation method for LCP crystallography	(*88*)
5X7D	β_2AR	Bound to Cmpd-15 intracellular allosteric antagonist	(*89*)

The first crystal structure bound to the inverse agonist, carazolol (2RH1) was expected to represent the traditional inactive R state. The inactive state in rhodopsin features a so called ionic lock between $Arg^{3.50}$ and $Glu/Asp^{6.30}$ in TM3 and TM6, which is understood to lock the receptors in the inactive conformation (superscript: Ballesteros-Weinstein GPCR residue numbering). This interaction is also expected to be present in β_2AR, as mutations of these residues render the receptor constitutively active (*90*). However the supposedly inactive crystal structure revealed the distance between $R135^{3.50}$ and $E268^{6.30}$ to be too large for ionic interactions. The lack of an ionic lock in the inverse agonist bound state may explain the basal activity of β2AR, but more significantly it can be considered as evidence that the receptor explores multiple conformational states, and these are only transiently stabilised by for example inverse agonist ligand binding.

The nanobody (Nb80) (3P0G) and $G\alpha_s$ (3SN6) stabilised active conformations of β_2AR have revealed two structural hallmarks of the active state. Firstly the cytoplasmic end of TM6 displays a 14 Å outward movement measured at $Glu268^{6.30}$ (Cα). Secondly a similar outward movement and extension by 7 residues was observed for the cytoplasmic end of TM5. The second intracellular loop (ICL2) was also seen to adopt an α-helical conformation. These changes are illustrated on **Figure 1.4**.

1.2.3 – Engineering the β₁AR for crystallisation

The turkey (*Meleagris gallopavo*) β_1AR ($t\beta_1AR$) has proved to be an ideal candidate for providing structural insight into the β_1-adrenergic receptor subfamily. It is ($t\beta_1AR$) 79% identical in sequence (dismissing N- and C-termini) to the human β_1 receptor ($h\beta_1AR$) and its stability is superior over the hβ1AR. The melting point for $t\beta_1AR$ was measured 10 °C higher at 31 °C, than for hβ₁AR observed in 0.01% DDM. In comparison, the wild type human β₂AR (hβ₂AR) was almost as stable as $t\beta_1AR$ with a melting point of 30 °C (*91*). Without further modification, the human β₁AR is hence unsuitable for crystallisation due to its instability.

To further improve stability, conformational and structural heterogeneity and expression levels, a number of modifications were introduced to the tβ1AR sequence (*92*). The N-terminal residues 3-32 were deleted, as well as residues 244-271 and 277-278 in intracellular loop 3 (ICL3), and the C-terminal was shortened between amino acids 368-483. The truncation of these flexible regions was aimed to assist crystal formation. To avoid heterogeneity arising from non-uniform palmitoylation, cysteine 358 was mutated to alanine to remove the post-translational modification site. The C116L substitution was found to increase functional expression levels. The feasibility of thermal stabilisation of the receptor was explored by the iterative mutation of all residues to alanine or leucine (where alanines are present originally). This is known as alanine scanning. Any increases in thermal stability were

assayed through measuring the apparent melting point of the protein (T_m); the temperature at which 50% of receptor molecules retain their ligand (antagonist in this case) binding ability after one hour. Residues showing potential in an initial alanine scan were then optimised by substituting a different class of amino acid to the hydrophobic alanine. On average, amino acid substitutions which brought about an increase in T_m did so by 1-3°C, however the Y227A mutation gave an 8°C shift. Mutations which increased stability were combined to achieve an overall T_m of 52.8°C. This construct, termed the β1AR-m23 mutant contained the thermostabilising mutations: R68S[1.59], M90V[2.53], Y227A[5.58], A282L[6.27], F327A[ICL3] and F338M[7.48], as well as the other sequence modifications mentioned above and provided the first crystal structure of the receptor.

Attempts to further increase receptor thermostability resulted in an ultrathermostable $β_1$AR-JM3 construct. The introduction of a salt bridge, which is present in $β_2$AR between residues Asp192 and Lys305 (equivalent $β_1$AR residues are Asp200 and Asp322) by the mutation D200E+D322K achieved a 5°C increase in T_m alone. Further substitutions identified by leucine scanning increased the receptor melting point to 61.7°C (in DDM micelles), when combined with the salt bridge forming mutation. Hence $β_1$AR-JM3 contained the stabilising mutations D200E, D322K, I129V and Y343L in addition to the $β_1$AR-m23 mutations (*93*).

The thermostabilised constructs were investigated for their agonist binding ability. Competition binding assays have shown $β_1$AR-m23 to bind the agonists norepinephrine and isoprenaline 2470 and 650 times weaker respectively than the non-stabilised, truncated $β_1$AR$_{32-424}$ construct. Similarly, the ultrathermostable JM3 construct bound noradrenaline and isoprenaline 4960 and 990 fold weaker than the non-stabilised receptor. (*93*, *94*). Furthermore, the thermostabilised constructs are also able to bind and signal through intracellular G-proteins, albeit with a higher predicted activation energy (*95*).

1.2.4 – Insights from $β_1$AR Crystal structures

Crystal structures of the thermostabilised turkey $β_1$-adrenergic receptor have been obtained in the presence of various ligands, such as agonists, partial agonists, inverse agonists, antagonists as well as in the absence of any ligand. The following table lists the published structures to date alongside the nature and identity of any ligand present.

Table 1.2 List of β$_1$AR crystal structures currently published

PDB ID	Bound Ligand	Ligand agonism	Reference
2Y02	Carmoterol	Full	(69)
2Y03	Isoprenaline	Full	(69)
2Y00	Dobutamine	Partial	(69)
2Y04	Salbutamol	Partial	(69)
2Y01	Dobutamine	Partial	(69)
2VT4	Cyanopindolol	Partial	(96)
2YCX	Cyanopindolol	Partial	(97)
2YCZ	Iodocyanopindolol	Partial	(97)
2YCY	Cyanopindolol	Partial	(97)
4AMI	Bucindolol	(Biased)‡	(23)
3ZPQ	20*	Antagonist	(98)
3ZPR	19†	Antagonist	(98)
4BVN	Cyanopindolol	Partial	(99)
2YCW	Carazolol	Inverse agonist	(97)
4AMJ	Carvedilol	Weak Partial (Biased)‡	(23)
4GPO	None	-	(100)
4BVN	Cyanopindolol	Partial	(101)
5A8E	7-methylcyanopindolol	Weak partial	(102)
5F8U	Cyanopindolol	Partial	(103)

*4-METHYL-2-(PIPERAZIN-1-YL)QUINOLINE
†4-(PIPERAZIN-1-YL)-1H-INDOLE
‡ Biased agonists show differential efficacies in the activation of G-protein mediated and alternative signalling pathways

The overall topology of the β$_1$AR was found to be very similar to the β$_2$AR, with the RMSD of 296 superimposed C$_\alpha$ atoms across all transmembrane helices being 0.25 Å (comparing carazolol bound β$_2$AR – PDB:2R4R, and cyanopindolol bound β$_1$AR – 2VT4) (96). The second extracellular loop in both β$_1$ and β$_2$ structures have adopted a partially α-helical conformation, which is proposed to be stabilised by the associated Na$^+$ ion visible in the β1AR structure.

Intracellular loop 2 also adopts a helical conformation, which has not previously been observed in the β$_2$ structure, probably due to perturbations owing to the presence of stabilising agents such as T4-Lysozyme or a Fab fragment. An interesting feature of this short helix is its interaction with the DRY motif (D138$^{3.49}$, R139$^{3.50}$, Y140$^{3.51}$) via a hydrogen bond between Y149 (ICL2) and D138. Mutagenesis studies of the M2 receptor have shown equivalent residues to be important for G-protein-receptor interaction, an observation that the presence of this hydrogen bond in β$_1$AR may confirm (104).

The antagonist cyanopindolol bound structure (2VT4) shows no evidence for the putative "ionic lock"; a salt bridge between residues Arg$^{3.50}$ (in the DRY motif) and Glu$^{6.30}$ which is observed in Rhodopsin.

The lock is proposed to be essential in stabilising the receptors in the inactive state by linking transmembrane (TM) helices 3 and 6. Upon agonist activation the lock breaks, permitting the characteristic outward shift of TM6. As the antagonist bound receptor is not considered to be in the active state, the absence of any ionic interactions between these residues questions its role in inactive state stabilisation.

However more recent structures of β_1AR with different crystal contacts revealed two consistent conformations of TM6, a bent form, which brings Glu285[6.30] to a distance of 3.7-3.9 Å of Arg139[3.50], compared to >6.2 Å in the straight conformation (*97*). It is only in the bent conformation that an ionic lock is observed, although the separation between the bonding residues is more than the equivalent distance in Rhodopsin, measured at 2.8-3.2 Å. These differences between a receptor with virtually none (Rhodopsin) and one with a relatively high basal activity (β_1AR) seem to agree with the view that GPCRs are highly plastic molecules which rapidly sample multiple active and inactive conformations.

Structures of β_1AR bound to the agonists isoprenaline and carmoterol, and to the partial agonists salbutamol and dobutamine have been compared to the antagonist cyanopindolol bound structure, to investigate the molecular basis of agonist activation of the receptor (*69*). All classes of ligands were bound to the cathecolamine binding pocket almost identically and none displayed the outward shift of TM helix 6, characteristic of GPCR active states observed so far.

Agonist binding causes Ser212[5.43] to adopt a rotameric conformation where it can hydrogen bond to Asn310[6.55], and agonists can also hydrogen bond to Ser215[5.46], which was not observed for partial agonists or inverse agonists. A 1 Å contraction of the ligand binding pocket (Asn329[7.39] to Ser211[5.42] C_α distance) was also observed upon agonist binding. Although no further structural change is observed in TM5 below Ser215[5.46], its hydrogen bond between Val172[4.56] is replaced with the agonist, weakening the TM4-TM5 interface. At the same time, the TM5-TM6 interface is strengthened by the Ser212[5.43]-Asn310[6.55] as well as by the ligand bridged link between Asn310[6.55] and Ser211[5.42]. Together these interactions may be sufficient to set up a structural transition to the active state conformation.

Figure 1.4 shows the superimposition of the antagonist (cyanopindolol) and full agonist (isoprenaline) bound β_1AR structures as well as the inverse agonist (carazolol), and full agonist and $G\alpha_s$ bound β_2AR. The active-like conformation adopted by the A$_{2A}$R is also shown. The comparisons demonstrate that agonist binding alone does not induce the structural changes observed in the presence of G-protein *and* agonist, which are considered key features of the active state. The three different states also give insight into a conformational energy landscape that is different for each receptor, i.e. ligand binding alone causes different degrees of receptor conformational change.

β₂AR
Fully active

A₂AR
Active-like

β₁AR
Inactive

TM5

TM6

TM5

TM6

5JQH: inverse antagonist carazolol (+Nb60)

3SN6: full agonist BI-167107 and Gαs

3PWH: inverse agonist ZM241385

2YDO: full agonist adenosine

4BVN: partial agonist cyanopindolol

2Y03: full agonist isoprenaline
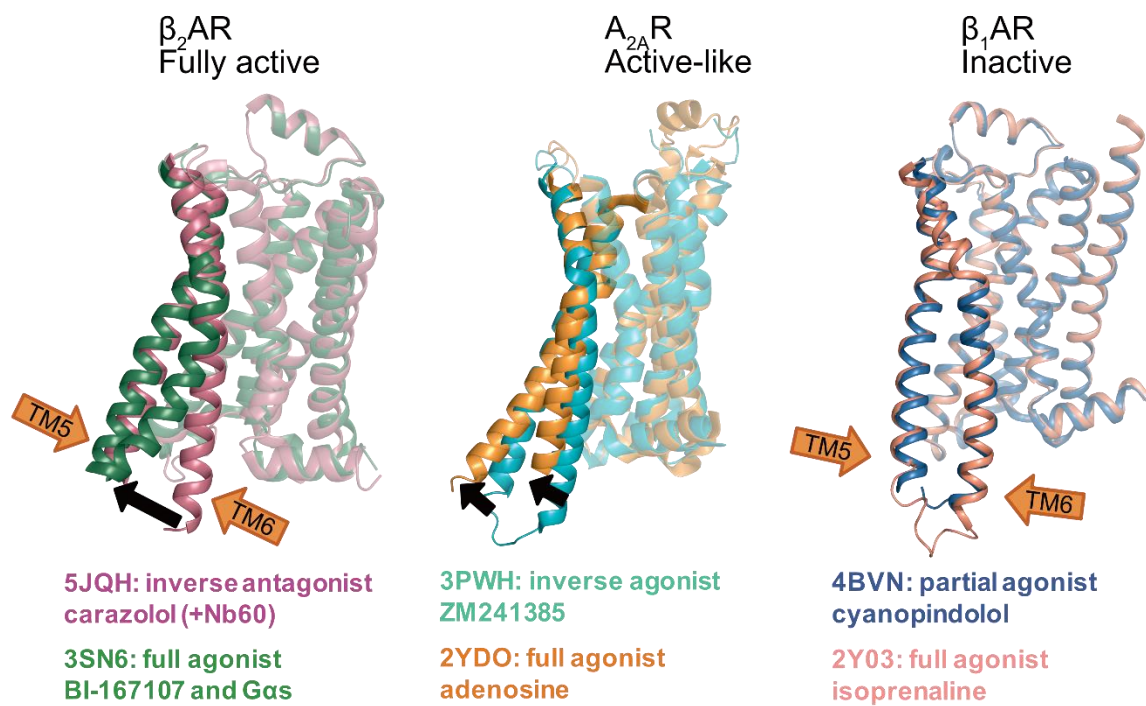
Figure 1.4 Structural superpositions of the β2AR, A2AR and β1AR

Structural superimposition of the inverse antagonist cyanopindolol bound β1AR 2YCY (blue) and the full agonist isoprenaline bound β1AR 2Y03 (pink), next to the A2AR bound to inverse agonist (cyan) and full agonist (orange). Overlay of the inactive carazolol and Nb60 bound β2AR (salmon) and the high affinity agonist and G protein bound β2AR (green).

1.3 – GPCRs are dynamic molecules

1.3.1 – Receptor plasticity

As mentioned above, the classical view of GPCR signalling envisaged a linear, agonist-binding induced transition to the R* active state, which enabled downstream G-protein mediated signalling. The desensitisation of the receptor by the interaction with arrestin, or the dissociation of ligand then causing the termination of receptor activity and a return to the inactive R state. This interpretation of GPCR activation/deactivation has been altered both by the emergence of a more complex picture of downstream signalling as well as the simultaneous emergence of high-resolution structural information about GPCRs. The discovery of alternative and tuneable signalling outputs of GPCR activation, such as cross pathway activation via arrestins, receptor desensitisation/re-sensitisation and degradation are difficult to explain via the switch on switch off R/R* transition model.

This simplistic view is easy to reconcile with rhodopsin, the earliest GPCR to be studied in molecular detail, which has a very large activation energy barrier and virtually no basal activity (*105*, *106*). Here the rhodopsin bound chromophore 11-*cis* retinal is isomerised to all-*trans* retinal upon photon absorption at an ultra-fast speed of 200 fs (*107*). The thermal barrier for this transition is around 45 kcal mol^{-1}, which translates to a single spontaneous receptor activation in 470 years (*108*). However other GPCRs such as β$_1$AR and β$_2$AR paint a very different energy landscape for the process of activation.

First of all their energy landscape is much more intricate than a two state system. The first piece of structural evidence to illustrate this is the conformational heterogeneity of TM6 in β$_1$AR in the inverse agonist bound state. This is described earlier in relation to two TM6 conformations, a bent one, where a salt bridge – typical of the inactive state rhodopsin family – is able to form compared to a straight conformation where this bridge is not observed. It is important to note that these variations were observed between receptors all binding the same ligand (cyanopindolol) and sometimes within the same crystal (*97*). This shows that the energy barriers between certain conformational states are very low and multiple states may be simultaneously sampled by a population of receptors. A second observation which suggests that the energy landscape is much more complex than proposed by a two state system is that none of the receptor structures crystallised in the presence of agonist *alone* and without intracellular binding partners show the hallmarks of the active state as described in section 1.2.2 based on the β$_2$AR-Gα$_s$ and β$_2$AR:Nb80 structures. A comparison of the β$_2$AR ligand binding sites across the Nb80 (3P0G), Gα$_s$ (3SN6) and covalent agonist bound structures (3PDS) shows a contraction of 1.2, 1.4 and 1.3 Å respectively, suggesting that although the agonist alone has not

induced transition to the active state, no further significant changes are expected to take place within the ligand binding pocket during the transition between these states (*109*).

Structures of the human adenosine A_{2A} receptor bound to agonists NECA, UK432097 and adenosine have also yielded insight into the dynamic landscape of receptor activation (*71*, *110*). Comparison to the inverse agonist (XM214385) bound structure has revealed that unlike for β_1AR, agonists make key interactions with TM7 residues Ser277$^{7.42}$ ad His278$^{7.43}$, and not with TM5 amino acids (Ser215$^{5.46}$ in β_1AR) (*111*). In addition agonist binding alone has caused the 2 Å shift of TM3 in the extracellular direction, a bulging and 3 Å outwards movement of the cytoplasmic end of TM5 and a 40° outward rotation together with a 5 Å opening of the cytoplasmic end of TM6. These changes set the receptor up in a very similar configuration to the activated β_2AR:Gα_s structure, except that the opening of TM6 is just too small to accommodate the insertion of the G-protein C-terminus (as observed in β_2AR:Gα_s) (*109*). This is a significant difference from the β_1AR and β_2AR structures, where agonist binding alone displayed no rearrangement on the cytoplasmic side. Therefore agonist binding positions the A_{2A}R on an unexpected intermediate on the activation landscape and suggests that this may be completely different to the landscape transition by the β-adrenergic receptors in the process.

Considering the insight that GPCR crystallography has lent to our understanding of agonist-induced receptor activation, a simplistic map of transition states can be drawn, coupled with a multi-minima energy landscape, which GPCRs have to traverse to change between inactive and active conformations. This is illustrated in **Figure 1.5**. Here agonist bound and agonist-free states change between active and G-protein bound states. Structures considered representative of each state obtained so far are highlighted.
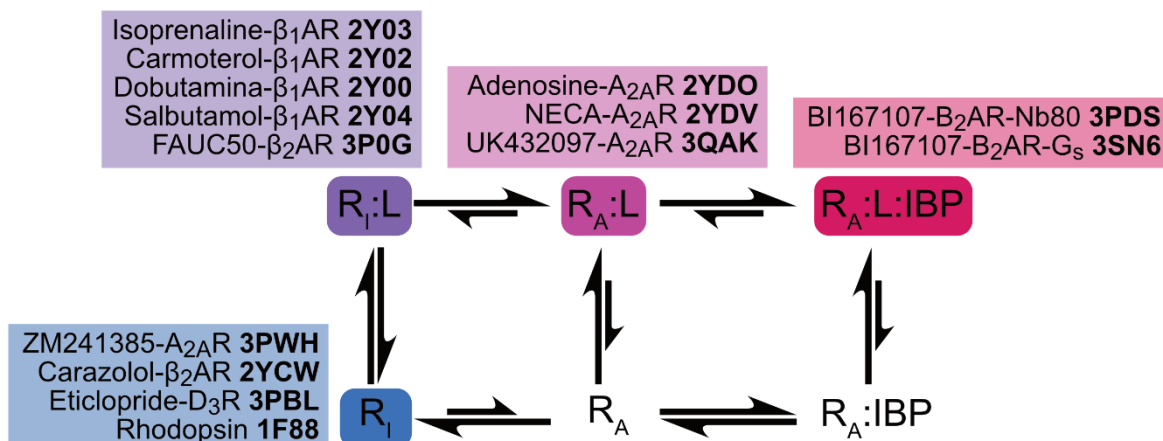
Figure 1.5 Extended ternary state model and representative GPCR structures

Selected examples of various receptor transition states between an inactive receptor conformation (as judged by dark state rhodopsin 1F88) and between the G-protein coupled state (β_2AR-G$_s$ 3SN6). Abbreviations: R$_I$: inactive state, R$_A$: active state, L: ligand, IBP: Intracellular Binding Partner.

It can be concluded that crystallography has provided valuable insight into the conformations GPCRs adopt *en route* from inactive to active states. However the field has opened up a gap for the study of dynamics between these observed states as well as for the characterisation of yet unobserved conformations, which may have been missed due to the tendency of lowest energy conformations to be represented in crystal structures.

1.3.2 – Solution state NMR can provide insights into the dynamic nature of β-adrenergic receptors

The study of GPCRs by NMR present a number of challenges. One of the most prohibitive is the apparent molecular size of detergent solubilised receptors. Large molecules suffer from long rotational correlation times, resulting in peak broadening and weak signals. Although the rate of molecular tumbling can be increased by temperature, this may not be permitted by receptor stability. Furthermore, residues undergoing conformational exchange on the μs-ms timescale in the intermediate time regime are likely to suffer first from the effects of peak broadening as peaks corresponding to such residues are already quite weak in intensity. Another factor is that the extensive α-helical nature of GPCRs gives rise to low spectral dispersion, in other words the chemical shift environment experienced by most TM helical residues is likely to be similar, causing spectral crowding, overlap (*112*).

NMR experiments rely on the presence of NMR active nuclei, and ones with half integral spin values such as ^{13}C, ^{15}N and ^{19}F are ideal. Unfortunately such nuclei have to be introduced to proteins due to their low natural abundance, through the process of labelling. Labelling strategies extend on one hand from the heterologous expression of fully labelled proteins through selective labelling strategies, where only certain residues contain NMR active nuclei, to post-translational incorporation of probes into the protein.

The expression of suitably labelled, functional receptors is also one of the challenges of studying GPCRs using NMR. Uniform ^{15}N, ^{13}C, ^{2}H labelling in *E. coli* or *E. coli* based cell-free expression systems have yielded complete structures of 7TM receptors using solution state NMR, namely of sensory rhodopsin II (pSRII) and of proteorhodopsin (*113*, *114*).

Most GPCR crystallographic studies have however relied on the expression of functional protein using the baculoviral expression system in insect cells, but uniform labelling in this system remained prohibitively costly and difficult until recently (*115*, *116*). Selective labelling approaches in insect cells have been successfully employed, however limitations due to the metabolic interconversion of amino acids, a phenomenon known as scrambling, must be carefully considered (*117*, *118*). An interesting alternative to uniform labelling in insect cells is presented by recent advances in isotopic labelling and expression in the nematode *C. elegans* (*119*).

NMR experiments of the β_2-adrenergic receptor have provided valuable insight into the dynamic nature of the receptor, revealing receptor conformations not observed through crystallisation. The study by *Nygaard et al.* employed a selective ^{13}C-methyl labelling of methionine residues in the β_2AR transmembrane region, an area not accessible to other biophysical methods or to ^{19}F tagging (*120*). The chemical shift spectra was compared across unliganded, inverse agonist bound, full agonist bound and full agonist *and* Nb80 (G-protein mimic) bound states. These spectra revealed that only the agonist together with Nb80 was able to stabilise a single active state and that the presence of agonist alone showed multiple dynamic conformations, but which were different from the ligand-free state chemical shifts.

A similar earlier study also relied on the selective labelling of methionine residues in β_2AR (*121*). Here a conformational equilibrium between two states was observed for both in the presence of partial agonists and of neutral antagonist.

A ^{19}F labelling strategy of two cytoplasmic cysteine residues in TM6 and TM7 has provided information about how the binding of biased agonists might induce G-protein or alternative signalling pathways.

A relatively slow conformational exchange between two states was observed for both TM cytoplasmic ends, which was differentially modulated by the binding of agonist or β-arrestin biased ligands (*122*).

Owing to the nature of the technique, these dynamic conformations are unlikely to be captured by conventional crystallography, hence NMR can be employed to give additional detail to the current repertoire of crystal structures. The opportunity presents itself to investigate the nature of dynamic intermediates of β1AR

Preliminary 2D [^{1}H, ^{13}C]-SOFAST-HMQC experiments recorded on ^{13}C-ε-methionine labelled Met2-$β_{1}$AR, purified by Dr Tony Warne (LMB, Cambridge) proved the feasibility of the approach to study the $β_{1}$AR using solution NMR. In such spectra, each peak corresponds to a ^{13}C-ε-methionine residue. A comparison of the number of peaks and their relative intensities across samples in the presence of different ligands or the G protein mimic nanobody, can provide information about structural differences between various ligand bound and activated states.

To interpret these spectra, is its essential that the peaks are assigned to the residue responsible for that peak. With this knowledge, these experiments can potentially yield information about local conformational changes during ligand or G-protein binding as well as about the relative timescales of these motions. It is possible to assign these peaks by mutagenesis, where one methionine is mutated at a time and changes in the spectra observed.

1.4 – GPCR-interacting antibodies

GPCR binding antibodies have attracted great interest for three main reasons. First, antibodies, especially single chain camelid nanobodies have played a significant role in the structure determination of GPCRs. They have been used to stabilise various receptor states, making them amenable to structural investigations by X-ray crystallography, cryo-EM and other methods. Second, antibodies have been exploited as therapeutic agents since the advent of the original hybridoma technology permitted the production of monoclonal antibodies (mAbs). Antibodies, as a form of biological therapeutics have been slow to enter the clinic in comparison to GPCR binding peptides, despite the R&D interest and a resulting rich pipeline of antibodies in development for the modulation of GPCR function. Third, antibodies are important components of the immune response against pathogens. Sometimes the pathogenic epitopes show resemblance to self-antigens, resulting in antibody cross reaction, which can manifest itself as part of an autoimmune disease. GPCR autoantibodies (including those against β_1AR) have emerged as an area of interest, as understanding the mechanism of chronic antibody mediated receptor agonism holds therapeutic promise for the treatment of conditions like idiopathic dilated cardiomyopathy (IDC) or chronic Chagas' disease to name two β_1AR related diseases.

I will discuss each of these three areas of interest in greater detail.

1.4.1 – Antibodies as tools for structural biology

Conventional mammalian IgG antibodies are formed of heterotetramers of two heavy (H) chains and two light (L) chains, linked by disulphide bridges. The light chains are composed of two immunoglobulin domains, a constant (CL) and a variable (VL) domain. The heavy chain is formed of three constant immunoglobulin domains (CH1-3) and of a variable domain (VH), with a hinge region linking CH1 and CH2. The N-terminal regions of the heavy and light chains form the variable domains, which are responsible for antigen binding via the complementarity-determining regions (CDR). The C-terminal CH domains are responsible for antibody mediated immune reactions, such as immune cell recruitment via Fc receptors, or complement activation via the classical pathway.

Antibodies can be engineered to adopt various formats, each with different sizes and properties. The conventional full IgG can be cleaved by papain protease to release two Fab fragments (**Figure 1.6**). Cleavage by pepsin generates the $F(ab')_2$ dimer. This dimer can be chemically reduced by breaking the disulphide bonds to generate a Fab' form. The V_L and V_H domains can be chemically linked (by heterologous single chain expression) to produce a single chain fragment variable (ScFv) domain. These can form dimers due to their chemical properties, to generate so called diabodies. Alternatively, ScFvs can be further linkered to generate di-ScFv molecules. Divalent antibodies can be generated by chemically cross linking two Fab fragments that have been generated from two different IgG molecules. These will have an affinity for two separate antigens. The smallest antigen binding domain of an antibody is a single V_L or V_H domain. This is called a nanobody. The CDRs or complementarity determining regions of an antibody determine its "paratope": the molecular surface of binding. Its antigen equivalent is called the epitope, present on the antigen surface. In summary the epitope and the paratope form the surface of interaction between antibody and antigen (*123*).
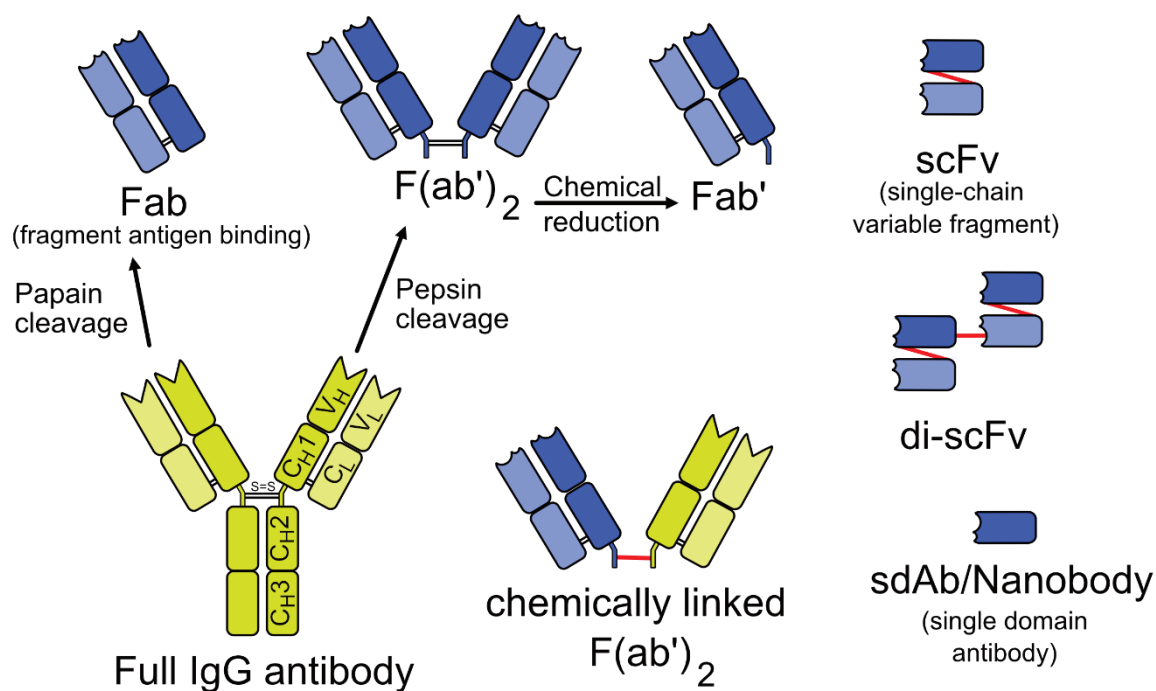
Figure 1.6 Schematic representation of IgG antibody formats

Depiction of IgG antibody formats. C_H: Constant Heavy, C_L: Constant Light, V_H: Variable Heavy, V_L: Variable Light. Double black lines represent disulphide bonds, red lines indicated chemical or peptide linkage. Modified from (124).

The serum of the family *Camelidae*, which comprises camels, llamas and vicuñas constitute an exception from the serum immunoglobulins of other mammals. In addition to the aforementioned conventional (IgG1) antibodies, they also possess heavy-chain only antibodies (HCAbs). They are different in that they lack the light chain component as well as the first constant heavy (CH1) domain. The antigen binding portion is hence only formed of a dedicated VHH domain, and not by two halves of a VL and VH pair. The VHH domain is equivalent to an ScFv construct made from an IgG1 antibody and is referred to as a single-domain antibody (sdAb) or a nanobody (*125*). One of the major structural differences that exists between conventional Abs and nanobodies, is that the antigen binding paratope is formed of only three loops as opposed to six. These three loops are generally longer, especially loop H3, which is often linked to loop H1 via disulphide bonds, though to reduce the entropic burden of binding associated with a large flexible loop (*126*).

Nanobodies have excellent stability and expression characteristics and high affinity binding clones can easily be screened for. Together with their small size (12-15 kDa), they have been exploited as tools in GPCR structural biology efforts. They have been used to lock GPCRs in selected conformations and to

stabilise transient states. A few examples where nanobodies and other antibody constructs have contributed to our understanding of GPCRs are outlined below.

The β2-adrenergic receptor was the first non-rhodopsin GPCR to be crystallised, overcoming a number of challenges associated with GPCR crystallography. In a nutshell, wild-type GPCRs like other integral membrane proteins need to be supplied with membrane mimetic factors such as detergents to enable their study in isolation from the plasma membrane. This presents a major challenge, as most GPCRs exhibit low stability when extracted, (expressed in melting temperatures) owing partly to the conformationally dynamic nature of these receptors. We now know that GPCRs transition through a continuum of conformational states, which translates to a degree of inherent instability. In order to restrict and so stabilise GPCRs, modifications and manipulations such as the combinatorial introduction thermostabilising point mutations, use of fusion partners, stabilising ligands, binding partners or membrane mimetics are required (*87*, *93*, *127*, *128*). The stabilising properties of the plasma membrane are also challenging to reproduce. For example detergents that are mild enough not to cause denaturation tend possess longer hydrocarbon chains, so the detergent micelle ends up obscuring polar and other contacts crucial for crystallogenesis.

A number of GPCR structures have been determined bound to antibodies, and a comprehensive list is given below (**Table 1.3**).

Table 1.3 Comprehensive list of GPCR structures bound to an antibody

Receptor	Class	Antibody (PDB code if complex structure exists)	Reference
US28	Class A	Nb7 (4XT1, 5WB1, 5WB2)	(*129*, *130*)
β2AR	Class A	Nb80, Nb6b9, Fab5 (2R4R, 2R4S) Nb35 (3SN6), Nb37, Nb60 (5JQH), A11	(*87*, *131–135*)
A2AR	Class A	Nb35 (6GDG)	(*136*)
κ-opioid receptor	Class A	(6B73)	(*137*)
M2	Class A	Nb8-9 (4MQS)	(*138*)
Glucagon receptor	Class B	mAb1 (5XEZ, 5XF1)	(*124*)
5-HT2B receptor	Class A	IgG P2C2-Fab (5TUD)	(*139*)
μ-opioid receptor	Class A	Nb39 (5C1M)	(*140*)
β1AR	Class A	Nb80, Nb6b9	(*141*)
PAR2	Class A	Fab3949 (5NJ6)	(*142*)

Figure 1.7 below depicts the binding interaction of some of the GPCRs listed in the table below. Antibodies have been used in various formats and their site of binding varies, depending on the intended application. Some of these complexes are discussed below in their molecular detail and functional properties.

PAR2 with Fab3949
(5NJ6)

μ-opioid receptor
with Nb39 (5C1M)

κ-opioid receptor
with Nb39 (6B73)

β_2AR with Fab5
(2R4R)

β_2AR with Nb80
(3P0G)

β_2AR with Gs heterotrimer
and Nb35 (3SN6)

Serotonin 2B receptor
and P2C2 (5TUD)

US28 with Nb7 and
CX3CL1 (4XT1)

M2 Muscarinic receptor
with Nb9-8 (4MQS)
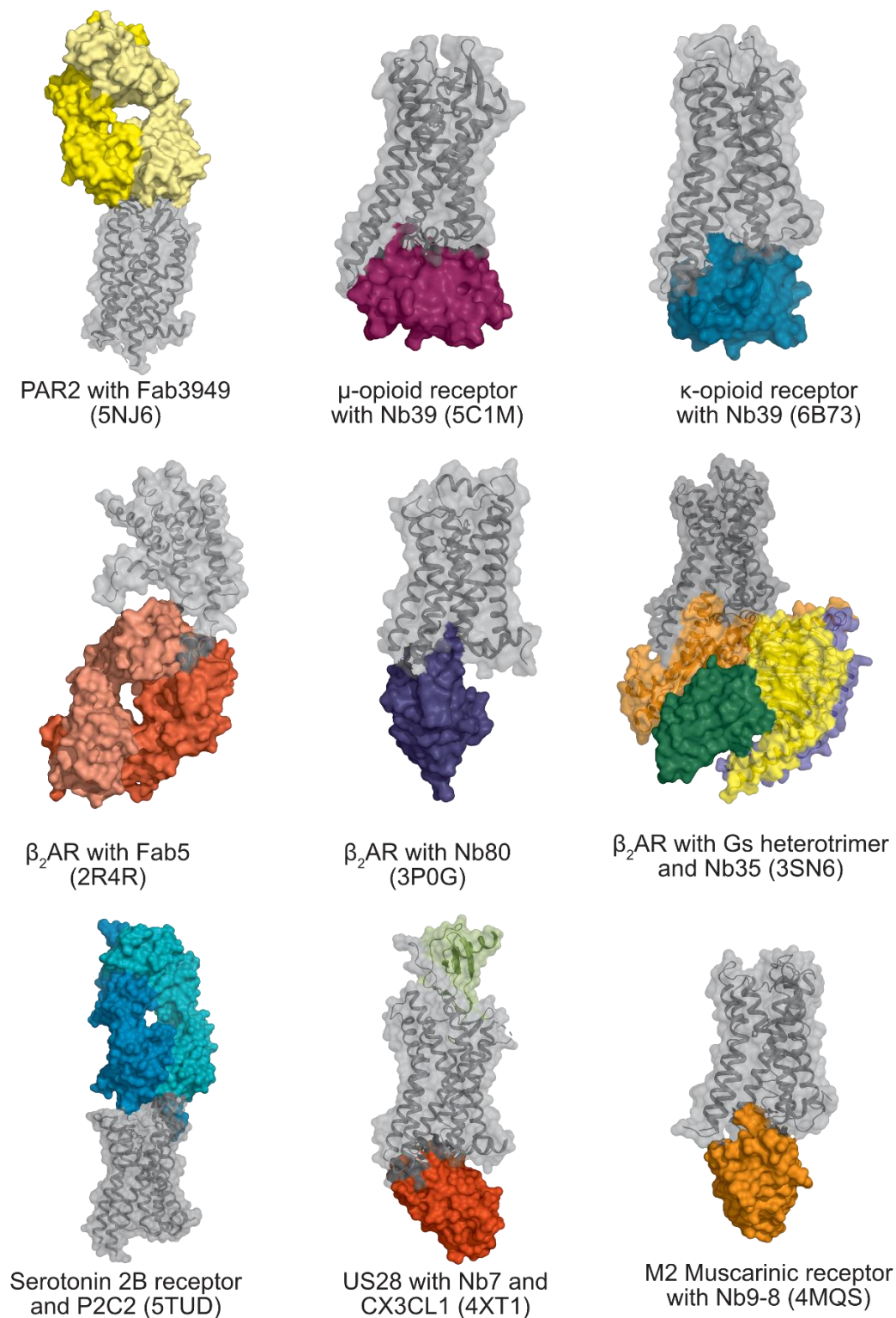
Figure 1.7 Examples of GPCR-antibody interactions.

The complexes show a range of antibody formats (Fabs and Nanobodies) and a range of binding sites. GPCRs shown in grey, antibodies and G-proteins (β_2AR 3SN6) in colour.

Initial approaches to determine the crystal structure of β_2AR followed two strategies. One strategy saw the insertion of the fusion protein T4-lysozyme into the loop between TM5 and TM6 (ICL3) and the other involved raising antibodies against a truncated receptor intracellular loop 3 (ICL3) and selecting immunoglobulins that bind preferentially to correctly folded, functional receptor. The aim of both was to provide hydrophilic packing contacts (*75, 76*). A suitable monoclonal antibody (mAb5) was identified, which was used in a Fab format (Fab5) to generate crystals. The resulting Fab5 and antagonist carazolol bound complex yielded a 3.4 Å model that revealed a well-ordered cytoplasmic side, however the extracellular regions beyond the membrane boundary above the ligand binding pocket were not possible to resolve.

Despite the limitations of this complex, the resulting structure has revealed key structural differences in comparison to the only previously known GPCR structure of bovine rhodopsin. A highly conserved region called the ionic lock, defined by the E/DRY motif on TM3 and a corresponding acidic residue on TM6 which form an important ionic interaction between TM3 and TM6, has been identified as a hallmark of the inactive state. In the inactive bovine rhodopsin a salt bridge is observed between R135 and E247, thereby restricting the movement of TM6 and locking it in an inactive conformation (*143*). In the active state opsin structure the arginine and glutamate residues are too far from each other to form an interaction, as TM6 is tilted outwards in the cytoplasmic direction (*144*). Similarly in the Fab5 stabilised β_2AR, the distance between R131 and E268 is 6.2 Å, too far to form a salt bridge. The lack of this association between TM6 and TM3 is in agreement with the high basal activity of β_2AR, in contrast to bovine rhodopsin (*131*).

Efforts to elucidate active state GPCR structures have historically proved challenging. The first β_2AR crystal structure bound to a covalent agonist probe FAUC50 revealed a receptor in inactive conformation, similar to the β_1AR in complex with the agonist levisoprenaline (*69, 81*). It seems that β-adrenergic agonists alone do not lend enough energy *in the crystallised state* (as opposed to in solution) to the receptor to stabilise a fully active state in the absence of any intracellular binding partners (IBPs). To raise nanobodies with the aim of stabilising an active state conformation, a high affinity, slow off-rate agonist (BI-167107) was identified and llamas were immunised with agonist complexed β_2AR in liposomes. The resulting G_s-mimic nanobody candidate Nb80 was identified for its G-protein like properties. A bimane fluorescent tag attached at $C265^{6.27}$, which reports differentially upon solvent exposure compared to being buried in a lipophilic phase showed a similar change to that observed with G_s. Furthermore, G_s binding has been shown to cause an increase in agonist binding affinity, and Nb80 elicited the same enhancement measured in competition binding assays with alprenolol. The T4L-β_2AR in complex with Nb80 and agonist yielded a 3.5 Å model, representing the first active state non-opsin GPCR structure (*80*). The Nb80 bound receptor revealed major changes on

the cytoplasmic surface, namely the outward movement of TM5 and TM6 and an associated inward shift of TM3 and TM7. The shifted helices accommodate CDR3 penetrating into the resulting space, with TMs 5 and 6 making further contacts to a short CDR1. The case of Nb80 highlights the utility of nanobodies and other antibodies in stabilising transient receptor states, so that they can be amenable to structure determination.

Similar to Nb80, nanobodies have been observed to bind conformational epitopes, or discontinuous segments of a primary sequence, that come into spatial proximity in various conformational states that might be higher or lower in potential energy and which as a consequence may be more or less transient. In contrast to conventional antibodies, nanobodies have been more frequently observed binding such conformational epitopes, due to a more convex paratope formed than that of six-CDR antibodies (*145–151*).

A question the use of nanobodies' raises is whether the conformations stabilised in nanobody-target complex structures are naturally occurring conformations or whether they are a result of induced fit caused by antibody binding. The B-cell mediated antibody response is a component of the adaptive immune system. The process of producing high affinity, secreted antibodies is the result of a process that starts with haematopoietic stem cell proliferation to produce progenitor B-cells, which undergo genetic recombination of the immunoglobulin loci to produce immunoglobulins with random variable regions. These immature B-cells then undergo a series of clonal selections. Negative selection ensures that cell surface immunoglobulins do not bind self-antigens in the marrow. Positive selection occurs when a B-cell engages its native antigen, causing that B-cell clone to proliferate. In the context of induced fit versus conformational stabilisation, B-cells that would bind an antigen in a forced conformation would have to pay a significantly higher energetic penalty in terms of enthalpy or entropy, than a B-cell-receptor that engages its antigen in a native conformation. This in turn would mean that B-cells which engage their antigen in an induced fit would have much lower chances of clonal selection (*123*, *152*). Hence it is argued that nanobody stabilised GPCR conformations represent organic intermediates within the broad structural ensemble of receptor activation.

This is further supported by the high degree of similarity between the Nb80 stabilised active state β_2AR structure and the same receptor in complex with the heterotrimeric G-protein G_s. The root mean squared deviation (RMSD) between the receptor components of the two structures is a mere 0.6 Å, with the biggest differences concentrating at the tips of TM5 and 6, where the contacts between the two binding partners vary (*54*). As Nb80 has been screened for eliciting functional changes akin to those induced by the presence of G_s, it can be argued that the remarkable conservation of receptor conformation across the two models supports the idea that nanobodies stabilise naturally occurring

intermediates, which don't necessarily occupy the lowest energy minima possible. This observation also validates Nb80 and similar IBP-mimetic nanobodies as tools for probing active-state and other intermediate conformations. This example reveals the intricate relationship between receptor structure and function, where a functional screen can identify a "stabilising scaffold protein" (i.e. the nanobody), which triggers the same molecular changes as does the native IBP (i.e. $G\alpha_s$).

The use of a nanobody in obtaining a crystal structure of the entire β_2AR-heterotrimeric-G_s complex demonstrates a different application of antibodies as tools for structural biology in comparison to Fab5 exploited for promoting crystal contacts or Nb80 as a stabiliser of a high energy state. Although it was initially difficult to form β_2AR-G_s complexes in detergent, the use of the non-specific purine phosphatase, apyrase was successful in alleviating the inhibitory effect of GDP to htG-protein-GPCR complex formation, released upon G_s association to the receptor (*153*). Heterogeneous stable complex were thus formed in a novel detergent called MNG-3 (commercially known as LMNG), however due to the large micelle size leaving inadequate amounts of extracellular surface exposed for contacts it was necessary to introduce a T4-lysozyme as an amino-terminal fusion. Despite these protein engineering interventions, significant structural heterogeneity was observed in the region of the nucleotide binding site, formed by the interface of the Ras-like GTPase domain (GαsRas) and the α-helical domain (GαsAH). The addition of the pyrophosphate analogue foscarnet achieved significant stabilisation, however the resolution of the resulting crystals could not be improved beyond 7 Å (*82*). To further stabilise the complex, a chemically cross-linked β_2AR-G_s-heterotrimer was used to immunise llamas, and a nanobody was screened for that was able to prevent the dissociation of the receptor from the heterotrimer upon the addition of GTPγS. The resulting Nb35 was both able to stabilise the complex and provide crystal contacts, achieving a 3.2 Å structure. Nb35 was found to bind at the interface of Gα and Gβ, with the long CDR3 interacting with both subunits. Nb35 also aided crystal formation through interactions of its scaffold region with Gα subunits of adjacent complexes in the crystal lattice.

Nb35 highlights a third use of an antibody in structural biology, where it is used to stabilise transient protein-protein interactions, making use of the unique propensity of camelid antibodies to be able to bind conformational or structural epitopes. The same strategy was employed to link the Gα and Gβ subunits in the cryo-EM structure of the adenosine A_{2A} receptor-TrxA fusion bound to an engineered MiniGs protein (6GDG) (*136*).

The value of conformational nanobodies in understanding receptor plasticity and the process of receptor activation is elegantly demonstrated by the β_2AR binding nanobody Nb60. In contrast to Nb80, which was shown to stabilise an active receptor conformation, Nb60 shows bias towards an

inactive state conformation. The cooperative effect of Gs and Nb80 on the affinity of the full agonist isoprenaline in comparison to the IBP (intracellular binding partner) free receptor has previously been demonstrated by the 75-fold increase in affinity caused by Nb80 binding and a 33-fold increase by Gs respectively. This was presumed to reflect the difference between an inactive receptor and an active receptor state, however Nb60 was further able to reduce isoprenaline affinity by 70 fold. This indicates that the form the receptor adopts in the absence of Nb80 is perhaps not the lowest activity state, or that it is only one of many inactive states that simultaneously exist. The nanobody-dose dependent modulation of ligand binding in opposite directions of affinity describe these nanobodies as classic allosteric modulators. Hence Nb60 can be referred to as a negative allosteric modulator, while Nb80 as a positive allosteric modulator. In pharmacological terms, the affinity of an agonist towards an uncoupled receptor (i.e. the absence of an IBP or effector) is referred to as a low-affinity state. However ^{19}F NMR work on the β_2AR describes this state as an ensemble of multiple low-affinity states, termed S1 and S2, which concurs with the aforementioned ability of Nb60 to further reduce full agonist affinity (*154*). The further decrease in affinity in the presence of Nb60 would then represent a very-low-affinity state, which resolves to one of the two observable spectroscopic states, namely S2 (*87*).

In structural terms, the S1 state represents an inactive state where residues R131$^{3.50}$ and E268$^{6.30}$ form a notable salt bridge referred to as the ionic lock, restricting the relative freedom of TM3 and TM6. The 3.2 Å crystal structure of β_2AR bound to Nb60 and the inverse agonist carazolol (5JQH) is remarkably similar to the receptor structure bound to carazolol alone (2RH1), with an RMSD of 0.3 Å in the transmembrane region and the ligand binding site. However the carazolol bound structure shows the ionic lock to be in a disengaged state, while the CDR3 of Nb60, protrudes into the space between TM3, 4 and 6 and a tyrosine and a threonine residue engages the ionic lock glutamate and arginine thereby mediating the engagement of the ionic lock. These structural and spectroscopic observations are however only part of the clue to understanding receptor allosteric modulation.

Radioligand binding experiments done on 12 ligands showed a strong positive correlation between Gs mediated efficacy and the Nb80 induced positive cooperative effect on affinity. In other words for these twelve ligands tested, the extent to which their binding affinity was increased by Nb80 was proportional to their efficacy. This observation implies that efficacy is a property that can stem from the modulation of ternary complex interactions alone, reflecting the allosteric effect of IBPs and ligands, and that it does not necessarily arise from the fine tuning of downstream signalling events. Unlike Nb80, Nb60 showed no correlation between efficacy and Nb60 enhanced negative cooperativity. It would make sense to observe an inverse correlation between Nb60 induced correlation and efficacy, however this was not the case with full agonists such as adrenaline,

noradrenaline, isoprenaline and others displaying both strong positive and negative cooperativity with Nb80 and Nb60, and several partial agonists such as clenbuterol, zinterol, salmeterol displaying only weak negative cooperativity with Nb60 but strong positive cooperativity with Nb80. This implies that the interaction of different ligands must vary depending on how the receptor conformational ensemble is biased by different IBPs.

To understand the relationship between negative and positive cooperativity the authors modelled the receptor as a simple two state system, interconverting between two allosteric conformations. This represents the receptor in an equilibrium between an Nb60 stabilised inactive conformation and an Nb80 stabilised active conformation, described by an equilibrium constant J and a ligand specific coefficient β, describing the effect each ligand has on the equilibrium. Agonists would hence be represented by larger β values, stabilising more of the Nb80 state. This two state assumption however fails to describe 30% of the ligands tested in terms of their effect on positive and negative cooperativity, and they cannot be described by an interconversion of two states.

The authors then explore a three state model, where a separate equilibrium constant describes the equilibrium between an Nb60 state and an intermediate (J_2), and the Nb80 state and that intermediate (J_1). Each of these constants are in turn modulated by the same ligands differentially through the coefficients β_2 and β_1 respectively. This model now allowed the cooperativity of agonists to be predicted, as by having a large β_1 coefficient, they stabilise active conformations, while having little effect on the stability of the inactive state as described by β_2. This allows for the correlation between efficacy and positive allosteric modulation, where full agonists have a large β_1, partial agonists have a small β_1. The cooperativity of other ligands could now also be explained through having a different effect on both J_1 and J_2, and the authors therefore class ligands into different groups, depending on their tendency to moderate these two equilibria, with each group described by the same coefficients. This means, that ligands of similar efficacy but belonging to different groups must vary in their ability to stabilise active and inactive states. In other words ligands such as clenbuterol, zinterol and procaterol which are similar in their efficacy with respect to Gs activation, the receptor conformations they stabilise to achieve receptor activation must be different, as their cooperativity cannot be described by the same coefficients.

In summary, the true value of conformationally selective nanobodies is demonstrated by this study of Nb60 and Nb80, in that they stabilise selected transient receptor conformations for long enough that they can be subjected to structural, spectroscopic and pharmacological investigation. In turn these allow the structure-function relationship of the receptor to be described by dissecting the conformational ensemble of receptor activation. These tools have moved our understanding of GPCRs

from being considered simple two state on-off systems, to much more complex mechanisms existing in a multitude of conformational states modulated by interaction partners such as ligands and IBPs. The value of this understanding is plain, as drug discovery efforts have provided a wide-range of tool compounds for the medicinal fine tuning of GPCRs without a clear understanding of the underlying structural processes, and now it is only fair to use these ligands and other tools, such as nanobodies, to improve this understanding for the betterment of the drug discovery process.

The human cytomegalovirus chemokine receptor US28 has been co-crystallised with the 77-residue chemokine domain of CX3CL1 with (4XT1) and without (4XT3) nanobody 7 (Nb7), which has been FACS selected from a yeast-display screen of alpaca nanobody library (*155*). The only difference between the two models, is that helix 8 adopts a different conformation due to spatial hindrance by the nanobody, but that the transmembrane helices take on an active conformation even with the chemokine peptide alone. This suggests that the Nb7 recognises a pre-formed active state and locks the receptor in that conformation from the intracellular side. While in this instance the same active conformation has been stabilised by the chemokine peptide alone, this scenario illustrates that should ligand binding alone not confer enough energy towards a lasting active state conformation, nanobodies can bind conformationally sensitive epitopes, stabilising a given receptor state. Indeed this was achieved for a later crystallographic model of US28. Here the receptor was C-terminally fused with Nb7, which allowed the ligand free apo-form to be determined (5WB1) (*130*). US28 shows promiscuity towards ligand binding, as it has been shown to associate with thousands of chemokines with high affinity (*156*).

The generation and use of the GPCR binding Nb7 has been indispensable to the understanding of US28 chemokine binding promiscuity. By stabilising the apo-state US28, Nb7 permitted experiments where the binding of different chemokine sequences was possible to evaluate. This built a picture of a unique protein-protein interaction, where unlike for other promiscuous interactions with a flat and exposed binding interface, the viral receptor accommodates thousands of human chemokines by allowing the distortion of a binding pocket formed by the TM helices, in a sequence independent manner while the receptor N-terminus also forms interactions with the ligand.

An additional observation afforded by the use of nanobodies is that another antibody, Nb11, showed impaired US28 binding to certain chemokine partners such as CCL5 and CCL3, evidencing intracellular allosteric modulation of ligand binding. Thus Nb11 and Nb7 appear to stabilise subtly different active states that preferentially couple to partially overlapping sets of ligands.

1.4.2 – Antibodies as therapeutic agents

GPCRs represent a significant group of targets for therapeutic exploitation. In a recent review Hauser et. al. have counted the number of GPCR-targeting FDA approved drugs to be 475, representing 34% of all FDA-approved drugs (*157*). This is in contrast to monoclonal antibody therapies, where out of the 83 mAbs approved in the EU or US, only Mogamulizumab (an anti-CCR4 IgG1 antibody) is approved in Japan for the treatment of CCR4 positive adult T-cell leukemia-lymphoma (*158, 159*). This was more recently followed by the approval of Erenumab (Aimoving) by the FDA at the time of this writing, an anti-CGRPR (calcitonin-gene related peptide receptor) mAb antagonist; the first and only FDA approved anti-GPCR mAb. This reflects the challenges associated with developing therapeutic antibodies against GPCRs, which primarily stem from the difficulty of preparing suitable antigens for immunisation and screening.

Despite the slow emergence of approved anti-GPCR therapeutic antibodies into the clinic, they possess some clear advantages over small molecule drugs (SMDs) that make this approach worth pursuing. Although monoclonal antibodies are more expensive to manufacture than SMDs, they show higher success rates in clinical trials, less susceptible to inter-patient variability and depending on the antibody format, they may have a longer duration of action, requiring less frequent dosing (*160*). Antibodies may also be better placed to expand the current repertoire of therapeutically exploited receptors, seemingly out of reach of drug traditional discovery approaches. The current 475 drugs acting on GPCRs target 108 receptors, representing only a mere 27% of all human non-olfactory GPCRs. High-throughput screening (HTS) methods have a relatively high hit-rate for Family A receptors, but four-fifths of these will fail due to their poor drug like properties. HTS methods have an even poorer track record for Family B GPCRs and HTS screening has proved inapplicable to Adhesion or Frizzled family receptors despite an emerging function in inflammation and oncology (*161*). Of the 399 non-olfactory human GPCRs 126 are classified as orphans, where no endogenous ligand has been identified to date (*162, 163*). Finding therapeutic molecules for this third of receptors remains an even greater challenge in the absence of tool compounds for functional screening, while they undoubtedly have a physiological role to play. This does not preclude them from the current drug development pipeline, as several GPCRs without endogenous ligand have been targeted in clinical trials. A few selected examples are listed below (**Table 1.4**).

Table 1.4 A table of most advanced anti-GPCR antibody therapies

Orphan GPCR	Indication	Compound	Clinical trial stage	Company
LGR5	multiple myeloma and colorectal cancer	BNC101: humanised mAb	Phase I	Bionomics Ltd.
GPR55	Spasticity related to multiple sclerosis and epilepsy	VSN16R (and GWP42003, GWP42006)	Phase II (VSN16R)	Canbex Therapeutics Ltd. (VSN16R)
GPR84	Ulcerative Colitis, Crohn's disease, Colon Cancer	GLPG1205	Phase II	Medical University of Graz, Galapagos NV
Source: clinicaltrials.gov and (*157*), orphan status: (*162*)				

Around 80 of these receptors are targeted with small molecules and about 30 or so with peptide biologics (*164*). Both of these types of drugs are often analogues of the natural ligands. The issue that many drugs face as a consequence of their similarity to natural ligands is poor selectivity between receptors in the same subfamily. Receptors within a GPCR subfamily, especially Class A receptors possess a large degree of sequence conservation in the region of the orthosteric ligand binding pocket, and the receptor-ligand interactions that are likely to be conserved when comparing natural ligands vs. small molecule drugs are the interactions between these conserved amino acids and the ligand (*165*). An emergent initiative of this issue of poor selectivity has been the search for allosteric modulators of GPCRs. These are molecules that bind to the receptor at a site away from the orthosteric ligand binding pocket. Allosteric ligands may exert a positive modulatory effect with regards to receptor signalling, in this case it would be referred to as a positive allosteric modulator (PAM). If a PAM is able to cause receptor activation in the absence of an orthosteric ligand, it is known as an ago-PAM. Conversely, if an allosteric molecule has a negative effect on receptor activity, it is called a negative allosteric modulator (NAM). In the case of PAMs and NAMs, a cooperativity exists between the orthosteric binding site and the allosteric site, so that the presence of the allosteric modulator may enhance or prevent activation via the orthosteric ligand. If a NAM is unable to completely block receptor activity, it may be referred to as a partial-NAM. Therapeutic antibodies can offer a solution to the issue of selectivity, if their mode of action (MoA) is independent of engaging residues in the orthosteric binding pocket.

The therapeutic utility of any raised antibodies depends on whether it is able to modulate receptor activity. The modulation of receptor function can take place either as positive process, where receptor signalling is enhanced, or as a negative process, where the extent of the signalling reaction is reduced.

The process of antibody-enhanced signalling may take place in different ways. An antibody may stabilise an active receptor conformation in the absence of a ligand, or it may trap the receptor in an active conformation with ligand, causing ligand binding to persist together with its associated signalling output. Antibodies formats such as divalent full length IgG, or diabody fragment formats (ScFv) may also cause receptor dimerisation and induce signalling through an active state homodimer (e.g. metabotropic glutamate receptor 7) or heterodimer (e.g. Class B GPCR in association with a RAMP) (*166, 167*).

Similarly, the negative modulation of signalling can be envisaged in various ways. An antibody may prevent ligand association either through competition within the active site or by binding to an allosteric site and inducing a conformation which is unfavourable to ligand binding. It may also block ligand entry by hindering access to the binding pocket. An antibody could also downregulate signalling without interference to ligand binding, for instance by stabilising an inactive receptor conformation or by triggering receptor internalisation (*168*).

Antibody therapeutics may offer an advantage over small molecules or peptides stemming from their limited ability to cross the blood brain barrier. Certain receptor targets such as opioid receptors are distributed throughout the body across both the central and peripheral nervous system. Long term morphine treatment of rats in the presence of paw inflammation has shown that tolerance does not arise from the stimulation of peripheral μ-opioid receptors, a key feature of addiction (*169*). Similarly, compounds selectively targeting peripheral κ and μ-opioid receptors were able to achieve 50-80% of the analgesic effect achieved with systemic administrations of opioid drugs (*170*). Hence antibodies with their limited CNS penetration offer favourable distribution, which may be exploited to limit undesirable side effects in CNS indications.

Antibodies are native parts of the immune system and as such they perform their function in concert with other components. This means that unlike other therapeutic molecules, antibodies can, in addition to binding and modulating the structure/function of their target ("blocking" mode of action – MoA), exert a therapeutic effect through the modulation of the immune system. This can take form as antibody-dependent cellular cytotoxicity (ADCC), where cells bound by certain antibodies are recognised by effector cells such as natural killer (NK) cells, macrophages, neutrophils or eosinophils to bring about cell death of the antigen coated cell. This is triggered by the recognition of the Fc portion of the antibody binding to Fc-receptors on the surface of effector cells (*171*). At least 8 mAbs employ ADCC as a mode of action, in addition to the first approved anti-GPCR mAb Mogamulizumab, where the Fc glycans have been devoid of fucose groups (afucosylation), in order to invite an ADCC response (*172*). Mogamulizumab binds to chemokine receptor 4 (CCR4) on the surface

of T-cells for the treatment of adult T-cell leukaemia-lymphoma. Another mode of action is complement dependent cytotoxicity (CDC). The circulating complement protein C1q is able to bind to the Fc region of antibody-antigen complexes, inducing a cascade that results in the formation of the membrane attack complex (MAC), which inserts into the lipid bilayer of the antibody-targeted cell, resulting in lysis. Together with ADCC, CDC is another MoA which lends therapeutic antibodies an anti-cancer potential (*173*). NK cells can also be triggered to phagocytose antibody bound target cells through a process known as antibody dependent cellular phagocytosis (ADCP). This is another Fc mediated immune function that has been exploited as an anti-tumour therapy and has been identified as the MoA for Elotuzumab for the treatment of multiple myeloma (*174*). Antibodies may also be chemically coupled to small molecule payloads that can serve as diagnostic markers or as cytotoxic agents in their own right for an enhanced therapeutic potential. The delivery of small molecules in this fashion also limits distribution and goes towards minimising off-target toxicity. A mAb linked to another molecule in this way is referred to as an antibody-drug conjugate (ADCC). There have been 6 ADCC mAbs approved to date predominantly against non-solid tumours such as lymphomas and leukaemia, with the exception of HER2 positive breast cancer (Ado-trastuzumab emtansine) (*159*, *175*).

The generation of antibodies against a specific target presents a number of prerequisites towards the nature and quality of the antigen used. The antigen must be as pure as possible, so that the antibodies raised are as specific to the target as possible, and it must be available in suitable quantities to permit the immunisation work to be carried out. The antigen has to be stable, so that the desired epitopes are present for long enough that an immune response can be raised, otherwise the latter will not be possible, or the antibodies generated will be against degraded or denatured versions of the target. An additional requirement is that the antigen is in the desired conformational state and that its structural distribution be as homogeneous as possible (*176*).

GPCRs are not ideal candidates for fulfilling any of these basic criteria. As membrane proteins, the majority of the receptor sequence is embedded in the cell membrane, and is rendered unsuitable as a solvent accessible antigen. Intracellular epitopes have been targeted by so called intrabodies: antibodies that have been delivered to the cytosol through physical manipulation such as electroporation or microinjection or by transfection through vector plasmids. However intrabodies have so far not proven amenable to therapeutic exploitation (*177*). Other attempts to achieve antibody cell penetration relies on intracellular targeting sequences or clathrin-mediated endocytosis and release from endosomes into the cytosol (*178–180*), although these have not been exploited in the context of targeting GPCRs. GPCRs with large cytoplasmic domains such as the large extracellular

domains (ECDs) of family B GPCRs, or those with large N-terminal domains (NTDs) serve well as antigens, and those with short NTDs or minimal extracellular loops (ECL) are less ideal (*181*).

The best source of antigen is a homologous preparation of purified protein, however for GPCRs this is not always easily achievable. Most native full length GPCRs exhibit a degree of instability when extracted from the native plasma membrane. The choice of detergent has to carefully balance the stability of the extracted protein versus the utility of the chosen system for the given application. For example short chain detergents such as SDS are highly potent in disrupting biological membranes and extracting membrane proteins, at the cost of losing tertiary structure. Longer chain, mild detergents such as dodecyl-maltoside (DDM), fare much better at preserving the integrity of the solubilised proteins, but they may occlude parts of the protein which are otherwise solvent accessible and would serve as useful antigens. When using detergents for the preparation of membrane proteins for immunisation, the critical-micelle-concentration (CMC) of detergents must be considered. The CMC is an empirically determined free detergent concentration above which micelles form under the given conditions (temperature, ionic strength, pH). The implication of this property is that detergent solubilised protein samples require the presence of some free detergent to allow the formation of protein bound detergent micelles. Free detergent can trigger an undesired immune reaction causing irritation and give rise to non-specific antibodies. GPCRs usually suffer from low levels of cell-surface expression and various approaches to protein engineering have been explored to boost both protein stability and expression levels. One method of thermostabilisation involves the introduction of a combination of stabilising point mutations into GPCRs to form StaR receptors, patented for use in drug discovery by Heptares Therapeutics (*182*, *183*). Another method of stabilising GPCRs has achieved a 26.8 °C increase in melting temperature (Tm) using a form of molecular evolution called CHESS (cellular high-throughput encapsulation, solubilisation and screening), for the neurotensin receptor 1 in the apo state (*184*). Modifying native receptors for the gain of enabling the production of purified GPCR may run the risk of losing biologically important features such as post-translational modification (PTMs) or loops, which may otherwise be susceptible to cleavage and receptor degradation. In essence, modifications which enhance protein stability and which have been instrumental in crystallogenesis may not be compatible with preparing purified proteins for immunisation.

Various alternative immunogen formats to purified protein have been explored. Peptide analogues of GPCR antigens are a convenient method of immunisation as peptides can be manufactured in milligram scale quantities with high purity. The structural conformation of peptide sequences may however not reflect that of the matching sequence as part of a larger protein, therefore the raised antibodies may recognise a linear epitope of the peptide as opposed to a structural epitope (*185*). This can be circumnavigated by training the peptide to a desired conformation as has been exploited by

Pepscan's CLIPS technology (*186*). Even with a representative conformation the use of peptides for raising anti-GPCR antibodies should consider the utility of the selected peptide sequence as an analogue for the portion of the GPCR that the raised antibodies will target. For example, using a receptor N-terminal sequence peptide will likely generate antibodies which are able to modulate GPCR activation, where the given N-terminus plays a role in the process, such as for chemokine receptors which bind their chemokine ligands at the N-terminus (eg. CXCR4), or for protease activated receptors (eg: PAR2), where N-terminal cleavage reveals the tethered activating ligand (*187*).

DNA immunisation is also an option. With an optimal choice of vector and the use of enhancing adjuvants (immune response enhancing compounds) DNA immunisation can trigger the host organism to translate the injected DNA at low but significant levels which can serve as antigens (*188*, *189*). Because of low levels of expression this method is often used in combination with other antigen formats in a so called prime-boost scheme.

Whole cells or prepared cell membranes without the cytoplasmic components can also be harnessed as antigens. Cell membranes preparations can be enriched for the overexpressed target protein in comparison to whole cells and are also more stable than intact cells which are susceptible to lysis. Membrane fractions on the other hand expose both intracellular as well as the desired extracellular epitopes, meaning that a counter-screening step with extracellular epitopes is necessary to eliminate these. Antibodies raised against intracellular receptor elements can be valuable for research and development purposes, but their utility as therapeutic agents remains limited for reasons of cell penetrance as discussed before (*168*).

1.4.3 – Antibodies as natural modulators of GPCR function

Antibodies are vital components of the immune system in binding foreign antigens for elimination by various immune processes. However sometimes this process gets dysregulated and autoimmune conditions arise from cross-reaction to self-antigens. This phenomenon has been noted in the context of GPCR binding autoantibodies in a wide range of indications (*190–202*). The β_1AR is no exception to this phenomenon, and an example of this process in the context of chronic Chagas' heart disease (cChHD) is discussed below.

1.4.3.1 – *Trypanosoma cruzi* ribosomal P proteins are major antigens in Chagas' disease

Chagas' disease is caused by a parasitic flagellated protozoan organism called *Trypanosoma cruzi*. Like many parasitic organisms, *T. cuzi* has a complex life cycle involving humans amongst other mammals as hosts. Blood sucking or haematophagous triatomine beetles of the *Triatominae* subfamily serves as the insect carriers of the parasite. The bugs acquire the parasite in its circulating trypomastigote form,

through taking a blood meal. Once inside the triatomine digestive tract it differentiates into a form called the epimastigote capable of reproduction through binary fission. As the parasite transits the gastric tract the differentiation continues in the hind-gut into metacyclic trypomastigotes. In this form the trypanosome is capable of being released in the faeces of triatomine bugs and can penetrate the mammalian host at the site of the insect bite. It is important to note that the protozoan does not penetrate its mammalian host through the triatomine mouth parts during the course of it taking a blood meal. It is the defecation onto the surface of the skin and the associated scratching behaviour of the host that "rubs" the trypanosome into the wound left by the bite. Consequentially *T. cruzi* penetrates local cells around the infection and takes on the amastigote form, replicating once again by binary fission. The resulting trypomastigotes are released into blood circulation, ready to be taken up by the carrier once again. Circulating trypomastigotes also carry out further tissue invasion, reproduction and cell lysis (*203*).

The disease manifests itself in an initial acute phase with symptoms such as general malaise, fever enlarged spleen and lymph nodes. The site of parasitic entry is often associated with the regions of ocular mucosa, resulting in a characteristic swelling of the area surrounding the eyes. This is an early manifestation of Chagas' disease called the Romaña sign (*204*). These early symptoms normally resolve spontaneously even in the absence of trypanocidal treatment. About 30-40 % of those infected go on to develop chronic Chagas' disease within 10-30 years. The nature of this complication can be categorised as either cardiac, digestive or cardiodigestive, leading to cardiomyopathy, megacolon, megaoesophagus or a combination of these. The cardiac manifestation of chronic heart disease is referred to as chronic Chagas' heart disease (cChHD). Globally there are 8 million people estimated to be infected with Chagas' disease, translating to a mortality rate of about 12,000 cases per annum. (*205*).

Antibodies that cross-react to self-antigens have been observed in the serum of chronic Chagas' disease (cChD) patients, which suggested an autoimmune component to the disease (*206–208*). The cardiac manifestation, typical of chronic adrenergic receptor stimulation has prompted the investigation of adrenergic receptor epitopes as cross-reacting antigens in the autoimmune response (*209*, *210*).

The major antigens in the immune response against *T. cruzi* infection have been identified as the ribosomal P proteins, which make up the stalk of eukaryotic ribosomes (*211–213*). The ribosomal stalk is formed of the 34 kDa protein P0 and two closely related, but smaller P1 and P2 proteins, both approximately 11 kDa in size. All three share a highly conserved acidic C-terminus. *Trypanosoma cruzi* possesses two P1 and P2 subtypes: P1α/P2α and P1β/P2β. Its P0 C-terminal end deviates from the

conserved eukaryotic sequence and instead bears resemblance to the Archeal P0 substitute L10 C-terminus (*214*). The presence of P proteins have been shown to be essential for the GTPase activity of the eukaryotic elongation factor 2 (e-EF2), which catalyses the transition of peptidyl-tRNA from the ribosomal A site to the P site (*215*). These proteins will be hereon referred to as TcP1/2/α/β.

A sequence alignment of the C-terminus of the TcP2α (UniProt: P23632), TcP2β (P26795) and TcP0 (P26796) proteins are shown below, together with the consensus sequence of the last 13 residues of these proteins referred to as the R13 peptide.

Table 1.5 T. cruzi Ribosomal P protein C terminal sequences

TcP2a	AEEEEDDDMGFG-LFD	107
TcP2b	EEEEEDDDMGFG-LFD	112
TcP0	EEEDDDDDFGMGALF-	323
R13	--EEEDDDMGFG-LFD	13

In eukaryotes this conserved C-terminus is suggested to mediate the toxicity of ribosome inactivating proteins (RIPs) such as sarcin, ricin and tricosanthin. These proteins have been shown to cause irreversible changes in the form of base purination in a region of the ribosome called the Sarcin/Ricin Loop (SRL), proximal to the GTP binding site. The crystal structure of trichosanthin in complex with the conserved P protein C-terminal sequence SDDDMGFGLFD revealed an electrostatic binding interface between the two proteins, through arginine and lysine residues (*216*). This is not at all dissimilar to the interactions between elongation factors and the P1/P2 C-termini, which extend up to 125 Å away from the N-terminal dimerisation domain, which links them to the P0 component of the stalk (**Figure 1.8**). The hydra-like P1/P2 protrusions, binding and thus generating a high local concentration of elongation factors are also suggested to mediate RIP toxicity, by recruiting such toxins to the region of the SRL (*217*).
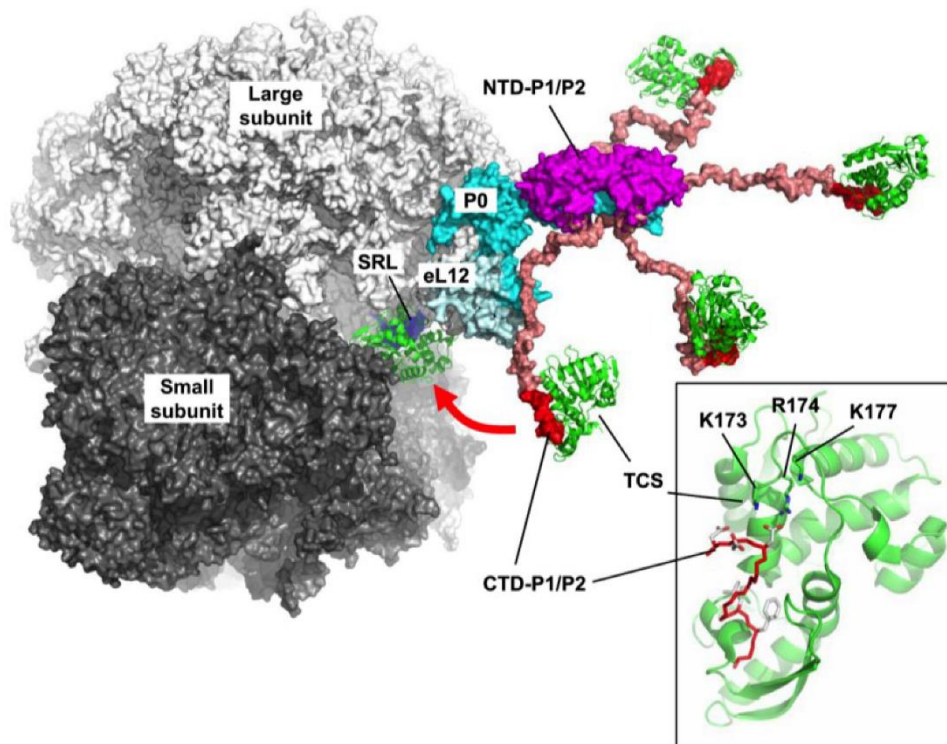
Figure 1.8 Structure and function of ribosomal P proteins

Proposed mechanism of ribosome inactivating protein recruitment to the Sarcin/Ricin Loop (SRL) via the P1/P2 C-terminal domains (CTD). The structure of the eukaryotic stalk complex is docked onto the yeast ribosome. A single unit of P0 (cyan) binds two P1 and P2 proteins each through their N-terminal domain (NTD – purple). The conserved P1/P2 CTDs (red) form electrostatic interactions between lysine and arginine residues of Trichosanthin (TCS – green) and the acidic residues of the CTD. The model therefore suggests that the P stalk is involved not only in elongation factor recruitment, but also the mediation of RIP toxicity. Figure taken from (*217*).

The last 13 residues of the *T. cruzi* P1 and P2 proteins (α and β variants) follow the conserved sequence of EEEDDDMGFGLFD, which has been recognised as the major epitope in the humoral response in Chagas' disease (*218*). Key studies establishing this are discussed in the section below.

1.4.3.2 – Anti-TcP antibodies cross-react to β₁AR extracellular loop 2 (ECL2)

The recognition that chronic Chagas' heart disease (cChHD) may be caused by an underlying autoimmune process was sparked by the histological observation of cardiomyocyte degradation alongside fibrotic lesions and mononuclear cell invasion (by T-cells, B-cells, NK cells) in the apparent absence of parasitic cells or tissue (*219*). Antibodies in the sera of cChHD patients were found to react

to β-adrenergic receptor (β1, β2) and to M_2 acetylcholine receptor epitopes (mAChR M_2), postulating a disease mechanism which is exerted through immune cross-reaction to these receptors (*220–223*).

From earlier studies it was understood that these GPCRs were recognised at their extracellular loops, but the exact epitopes and the mechanism of action of these cross-reacting antibodies has not been fully understood. *Elies et al.* investigated the epitopes relevant to $β_1AR$, $β_2AR$ and muscarinic acetylcholine receptor M_2, by looking at cChHD patient blood serum: the supernatant left-over from clotted blood lacking clotting factors, but containing blood immunoglobulins. In the study, synthetic peptides from the extra cellular loop 2 (ECL2) of these receptors were prepared, and the total serum immunoglobulin fraction and the IgG fraction were tested for reactivity to each of these peptides (*224*). The table below shows a summary of ECL and other peptides. The nomenclature of these peptides is conserved throughout much of the relevant literature.

Table 1.6 Peptides relevant to cChHD autoantibody cross reaction

Peptide	Protein of origin	Sequence
H26R	$β_1AR$	HWWRAESDEARRCYNDPKCCDFVTNR
H26Q	$β_2AR$	HWYRATHQEAINCYANETCCDFFNTQ
V25I	mAChR M_2	VRTVEDGECYIQFFSNAAVTFGTAI
A7D	$β_1AR$	ARRCYND
P0β	*T. cruzi* P0	AESEE
R13	TcP C-terminus	EEEDDDMGFGLFD

Enzyme immunoassays (ELISA) performed on immobilised H26R, H26Q and V25I peptides showed a significant response for all three peptides for sera of Chagas' positive patients (diagnosed by electrocardiogram tests – ECG) compared to healthy donors. The strongest reaction was observed for the $β_1AR$ peptide, followed by $β_2AR$ and for the M_2 peptide. All three peptides also elicited a reaction for an isolated IgG fraction of the serum. The binding to both the $β_1AR$ and the M_2 peptide was sensitive to the strength of the ionic buffer, with decreased binding observed for an increasing strength of salt, suggesting a strong electrostatic component to the peptide-immunoglobulin interaction. To evaluate direct binding the authors used surface plasmon resonance (SPR) with immobilised peptides and serum IgG fractions. The binding observed to the H26R $β_1AR$ peptide was much higher than for the $β_2AR$ or M_2 peptides, which were both slightly higher than to the polyanionic control chips, suggesting that while cChHD patients' sera contains higher levels of anti-polyionic IgG antibodies than that of healthy controls, but the major immune response is against the $β_1AR$ ELC2 H26R peptide, with a minor component against the other two receptors.

To evaluate the physiological effects of the antibodies the authors incubated the serum IgG fractions on spontaneously beating neonatal rat cardiomyocytes in culture. Five out of eight sera caused an

increase in the basal beating frequency, however in the presence of atropine, an M_2 receptor antagonist, all eight patient's sera converted to a positive chronotropic response. M_2 agonism results in a decreased chronotropic effect, meaning that the abolition of this allows the positive effect caused by adrenergic receptor antibodies to manifest itself. Consequently, when the serum IgG pool was affinity purified using the H26R peptide, the resulting β_1AR specific increase in beating frequency could be abolished by the β_1AR specific antagonist bisoprolol, as well as by the *T. cruzi* P0 peptide P0β. This work has shown biophysical and physiological evidence of an interaction between cChHD patient serum and GPCR ECL2 peptides.

The GPCR epitope sequences shown in the table above, which have been implicated by this study suggests that it is the charged stretch of residues which bear resemblance to the AESEE TcP0 protein and to the TcP2β R13 sequence component EEEDDD which are primarily responsible for the immune cross-reaction. This is homologous to the β_1AR AESDE motif and to the mAChR M_2 EDGEC sequence in reverse (CEGDE) (*224*).

A subsequent study by Kaplan *et. al.* looked at the sera of 15 further cChHD patients and compared it against acute Chagas' disease, which lacks a cardiac indication, idiopathic dilated cardiomyopathy (IDC) a hallmark of which is chronic heart stimulation by β_1AR autoantibodies, systemic lupus erythematosus (SLE), which features anti human-ribosomal P-protein antibodies, all against reactivity to R13 peptide. The study aimed to confirm whether it is the humoral anti-R13 response that is responsible for the chronic heart disease component of Chagas' disease (*225*). Additionally, the authors wished to discern, whether the anti-human ribosomal P0 C-terminus (hP0) autoantibodies prevalent in SLE contribute to a shared disease mechanism with cChHD. This was postulated by several preceding studies, which upon finding that SLE anti-hP0 antibodies showed a shared specificity to the homologous TcP0 peptide suggesting that the two diseases may share a common mechanism for generating the anti-P response.

The results of this work showed that the anti-P antibodies in SLE reacted equally to TcP0 as well as to hP0, while anti-P antibodies in cChHD showed a clear preference for TcP0 whole protein and also did not bind to *T. brucei* P-protein. These observations translated to reactions against the R13 and H13 C-terminal peptides of *T. cruzi* and human origins respectively, indicating that the parasite ribosomes are the most likely to be responsible for the cross-reacting anti-P antibodies in chronic Chagas' disease. Furthermore acute patient sera showed no reaction to the R13 peptide. The difference in terms of species preference for autoantibodies in cChHD and SLE suggests that the two diseases do not share a common disease mechanism, as the serum response of cChHD did not recreate the typical anti-hP0 response characteristic of SLE.

With regards to the comparison of the humoral antibody response in cChHD, SLE and IDC, only insignificant amounts of immune cross-reaction was observed for IDC sera, tested both with recombinant TcP1, TcP2β, hP1 and hP2 protein and with equivalent peptides. The physiological response of H26R and R13 affinity purified serum antibodies tested on spontaneously beating cardiomyocytes was as high as that elicited by the adrenergic receptor full agonist isoprenaline and showed significant reciprocal abrogation by H26R (for R13 purified Abs) and by R13 (for H26R purified Abs), as well as by bisoprolol. These findings establish the anti-P response as a critical biomarker for cChHD, and pinpoint the polyanionic, highly conserved C-terminus of the *T. cruzi* ribosomal P proteins as the major epitope for the humoral response raising cross-reacting antibodies (*223*, *225*).

Sepulveda *et. al.* present *in vivo* evidence to further validate the conserved C-terminus of TcP2β as the major antigen responsible for the cardiopathic element in chronic Chagas' disease. They used mice to compare the immune response that arose from *T. cruzi* epimastigote infection and from immunisation with TcP2β purified protein antigen (*226*). The humoral response was probed with 25 twelve residues long fragments of TcP2β using an ELISA, to evaluate the relevant immunogenic epitopes. The authors observed that antibodies raised against the recombinant His-tagged TcP2β recognised peptides from several part of TcP2β, with two major clusters of epitopes at the hinge region and at the C-terminus. In contrast, the epitopes from the *T. cruzi* infected mice matched almost exclusively to the conserved C-terminus. This implied that the acidic C-terminus of TcP2β is the immunodominant epitope in cChHD.

1.4.3.3 – A monoclonal antibody against cChHD relevant epitopes that cross-reacts to β$_1$AR

As discussed earlier, monoclonal antibodies are great tools as they allow the uniform modulation of their target for either therapeutic or research purposes. To investigate the disease mechanism mediated by cross-reacting antibodies in cChHD, Mahler *et. al.* immunised mice with recombinant MBP fusion TcP2β protein (*227*). A resulting monoclonal antibody candidate mAb 17.2 was identified and evaluated. mAb 17.2 was able to recognise all *T. cruzi* ribosomal proteins as well as human ribosomal proteins. Cross reaction was shown to β$_1$AR in immunohistochemistry experiments. The antibody showed agonistic properties on spontaneously beating rat cardiomyocytes. It also caused chronotropic alterations of mice determined by ECG measurements (*214*, *225–232*). The sequence of an agonistic ScFv format of mAb17.2 (ScFv C5) as well as a non-agonistic version (ScFv B7) are published, that have been used as a starting point for the investigation of this disease mechanism.

1.5 – Aims and objectives

The specific question this study seeks to answer is: what information can we gain about the activation mechanism of β_1AR, by studying the conformational signatures of the process? The labelling strategy relied on ^{13}C-methyl-methionine residues, as methionines do not show metabolic interconversion or scrambling in insect cells (*233*). Methyl groups are also highly sensitive reporters of changes in their chemical environment. Their flexibility in methionines offers a degree of relief from the long rotational correlation times experienced by micelle solubilised membrane proteins, which causes peak broadening.

The objectives of this project are hence threefold:

Firstly, to express ^{13}C-ε-methionine labelled functional β_1AR using the baculoviral insect cell method.

Secondly, to establish a method for the purification of such expressed β_1AR constructs.

And thirdly to record 2D HMQC experiments on the isotopically labelled purified protein. Resonance peaks are to be assigned by mutagenesis.

In the broader biological context this project aims to characterise the dynamic nature of β_1AR activation, by describing the conformational transitions that take the receptor through its various activation states induced by interactions with ligands and other binding partners. In addition to these objectives, the preparation of an agonistic anti-β_1AR antibody is sought, for studies of the ligand independent mechanism of activation.

[PAGE INTENTIONALLY LEFT BLANK]

Chapter 2 – Labelling, Expression and Purification for NMR experiments on the β₁-adrenergic receptor (β₁AR)

2.1 – Aim

The aim of the experiments detailed in the following section is to successfully express and purify thermostabilised β1-adrenergic receptor for the purpose of studying receptor dynamics using solution state NMR spectroscopy. Thermostabilisation makes the receptor suitable for spectral acquisition at higher temperatures which is necessary to counter adverse effects due to fast spin relaxation that are related to the slow tumbling experienced by large proteins. Baculoviral expression in insect cells has been successful in producing sufficient quantities of functional receptor for structural studies, for a number of different GPCRs, including the β2-adrenergic receptor. However uniform labelling in insect cells remains costly and expensive. In addition to these challenges, detergent solubilised membrane proteins experience very long rotational correlation times and fast relaxation which is particularly pronounced for the backbone positions. The extensive alpha helical nature of GPCRs mean a large degree of similarity in the chemical environment experienced by backbone amide positions, causing spectral crowding of the amide region. These factors contribute to the difficulties in backbone resonance experiments, making selective side chain labelling a more attractive approach to studying receptor dynamics. In this study ¹³C-methyl labelled methionine was used to this end. Preliminary NMR experiments (2D [¹H, ¹³C]-SOFAST-HMQC) were recorded on ¹³C-ε-methionine labelled Met2-β₁AR, in our lab by Dr. Duncan Crick (PhD thesis, Cambridge, 2014), which showed intense but overlapping peaks in the region of the methyl proton chemical shift region (1.0-2.2 ppm). The narrow range of chemical shifts experienced by methyl groups can cause spectral overlap and to avoid this five methionines were mutated to other residues. This is referred to as the Met2-Δ5 construct. The remaining methionines were hence labelled.

The construct used in this work is a variation of the β₁AR-β44-m23 construct that was used for receptor crystallisation (*69*, *96*, *234*). The creation of the Met2 construct saw the reversal of selected thermostabilising mutations: V90M[2.53], A227Y[5.58], L282A (ICL3). As a result, only the following thermostabilising point mutations remained: R68S[1.59], E130W[3.41], F327A[7.37] as well as mutations C116L and C358A for the removal of palmitoylation and for improved yield. Five methionines were mutated from the Met2 construct to alternative amino acids to reduce spectral overcrowding. These were: M44L, M48L, M179L, M281A, M338A, leaving the following methionines available for labelling: M1, M90[2.53], M153 (ICL2), M178[4.62], M223[5.54], M283[6.28], M296[6.41]. This generated the Met2-Δ5 construct, sometimes referred to simply as Δ5.
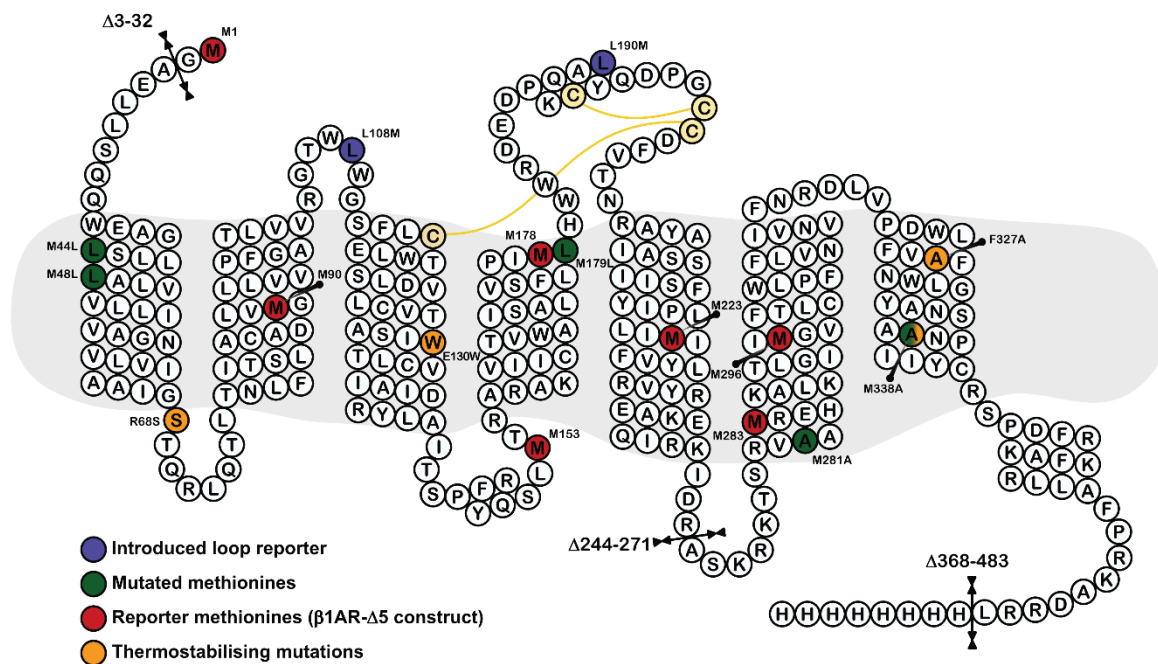
Figure 2.1 Snake diagram of the β₁AR Met2-Δ5 construct

Positions of truncations, mutations, reporters and thermostabilisation as well as the positions of native disulphide bridges are highlighted.

The six methionines (not including the N-terminal M1) retained are located in key positions of the receptor. These residues should be able to give information about local receptor dynamics and changes occurring in the presence of different ligands describing various receptor activation states. $M90^{2.53}$ in helix 2 is just beneath the ligand binding site. M153 is in the intracellular loop 2 (ICL2) near the N terminal end of helix 4. $M283^{6.28}$ is located towards the N terminal end of helix 6, which is observed to undergo a 14 Å translation upon G-protein binding. $M296^{6.41}$ is in the intramembrane part of helix 6 above $M283^{6.28}$, facing helix 5. Located opposite in helix 5 is $M223^{5.54}$, facing helix 6. The role of extracellular loops was investigated through the introduction of methionines at the L108 position in extracellular loop 1 (ECL1) and in the L190 position for ECL2. These are referenced as (Met2-)Δ5-L108M or Δ5-L190M.

To assign which peaks originate from which methionine, a range of constructs were generated with one residue mutated at a time. These are referred to as Met2-Δ6 mutants.

Another important consideration for the purification of membrane proteins is the choice of detergent used for solubilising the expressed protein. Lauryl Maltose Neopentyl Glycol (LMNG) is a novel non-ionic detergent, with a low critical micellar concentration (CMC) of 0.001% and a slow off rate. It has displayed superior stabilising properties for the β₂AR and other GPCRs compared to other amphiphiles (*49*, *235*). The low CMC reduces the amount of free detergent present in NMR samples, which reduces

the intensity of detergent associated peaks. The low off-binding rate is also preferential for NMR experiments as fast detergent dissociation has been shown to cause non-physiological conformational changes which are likely to result in peak broadening.

2.2 – Baculoviral expression

2.2.1 – Viral generation

Baculoviral stocks can be stably stored at 4°C for extended periods of time. Should stocks get infected or degrade, viral reamplification is possible from P1 or P2 stocks. P3 viral stocks were prepared for the following constructs: Met2, Met2-Δ5, Met2-Δ6 constructs: M153L, M178A, M178S, M223I, M283A, M296V, M296A, M90A, M90V, L190M, L108M, M153A. Alternative substitutions of alanine, valine, isoleucine or serine in place of methionine were generated, should a given substitution cause serious instability issues.

Sf9 insect cells were transfected both in suspension and in monolayer culture to generate a P0 viral stock. A separate GFP-only construct under the control of a late phase promoter was used as a positive control to monitor the rate of transfection. Overlaying a bright field image of the cultures with a GFP excitation image permitted the comparison of the ratio of GFP producing, i.e. cells that have taken up the bacmid DNA and started producing GFP, versus non-transfected cells. This is shown in **Figure 2.2.**

In monolayer cultures the proportion of transfected (green fluorescing) to non-transfected cells (dark appearing cells) was much lower (**Figure 2.2A**). Cells which have not produced the virus continued to multiply, producing a very dense carpet of cells. In suspension culture, the ratio was much higher, partly because there was a lower density of cells present. This was in spite the fact that the starting culture density for suspension cultures was 0.8×10^6 cell/mL, double that of the monolayer culture; 0.4×10^6 cells/mL. This implies that the culture has either stopped growing earlier or that a large proportion of cells have died, potentially due to viral lysis. Such cell lysis would release further virus particles, hence it is highly probable that the viral titre is much higher in the suspension culture than in the monolayer culture.

P0 viral stocks are used to infect further cells, to generate a P1 stock, which is in turn used to make P2. This amplification is continued to P3. The effect of media types, namely SF4 and SF900III (both supplemented with 1% foetal calf serum), were compared on this viral amplification process (**Figure 2.2B**). The ratio of GFP producing cells to non-fluorescent cells was noticeably higher in SF900III medium compared to SF4, making the former more desirable as a choice of medium for viral amplification.
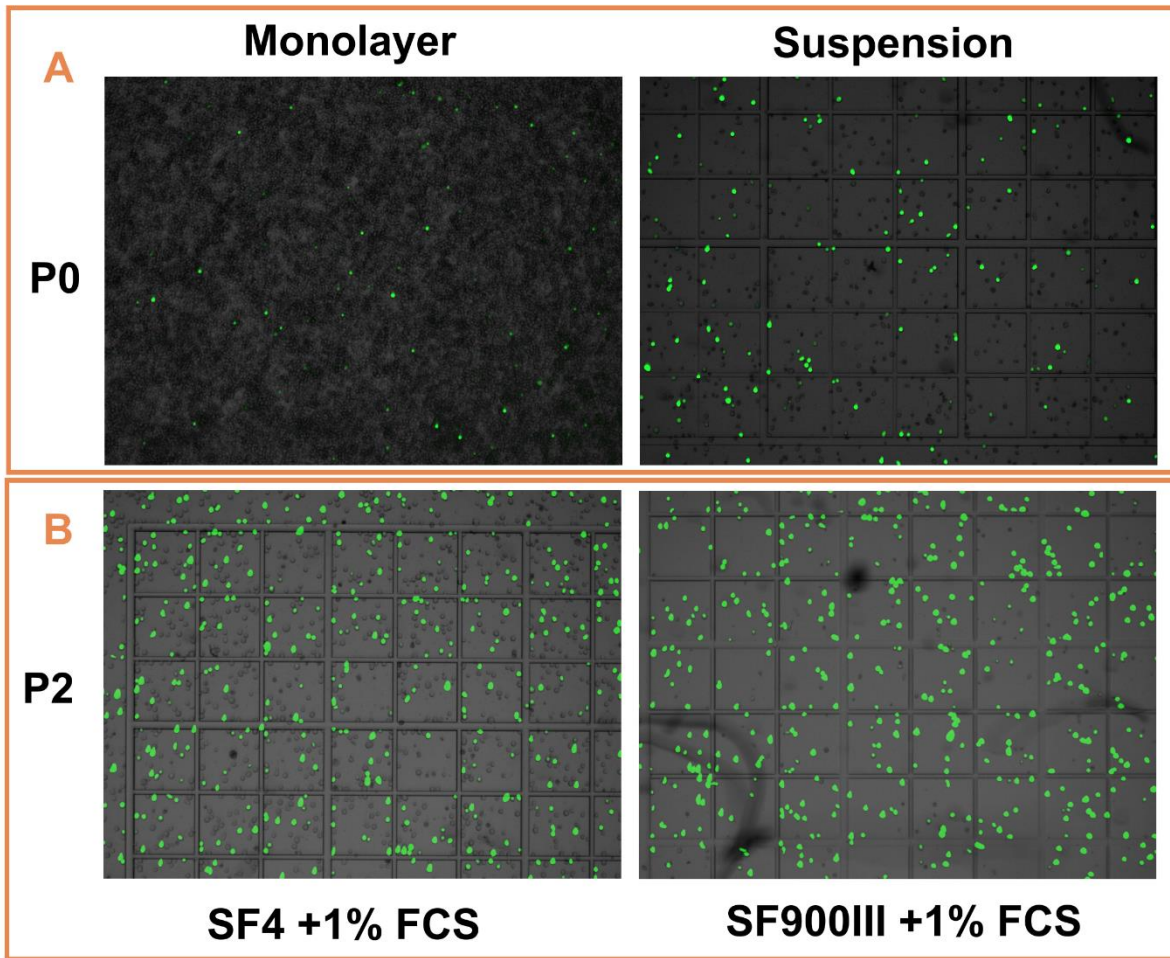
Figure 2.2 Sf9 transfection with GFP control plasmid

Sf9 cells were transfected with bacmid DNA, containing a GFP construct, as a positive control alongside other constructs. The bright field image was superimposed to a UV excitation image, which causes GFP fluorescence. A) Transfection cultures (P0) were seeded in suspension and in monolayers. Image taken five days after transfection. B) Viral stocks generated in suspension stocks were amplified to P1 and P2 amplification was compared in SF4 and SF900III media. Image taken 48 h after infection.

2.2.2 – Protein expression

Infection was carried out with 5mL of Sf9 cells at a dilution of 1×10^6 cells/mL being infected with 300µL of P3 virus stock (60 µL P3 virus/mL). **Figure 2.3** shows that over the course of 48 hours, all three viral stocks (M296A, M153L, M283A) had a very similar effect on the cells. Cell viability dropped from 95% to 60-70% and a marked 2-3 µm increase in cellular diameter was observed. These are hallmarks of baculoviral infection. It is not uncommon for cells to keep dividing for the initial 24 hours after the infection, therefore the persistent drop in viability suggests that the virus to cell ratio was relatively

high.
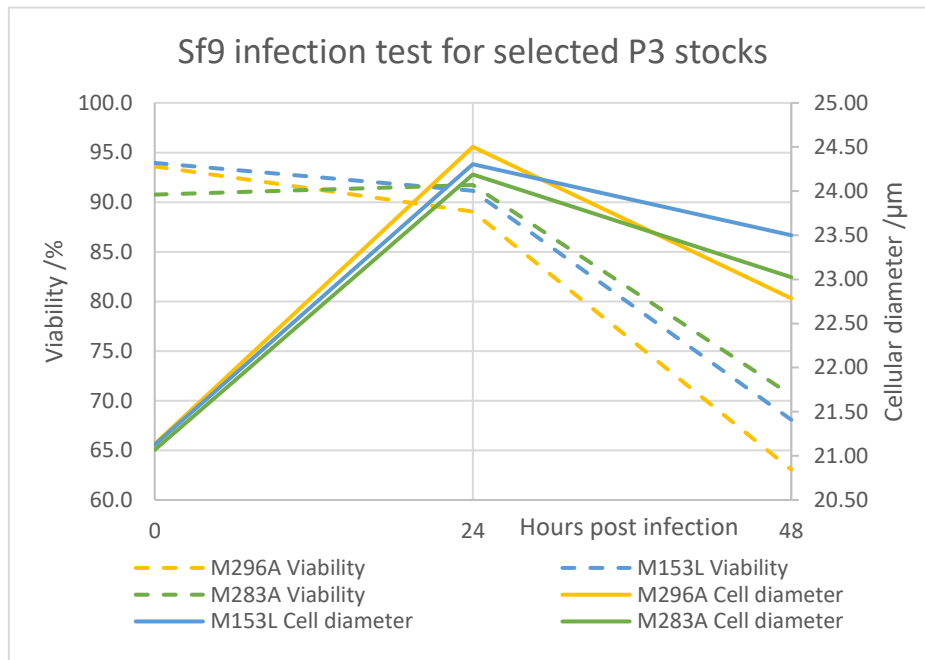
Figure 2.3. Sf9 cells infection by P3 baculoviral stocks

The effect of the infection on cell viability (dashed lines) and diameter (solid lines) is shown across a period of 48 hours.

2.3 – Protein purification

2.3.1 – Initial Purification Strategy

The initial purification strategy can be summarised in the following diagram.

Solubilisation
- **Cell disruption** in 20 mM Tris-HCl pH 8.0, 2 mM EDTA
- **Solubilisation buffer:** in 20 mM Tris-HCl pH 8.0, 3 mM Imidazole, 350 mM NaCl, 2% LMNG detergent

HisTrap
- **Buffer A:** 20 mM Tris-HCl pH 8, 10 mM NaCl, 3 mM Imidazole, 0.02% LMNG
- **Buffer B:** 20 mM Tris-HCl pH 8, 350 mM NaCl, 250 mM Imidazole, 0.002% LMNG
- **Gradient elution**

Gel filtration
- Superdex S200 (10/300) column
- **Buffer C:** 20 mM Tris-HCl pH 8.0, 10 mM NaCl, 1 mM EDTA, 0.001% LMNG

Figure 2.4 Initial purification strategy

Strategy based on a protocol by a previous lab member. Note the low NaCl concentration in the HisTrap Buffer A and the Gel filtration buffer.

The thawed Sf9 cell pellets were homogenised and centrifuged three times in cell disruption buffer, then the membrane fraction solubilised in 2% LMNG detergent. The resulting solution, was loaded onto a nickel affinity column and eluted with a linear gradient of imidazole. The desirable fractions were then clarified over a size exclusion column into a 10 mM NaCl, low salt buffer.

This initial approach was beset with two main issues that had to be overcome:

1. Elution off the HisTrap column did not achieve reproducible separation under a linear gradient.
2. Protein samples precipitated out of solution after gel filtration.

2.3.1.1 – Nickel affinity chromatography

Figure 2.5 shows representative UV 280 nm absorbance chromatograms of nickel affinity chromatography of M23-β1AR. The gradient length, that is the time it takes to go from 5% elution buffer B to 100%B in minutes, was varied and is shown above each curve. The flow rate in each case was 0.5 mL/min. In other words, a longer time implies a shallower gradient.
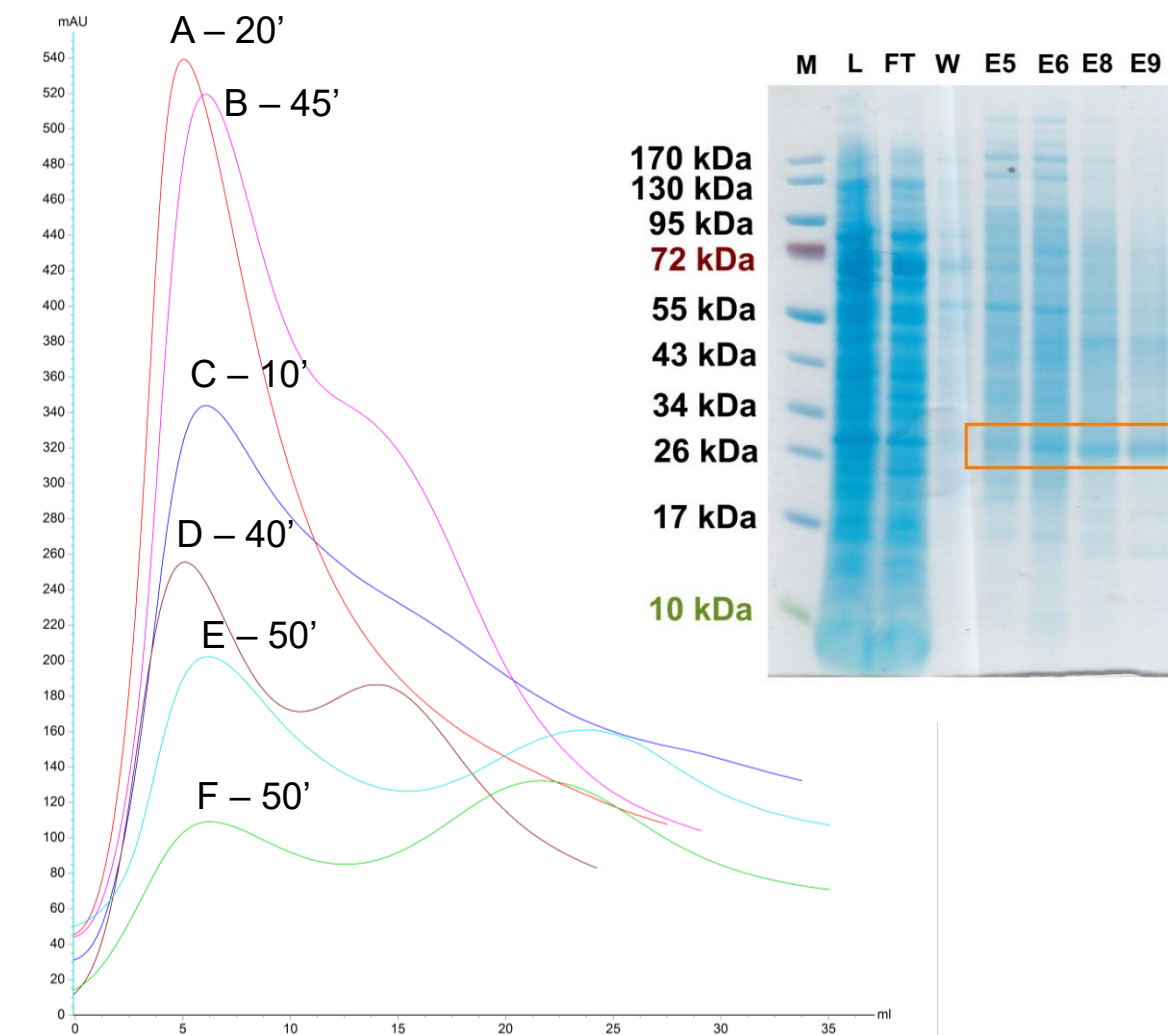
Figure 2.5 Linear gradient HisTrap elutions in 10 mM NaCl

280 nm absorbance chromatograms from Met2-β1AR HisTrap purifications with a 10-50 minute, 5-100% Buffer B linear gradient elution. Gel inset: SDS PAGE gel from curve D. Elution fractions E5, E6 correspond to the first peak, and E8, E9 to the second peak (both highlighted in orange).

The effect of a gradient of 50 minutes on separation, compared to shorter lengths (**Figure 2.5** traces A, C, D) is noticeable with the evolution of a second peak (traces E and F). However in no case has complete separation been achieved as the absorbance is not given time to reduce to baseline levels between the two peaks. This has seriously hindered the efficiency of chromatographic separation.

In addition to the poor separation, the presence of a 26 kDa band in both peaks (**Figure 2.5** SDS-PAGE gel, boxed in orange) would suggest that both peaks contain the target protein in sufficient quantities as to validate further purification of both peak fractions. However the inclusion of the first peak consistently resulted in aggregation, with the sample turning a yellow colour upon concentration for gel filtration.

Furthermore, by analysing the two peaks on an anti-his-tag western blot, a signal was only observed for the second peak. This observation held true for both Met2-β1AR, Met2-β1AR-Δ5 constructs purified by gradient elution. A representative example is shown in **Figure 2.6** below.

The chromatogram shown in **Figure 2.6C** is typical in profile and presents two poorly resolved peaks, with peak 1 represented by fractions 8-12 and peak 2 by 13-20. The SDS PAGE gel (**Figure 2.6A**) shows a near 26 kDa band present in all fractions, with a near 34 kDa band trailing from peak 1 to 2. The western blot (**Figure 2.6B**) only shows bands in fractions 14-20, corresponding to the second peak. This implies that the 34 kDa band is either a misfolded or a co-purifying contaminant and should not be included in further purification steps. This was often the case in early purifications where the second peak was absent or minimal.
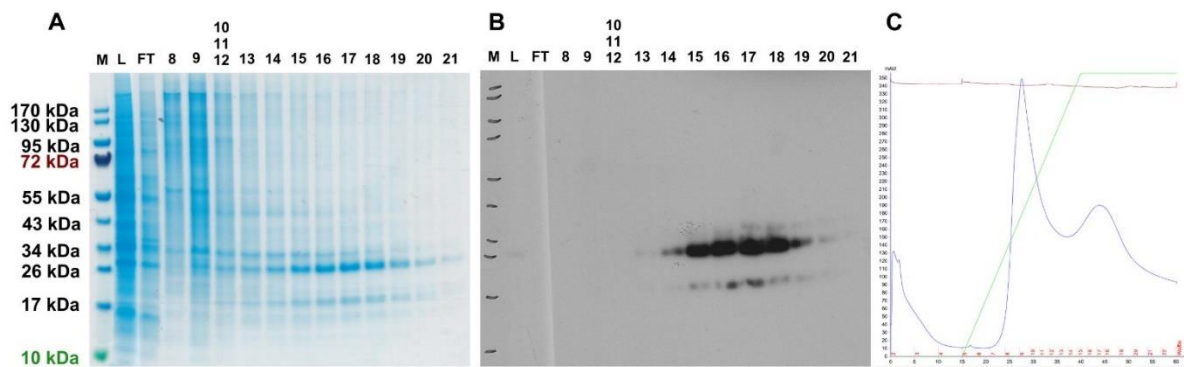
Figure 2.6 SDS-PAGE and western of gradient elution product of HisTrap chromatography.

Met2-β1AR-Δ5 purification by nickel affinity. **A)** SDS-PAGE gel. M: Marker, L: Load, FT: Flow through. Fractions 10-12 have been pooled and is represented in a single lane **B)** Western blot of the same gel as in A **C)** UV 280 nm chromatogram. Peak 1 corresponds to fractions 8-12, peak 2 to 13-20.

2.3.1.2 – Size exclusion chromatography – observations and salt dependency

The initial purification method made use of gel filtration to simultaneously remove any large sized aggregates and to exchange the protein into a low salt (10 mM NaCl) buffer. As mentioned above, this always resulted in protein precipitation from the solution. The major protein peak containing a 34 kDa construct (by SDS-PAGE, not shown) was consistently observed to elute off the column in a non-size typical fashion, very close to the salt fraction (**Figure 2.7**). This indicates that the protein is interacting with the column under these buffer conditions (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 1 mM EDTA, 0.001% LMNG).
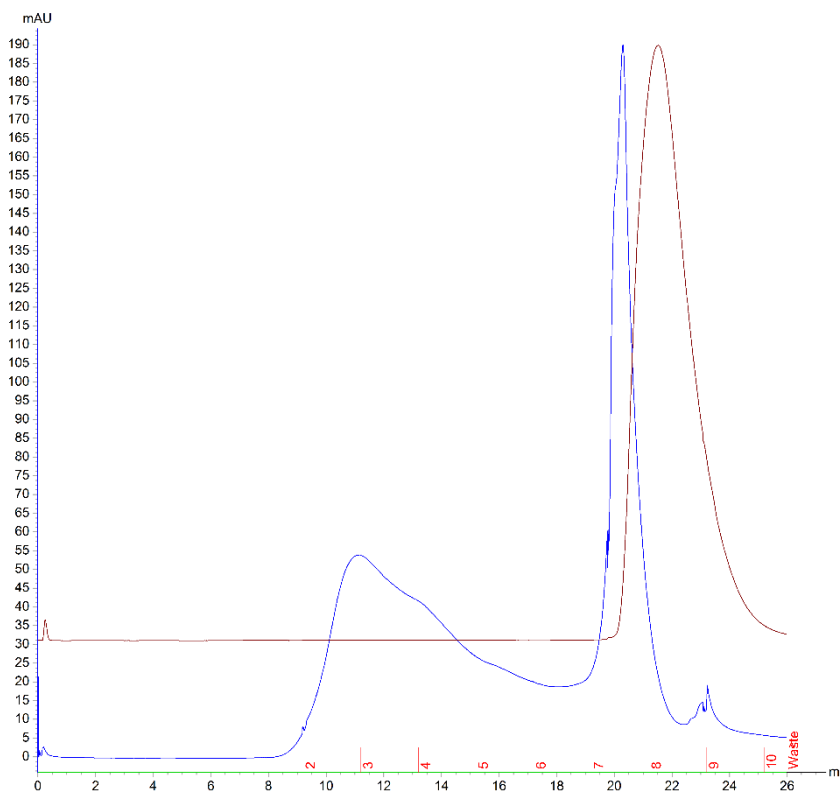
Figure 2.7 Gel filtration of HisTrap purified M23-β1AR.

Blue curve: UV 280 nm, Brown: conductivity

In an attempt to prevent protein precipitation, the gel filtration step was carried out with an increased salt concentration of 100 mM and 350 mM NaCl.
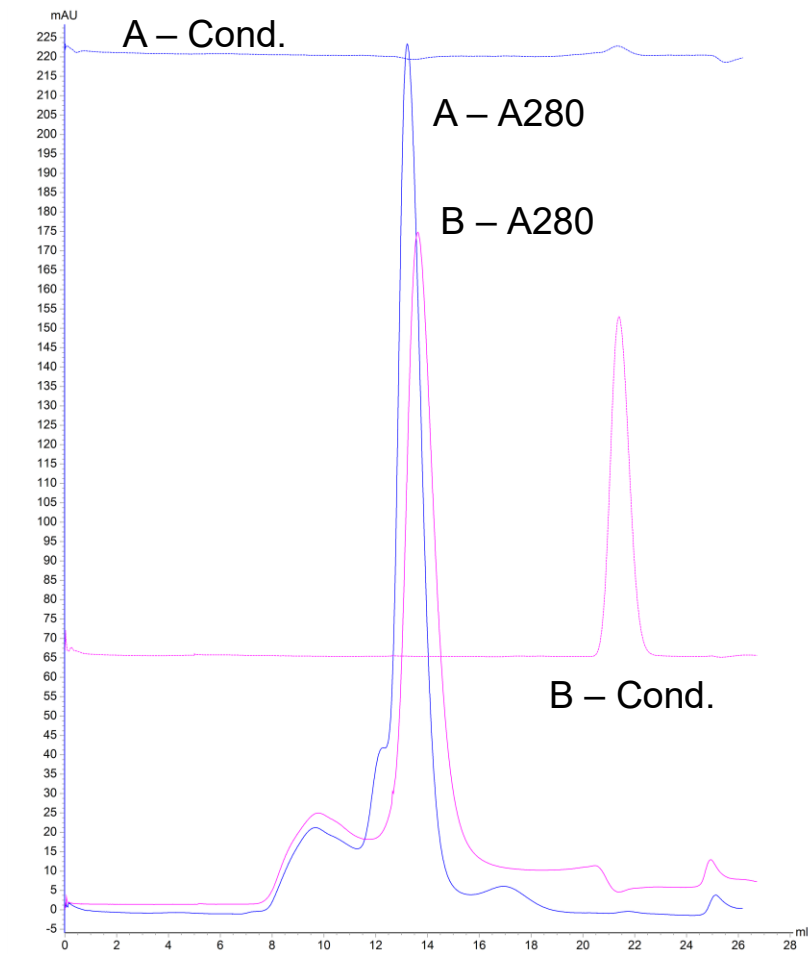
Figure 2.8 Size exclusion chromatography into increased NaCl (100 mM) buffer

Overlay of gel filtration chromatograms of M23-β1AR in 350 mM (A-blue) and 100 mM NaCl (B-pink) buffer (10 mM Tris-HCl pH8.0, NaCl, 0.001% LMNG). Conductivity traces are labelled with "Cond" and UV 280 nm traces with "A280".

Figure 2.8 shows UV absorbance ("A280") and conductivity ("Cond.") chromatograms for 350 mM NaCl (blue) and for 100 mM NaCl (pink) gel filtration. The protein sample loaded is present in 350 mM NaCl, carried forward from nickel affinity purification, which explains the level conductivity trace throughout 350 mM gel filtration and the conductivity spike at 23 mL (1 column volume) in 100 mM NaCl. It should also be noted that there is a shift from the elution volume of 20mL in 10 mM NaCl to 13.5mL under higher salt concentrations. Consequently the purified protein was stable and non-aggregating.

The structural significance of Na$^+$ ions in β1AR have recently been observed in high resolution (2.1 Å) crystal structures (*99*). Sodium ions have been found coordinated by residues Asp87$^{2.50}$, Ser128$^{3.39}$ and by three water molecules within a network of interhelical water molecules. The absence of Na$^+$ has been shown to decrease receptor stability, signified by a drop in receptor T_m of 7.5 °C. Equivalent residues coordinating sodium in β1AR have been found to form hydrogen bonds with each other in

the β2AR:Gα$_s$ structure, suggesting a loss, or rearrangement of Na$^+$ coordination. This proposes that Na$^+$ plays a role in stabilising the ligand-free state.

Salt concentration is worthy of discussion not only for its effect on protein stability, but also on NMR experiments. NMR spectrometers can be equipped with cryoprobes, where the receiver and pre-amplifier coils are cooled to cryogenic temperatures, thus reducing thermal noise in the measurement. Salts are critical components of buffers solutions both for maintaining pH and for stabilising proteins in a soluble state. However the increased dielectric conductivity caused by the presence of salts results in the attenuation of radio frequency power and the appearance of ring currents in the sample. This can lead to sample heating and increased pulse lengths. These effects abate the increased sensitivity offered by cryoprobes. The weight of this effect also increases with the strength of the magnetic field (*236, 237*). Therefore it is desirable to find a concentration of salt where a favourable compromise exists between protein stability and NMR sensitivity.

Met2-β1AR and Met2-β1AR-Δ5 constructs have shown stability in 100 mM salt, however the Met2-β1AR-Δ5-M90A construct precipitated after gel filtration in this condition. To find a suitable salt concentration, dynamic light scattering was used to detect aggregation in different concentrations of salt. The purified construct was dialysed from 350 mM salt to a range of salt concentrations and the ratio of various particle sizes was quantified and recorded as a portion of the total (**Figure 2.9**).
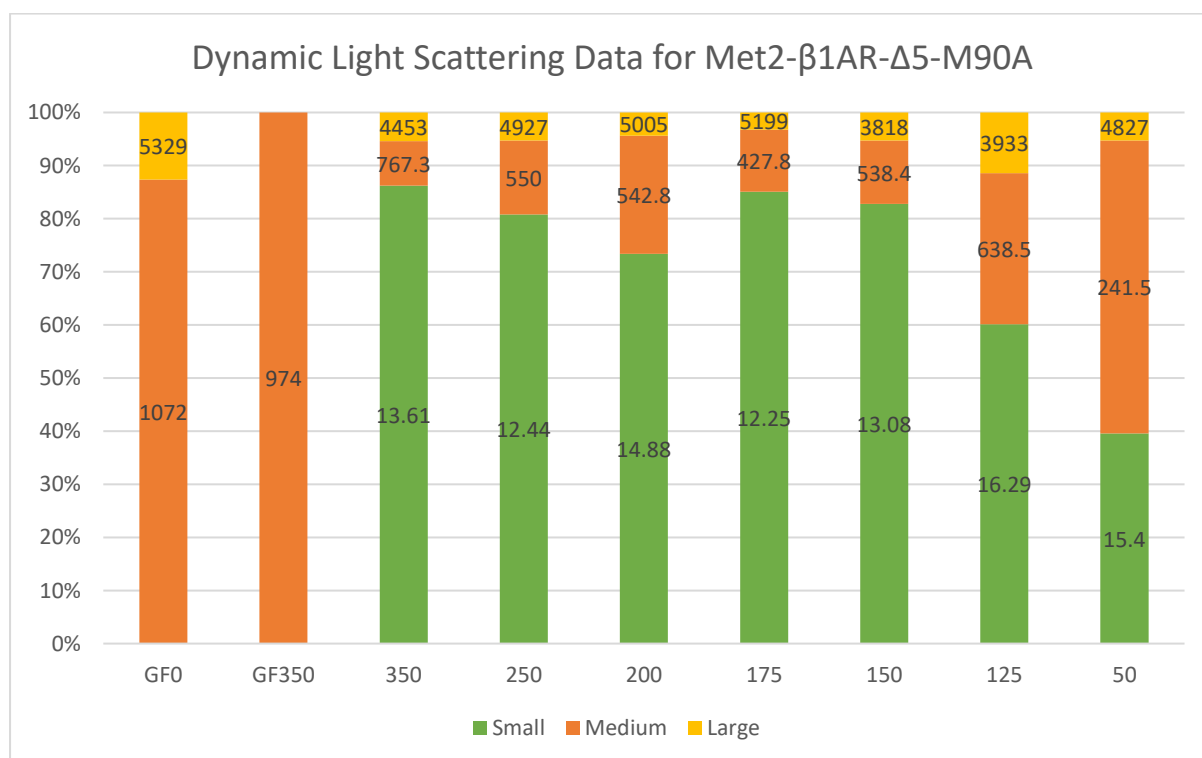
Figure 2.9. Dynamic light scattering to determine optimal NaCl concentration

Particles were classed as "small", "medium" or "large" on the basis of average particle diameter (nm) and the variation in ratios was monitored in different salt concentrations. Blank buffer of 0 mM and 350 mM NaCl was also measured (GF0 and GF350). The absence of small particles in buffer alone suggests that these correspond to receptor micelles. An increase in the 200 nm-800 nm particle ratio (medium particles) was observed under 150 mM NaCl, which correlates with the aggregative behaviour of the construct in 100mM salt. Hence 150 mM NaCl was trialled for β1AR-Δ5-M90A. No precipitation was observed. 90° pulses were in the range of 11.7 µs to 12.3 µs at -11dB.

Purification of the β1AR-Δ5-M90A construct in 350 mM NaCl buffer achieved marginal increase in purity at the cost of time invested in the gel filtration process, coupled with losses associated with preceding protein concentration and the chromatography step itself. This can be illustrated by comparing the purity of the load fraction to the eluted protein on an SDS PAGE gel as shown below in **Figure 2.10**. The associated UV 280 nm absorbance trace was representative of a soluble non-aggregating trace (see **Figure 2.8** above).
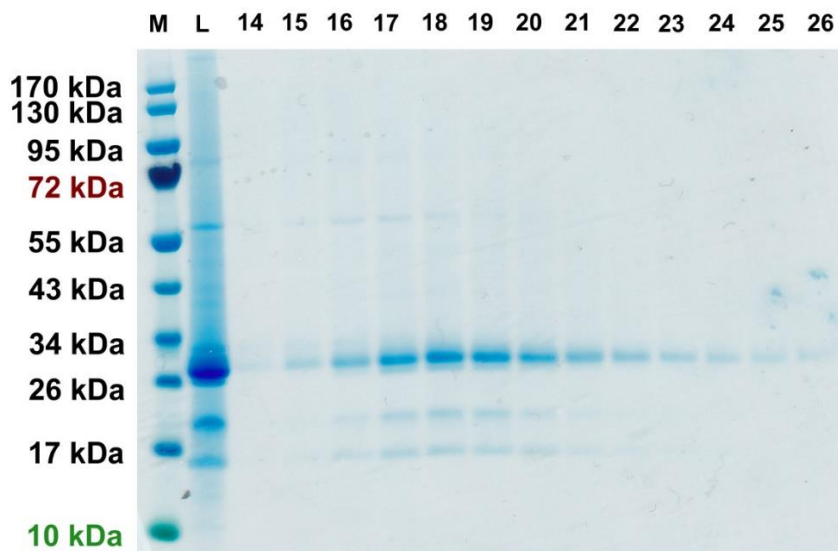
Figure 2.10. SDS PAGE of gel filtration fractions of Met2-β1AR-Δ5-M90A in 350 mM NaCl buffer.

M: Marker, L: Load. Fraction numbers refer to collected elution fractions from a single symmetrical absorbance peak.

This therefore warrants the elimination of the gel filtration step from the purification process as a means of increasing purity. To achieve higher purity, an alternative, optimised nickel affinity chromatography was used.

2.3.2 – Refined purification strategy

Although the issue of protein stability in relation to salt concentrations have been explored resulting in stable protein preparations, the chromatographic separation of receptor from other impurities remained a challenge using the linear gradient conditions described above. This issue could have been resolved in one of two ways. One method would have been to extend the length of the gradient in time, where the elution conditions change slowly allowing more time for different species to separate. An inadvertent side effect of shallower gradients is peak broadening, in other words the separated proteins are diluted considerably. Protein concentration can take a considerable time and therefore it is preferable to achieve sharp elution peaks.

Another method of improving resolution, which can also achieve sharp elution, is through stepwise gradients. The basic concept of chromatography is to separate individual mixture components through the induction of conditions which affect these components differentially. Increasing the imidazole concentration during elution rapidly, to a precise predefined level will likely achieve such differential conditions. Moreover, maintaining that condition until all proteins dissociated by it from the column have eluted, will avoid any unwanted mixing with species eluting at higher imidazole concentrations. This cannot be achieved by a linear gradient which changes continuously and too rapidly.

The key to the successful implementation of a stepwise elution profile is the understanding of how much imidazole is needed to elute each mixture component. To this end an initial regime of consecutive 5, 10, 15, 20, 25, 50, 75 and 100% imidazole buffer steps were trialled (data not shown). This has shown that at 20% and above, receptor elution is observed (also observed on **Figure 2.11B**, 25% corresponding to 52 mM imidazole). It is also important to allow the UV 280 nm trace to return to baseline after each elution step, so that all proteins dissociated under that step elute prior to the elution of tighter binding species in the next step.

Elution from the affinity column in the up-flow direction, together with a 40% reduction in flow rate compared to loading flow rates can also achieve increased chromatofocusing, resulting in a sharper elution peak, hence faster elution, lower buffer use and more concentrated protein.

This HisTrap strategy was hence followed with on column desalting of the sample into a lower salt buffer, suitable for NMR experiments.

The construct Met2-β1AR was purified with the above strategy. The HisTrap elution profile and the SDS PAGE gel is shown below (**Figure 2.11**).
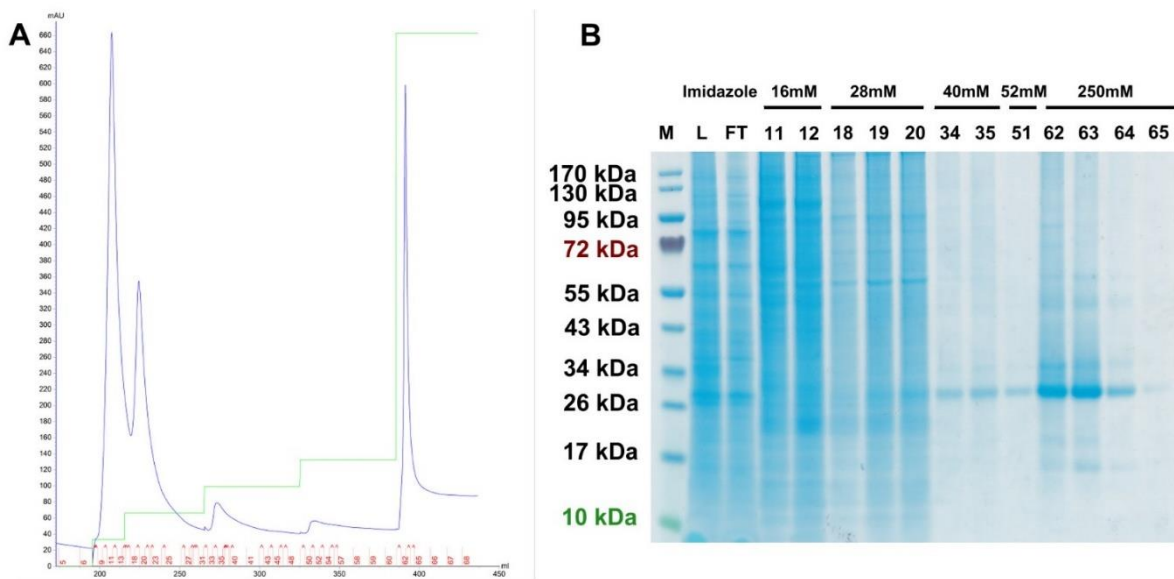
Figure 2.11 HisTrap purification of Met2 with a step gradient elution

HisTrap purification of Met2-β1AR. (A) UV 280 nm absorbance spectrum shows that impurities are separated by a 16 mM, 28 mM, 40 mM and 52 mM Imidazole wash. The peak of interest is eluted with 250 mM imidazole and has an absorbance maximum of 610 mAU. (B) SDS PAGE gel of the elution fractions. Fractions 62-64 show highly concentrated protein with minimal impurities.

The isocratic elution allowed the UV trace to return to baseline, before the next step was initiated, hence maximising separation. The column upflow elution with a reduced flow rate (0.5 mL/min) compared to loading flow rates (1 mL/min) achieved a narrow elution peak. This was maximised by a pump fill prior to step initiation; this is where the required proportion of Buffer A and Buffer B is mixed and applied onto the pumps prior to injection onto the column. Fractions from a 40 mM and 52 mM imidazole wash also produced a highly pure preparation of the receptor (near 26kDa band).
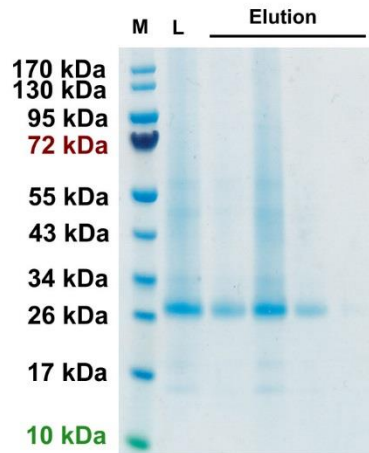
Figure 2.12 Receptor buffer exchange over a HiPrep Desalting column.

L: load fraction

Buffer exchange was achieved by loading the HisTrap elution fractions onto a 150 mM NaCl buffer equilibrated HiPrep Desalting column. **Figure 2.12** shows that purity remains unaffected.

The final purification strategy is illustrated below.

Solubilisation
- **Cell disruption** in 20 mM Tris-HCl pH 8.0, 2 mM EDTA
- **Solubilisation:** 1 hour on ice in 20 mM Tris-HCl pH 8.0, 3 mM Imidazole, 350 mM NaCl, 2% LMNG detergent

HisTrap
- **Buffer A:** 20 mM Tris-HCl pH 8, 350 mM NaCl, 3 mM Imidazole, 0.02% LMNG
- **Buffer B:** 20 mM Tris-HCl pH 8, 350 mM NaCl, 250 mM Imidazole, 0.002% LMNG
- **Stepwise elution:** 5%, 10%, 20%, 100% Buffer B

Desalting
- HiPrep Desalting Column
- **Buffer GF150:** 10 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.001% LMNG

Figure 2.13 Optimised β_1AR purification strategy.

2.4 – Protein quality

2.4.1 – Mass spectrometry

To confirm the identity of the purified protein, mass spectrometric analysis was used to assess the major 26 kDa band from an SDS PAGE resolved M23-β1AR construct by two alternative methods. The trypsin digested protein was analysed by:

1. LC-MS, MS-MS analysis by Dr Mike Deery at the Cambridge Centre for Proteomics. The submitted band was trypsin digested and resolved by liquid chromatography. The resulting peptides were analysed for amino acid sequence by tandem MS.

2. Peptide identification by MALDI-Fingerprinting, by Dr Len Packman at the PNAC facility. Here the resulting peptides are matched to expected protein sequences or candidates from the NCBI database.

Both methods have identified expected fragments from the M23-β1AR protein sequence, and the results are summarised in the following table:

Table 2.1 Summary of mass-spectrometry results.

	LC-MS MS/MS	MALDI-Fingerprinting
Coverage	16.29%	24%
Protein Score (-10*Log(P))	N/A	98
_p_<0.05 significance level	N/A	81

Protein scores are given to peptide matches against a database, where a calculated probability threshold value is given to a match which occurs at random. Any scores above this are considered significant at that significance threshold. For MS/MS data, protein scores are calculated from ion scores to give a non-probabilistic tool for ranking the fragment hits.

The following diagram shows a comparison of the peptide fragments identified by the two different techniques. Yellow residues were identified by MS/MS only, blue residues by MALDI Fingerprinting only and green residues by both methods.

β44-m23 RECEPTOR CONSTRUCT

Detected by MALDI fingerprinting only

Detected by MS-MS only

Detected by both either method
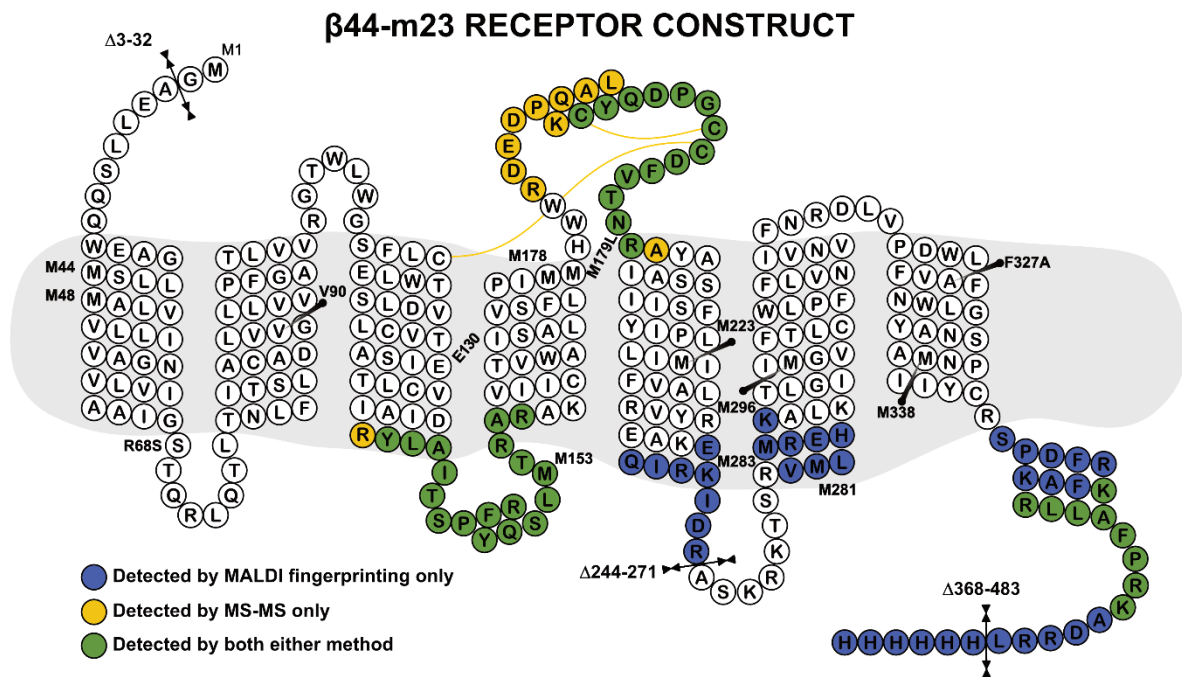
Figure 2.14 Snake diagram of the β44-m23-β1AR construct for MS

This construct was sent for mass spectrometric analysis.Blue residues were detected by MALDI fingerprinting only, yellow residues by tandem MS only. Green residues covered by both methods. Yellow lines link cystine pairs.

Peptide fingerprinting has identified two additional fragments compared to MS/MS, and the C-terminal end has also been successfully ionised (**Figure 2.13**). This establishes, that albeit with <25% coverage, both methods are suitable for identification of the purified receptor.

2.4.2 – Ligand Affinity Chromatography

Sepharose and agarose conjugated to the non-selective β-adrenergic antagonist alprenolol has been demonstrated as a tool for affinity purification of the β1 and β2 receptors (*238*).

In this project, this tool has been used both preparatively and analytically. In a preparative setup, Met2-β1AR from a preceding HisTrap purification was loaded onto a packed 5 mL Alprenolol-Sepharose column. Bound protein was eluted with 0.8 mM alprenolol.

The column has demonstrated a very high purification factor, with all impurities eluting in the sample flow-through (**Figure 2.14**). The resulting population of purified protein has been selected on the basis of its ligand binding ability and is hence expected to be highly functional. The drawback of this method is that it generates alprenolol bound samples. This means that a dialysis step must be included to exchange the receptor to a ligand free state, ready for the addition of alternative ligands. Alprenolol has an absorbance spectrum that ranges from 225-290 nm with two major peaks at 230 nm

and at 280 nm. Hence the purification process cannot be monitored by absorbance at 280 nm. Instead the sample must be visualised on by SDS PAGE, or at alternative wavelength, such as 220 nm – indicative of peptide bonds.
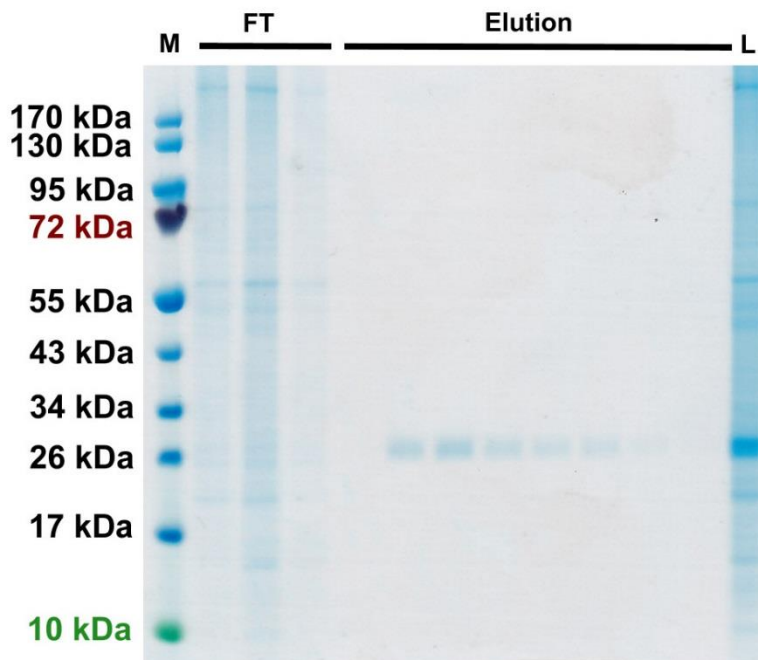
Figure 2.15 Ligand affinity chromatography SDS-PAGE

SDS PAGE of alprenolol sepharose column purification of M23-β1AR. Bound protein is eluted with 0.8mM alprenolol. The step results in highly pure, ligand bound receptor. M: Marker, FT: Flow through, L: Load

Under the assumption that binding to the ligand conjugated resin is a sign of functionality, alprenolol resin packed into spin columns can also be used analytically. This application is demonstrated below (**Figure 2.15**) for Met2-β1AR purified by nickel affinity chromatography. 500 µL of the selected 2 mL elution fraction was loaded onto 500 µL of alprenolol sepharose equilibrated with binding buffer. The sample was incubated in the capped spin column for 1 hour with mixing. Unbound sample was washed out with binding buffer under brief centrifugation. Bound protein was eluted with 0.8 mM alprenolol.
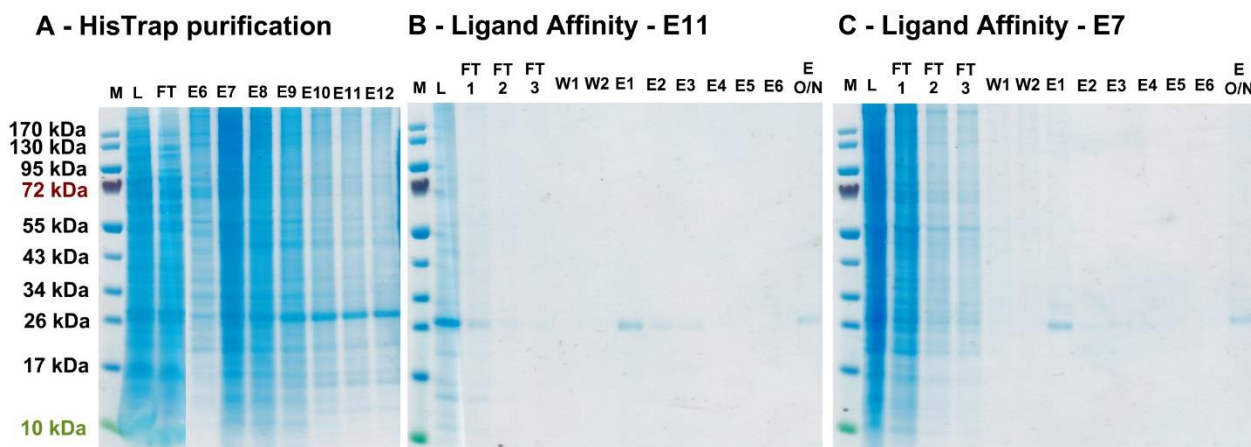
Figure 2.16. HisTrap followed by Ligand Affinity

A) HisTrap purification of Met2-β1AR. Bound protein eluted as a single peak with 175 mM Imidazole. B) Alprenolol spin column application of elution fraction E11 from HisTrap purification (panel A) C) Alprenolol spin column application of elution fraction E7 from HisTrap purification (panel A). M: Marker, L: Load fraction, FT: Flow through, W: Wash fractions E: Elution fraction, O/N: Elution after overnight incubation with 0.8 mM alprenolol.

Figure 2.15 A and B confirm the presence of ligand binding receptor in both highly impure (panel A fraction 7) and relatively pure (panel A fraction E11) receptor fractions. Bound receptor was incubated with alprenolol for 15 minutes before elution. Six rounds of elution-centrifugation were applied, with a further overnight incubation with alprenolol. The absence of a protein band by elution fraction 6, and a reappearance thereof after an overnight incubation (panel B and C, "E O/N") indicates that a period of incubation with free ligand before elution is desirable for maximum yield. These results confirm the use of the alprenolol resin in a spin column method as a quick tool for confirming ligand binding in purified samples. Furthermore, the samples applied are recoverable, albeit in a ligand bound state.

Assignment of Met2-Δ6 constructs

In order to assign each methionine to the observed peaks in the chemical shift spectrum the following Δ6 mutants were expressed, purified and recorded: (Δ5)-M90A, M153A, M178A, M223L, M283A, M296A, L108M, L190M. The resulting spectra were compared to the Δ5 equivalent, where the assignment residue (which was mutated in Δ6 and therefore its associated peak absent) was present. This comparison was repeated in the presence of ligand alone, with Nb6b9 or Nb80 alone or in a ternary complex with either nanobody, until every residue could be unambiguously assigned in all receptor states. This translated into a large number of possible assignment scenarios as represented below (**Figure 2.17**).
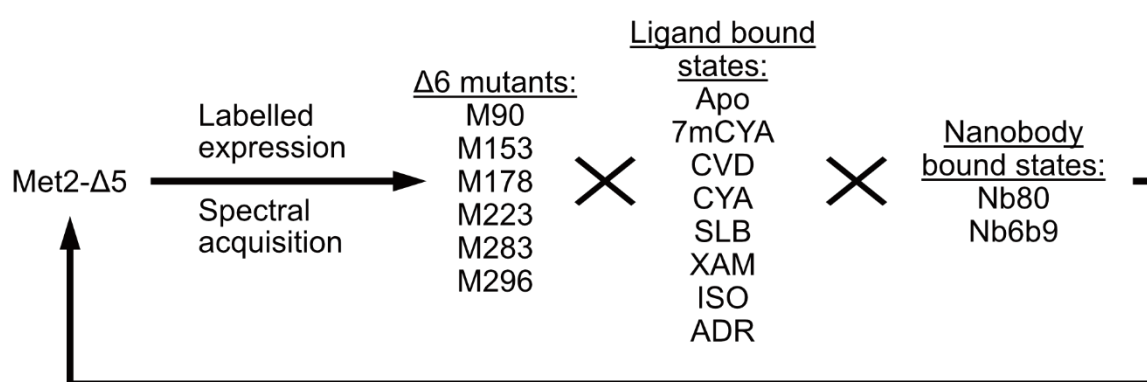
Figure 2.17 Assignment possibilities: constructs, ligands and nanobodies

The 6 methionines that were mutated for assignment were to be recorded bound to each ligand, and to nanobody. This created the possibility of 6 x 8 x 2, or 96 potential assignment states. This was reduced as much as possible depending on spectral similarities.

Two main challenges made the assignment work more difficult. One was the spatial proximity of residues M223 and M296. The other was the spectral overlap of some residues, such as M283 and M1 (**Figure 2.19 e, g, n**). The issue of spatial proximity is illustrated below in **Figure 2.18**. The methyl groups of M223 and M296 side chains in 4BVN are modelled at 6.3 Å away, close enough for changes in one residue to exert an effect on the chemical shift environment of the adjacent residue. This was also true for when either methionine was mutated for assignment, resulting in a change in the overall spectrum that was not solely due to the absence of the mutated assignment residue. This is illustrated in **Figure 2.18 b, c and d**, where the peaks indicated by an arrow assigned to M223 have moved chemical shift positions, in addition to the disappearance of the peak assigned to M296.
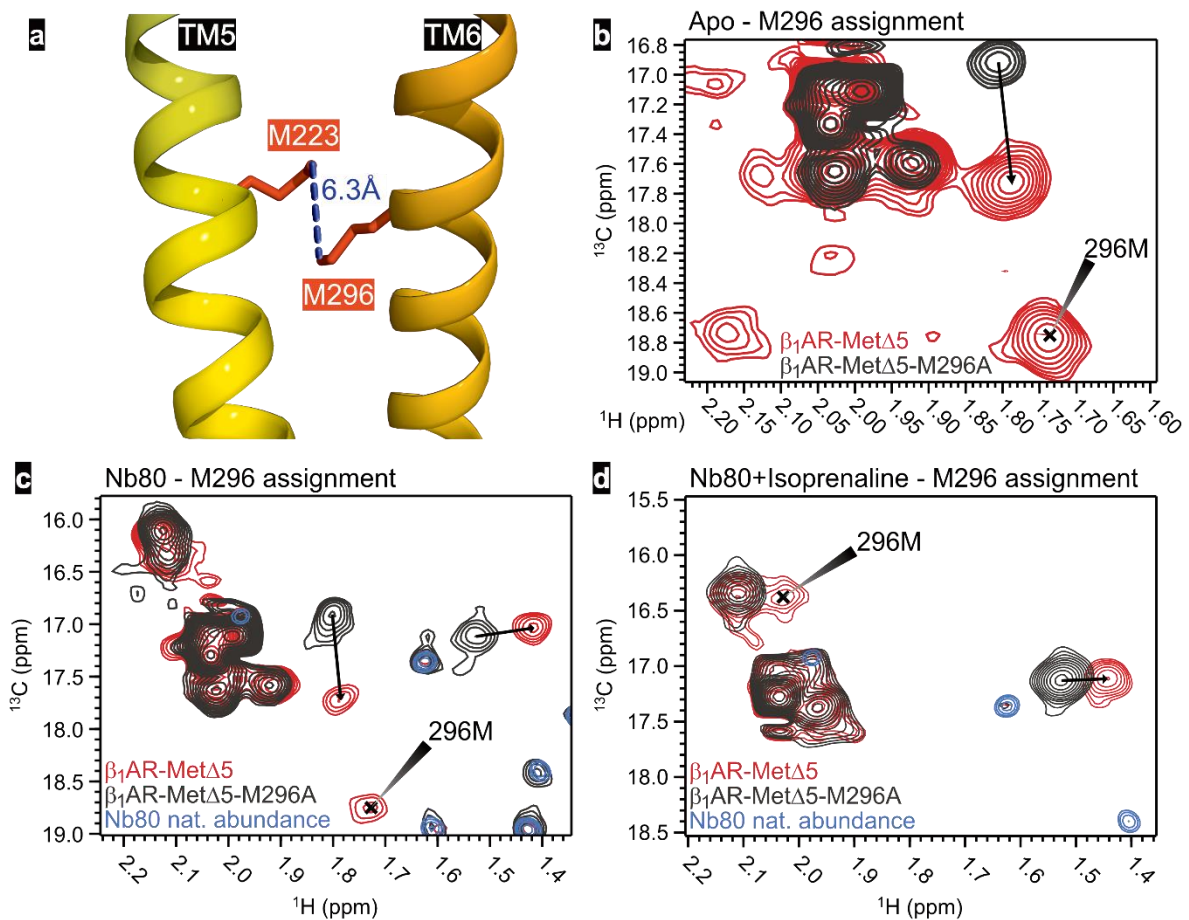
Figure 2.18 Spatial proximity of M223 and M296 and its effect on assignment

(a) shows the side chains of the two residues in the structure of 4BVN. (b) assignment of Δ5-M296A in the apo state. (c) assignment of Δ5-M296A in the Nb80 bound state. (d) assignment of Δ5-M296A in the Nb80 and isoprenaline bound ternary complex state. Figure adapted with permission from Solt et al. (*127*).

The full assignment panel is shown below in **Figure 2.19 a-p**.

Figure 2.19 Methionine assignment spectra.

(Shown overleaf across 2 pages.)

The reference (Δ5) spectra is shown in red, and the mutated construct in black except where otherwise indicated (such as g, h). Depicted this way, the assignment residue shows as a red peak without any black overlap. Inset (e) shows a cross section of the spectrum in the indirect dimension along the dashed line projected as a 1D spectrum. Blue peaks indicate natural abundance peaks from a large excess of unlabelled nanobody (k-p). (p) inset shows M296 in the Nb80 bound state from a different experiment of the same state, where the receptor has been fully saturated by Nb80 as judged by the presence of M153c and no M153a peak. This required a 30-fold molar excess of Nb80, resulting in strong natural abundance peaks, hence (p) shows a non-saturated state for Nb80 for clarity.

i **Cyanopindolol - M90 assignment**

90M

β1AR-MetΔ5
β1AR-MetΔ5-M90A

j **Salbutamol - M90 assignment**

90M

β1AR-MetΔ5
β1AR-MetΔ5-M90A

k **Nb80 - M153 assignment**

153M

β1AR-MetΔ5
β1AR-MetΔ5-M153A
Nb80 nat. abundance

l **Nb80 - M296 assignment**

296M

β1AR-MetΔ5
β1AR-MetΔ5-M296A
Nb80 nat. abundance

m **Nb80 - M223 assignment**

223M

β1AR-MetΔ5
β1AR-MetΔ5-M223L
Nb80 nat. abundance

n **Nb80+Isoprenaline - M296 assignment**

296M

β1AR-MetΔ5
β1AR-MetΔ5-M296A
Nb80 nat. abundance

o **Nb80+Isoprenaline - M178 assignment**

178M

β1AR-MetΔ5
β1AR-MetΔ5-M178A
Nb80 nat. abund.

p **Nb80 and Nb6b9 bound - comparison**

153Ma
153Mb
153Mc
190M
223Mb
296Mb
1M
283/1M
223Ma
296Mb
178M
296Ma

β1AR-MetΔ5
β1AR-MetΔ5-L190M

Chapter 3 – Molecular signatures of β_1AR activation

Declaration

This work and the results presented in this chapter have been published (*127*). The raw data was the result of my own work, and the analysis was carried out in collaboration with a number of the co-authors named. Their contributions are specifically highlighted throughout this chapter where relevant.

3.1 – Aims

GPCR crystallography has provided significant insight into the activation mechanism of these receptors. G-protein bound crystal and cryo-EM structures as well as arrestin bound and antibody bound fully active structures (*54*, *74*, *80*, *139*, *239*, *240*) have shown that receptors require the presence of an intracellular binding partner (IBP) to achieve a fully active state. Although these structural biology techniques have made significant progress in terms of our understanding of GPCR activation, they are limited in their ability to report on the dynamic transitions between these structural snapshots. NMR spectroscopy has proven itself as a valuable tool in the arsenal for the study of the dynamic processes and structural ensembles that form part of the complex conformational landscape of GPCR activation (*120*, *122*, *154*, *241–244*). GPCRs have been demonstrated to be highly dynamic molecular entities, exhibiting a great degree of molecular plasticity.

The aim of this work is to evaluate the molecular signatures that accompany receptor activation by small molecule ligands of varying efficacy, as well as by G-proteins using the G-protein mimic nanobodies Nb6b9 and Nb80, which have been shown to elicit the same structural changes as G_s-binding. In the absence of an active state crystal structure for the β_1-adrenergic receptor, the aim was to gain an understanding of the dynamic conformational changes that accompany receptor activation, by using ^{13}C-methyl-methionine labelling and methyl detected ^{13}C HMQC NMR experiments.

3.2 –β$_1$AR-receptor construct used

The receptor construct used in this work is a variation of the β$_1$AR-β44-m23 construct that was used for receptor crystallisation (*69*, *96*, *234*). The construct used, called Met2 and variations thereof, saw the reversal of selected thermostabilising mutations: V90M$^{2.53}$, A227Y$^{5.58}$, L282A (ICL3). As a result, only the following thermostabilising point mutations remained: R68S$^{1.59}$, E130W$^{3.41}$, F327A$^{7.37}$ as well as mutations C116L and C358A for the removal of palmitoylation and for improved yield. This generated the Met2 construct, from which five methionines were mutated to alternative amino acids to reduce spectral overcrowding. These were: M44L, M48L, M179L, M281A, M338A, leaving the following methionines available for labelling: M1, M90$^{2.53}$, M153 (ICL2), M178$^{4.62}$, M223$^{5.54}$, M283$^{6.28}$, M296$^{6.41}$. This generated the Met2-Δ5 construct, sometimes referred to simply as Δ5. The role of extracellular loops was investigated through the introduction of methionines at the L108 position in extracellular loop 1 (ECL1) and in the L190 position for ECL2. These are referenced as (Met2-)Δ5-L108M or Δ5-L190M. The spatial and sequential distribution of the labelled methionines as well as the position of other loop truncations retained from the m23 crystallisation construct are shown below on the snake diagram (**Figure 3.1a**) and on the crystal structure (PDB: 4BVN, **Figure 3.1b**).
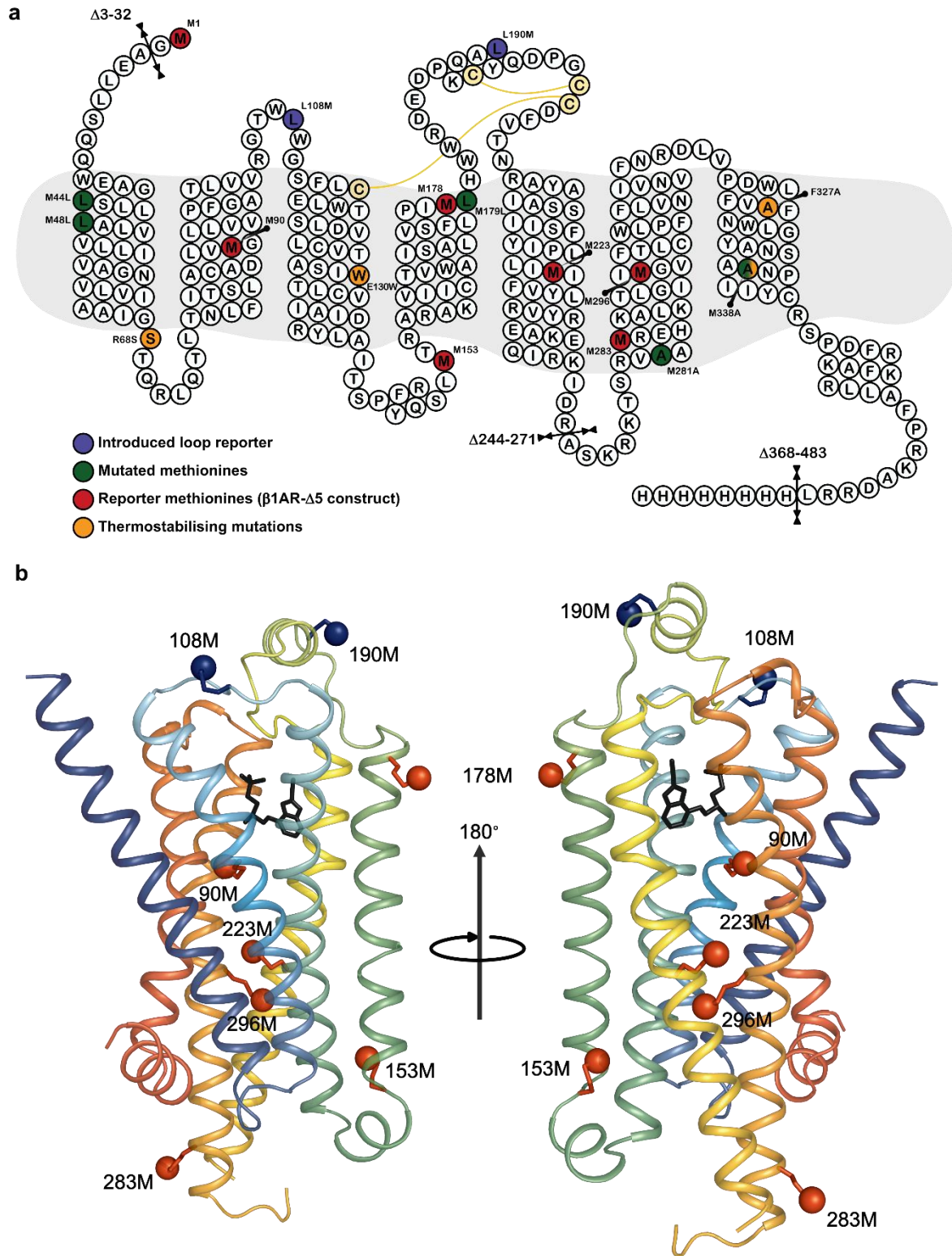
Figure 3.1 Snake diagram of β1AR-Met2-Δ5

This shows the distribution of labelled methionines, truncations, thermostabilisation mutations in the Met2-Δ5 construct. The crystal structure of the closest construct (β44-m23, PDB: 4BVN) highlights the spatial distribution of labelled methionines in red spheres. Figure adapted with permission from Solt et al. (*127*).

3.3 – NMR Timescales

In order to understand the significance of the chemical shift changes that take place as a result of ligand or G-protein mimetic nanobody addition (hereafter simply nanobodies) it is worth noting how multiple resonances in equilibrium manifest themselves in an NMR spectrum. NMR chemical shift denotes the difference between the frequency of the specific nucleus and a reference frequency divided by the reference frequency. This is given in values of ppm. When two molecular states exist in equilibrium, the magnitude of the exchange constant (k) for the equilibrium in relation to the chemical shift difference between the two states (ω given in Hz) defines the way the two states appear in a chemical shift spectrum. In other words, the comparison of the exchange rate and the chemical shift difference of the states involved is known as the NMR timescale. In practical terms, if k is much lower than ω, then it means that the exchange process is slow enough for each state to be resolved in a spectrum, with the intensity of the peaks proportional to the population of each. This is then referred to as slow exchange. At the other extreme, where k is much greater than ω, the amount of time spent in each chemical shift environment is too short to resolve both peaks, and a single peak appears at the weighted average of the two populations. As the distribution of the two populations changes relative to one another, in other words the amount of time spent in each environment (a shift in the equilibrium), while $k \gg \omega$ remains true, the chemical shift of the averaged resonance changes to reflect the new average. This shift in the equilibrium can occur due to changes in temperature, presence of ligand, or other factors which govern the equilibrium. Lying in between these two extreme scenarios, where k approaches ω, so called intermediate exchange occurs. This is represented by peak broadening and a loss of peak intensity.

In the following discussion and description, the equilibria that are identified are hence described as slow, fast or intermediate in terms of this NMR timescale.

3.4 – Molecular signatures of ligand bound β_1AR

^{1}H, ^{13}C HMQC spectra of labelled receptor samples prepared in LMNG detergent micelles were recorded in the presence of small molecule ligands of varying efficacy. The receptor was recorded in its ligand free (apo) state, and bound to the very weak partial agonist 7-methylcyanopindolol, the weak partial agonists carvedilol and cyanopindolol, the partial agonists salbutamol and xamoterol, the full agonist isoprenaline and natural full agonist adrenaline (**Figure 3.2**).
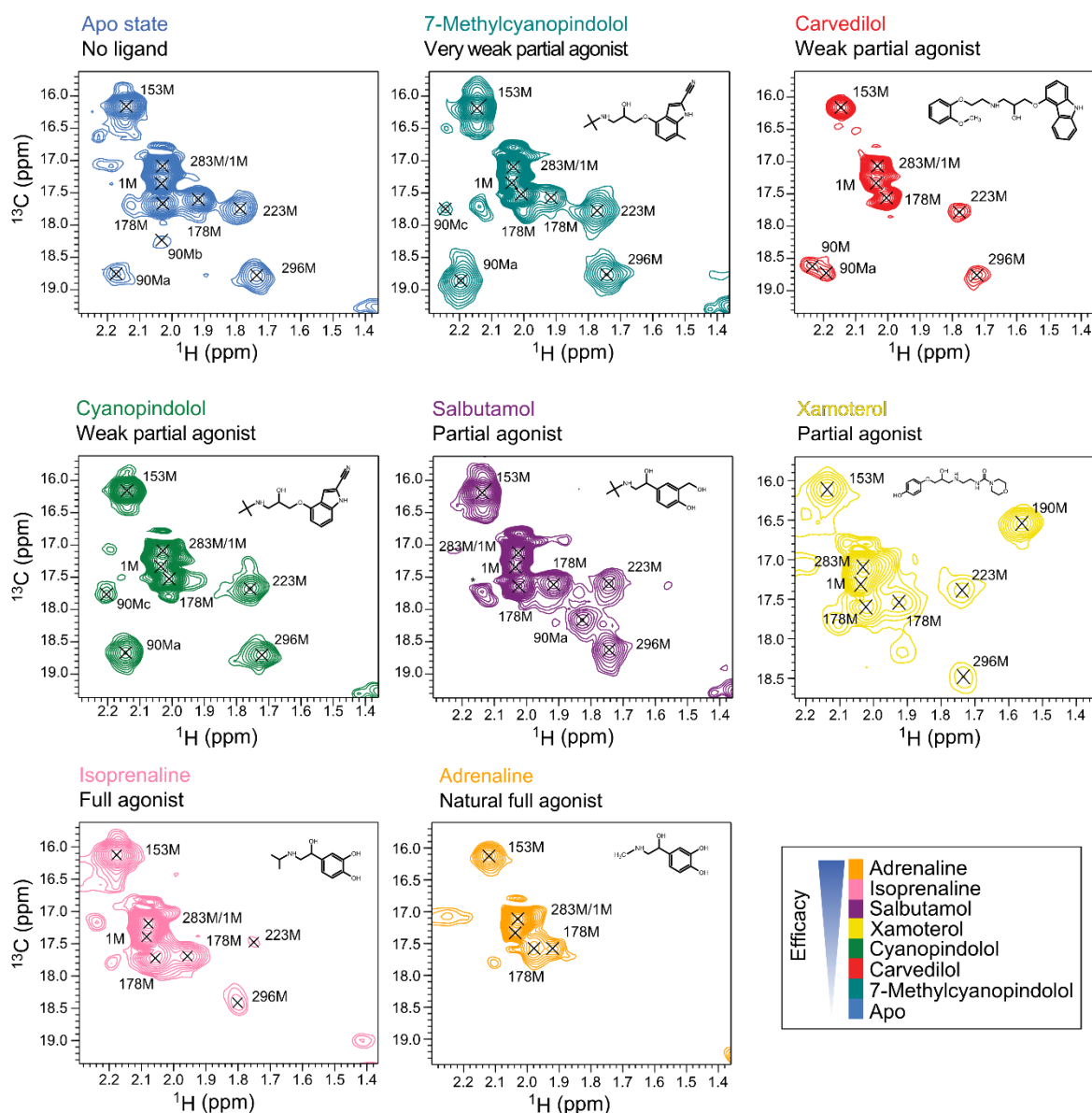
Figure 3.2 Ligand bound spectra ordered according to increasing ligand efficacy.
Figure adapted with permission from Solt et al. (*127*).

These ligand bound spectra were fully assigned as discussed before. An overview of these ligand bound states immediately shows some chemical shift differences. For example, methionine $90^{2.53}$ (M90) presents two conformations in the apo state, which resolves to a single state for ligands such as 7-methylcianopindolol and cyanopindolol. Salbutamol presents a unique conformation for $M90^{2.53}$, unobserved for any of the other ligands, and carvedilol resolves this residue into two peaks: the 90Ma conformation seen in the apo state and to a new conformation. $M90^{2.53}$ undergoes intermediate exchange for the full agonists, isoprenaline and adrenaline as well as the partial agonist xamoterol, and is not resolved at all.

73

M178[4.62] in ECL2 sits in proximity to the core transmembrane bundle on the top of TM4. With the exception of cyanopindolol and carvedilol it resolves to two distinct peaks, implying that this portion of ECL2, tethering it to TM4, samples at least two conformations.

The pharmacological properties of each of these ligands are given below in **Table 3.1** as they have been identified by previous studies, in terms of their efficacy in comparison to isoprenaline (taken as 100%), and their affinity to the thermostabilised turkey receptor (β36-m23, see Appendix A) (*234*, *245*, *246*).

Table 3.1 Pharmacological properties for ligands used in this study.

Ligand name	Abbreviation	Classification	Efficacy, ^{3}H-cAMP accumulation assays in CHO (% ISO activity)[‡§]	pK_D from ^{3}H-CGP12177 whole cell binding studies in CHO cells[†§]
7-Methylcyanopindolol	7mC	Very weak partial agonist	2.3±0.3[§]	10.37±0.03[§]
Carvedilol	CAR	Weak partial agonist	12±0.4	9.43±0.05
Cyanopindolol	CYP	Weak partial agonist	39±1	10.89±0.06
Xamoterol	XAM	Partial agonist	61±7	6.58±0.02
Salbutamol	SAL	Partial agonist	97±2	4.99±0.03
Isoprenaline	ISO	Full agonist	100	6.86±0.08
Adrenaline	ADR	Full agonist	110±2	6.01±0.04

[‡]Data taken from Baker et al. 2011. (*234*)

[†]Data taken from Baker et al. 2010. (*245*)

[§]Values for 7-methylcyanopindolol are taken from Sato et al. 2015. (*246*)

The molecular structure of the applied ligands is shown below in **Figure 3.3**. Adrenoreceptors natively bind catecholamine ligands such as adrenaline, and with the exception of this latter ligand all others are synthetic molecules.
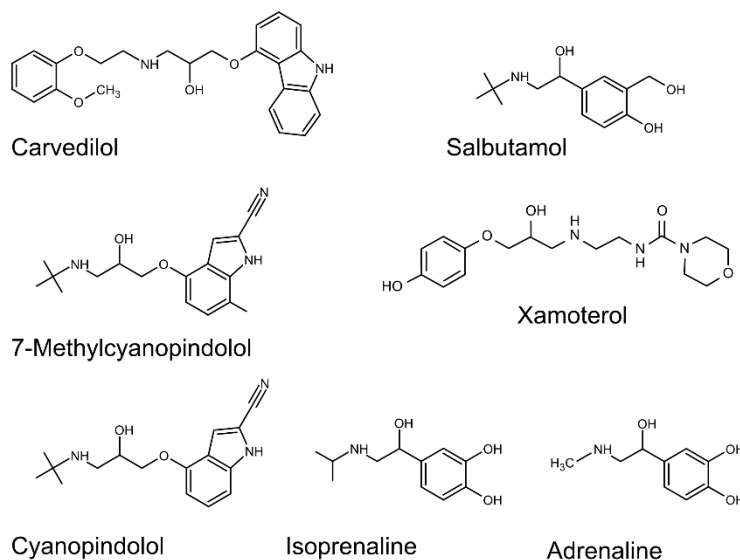
Figure 3.3 Chemical structures of the ligands used in this study.
Figure adapted with permission from Solt et al. (*127*).

The figure below shows a detailed comparison of residues M90[2.53], M153 (ICL2), M223[5.54], M296[6.41] and M283[6.28] in their apo and ligand bound states (**Figure 3.4**). M223[5.54] and M296[6.42], both undergo significant chemical shift changes with the addition of ligands, both showing significant correlation between linear combined chemical shift changes (taking into account changes in both ^1H and ^{13}C dimensions) and ligand efficacy with R^2 values of 0.830 and 0.683 respectively. The efficacy correlation plots were calculated as a linear fit between the ^{13}C and ^1H chemical shift of residues M223[5.54] and M296[6.41] in the ligand and apo bound states to the reported efficacy values (**Table 3.1**) using the equation: $a(^1H) + b(^{13}C) + c$.

M223[5.54] and M296[6.41] residues are situated in the region of TM 5 and 6, which undergo an outward extension in the unified class A activation mechanism, hence this efficacy linked transition seems indicative of a receptor undergoing structural changes towards an active-like state. The fact that both M223[5.54] and M296[6.41] resolve to a single peak transitioning in the chemical shift spectrum with the addition of different efficacy ligands is indicative of a fast exchange process. Combined chemical shift calculations were carried out by Dr. Mark Bostock.

Residue M90, which is located on TM2 facing the ligand binding pocket exhibits two conformations in the apo state denoted 90Ma, 90Mb. In the presence of cyanopindolol and 7-methylcyanopindolol only

90Ma remains, and a new conformation represented by 90Mc appears. Carvedilol and salbutamol trigger yet two different states. As mentioned earlier, full agonists and xamoterol did not resolve any peaks associated with M90$^{2.53}$, which is likely in an exchange broadened state. The equivalent residue to M90$^{2.53}$ in β_2AR is M82$^{2.53}$, and this has been reported on by Kofuku *et. al* (*241*). M82$^{2.53}$ was also observed to resolve to multiple major and minor conformations, referred to as M82^D and M82^U. Unlike for M90$^{2.53}$, M82$^{2.53}$ demonstrated a pattern of efficacy linked chemical shift changes. Furthermore while M90$^{2.53}$ was broadened beyond detection in the presence of xamoterol and the full agonists isoprenaline and adrenaline, M82$^{2.53}$ resolved to a distinct and separate third state upon the addition of the β_2AR full agonist formoterol. These comparisons set apart the observations near the ligand binding site for β_2AR and β_1AR, where for the latter the presence of multiple slow exchange conformations is observed with partial agonists, that change to intermediate exchange in the presence of full agonists (and xamoterol), giving rise to significant peak broadening.

Methionines 153 in ICL2 and M283$^{6.28}$ at the cytoplasmic end of TM6 showed much more modest chemical shift changes upon ligand addition, indicating that ligand binding on the extracellular surface does result in small but observable changes on the cytoplasmic end of the receptor, which are propagated through the transmembrane region.
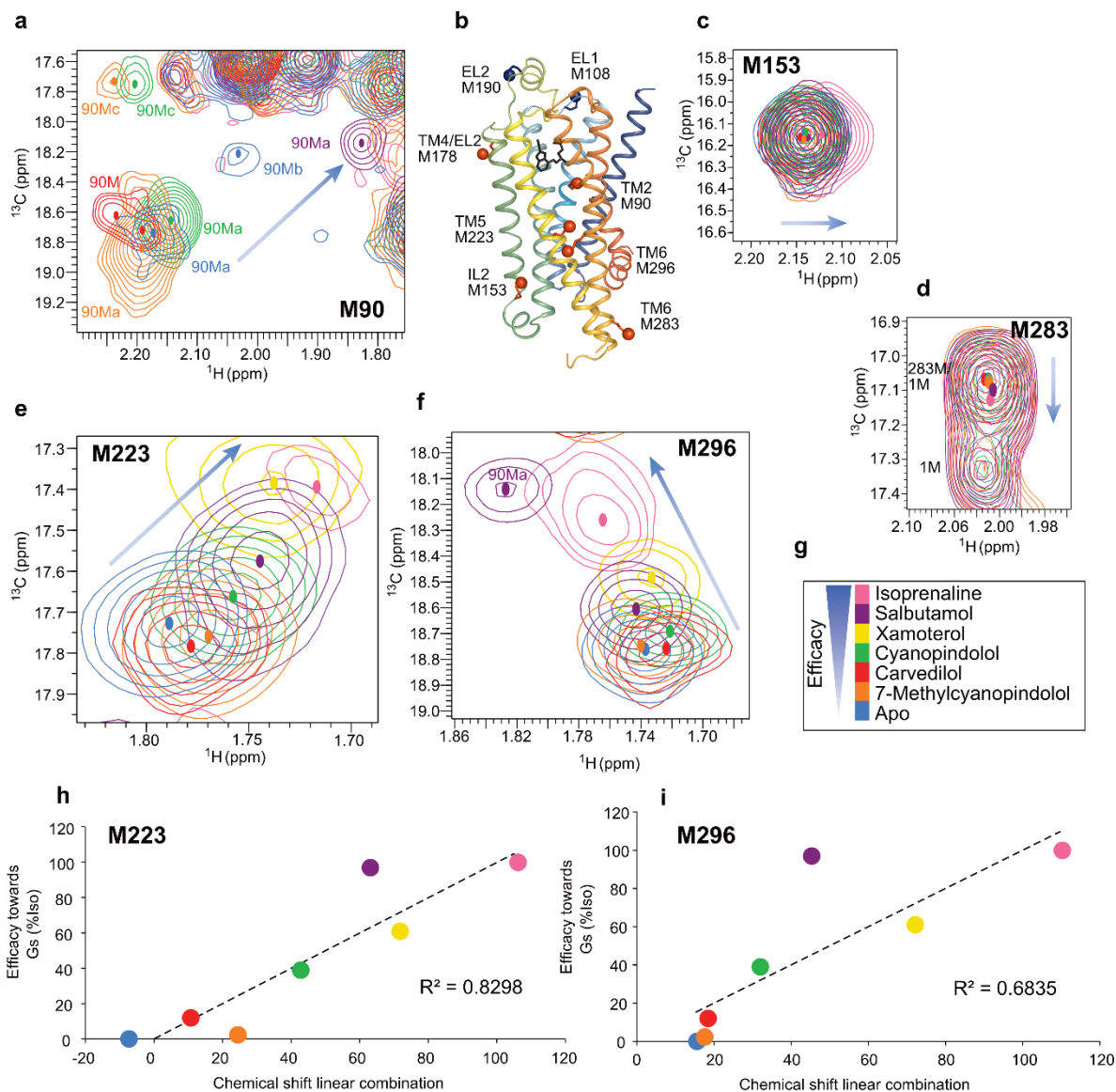
Figure 3.4 Ligand bound states and equilibria

Focus on ligand bound states and correlations between linear chemical shift and efficacy. b) Crystal structure (4BVN) showing the spatial distribution of methionine reporters. Chemical shift changes observed in the presence of different ligands for M90 (a), M153 (c), M223 (e), M296 (f), M283 (d). Ligands in order of efficacy: (g). Panels (h) and (i) illustrate the correlation between ligand efficacy and combined linear chemical shift change for residues M223 and M296. Figure adapted with permission from Solt et al. (*127*).

Isogai *et*. *al.* carried out a ^{15}N-valine backbone labelled study of a more thermostabilised (see Appendix A for a comparison between TS-β1AR and Met2-Δ5) turkey β$_1$AR construct, with an equivalent reporter to M223$^{5.54}$ (V226$^{5.57}$) showing ligand efficacy related correlation to chemical shift.

Reporters in TM6 (V280[6.25] and V298[6.43]) the latter being closest in position to M296[6.41] did not show this correlation (*244*). Given that the cytoplasmic side of TM6 is extensively demonstrated to be undergoing conformational changes, this result may be due to the relative insensitivity of backbone reporters compared to side chain methyl groups. Alternatively the greater degree of thermostabilisation of the construct used may have contributed to this. In the aforementioned methyl-methionine labelling study by Kofuku on β$_2$AR, M279[6.41] did show efficacy related changes in the presence of ligands (*241*).

3.4.1 – Temperature dependent effects reveal µs-to-ms timescale contributions to isoprenaline bound full agonist state

The full agonist, isoprenaline bound receptor was measured at 308 K, as were all the other spectra presented here, but also at 298 K and 288 K. As established above, the nature of the chemical shift changes of residues M223[5.54] and M296[6.41] are indicative of a fast exchange equilibrium on the NMR timescale. Lowering the temperature has the effect of reducing the rate constant *k*. Since the exchange regime results from the relative values of *k* and the chemical shift difference (δω) this has the effect of pushing exchange processes towards a slower timescale. Hence fast exchange processes are pushed towards intermediate and slow exchange and intermediate exchange processes are pushed towards slow exchange. For instance, if a fast exchange process can be slowed down sufficiently a single peak separates to two overlapping peaks, where the nuclei now spend enough time in each environment for the two states to be identified.

This proved to be the case for the full agonist bound state, where upon reducing the temperature to 298 K, the M223[5.54] and peak resolved to two separate ones, indicating the presence of intermediate exchange at 308 K (**Figure 3.5**). This was also accompanied by an increase in intensity, meaning that exchange broadening must be taking place at 308 K, and that at the lowered temperature this effect is reduced enough to identify two peaks. A further reduction in temperature to 288 K resulted in the retention of one of the two signals, and the signal intensity was still higher than that at 308 K. These effects are shown below, with the three spectra referenced to the signal of unbound isoprenaline.
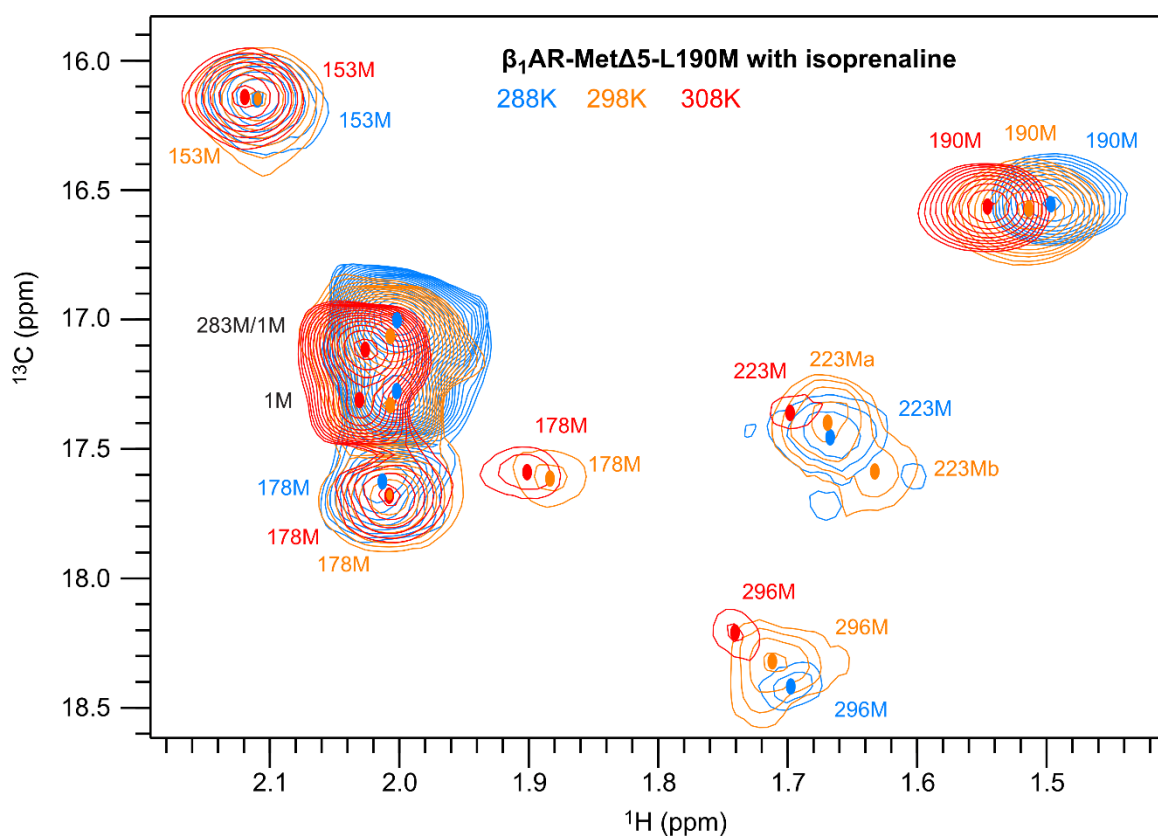
Figure 3.5 Temperature dependence of the full agonist isoprenaline bound state. The Met2-Δ5-L190 construct was recorded at 288 K (blue), 298 K (orange) and 308 K (red). The spectra were aligned using the isopropyl methyl signal of unbound isoprenaline. Figure adapted with permission from Solt et al. (*127*).

The qualitative parameters which describe these changes can be quantified in terms of signal to noise ratio (SNR). Peak broadening results in weaker signals, decreasing the SNR. Temperature also affects the molecular tumbling of the receptor and solution viscosity, factors which result in a decrease of SNR in their own right and need to be accounted for in order to quantitatively compare SNR values across different temperatures. This way the experimentally measured SNR at 298 K and 288 K can be back calculated to what it should be at 308 K, if no other major processes are at play which cause signal broadening, in this case that being the presence of μs-ms timescale exchange processes.

The diagram below shows the measured and corrected SNR values for M223[5.54] and M296[6.41] in the isoprenaline bound state at 288 K, 298 K and 308 K (**Figure 3.6**). When corrected for the aforementioned effects of temperature and viscosity, the SNR shows that a drop in temperature reduces the effects of peak broadening leading to a loss of signal, thereby revealing the presence of μs-to-ms timescale exchange contributions at 308K.
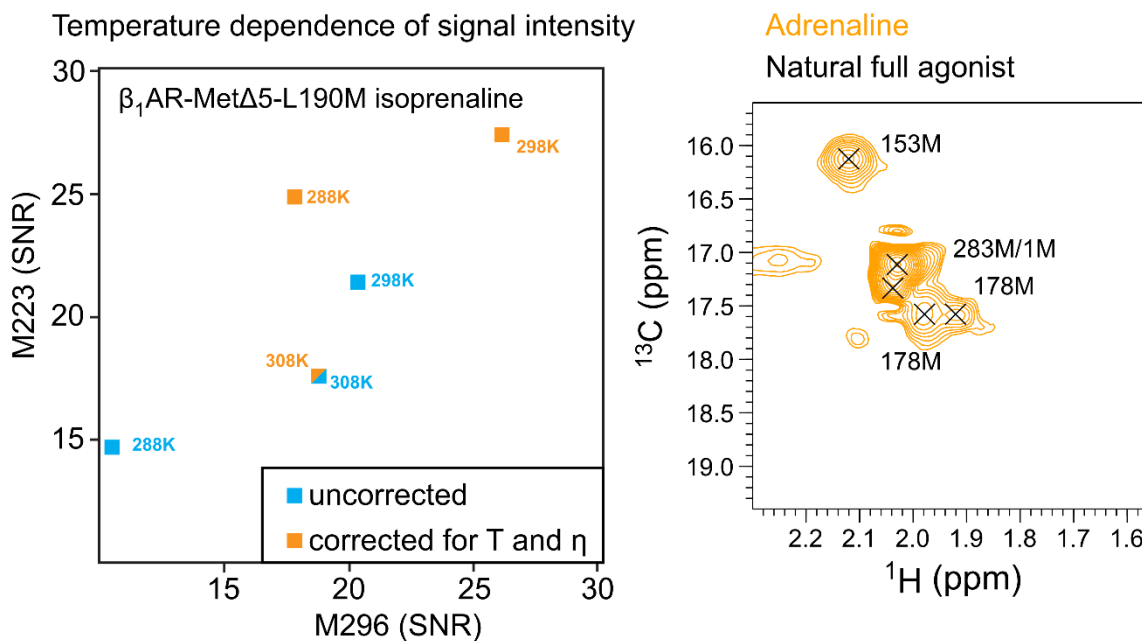
Figure 3.6 Temperature effects on signal-to-noise in the full agonist state

Temperature dependence of signal intensity in the full agonist isoprenaline bound state. Intensities shown as signal to noise (SNR) are plotted as M223 vs, M296. Corrected values (blue) take into account temperature effects and temperature associated changes in viscosity (η). In the natural full agonist adrenaline bound state no peaks were visible for M223 and M296. Figure adapted with permission from Solt et al. (127).

This is significant as this indicates that μs-to-ms timescale exchanges take place in the full agonist bound state, in addition to the ligand-modulated fast exchange equilibrium demonstrated by the efficacy dependent chemical shift changes. These exchange contributions are even more marked for adrenaline, where M223[5.54] and M296[6.41] are exchange broadened beyond resolution at 308 K. This was a unique feature of adrenaline, as no other ligands showed such exchange broadening for these two residues (**Figure 3.6**).

The SNR data was extracted by and the correction calculations were performed by Dr. Daniel Nietlispach.

3.5 – Receptor basal activity

β-adrenergic receptors have been shown to display a degree of basal activity (*247*), giving physiological relevance to investigating the interaction between the receptor and G-protein mimetic nanobodies in the absence of orthosteric ligand, with the aim of gaining insight into the receptor basal activity complex. The allosteric enhancement of ligand binding affinity in the presence of G-protein and *vice versa* is a well-known phenomenon, and therefore the nanobody's binding affinity was expected to be significantly reduced by the absence of agonist (*87*). In fact to achieve receptor saturation, a 2.6 fold molar excess of Nb6b9 and a 31 fold excess of Nb80 was used in titrations.
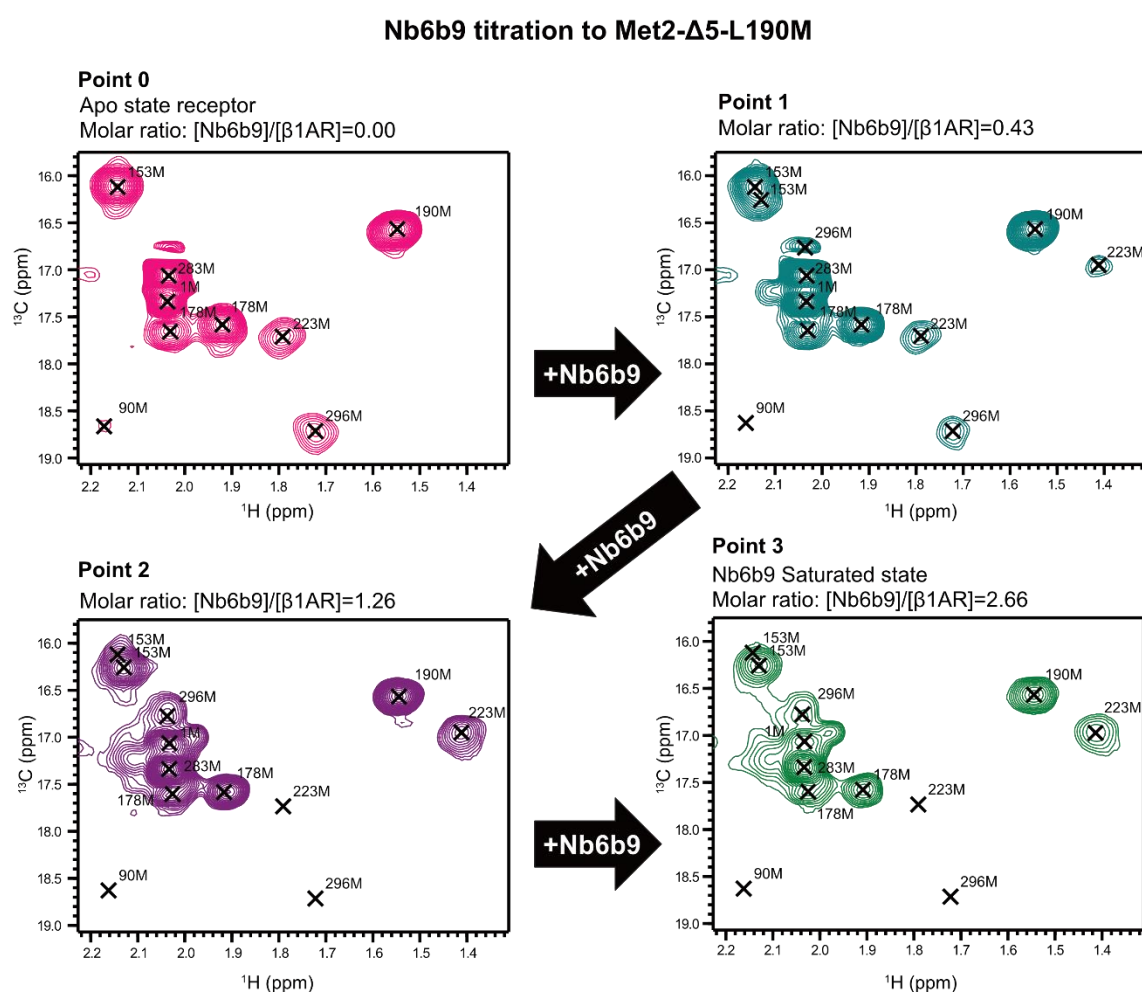
Figure 3.7 Nb6b9 titration experiment.
Nb6b9 was added to the apo state receptor until no further chemical shift changes were observed.

Titration of Nb6b9 to Met2-Δ5-L190M was able to reveal the molecular signatures of saturated and partially saturated states, as well as the amounts of nanobodies needed to achieve a saturated state for emulating the basal activity state (**Figure 3.7**). A Nb6b9 molar ratio of 0.43 to receptor (Titration

Point 1) reveals the slow exchange process of residues M223[5.54], M296[6.41] and M153, since two peaks are observed for each of these residues. The presence of a set of two peaks assigned to these residues corresponding to the apo and Nb6b9 saturated states at the same time (titration "start" and titration "endpoint peaks") is indicative of two states, that change only in intensity as the titration progresses, each resolvable as a separate peak. This bears the hallmark of an exchange process where the exchange rate k is significantly lower than the chemical shift difference ($\delta\omega$) between the two states. This change therefore represents a slow exchange process on the NMR timescale as discussed earlier under NMR Timescales. These observations establish nanobody binding as a slow exchange process.

The figure below shows the spectrum of Nb6b9 saturated β_1AR-Met2-Δ5-L190M in comparison to the apo state, both in the absence of ligand (**Figure 3.8**).
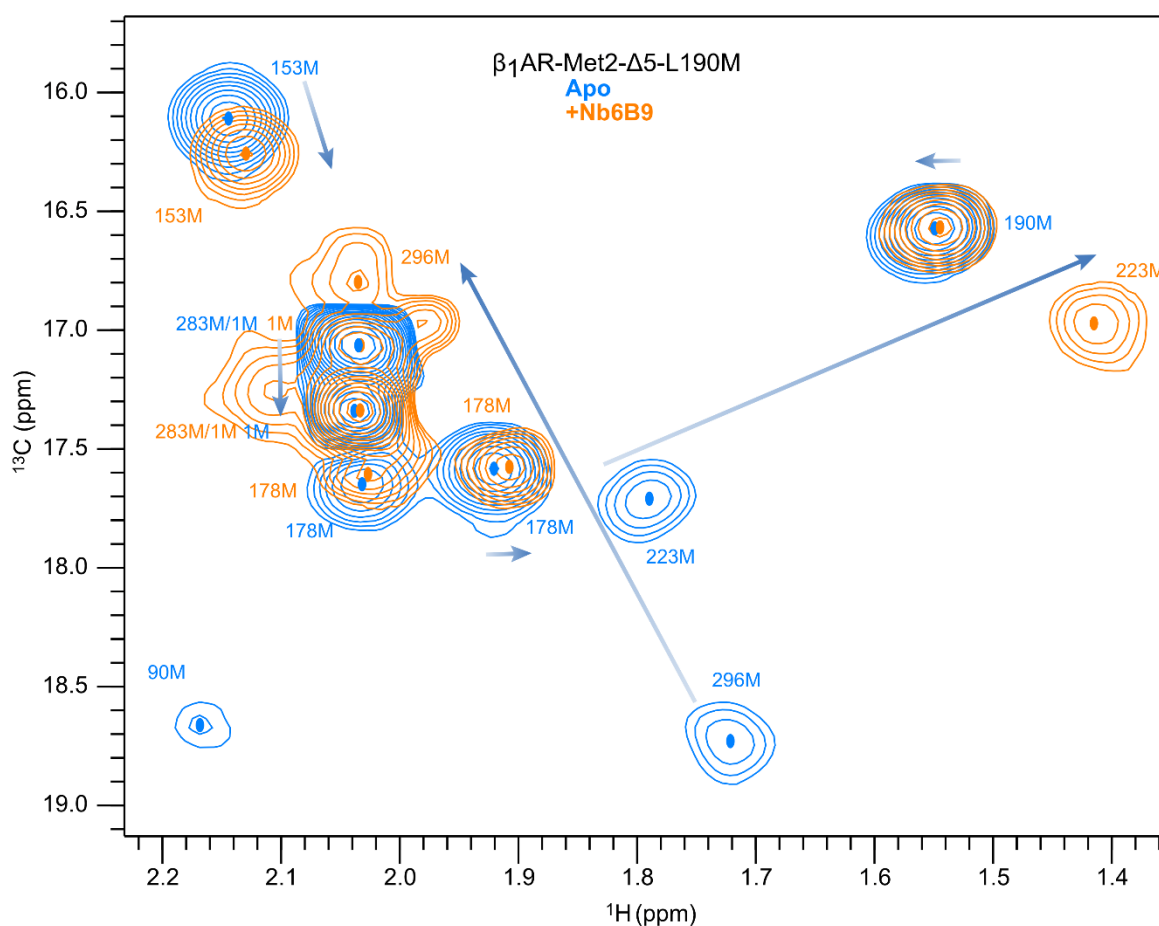
Figure 3.8 Chemical shift changes upon Nb6b9 addition to the apo state receptor. Figure adapted with permission from Solt et al. (127).

This shows significant chemical shift changes for residues on both extracellular and intracellular sides of the receptor. On the extracellular side, M190 and M178[4.62] show minor chemical shift differences,

while M223$^{5.54}$, M296$^{6.41}$, M283$^{6.28}$ and M153 on the intracellular side, show major chemical shift changes on the slow exchange regime, on the NMR timescale, as explained above.

3.6 – Ligand-nanobody ternary complexes

The addition of ligand and G-protein mimic nanobodies together was used to observe receptor states that would correspond to a fully active G-protein bound state.

The addition of saturating amounts of Nb6b9 to achieve this resulted in significant chemical shift changes compared to the full agonist bound receptor. Residues on the intracellular side, particularly M223$^{5.54}$ and M296$^{6.41}$, on TM5 and TM6 underwent large chemical shift changes similar, but slightly different to those observed for the Nb6b9-only bound basal active state. Residue M153 on ICL2, in the vicinity of the nanobody binding site also displayed a notable change characteristic of nanobody binding. M283$^{6.28}$ on TM6 was largely obscured by the peak assigned to the N-terminal methionine, however it too shows a chemical shift change upon nanobody (Nb6b9) addition. Of the methionines on the extracellular side, M178$^{4.62}$ on the base of ECL2 showed the largest difference, which has now resolved to a single peak as opposed to the two peaks present for the full agonist bound state. Extracellular loop residues M108 (ECL1) and M190 (ECL2) showed minor changes (**Figure 3.9**).
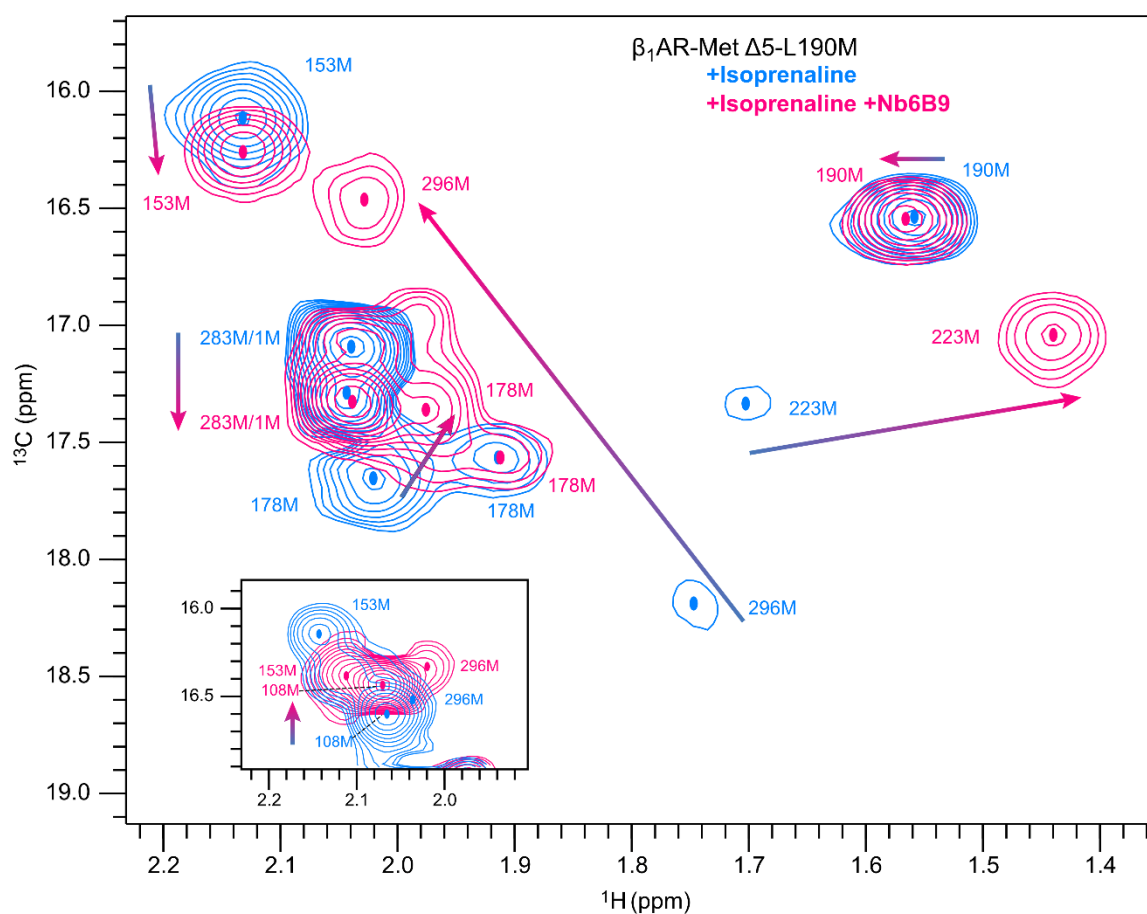
Figure 3.9 Full agonist and full agonist ternary complex states

Spectra recorded in the presence of isoprenaline and isoprenaline and Nb6b9 for the Δ5-L190M construct. Inset shows Δ5-L108M construct for the region of difference. Figure adapted with permission from Solt et al. (127).

These spectral differences between the full agonist bound state and the full agonist ternary complex state show that nanobody binding induces significant and widespread conformational changes throughout the receptor, that propagate to both the intracellular and extracellular sides (**Figure 3.10**). These changes are also more substantial than those observed for ligands alone, as represented by the full agonist bound spectra below (orange spectrum). This is typified by the differences observed for M153, M178A$^{4.62}$, M223$^{5.54}$, and M296$^{6.41}$ upon Nb6b9 addition (pink spectrum) to the isoprenaline bound state (orange spectrum), in comparison to the changes observed for these residues when isoprenaline is added the apo state (blue spectrum). These signatures are characteristic of a significantly different receptor conformational ensemble to those that exist in the presence of orthosteric ligands alone. The largest changes seen for M223$^{5.54}$ and M296$^{6.41}$ are not unexpected, given the well described structural changes that occur on TM6 and TM5 upon G-protein binding.
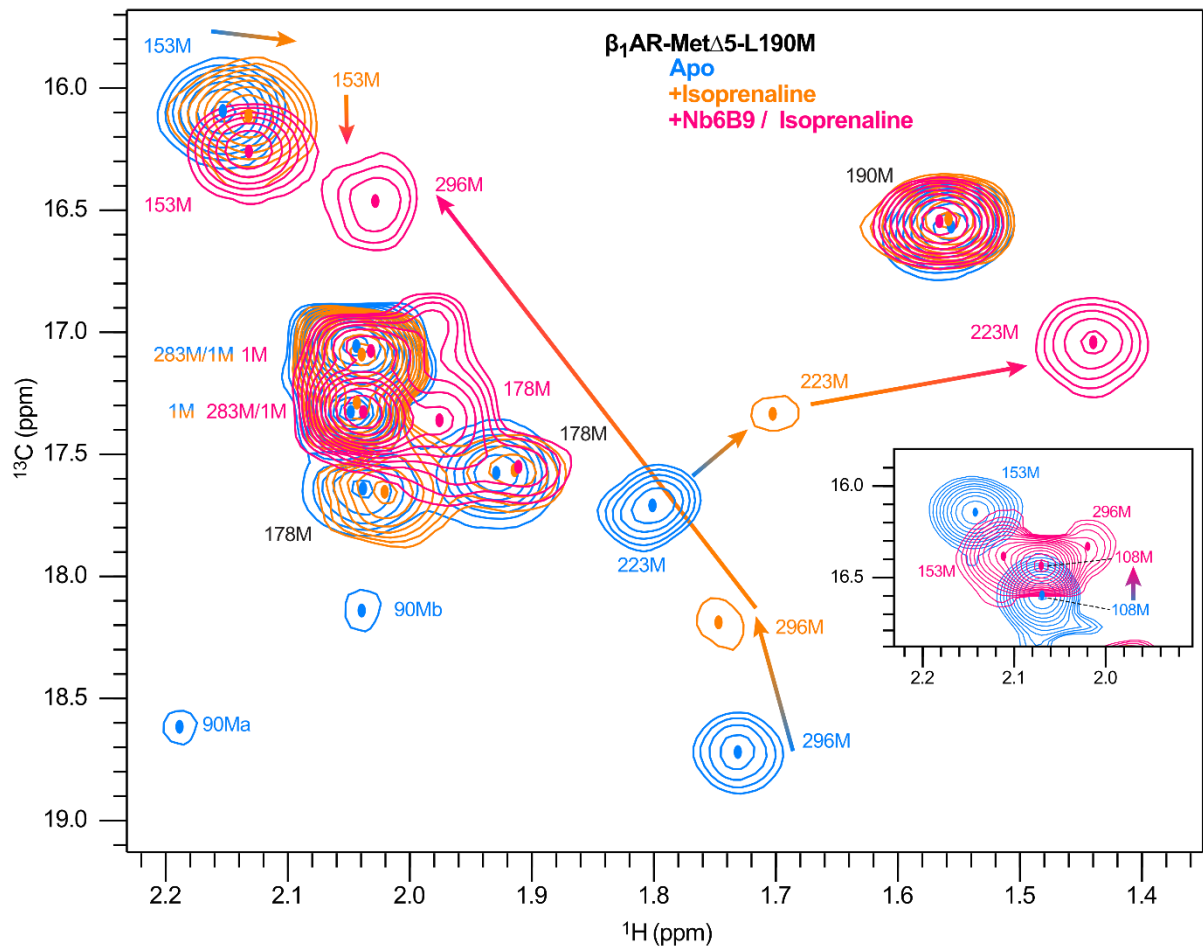
Figure 3.10 Comparison of the apo, basal activity and full-agonist ternary complex. Most significant chemical shift changes are shown by arrows. Inset shows Δ5-L108M construct for the region of difference from the Δ5-L190M construct. Figure adapted with permission from Solt et al. (127).

Comparing the ternary Nb6b9 bound basal active state (green spectrum) to the ternary complex state (pink spectrum) and the apo state (blue spectrum) to the full agonist isoprenaline bound state (orange spectrum), it becomes apparent that the basal activity state bears close similarity to the ternary complex (**Figure 3.11**). This is discussed in more detail below, but the similarities between the apo and ligand bound states and the basal active and ternary states are symptomatic of conformational equilibria that link these states.
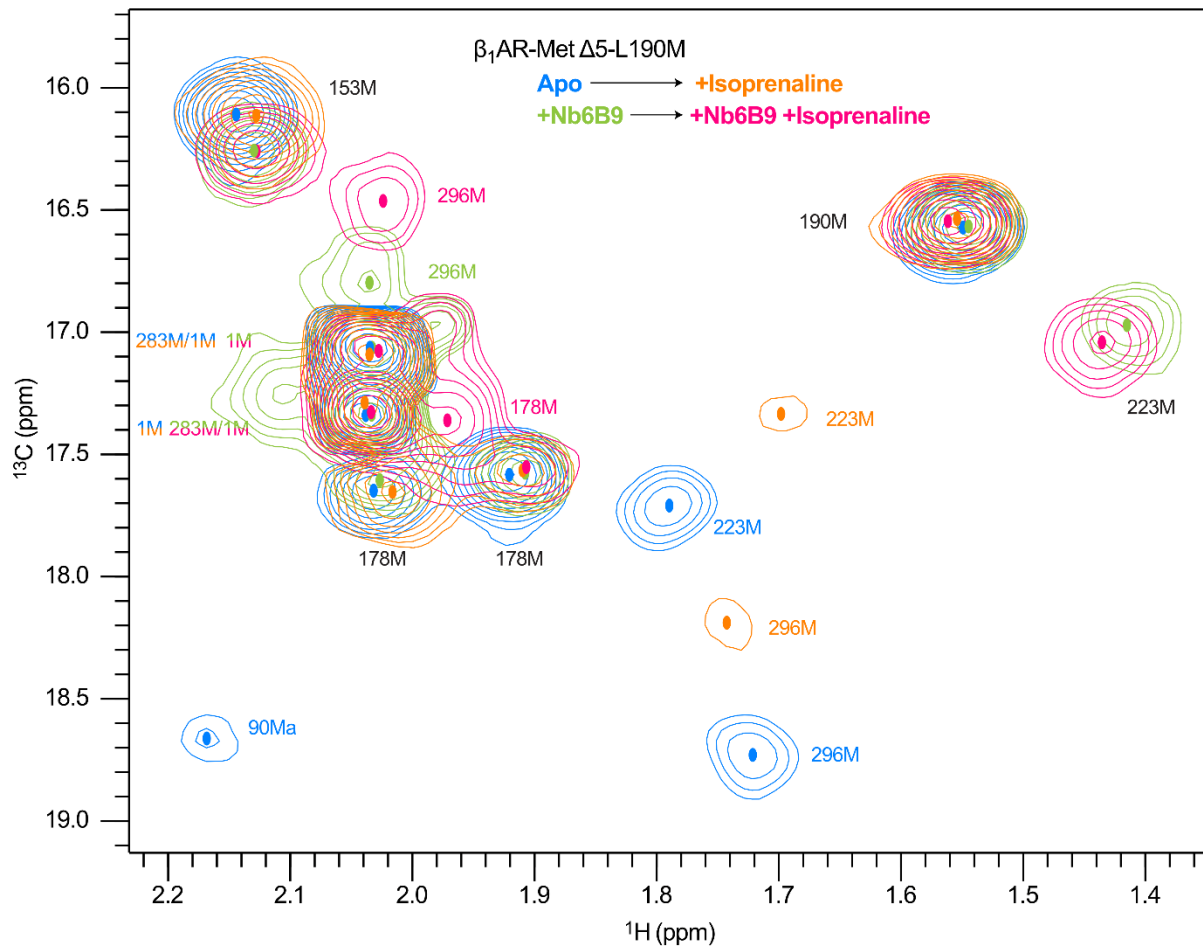
Figure 3.11 Basal activity state is shows similarities to the ternary complex state. Figure adapted with permission from Solt et al. (127).

Experiments varying the order in which agonist and nanobody was added to generate the ternary complex showed that it does not matter whether the receptor is stimulated from the extracellular side by an orthosteric ligand or from the intracellular side by a nanobody first, *with respect to* the conformation of the eventual ternary complex. This does not imply that there can be no variations in the allosteric effects of stimulation from either receptor side, but simply that once both binding partners are present the conformational ensemble resulting bears no relevance to the pathway to its formation (**Figure 3.12**).
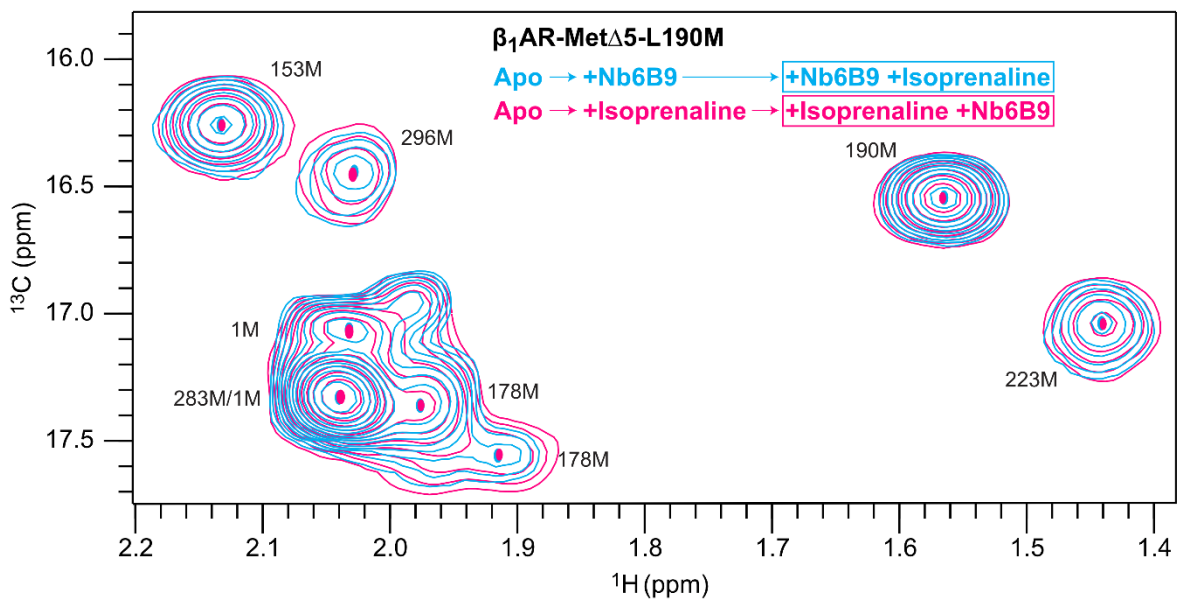
Figure 3.12 Order of ternary complex formation

The route to ternary complex formation does not bear relevance to the end state. Nb6b9 and isoprenaline added to Met2-Δ5 construct in varying order of sequence. The resulting ternary state is identical irrespective of the order of ligand and Nb6b9 addition. Figure adapted with permission from Solt et al. (127).

3.6.1 – Ternary complexes show correlation between efficacy and chemical shift

In addition to the full agonist bound ternary complex shown and discussed above, additional complexes were formed bound to adrenaline, salbutamol and cyanopindolol. An overlay of the ternary complex states of Met2-Δ5-L190M with Nb6b9 are presented below (**Figure 3.13**). Variations in ligands (including the apo – Nb6b9-only – state) resulted in marked chemical shift changes for residues M178A[4.62] on the extracellular side, as well as for residues M223[5.54] and M296[6.41] on the intracellular side. Smaller but significant chemical shift changes took place for M153 also on the intracellular side.
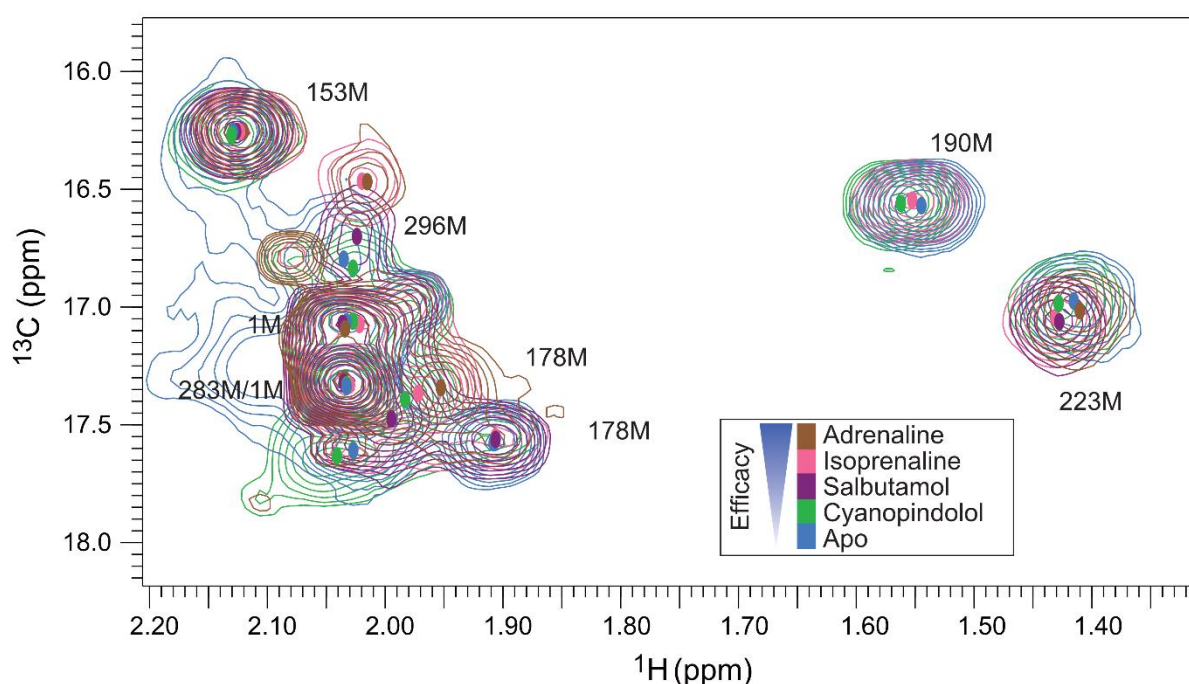
Figure 3.13 Ternary complex states bound to Nb6b9 and different ligands.
Ligands: adrenaline, isoprenaline, salbutamol, cyanopindolol. Figure adapted with permission from Solt et al. (127).

Similar to the agonist-only bound receptor states, a plot of linear combination of chemical shift changes (i.e. combined chemical shift change in the direct and indirect dimensions) against ligand efficacy revealed a strong correlation for some residues (**Figure 3.14**). One of the peaks assigned to M178[4.62] showed a strong efficacy dependent chemical shift correlation, with a coefficient of correlation (R^2) of 0.817. The other M178[4.62] peak did not show significant changes upon ligand addition to the ternary complex. M296[6.41] also showed a clear and marked correlation between chemical shift change and efficacy, with a R^2 value of 0.879. While the shift changes were much smaller in magnitude the correlation persisted for M153 (ICL2) and M223[5.54], with coefficients of 0.805 and 0.797 respectively. This is shown on the figure below, with (a) showing the two extreme states: the

Nb6b9 bound basal activity state and the full agonist isoprenaline bound complex. Diagrams (b, c, f-h) show each of the mentioned residues in detail. The combined chemical shift calculations were carried out by Dr. Mark Bostock.
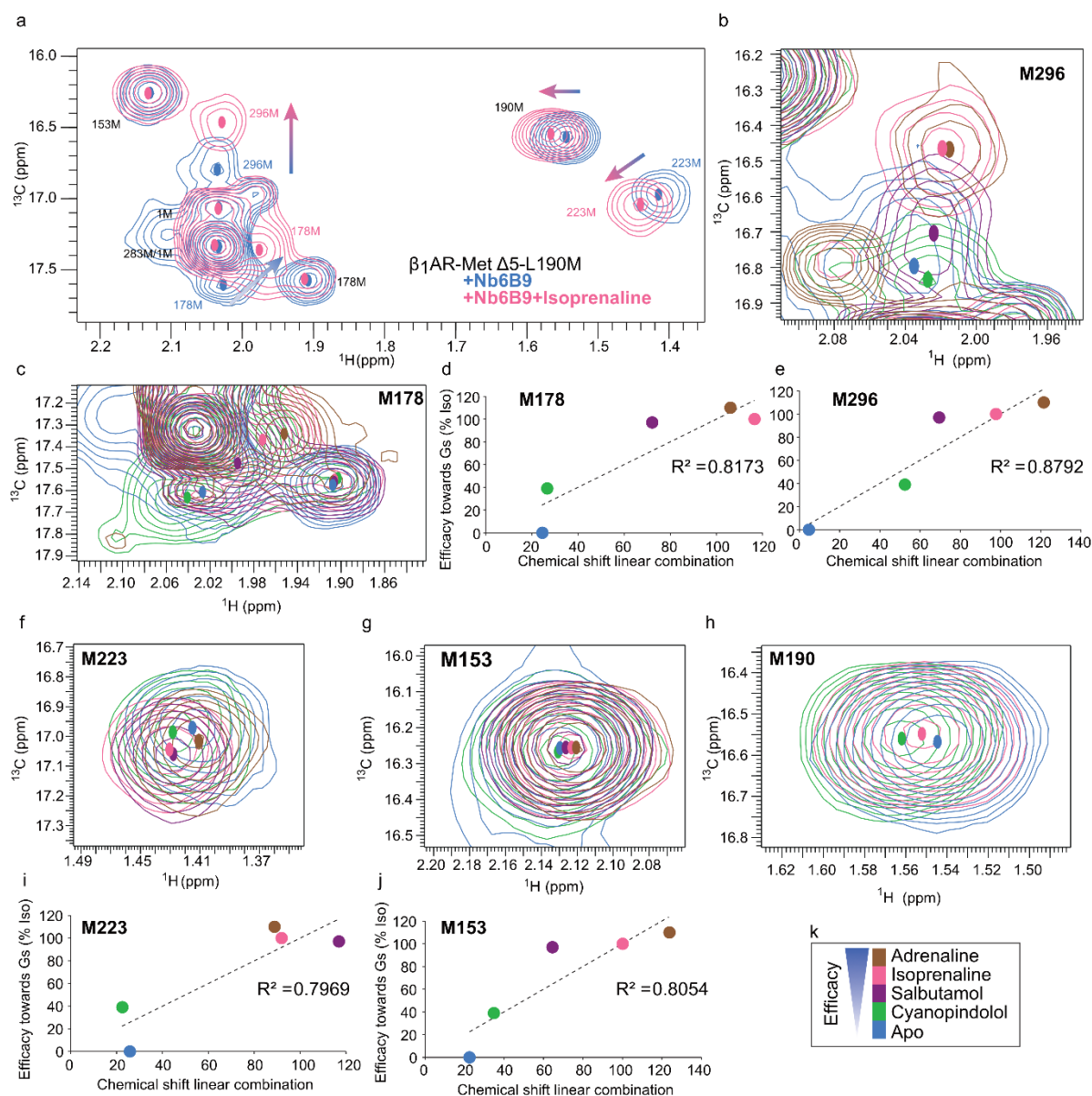
Figure 3.14 Ternary complex states and efficacy correlations

Ternary complex states and efficacy correlation to combined linear shift changes. The overall spectral change between Nb6b9 bound and isoprenaline bound ternary states are shown in (a). Chemical shift changes for residues M296 (b), M178 (c), M223 (f), M153 (g), M190 (h) are shown independently in the presence of various ligands of different efficacies (k). Correlations between linear chemical shift change and ligand efficacy are plotted for M178 (d), M296 (e), M223 (i) and M153 (j). Figure adapted with permission from Solt et al. (127).

3.6.2 – Lower temperature experiments indicate a ligand dependent ternary complex equilibrium in fast exchange

As discussed earlier in the context of temperature dependent effects on ligand-only bound receptor states, lowering the experiment temperature has the potential to alter the rate of exchange processes, changing the relationship between the exchange rate constant k and the chemical shift difference between the two exchange states (ω). In this way, the presence of processes leading to exchange broadening, taking place on the intermediate NMR timescale were revealed for the full agonist bound receptor in the presence of isoprenaline and adrenaline.

In the case of ternary complexes, recording the partial agonist salbutamol and Nb6b9 bound ternary complex at 308 K and at 293 K showed evidence indicating that the ligand-modulated equilibria present in the ternary complex receptor state takes place on the fast exchange regime on the NMR timescale (**Figure 3.15**). The two spectra below show the salbutamol ternary complex at 308 K and at 298 K referenced to M153 (ICL2). This shows that M223[5.54] and M296[6.41] both shift noticeably towards the basal activity state, away from the full agonist ternary complex chemical shift value.
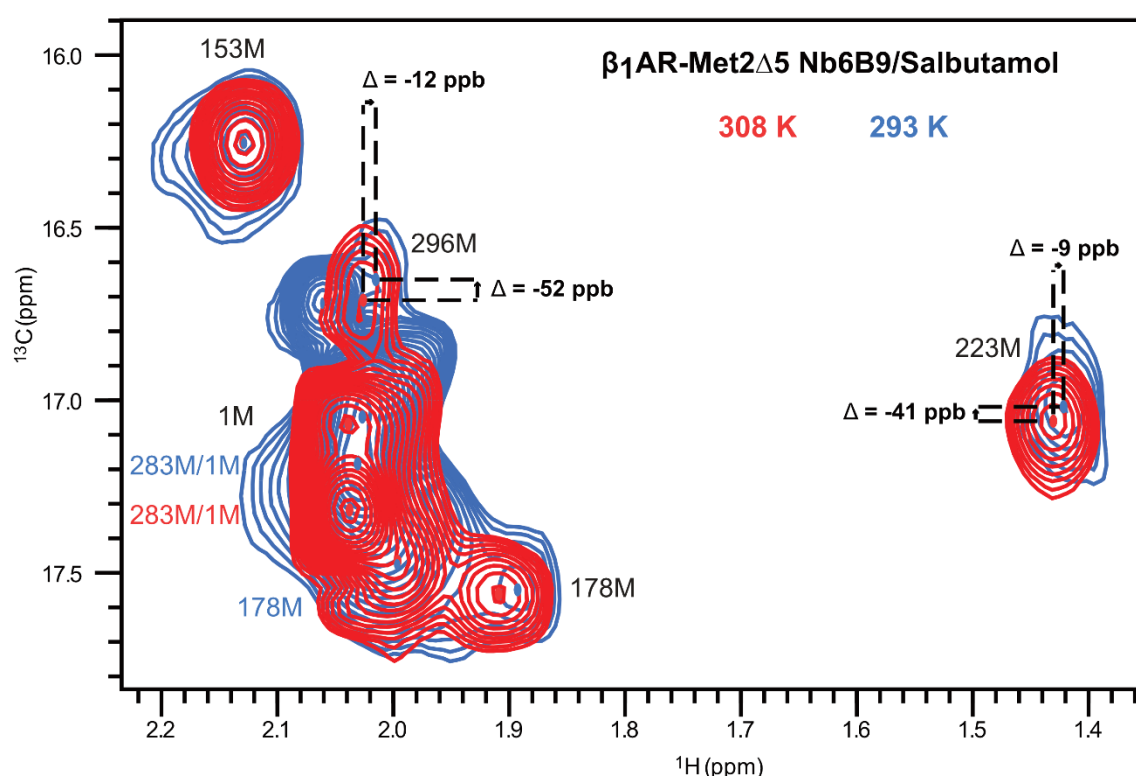
Figure 3.15 Temperature dependence in the ternary complex state
Receptor bound to Nb6b9 and salbutamol. Figure adapted with permission from Solt et al. (127).

The relative intensity values were measured for these two residues, taking M153 as a reference (with a set value of 1). This shows a degree of peak broadening at the lower temperature. The signal to noise ratios (SNR) were also measured at both 308 K and at 293 K. A number of factors play a part in modulating the SNR upon temperature change, such as viscosity, and the presence of intermediate timescale exchange processes, which result in peak broadening. The effect of temperature *per se* and viscosity can be corrected for, by back compensating for it through a back calculation of what the SNR value should be at 308 K based on the 293 K data, *if* no other processes influence SNR. In a simplified context it was assumed that factors such as the presence of detergents and protein concentration differences did not influence SNR. The 298 K SNR values were multiplied by the ratio of viscosity values of pure water at 308 K (η_{308}=0.7192 cP) over that at 293 K (η_{293}=1.002 cP), multiplied by the temperature difference: (η_{308}/η_{293})*(308/293), giving a compensation factor of 1.465. The ratio of the back-calculated SNR value at 308 K compared to the measured SNR at 308 K gives an indication of the presence of exchange contributions on the µs-ms timescale, corresponding to the intermediate regime on the NMR timescale. If these ratios are close to 1, it indicates the lack of such processes. If the 293 K/308 K ratio is less than 1, it indicates the presence of exchange contributions at this timescale. For M153, this ratio was close to 1, meaning that over this temperature range no significant, detectable change has occurred in the exchange regime. For M223[5.54] and M296[6.41] however this ratio was now less than 1, indicating that the exchange rate has been slowed down, with the appearance of some intermediate timescale contributions. This infers that at 308 K the exchange process likely takes place on the fast exchange timescale. This provides evidence, that the ligand modulated efficacy dependent equilibrium that has been identified in the receptor ternary complex follows a fast exchange regime.

The table below lists the SNR and relative intensity measurements corresponding to the salbutamol ternary complex. This analysis, together with the SNR and intensity measurements were done by Dr. Daniel Nietlispach.

Table 3.2 SNR values for ternary complex state

Temperature	β₁AR-Met2Δ5 with salbutamol / Nb6B9							
	relative intensity (reference M153)		signal-to-noise ratio (SNR)					
			uncorrected			corrected for T and η (back-calculated to 308 K)		
	M223	M296	M223	M296	M153 (ref)	M223	M296	M153 (ref)
308 K	0.59	0.36	130	80	219	130	80	219
293 K	0.46	0.29	64	45	145	94	65	212
ratio 293 K vs 308 K	0.78	0.81				0.72	0.81	0.97

3.7 – Peak intensity as a measure of μs-ms dynamics

GPCRs have been shown to be highly dynamic molecular entities, undergoing motion on a range of timescales [refs]. The exchange broadening observed for several residues (M90$^{2.53}$, M223$^{5.54}$, M296$^{6.41}$) upon the addition of agonists is symptomatic of motion on the μs-to-ms timescale, reflecting a high degree of receptor motility reflecting the functional state of the receptor. It is therefore valuable to be able to describe and compare the nature of receptor dynamics that take place as a consequence of receptor activation by ligands or G-protein mimetic nanobodies (and via inference by other native intracellular binding partners). NMR spectroscopy offers itself as an ideal toolkit for the investigation of protein dynamics in the form of spin relaxation techniques, however several parameters of the protein used here confer severe limitations to these types of experiments. Relaxation experiments require a high degree of sensitivity, which demand extended measurement times. The reduced thermostabilisation and low protein concentrations in the NMR samples of β_1AR measured both pose limitations to sample stability, meaning that sample lifetimes are insufficient for these experiments. The size of the receptor in LMNG micelles has been measured at around 80 kDa (by calibrated SEC). Larger particles such as these experience longer rotational correlation times, reducing T_2, translating to a loss of signal. Additionally, in the absence of deuteration, undesirable dipole-dipole coupling results in fast T_2 relaxation. Deuteration is able to counter short T2 relaxation. However this is very limited in insect cell expression, and is not a widely available tool for the study of GPCRs, which do not express well in other expression systems where perdeuteration is technically feasible (such as *E. coli* expression). Insect cells do not have a high tolerance to growing in D$_2$O. Therefore, the effect of dipole-dipole interactions and short T_2 timescales result in rapid signal decay and low sensitivity, making it impossible to use techniques such as CPMG relaxation experiments for a direct quantitative assessment of receptor dynamics.

Another method of gaining information on μs-to-ms timescale dynamics, which result in line broadening is the measurement of peak linewidths. Unfortunately, in the ^{13}C HMQC experiments conducted, the resolution in the ^{13}C dimension (indirect dimension) was reduced in order to allow an increased number of scans, improving signal-to-noise. This resulted in significant signal overlap in the indirect dimension, which is immediately obvious from the overview of the ligand bound spectra alone as shown above. This made it impossible to obtain precise linewidths for use in the estimations of dynamics on the timescales resulting in peak broadening.

As a substitute, the peak intensity values of the assigned residues were still able to be used as a means of observing methionine methyl dynamics. This provided a semi-quantitative assessment of how the overall receptor dynamics change as a consequence of ligand and nanobody binding.

Methyl groups are particularly useful tools for the study of large macromolecules, as the rapid rotation around the methyl head group causes a partial decoupling of the group from the overall tumbling of the larger macromolecule. This is described by low S^2 order parameters (high degree of local motion) and results in favourably longer T_2 relaxation times, which manifest themselves in sharper signals. The long spin lifetime of methyl residues makes them sensitive to changes in dynamics on the slower timescales, such as those on the μs-to-ms timescale. In this way methyl signals are qualified reporters of conformational exchange processes which result in exchange broadening and the loss of signal intensity, which is used in this work. These advantages however are not enough to overcome the challenges of relaxation measurements.

It is important to note, that when discussing receptor states as being more or less dynamic, based on the estimates obtained from signal intensities, these descriptions refer to motions on the NMR timescales that cause a change in signal intensity. These NMR timescales coincide with μs-to-ms timescale motions. It is crucial to note that the description of motions on this timescale is separate from motions on a much faster timescale, which cause a sharpening of signals *either* because of lower internal correlation times (lower S^2) *or* because of a reduction in motional order parameters when the molecule becomes more dynamic.

Peak intensities for a specific residue can be compared in one of two ways. In a relative comparison, the peak intensity for a specific methionine was compared to a given reference in a spectrum, such as residue M153, and the intensity of other selected residues measured in relation to the reference (set as a value of 1). This allows the comparison of residue intensities across the same spectra. Alternatively the relative intensities can be normalised against each other across different sets of spectra (different experiments on the same construct), with the most conformationally restricted state set a value of 1, and all other intensities scaled to this. In either approach, peaks with a lower intensity value display increased peak broadening, showing increased dynamics on the μs-to-ms timescale, which is responsible for this effect as discussed.

3.7.1 – Dynamics in the apo, ligand bound and ternary complex states

The observation of peak intensity changes showed a large variation for residues M90[2.53], M223[5.54] and M296[6.41]. As discussed earlier, M90[2.53] is completely undetectable from the spectrum of xamoterol, isoprenaline and adrenaline bound receptor, indicating the presence of conformational exchange at a similar rate to the NMR timescale. The tables below show the intensity values obtained for residues M223[5.54] and M296[6.41] in receptor bound to a range of ligands with and without Nb80 or Nb6b9. All relative intensities were measured from ^{13}C, ^{1}H HMQC experiments recorded at 308 K at 800 MHz. Intensity values were calculated relative to the well resolved M153 residue. For each construct (Met2-

Δ5-L109M and Met2-Δ5) the intensity values were normalised against the highest intensity peak for that construct, and this was given an intensity score of 1. In this ranking, lower values represent increased peak broadening and increasing μs-ms dynamics. Peak intensities were measured and graphed by Dr. Daniel Nietlispach.

Table 3.3 Peak intensities for M223 and M296

Relative peak intensity of methyl resonances M223[5.54] and M296[6.41] for β₁AR-Met2-Δ5-L190M, β₁AR-Met2-Δ5								
	Relative peak intensities				Normalised peak intensities			
	Met2-Δ5-L190M		Met2-Δ5		Met2-Δ5-L190M		Met2-Δ5	
Ligand	M223	M296	M223	M296	M223	M296	M223	M296
Adrenaline (ADR)			0.06	0.04			0.1	0.11
Isoprenaline (ISO)	0.08	0.12	0.05	0.10	0.14	0.31	0.08	0.28
Salbutamol (SLB)			0.16	0.22			0.27	0.61
Cyanopindolol (CYA)	0.28	0.24	0.25	0.23	0.49	0.62	0.42	0.64
7-Methylcyanopindolol (7mC)			0.27	0.21			0.45	0.58
Apo	0.26	0.21	0.27	0.22	0.46	0.54	0.45	0.61
Carvedilol (CVD)			0.27	0.25			0.45	0.69
Apo/Nb80	0.38	0.36	0.44	0.29	0.67	0.92	0.73	0.81
Apo/Nb6B9	0.43	0.36			0.75	0.92		
ADR/Nb6B9			0.60	0.35			1	0.97
ISO/Nb80	0.57	0.39	0.59	0.35	1	1	0.98	0.97
ISO/Nb6B9	0.57	0.39			1	1		
SLB/Nb6B9			0.59	0.36			0.98	1
CYA/Nb6B9	0.47	0.37			0.82	0.95		

The figure below shows the relative intensities of M223$^{5.54}$ and M296$^{6.41}$ plotted against each other (**Figure 3.16 c, d**). This allowed the direct comparison of the two residues in terms of their conformational dynamics. The normalised intensities (**Figure 3.16 a, b**) are also shown, and this allows the comparison of different receptor states and constructs.

The full agonists adrenaline and isoprenaline bound receptors show consistently lower peak intensities for M223$^{5.54}$ and M296$^{6.41}$ than they do for partial agonist cyanopindolol, 7m-cyanopindolol, salbutamol, carvedilol or the apo form. The adrenaline bound state was even more mobile than the isoprenaline bound state, even though both exhibited a significant conformational exchange.

Peak intensity ratios for nanobody bound ternary complexes showed significant increases in signal intensity, showing that these receptor states correspond to a more rigid form, experiencing much less conformational exchange on the µs-ms timescale.

Nanobody-only bound receptor states showed more rigidity than the ligand bound or apo states, but showed more dynamics resulting in peak broadening than did the ligand plus nanobody bound ternary complexes. Both Nb80 and Nb6b9 showed this effect.

The pattern revealed by normalised peak intensities and by relative peak intensities persisted across the two types of analyses. Similarly, the construct used (Met2-Δ5 and Met2-Δ5-L190M) did not change this either.
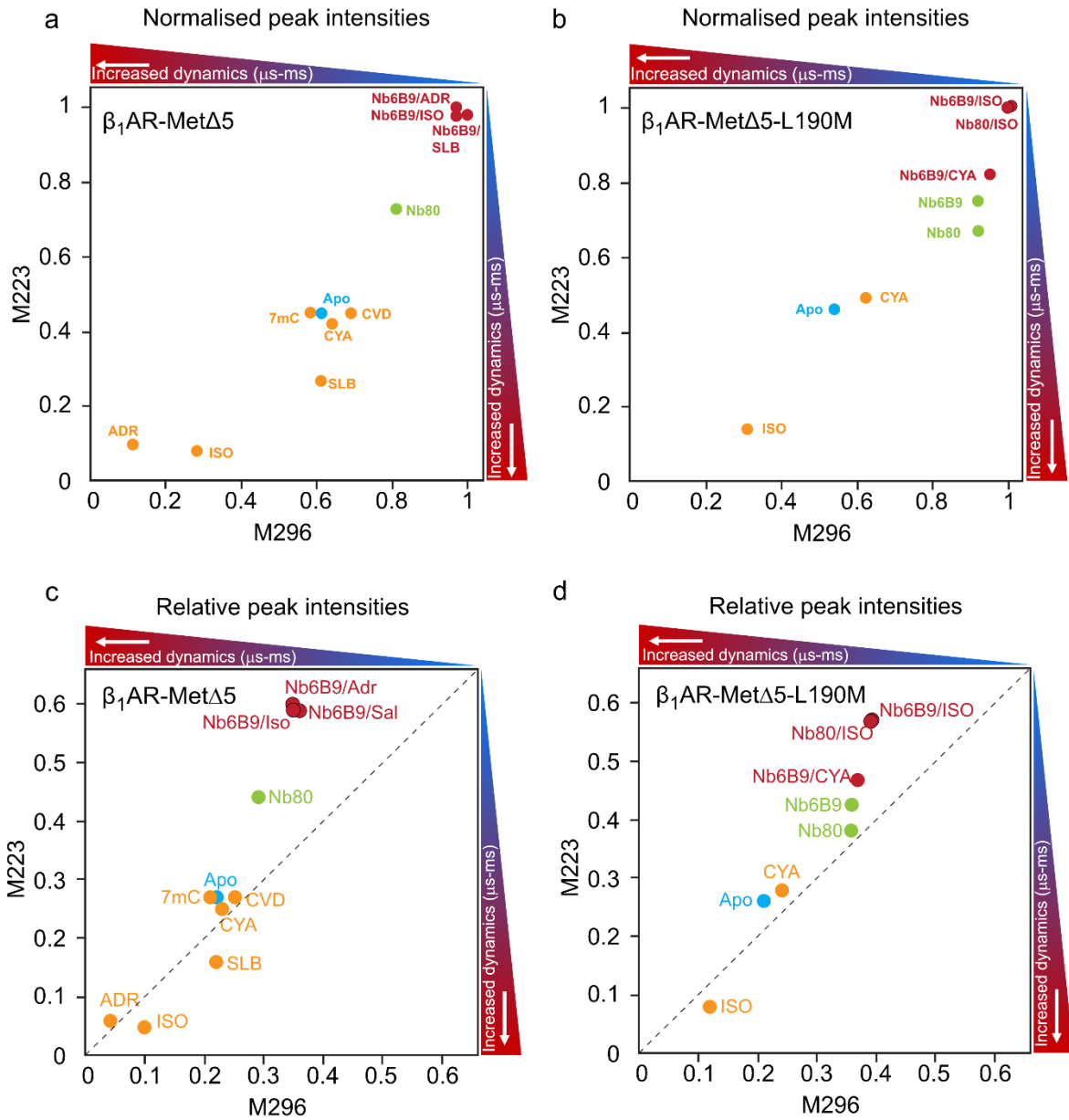
Figure 3.16 Peak intensity analysis

Peak intensities show changes in receptor dynamics on the intermediate NMR timescale. Relative peak intensities (a,b) were calculated relative to the M153 signal in each spectrum. These can be converted to normalised peak intensities (c,d) by assigning a value of one to the most intense peak in each spectra. Higher values represent an absence of conformational fluctuations leading to peak broadening and lower values correspond to increased dynamics on the μs-to-ms timescale. Figure adapted with permission from Solt et al. (127).

3.8 – Effect of thermostabilising M90$^{2.53}$A mutation

The mutation of residue M90$^{2.53}$ to an alanine has been established to increase the melting temperature (T$_m$) of the antagonist alprenolol bound receptor (construct: β$_1$AR$_{34-324}$, see Appendix A) by 8 °C ([94]). The introduction of this mutation not only assisted the assignment of residue M90$^{2.43}$, but also showed insight into the effect of thermostabilisation. NMR spectra were obtained for the Δ5-M90A mutant in a number of receptor conformations. Comparing the apo state Δ5-M90A and Δ5 constructs (**Figure 3.17a**) showed an identical picture implying little change in the apo state conformation due to the mutation. This was in contrast to the ligand bound states (**Figure 3.17b**). The signals for residues M296$^{6.41}$ and M223$^{5.54}$ in the M90A construct when bound to the full agonist isoprenaline (red and light blue spectra) are shifted in contrast to the Δ5 construct signals in the direction of chemical shift displayed by lower efficacy agonists. These effects persisted for the lower efficacy agonist salbutamol (purple and green signals).

Peak intensities (**Figure 3.17c**) of the Δ5-M90A construct in the presence of salbutamol and isoprenaline were compared. This showed significantly higher intensity values for Δ5-M90A compared to Δ5, indicating a more rigid and less exchange broadened receptor state attributable to the thermostabilisation. The magnitude of intensity difference also seemed to increase with efficacy; the change in the relative intensity for the apo states is the smallest, followed by the difference for the salbutamol bound states, with the full agonist bound states showing a considerable difference in relative intensity.

A comparison of the isoprenaline bound ternary complexes also showed that the peaks for M296$^{6.41}$ and M223$^{5.54}$ are slightly shifted in the direction of the partial agonist ternary complexes (**Figure 3.17d**). Peak intensities for the ternary complexes show that the Δ5-M90A construct adopts a less rigid conformation in the full agonist ternary complex, compared to the Δ5 full agonist ternary complex. This has previously been observed to be the case for the low efficacy ternary complexes, such as the basal activity state or the cyanopindolol bound complex. Peak intensity values were measured by Dr. Daniel Nietlispach.
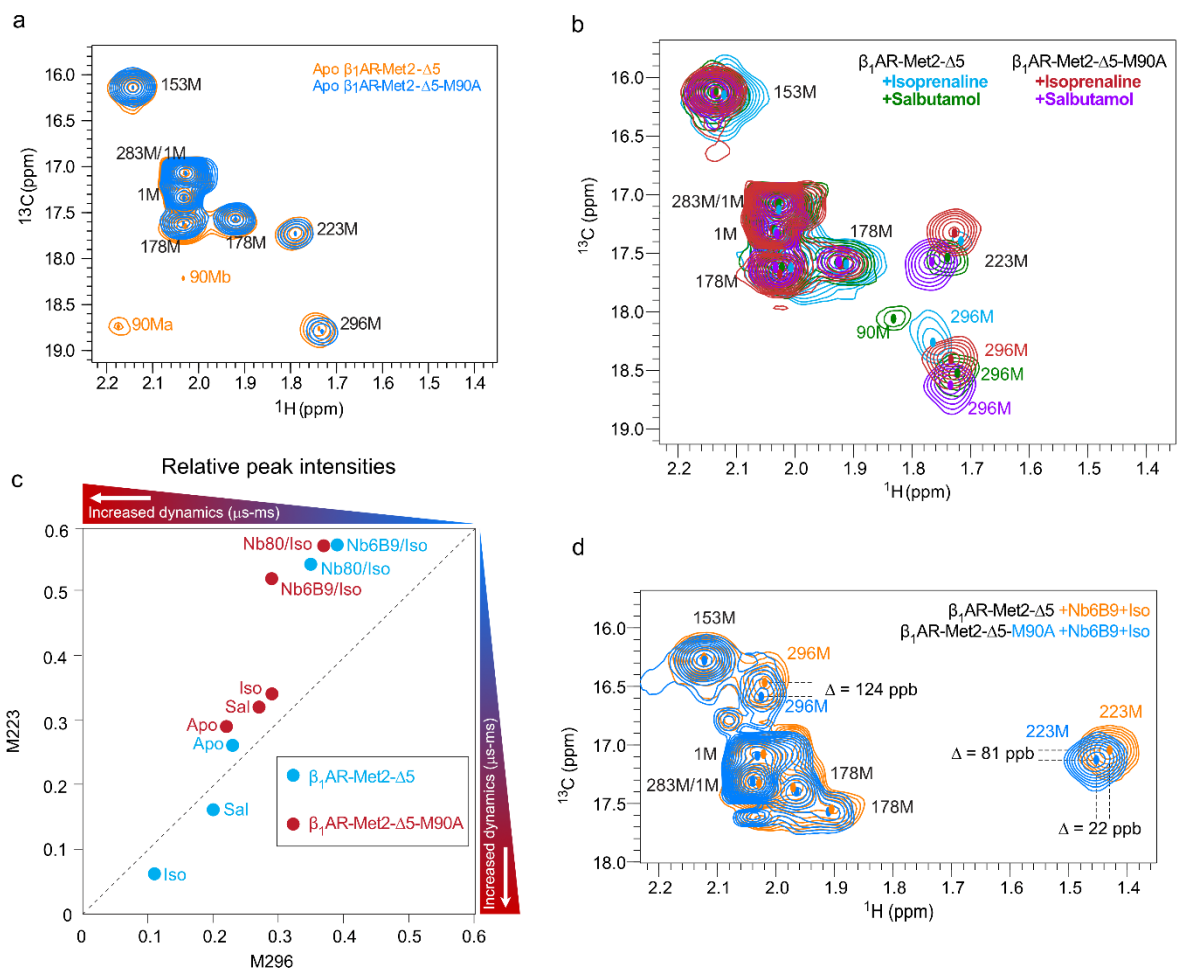
Figure 3.17 Effects of the M90A thermostabilising mutation

(a) Apo state spectra of Δ5 and Δ5-M90A constructs. (b) Comparison of isoprenaline and salbutamol bound states. (c) comparison of peak intensities across various states of the Δ5 and Δ5-M90A constructs. (d) comparison of isoprenaline+Nb6b9 bound ternary complex states of Δ5 and Δ5-M90A. Figure adapted with permission from Solt et al. (127).

These observations suggest that increased thermostabilisation affects the ligand dependent equilibria that has been observed in the ligand bound and ligand-ternary complex states, in a way that they are shifted towards the less active (I) and (A^{G-}) states (**Figure 3.19**).

3.9 – Comparison of βAR crystal structures with methyl rotational conformations

Unlike structural biology techniques such as crystallography or cryo-EM, which generate structural models of receptor conformations, the ^{13}C-HMQC NMR experiments used in this work report on the molecular signatures of conformational ensemble of various receptor states. This means that it is not possible to directly back-calculate a structural model of the detected conformational ensembles to a structural model from the information provided by these experiments. NMR spectroscopy can achieve this through other more suitable experiments performed on uniformly labelled (^{13}C, ^{15}N) protein samples, where through the sequential assignment of residues and through measurements of NOE or other through-space constraints a structure can be calculated. The nature of the system used in this work and the limitations posed by this do not make this a possibility and to date the only GPCR or GPCR analogue structures that have been determined purely by NMR data alone are that of rhodopsin (*248*, *249*).

However, ^{13}C methyl resonances are sensitive reporters of the methionine χ^3 rotamer conformation, while the methyl ^{1}H chemical shifts are sensitive to perturbation by ring currents from nearby aromatic residues. Rotations around the χ^3 bond position the CH_3 head group in what is referred to a *gauche* conformation (between χ^3 angles of -67 ° and +67 ° with respect to the C_γ-S bond plane) or a *trans* conformation (χ^3 near 180 °) (**Figure 3.18c**). The effect of the χ^3 rotation angle on the ^{13}C chemical shift has been modelled by Butterfoss *et al*. and is shown below (**Figure 3.18b**) (*250*).

The chemical shift values of M223$^{5.54}$ and M296$^{6.41}$ in the apo state, isoprenaline bound state and isoprenaline-Nb6b9 ternary complex state receptor are plotted. This shows M296$^{6.41}$ undergoing a *trans* to *gauche* conformational change as a consequence of ternary complex formation, predicting a χ^3 rotamer angle of -67 ° (**Figure 3.18a**). The aromatic residues which can have an effect on M223$^{5.54}$ and M296$^{6.41}$ are Tyr$^{5.58}$, Phe$^{6.44}$ and Tyr$^{7.53}$. In all of the published β$_1$AR structures Y227$^{5.58}$ is mutated to an alanine, however our construct contains this tyrosine (See Appendix A). An observation of the β$_2$AR crystal structures, comparing the inverse agonist carazolol bound state (2RH1 2.4Å resolution) (**Figure 3.18d**) and the full agonist BI-167107 and G$_s$-bound structure (3SN6 3.2Å resolution) (**Figure 3.18e**) shows marked changes in the position of the mentioned aromatic residues in relation to M223$^{5.54}$ and M296$^{6.41}$. In the NMR spectra of the ternary complex, M223$^{5.54}$ experiences a significant change in ^{1}H chemical shift, which can be explained by the increased proximity of aromatic residues (distances measured in Å are shown in yellow) to M$^{5.54}$. These observations link those obtained by NMR spectroscopy to the crystallographic evidence of conformation changes upon receptor activation.
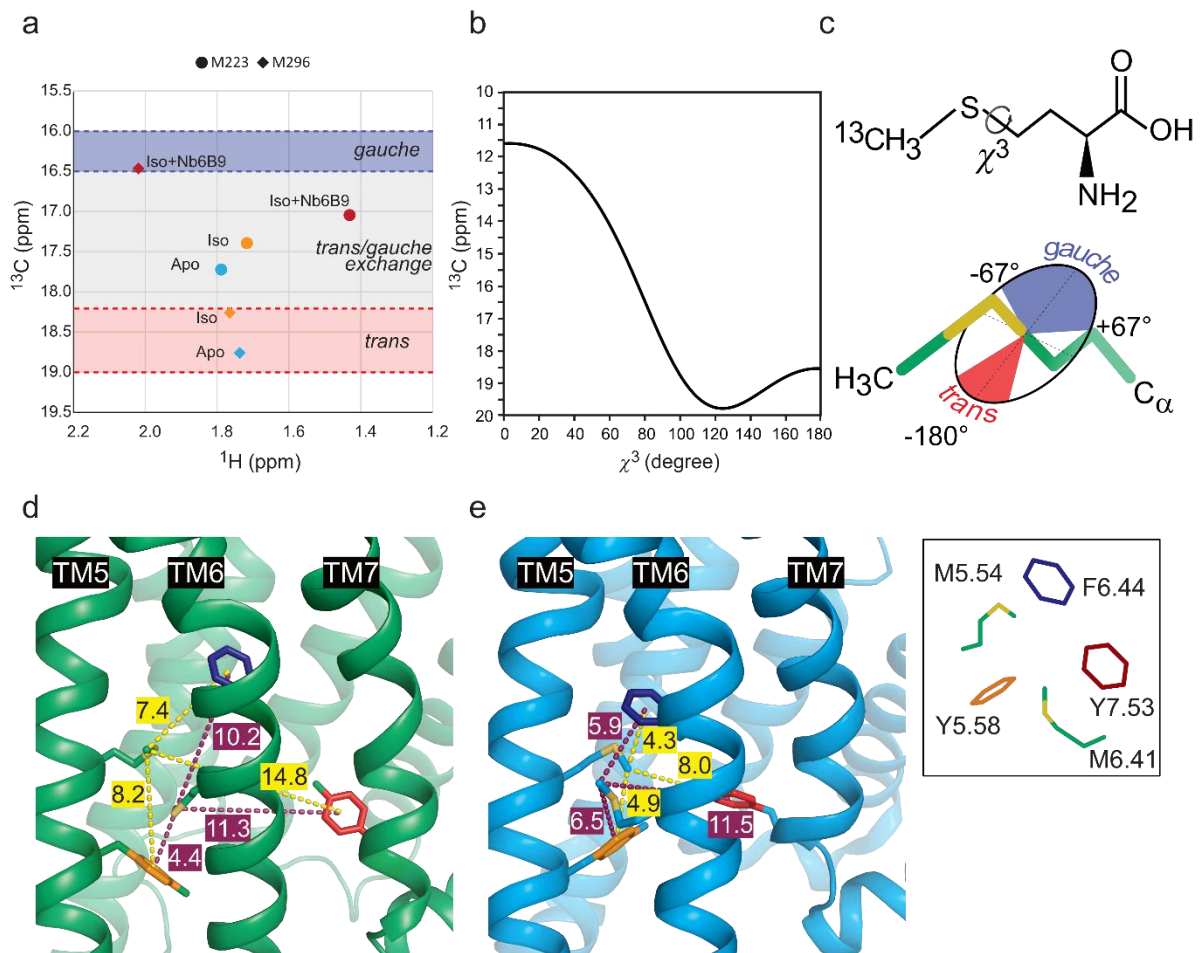
Figure 3.18 χ^3 rotamer associated chemical shift changes

M223 and M296 χ^3 rotameric changes and effects of nearby aromatic residues. Methyl ^{13}C chemical shift positions are sensitive reporters of χ^3 rotamer conformations, while ^1H shifts are influenced by nearby aromatic ring currents. (a) shows a [^1H, ^{13}C] correlation spectrum for residues M223 and M296 in various receptor states together with the approximate trans, trans/gauche exchange and gauche rotamer positions as expected from the ^{13}C chemical shift signature and as shown in (b). A simulation of the ^{13}C chemical shift change associated with the χ^3 rotamer conformation is shown in (b) as taken from Butterfoss et al (*250*). The possible χ^3 conformations are shown in (c) in the context of methionine side chains with the preferred gauche and trans regions highlighted. (d) shows the inverse agonist carazolol bound β_2AR (2RH1) while the active state full agonist BI167107 and G_s-bound β_2AR (3SN6) is shown in (e). The equivalent methionines of M215$^{5.54}$ (to M223$^{5.54}$) and M279$^{6.41}$ (to M296$^{6.41}$) are shown on both (d) and (e) in green and yellow. The aromatic residues which are in proximity to influence ^1H chemical shifts are Tyr$^{5.58}$ (orange), Phe$^{6.44}$ (blue) and Tyr$^{7.53}$ (red). In all available β_1AR crystal structures Tyr$^{5.58}$ is mutated to alanine, while it is present in Met2-Δ5. Figure adapted with permission from Solt et al. (127).

3.10 – Proposed model of β₁AR receptor activation

The following diagram condenses the results described into the framework of the ternary complex model of receptor activation.
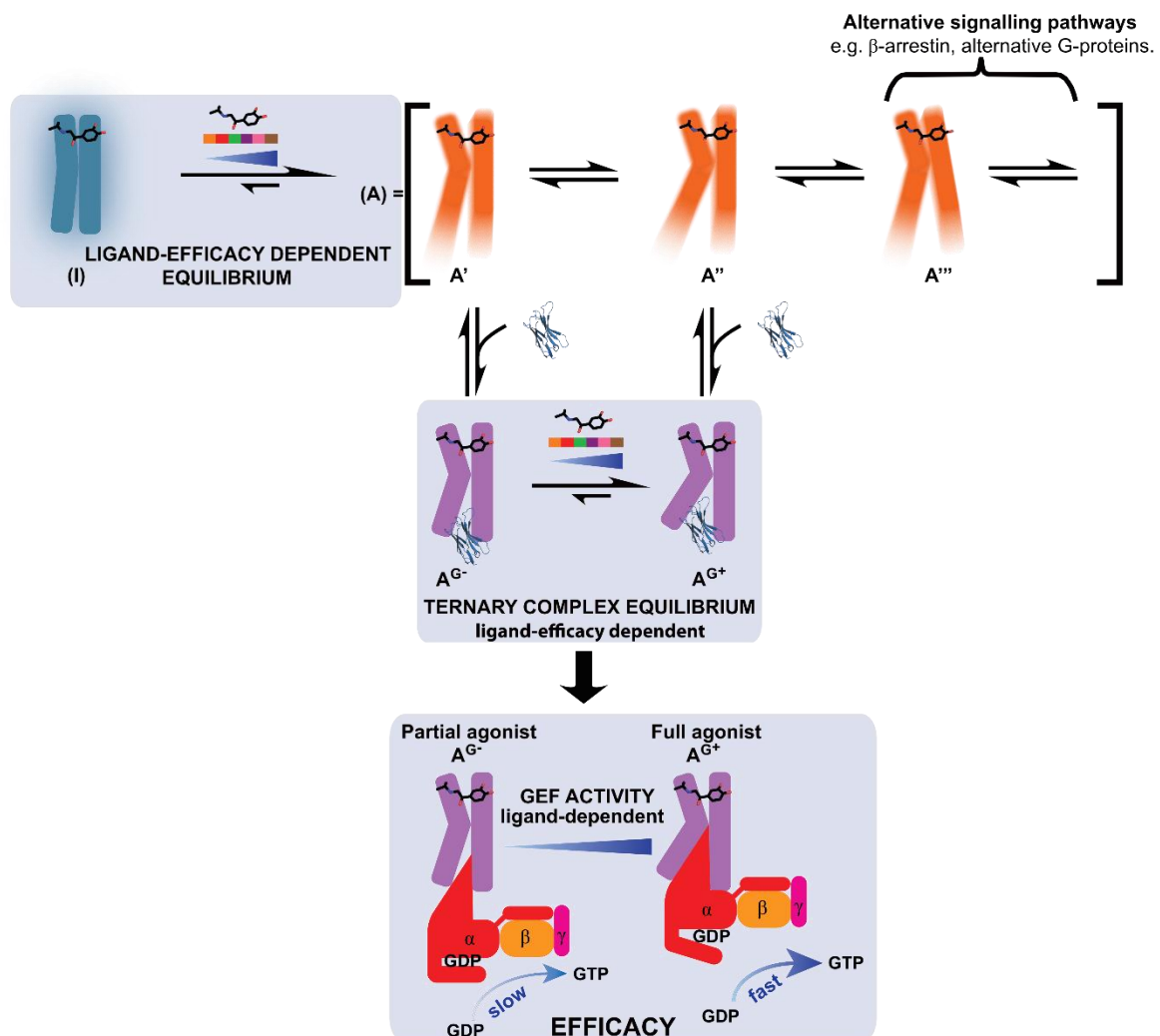
Figure 3.19 Proposed ternary complex model based

Ligand binding modulates equilibria between an inactive (I) state and a more active (A) state that is more predisposed to coupling to Intracellular Binding Partners (IBPs) such as G_s. Upon formation of a ternary complex, ligands modulate a second equilibrium between a more signalling competent and a less competent AG^+ and AG^- states. Figure adapted with permission from Solt et al. (127).

Firstly, a ligand efficacy dependent equilibrium was observed, which was particularly pronounced for residues $M223^{5.54}$ and $M296^{6.41}$ on the cytoplasmic side of the receptor. This establishes an equilibrium that takes place in the fast exchange regime in terms of the NMR timescale, between two states: an inactive state – termed (I) – and a pre-active state, the (A) state. In this manner, the (I) state is a

signalling incompetent, state, unable to bind G-protein. The pre-active (A) state represents an "active-like" conformation with a highly dynamic cytoplasmic side. The apo state would represent a conformational equilibrium predominantly in the (I) state. Ligand efficacy determines the shift in this equilibrium, and this is indicated by the averaged chemical shift of the population distribution across the (I) and the (A) states.

Peak intensity analysis as a proxy to estimating μs-to-ms scale exchange dynamics revealed a highly exchange broadened full agonist-bound state and a series of less dynamic partial agonist bound states. The recovery of signal intensity at lower temperatures in the full agonist state showed that in addition to the fast exchange (I) $\rightleftharpoons$ (A) equilibrium, governed by ligand efficacy, the pre-active state undergoes additional exchange that is closer to the NMR timescale in its exchange rate. This means that the pre-active state represents a highly dynamic (on the μs-ms timescale) conformational ensemble, sampling a number of active-like states. These are denoted as (A'), (A") and (A'''), which may be predisposed to coupling to different intracellular binding partners (IBPs). This implies that the ability of the pre-active state to explore different conformations may be significant for the receptor to engage with different IBPs. Lower efficacy ligands or the apo state maintain a dynamic receptor ensemble, but this is more rigid in comparison to the full agonist complex.

The binding of nanobody on the intracellular side caused significant chemical shift and intensity changes throughout the receptor, extending all the way to the cytoplasmic side. This demonstrates the allosteric outward intramolecular signalling effect of IBP coupling, which contributes to the increase in orthosteric ligand affinity (*87*). A similar inwards allosteric effect has also been observed, where in the absence of agonist nanobody affinity was substantially reduced in comparison to the reported nM affinity (*84*).

Nanobody and ligand bound ternary complexes have revealed a second fast exchange equilibrium modulated by ligand efficacy. Two predominant models have been described for ligand dependent efficacy (*251*, *252*). Kobilka & Deupi propose that partial agonists induce distinct states in the ternary complex which are different from the inactive and fully active states and achieve a lower level of activity than full agonists (*251*). Lape *et al.* suggest that partial agonists modulate an equilibrium between an inactive and an active ternary complex form (*252*).

The results presented here show that partial agonists govern both the equilibria between the inactive and pre-active states, as well as between a less active ternary state (A^{G-}) and a fully active state (A^{G+}), both in fast exchange. In this form, the (A^{G-}) conformation is represented by the ligand free nanobody bound conformation, i.e. the basal activity complex. Signal intensity analysis has shown these ternary

complexes to be more rigid than the basal active state, suggesting a more conformationally restricted state (though not absent of dynamic structural changes) appropriate to a given IBP interaction.

The phenomenon of basal activity is accommodated in this model though the assumption that the nanobody-only bound receptor corresponds to a very weak (A^{G-}) conformation, owing to the good fit of the chemical shift of this state to the ternary complex efficacy correlation. This implies that the apo form, predominantly in the (I) state must sample the (A) state at least a fraction of the time, to allow complex formation to take place. This would explain the significant nanobody affinity reduction in the absence of orthosteric ligand.

Hence this model takes into account the phenomena of partial agonism and basal activity, as well as the additional exchange parameters of the active-like state. Accordingly, ligand efficacy is converted into a distinct conformational output on the cytoplasmic side of the receptor, affecting the quality and nature of interaction with the IBP. In this way it could be envisaged that orthosteric ligands can regulate the guanosine-nucleotide exchange factor function of Gαs through modifying the quality of this interaction, thereby directly modulating the downstream signalling output.

3.11 – The proposed model in the context of other NMR investigations of GPCRs

To date, six distinct GPCRs have been studied by NMR spectroscopy: the $A_{2A}AR$, β_1AR, β_2AR, BLT2, NTS1 and µOR, in a range of detergent and lipid nanodisc settings. These works are summarised below in **Table 3.4**.

Table 3.4 List of NMR studies of GPCRs

Table adapted with permission from Bostock, Solt, Nietlispach, COSB, 2019 (*253*).

Receptor	Expression system	Solubilisation system	Labelling	NMR experiments	Reference
$A_{2A}AR$	*P. pastoris*	LMNG	^{19}F-BTFMA	^{19}F; 1D, STD	(*254*)
$A_{2A}AR$	*P. pastoris*	LMNG	^{19}F-BTFMA; metal ions	^{19}F, ^{23}Na$^+$, ^{25}Mg$^+$; 1D, CPMG	(*255*)
$A_{2A}AR$	*P. pastoris*	LMNG/CHS	U-^{15}N, ~70% ^{2}H	^{1}H,^{15}N TROSY	(*256*)
$A_{2A}AR$	*P. pastoris*	LMNG/CHS	U-^{15}N, ~70% ^{2}H	^{1}H,^{15}N TROSY	(*257*)
$A_{2A}AR$	*P. pastoris*	DDM	^{13}CH$_3$ Ile δ1 / ^{2}H	^{1}H,^{13}C HMQC, 3Q-relaxation	(*258*)
β_1AR	Sf9 or Sf21	LMNG	^{13}CH$_3$-Met	^{1}H,^{13}C-HMQC	(*259*)
β_1AR	High five	DM	^{15}N-Val	^{1}H,^{15}N HSQC	(*260*)
β_2AR	β_2AR: expressSF+ C-terminal tail: *Escherichia coli*	POPC/POPG nanodiscs	β_2AR [^{2}H-9AA, αβγ^{2}H-,^{13}CH$_3$-Met] C-tail: U-[^{2}H, ^{13}C, ^{15}N] or ^{13}CH$_3$ Thr γ2 & Ile δ1	^{1}H,^{13}C-HMQC; ^{1}H,^{15}N HSQC; cross-saturation	(*261*)
β_2AR	Sf9	DDM	^{13}CH$_3$-Met	^{1}H,^{13}C HSQC	(*262*)
β_2AR	Sf9	DDM/CHS	^{19}F-TET	^{19}F; 1D	(*263*)
β_2AR	expressSF+	DDM	^{13}CH$_3$-Met or α,β,β-^{2}H$_3$-,^{13}CH$_3$-Met	^{1}H,^{13}C-HMQC	(*264*)
β_2AR	expressSF+	POPC/POPG nanodiscs	^{2}H-9AA, αβγ^{2}H-, ^{13}CH$_3$-Met	^{1}H,^{13}C-HMQC	(*265*)
β_2AR	Sf9	LMNG	^{19}F-BTFA	^{19}F; 1D, T_1, T_2	(*266*)
β_2AR	Sf9	DDM/CHS	^{19}F-TET	^{19}F; 1D, 2D EXSY	(*267*)
β_2AR	Sf9	DDM/CHS or LMNG	^{19}F-3-bromo-1,1,1-trifluoroacetone	^{19}F; 1D, T_1, T_2	(*268*)
β_2AR	Sf9	DDM	^{13}CH$_3$-Lys (reductive methylation)	STD-filtered ^{1}H,^{13}C HMQC; ^{1}H,^{13}C HSQC	(*135*)
β_2AR	Sf9	LMNG	^{19}F-BTFA	^{19}F; 1D, CPMG, STD	(*269*)

BLT2	*E. coli*	DMPC/CHS nanodiscs	u-^{2}H, ^{12}C Ile-[δ1-^{13}CH$_3$], ^{12}C Met-[ε-^{13}CH$_3$]	^{1}H,^{13}C-HMQC	(*270*)
NTS$_1$	*E. coli*	DDM	^{13}CH$_3$-Met	^{1}H,^{13}C-HMQC	(*271*)
µOR	Sf9	LMNG/CHS	^{13}CH$_3$-Lys (reductive methylation)	^{1}H,^{13}C-HMQC	(*272*)
µOR	Sf9	LMNG/ CHS	^{2}H-8AA, αβ-^{2}H-^{13}CH$_3$-Met	^{1}H,^{13}C-HMQC	(*273*)

The pathways of receptor activation and interactions with ligand or IPBs are explained as a series of equilibria forming the ternary complex model (*109*, *274*). This is illustrated below in **Figure 3.20**.
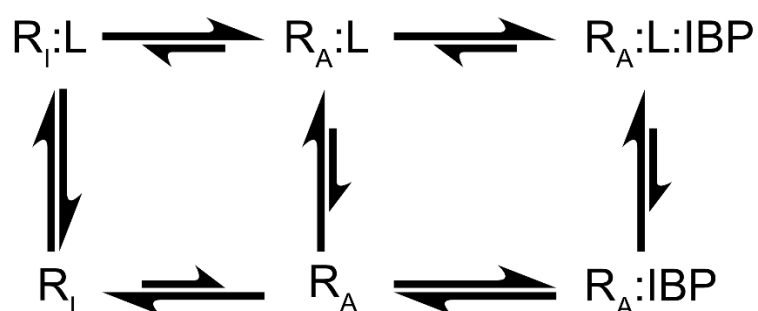
Figure 3.20 Ternary complex model: a series of equilibria

Abbreviations: R: receptor (A: inactive, I: active), L: ligand, IBP: intracellular binding partner.

The model proposed in this thesis, described in **Figure 3.19**, offers an explanation for partial agonism in the way the equilibria is modulated by ligand efficacy between the R$_I$-R$_A$:L equilibrium and between the R$_A$:IBP-R$_A$:L:IBP equilibrium. Observations of efficacy correlated and ligand induced changes in conformation have been observed in other studies of the β$_2$AR and β$_1$AR (*121*, *275*). Such observations have been made for V226$^{5.57}$ (β$_1$AR, backbone amide, **Figure 3.21d**) and for M82$^{2.53}$ (β$_2$AR, methionine methyl, **Figure 3.21b**). These results revealed a fast exchange equilibria on the NMR timescale, while studies of the β$_2$AR in high density lipoprotein (rHDL) nanodiscs, showed a slower exchange rate on the order of milliseconds (*241*). Similarly, exchanging conformations were also observed on the cytoplasmic side of β$_2$AR studied by ^{19}F tagged cysteines on TM6 and TM7; here the population of two peaks, assigned to an more inactive or more active receptor state, were modulated by the efficacy of the bound ligand (**Figure 3.21e**) (*276*).
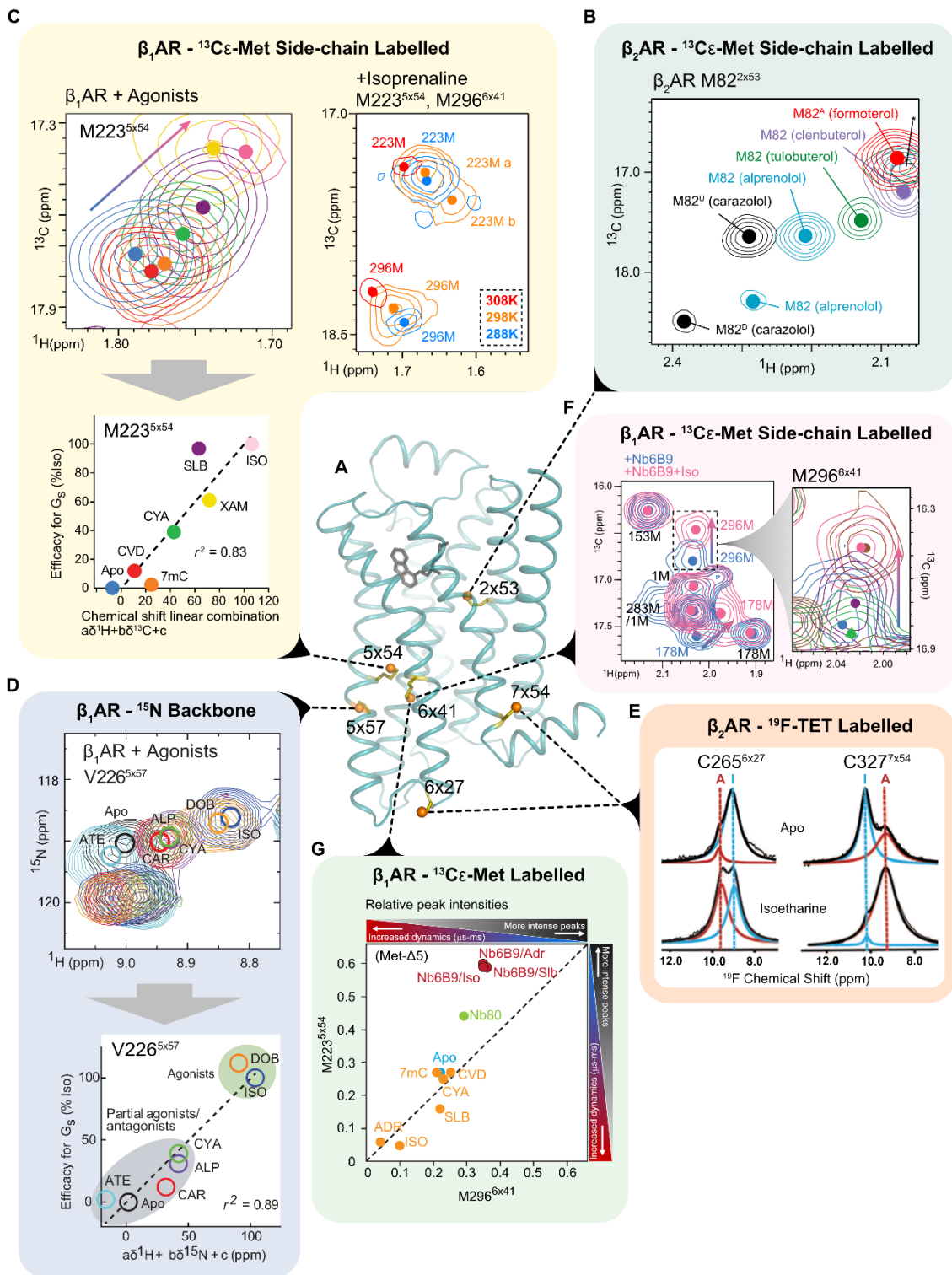
Figure 3.21 Ligand induced conformational changes observed by NMR spectroscopy

A) The region of a generic GPCR showing the location of reporters studied by each case study (B-G). Ligand abbreviations: DOB (dobutamine), ISO (isoprenaline), CYA (cyanopindolol), ALP (alprenolol), SLB (salbutamol), CAR (carvedilol), ATE (atenolol), 7mC (7-methylcyanopindolol), CVD (carvedilol), XAM (xamoterol). Figure adapted with permission from Bostock, Solt, Nietlispach, COSB, 2019 (*253*).

Various IBP mimetic proteins have been used to study the structural changes induced by IBP binding to GPCRs. These include nanobodies (e.g. Nb6b9, Nb80, Nb60, Nb35 or the G_α-c-terminal peptide). While small ligand specific changes have been observed in ternary complex structures, they are overall very similar (*84*, *277–279*).

NMR studies which observed ternary complex states have all converged on noting large chemical shift changes between a ligand bound and a ternary complex state. One of the major contributions of the work presented in this thesis is the observation of a ligand efficacy modulated equilibrium in the ternary complex state, which was hitherto not demonstrated. This demonstrated a mechanism for modulating the nature and quality of the receptor-IBP interaction, which may have repercussions with regards to downstream signalling.

Future studies building on these findings should expand into looking at the effect of the membrane environment in modulating the dynamics and equilibria of receptor conformations.

Chapter 4 – Preparation of ScFv C5, an anti-TcP2b antibody with agonistic immune cross-reactive properties to β₁AR

4.1 – Aim

The aim of the work below was to clone, express, purify and evaluate the ScFv C5 construct, an ScFv version of the mAb 17.2 antibody discussed in the Introduction chapter. The longer-term outlook of this work seeks to use the prepared antibody as a tool in evaluating the activation mechanism of the β_1AR in the absence of an orthosteric small molecule agonist in the context of Chagas' disease and in broader terms too.

4.2 – Construct Design

The construct was synthetically ordered (Biomatik) based on the published sequence and in consultation with Dr. Karina Gomez (INGEBI CONICET, Buenos Aires, Argentina) (*280*). The published protein sequence is shown below:

```
EVQLEESGGRLVQPKGSLKLSCAASGFSFNTNAMNWVRQAPGKGLEWVARIRSKINNYSTYYADSVKD
RFTISRDDSQSMLYLQMNNLKTEDTAMYYCVRGTTYWGQGTLVTVSGGSGGELVMTQSPLTLSVTIGQ
PASISCKSSQSLLDSDGKTYLNWLLQRPGQSPKRLIYLVSKLDSGVPDRFTGSGSGTDFTLKISRVEA
EDLGVYYCWQGSHFPYTFGGGTKLEIKRAAAHHHHHHGAAEQKLISEEDLNGAA
```

Segments underlined and in bold indicate the 6 CDR sequences. The VL and VH chains are connected by a 5 residue GGSGG sequence (green). The construct is tagged by a His-6 (red) and a Myc tag (blue).

The construct was to be cloned into the pMAL1-Nb vector, with the aim of using a maltose binding protein (MBP) tag for affinity purification as well as for its periplasmic export sequence. The oxidising conditions of the gram negative bacterial periplasm in contrast to the generally reducing environment of the cytoplasm allows for the formation of necessary disulphide bonds. Together with the lower levels of protease activity, the periplasm constitutes an ideal compartment for the soluble expression of antibody fragments (*281*). The pMAL1-Nb vector has been used to increase the yield of poorly expressing nanobodies (*282*).

The synthetic construct was codon optimised for *E. coli* expression and was designed with an N-terminal HRV 3C (Prescission protease) cleavage site, a C-terminal reversed TEV cleavage site and a His-8 tag, flanked by recognition sequences for the restriction endonucleases *Eco*RI and *Hin*dIII at the 5' and 3' ends (N and C termini) respectively.

The final construct produces a 71.8 kDa protein, where the MBP fusion is linked to the ScFv construct via a poly-N linker, an enterokinase and a HRV 3CP cleavage site. The intention is to produce the ScFv through 3CP cleavage, producing a 43.68 kDa MBP fragment and a 28.13 kDa ScFv fragment as shown below (**Figure 4.1**).
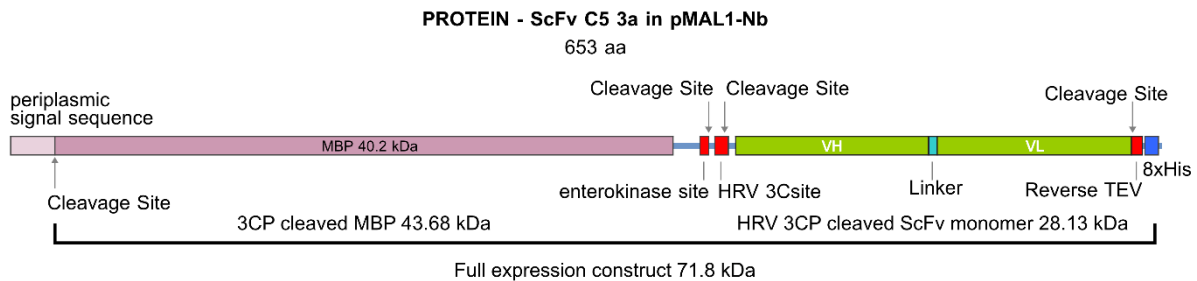
Figure 4.1 Schematic representation of the ScFv C5 MBP fusion construct

The final sequence of the synthetised ScFv segment is shown below.

```
LEVLFQGPGGSGEVQLEESGGRLVQPKGSLKLSCAASGFSFNTNAMNWVRQAPGKGLEWVARIRSKIN
NYSTYYADSVKDRFTISRDDSQSMLYLQMNNLKTEDTAMYYCVRGTTYWGQGTLVTVSGGSGGELVMT
QSPLTLSVTIGQPASISCKSSQSLLDSDGKTYLNWLLQRPGQSPKRLIYLVSKLDSGVPDRFTGSGSG
TDFTLKISRVEAEDLGVYYCWQGSHFPYTFGGGTKLEIKRAAGQFYLNEAHHHHHHHH*
```

4.3 – Expression of the MBP fusion construct

Bacterial protein expression can yield high quantities of heterologously expressed protein, which may either be in a soluble form, or packed into inclusion bodies in an often denatured, insoluble form. The latter requires refolding, which may or may not be feasible. Soluble expression can take place in either the cytoplasm or the periplasm as mentioned above. When a protein is synthesized by the ribosome it is the conditions experienced by the nascent chain which determine its propensity for adopting the correct folded conformation we refer to as the soluble form. If conditions of the environment such as pH, temperature, hydrophobicity, ionic potential, presence of chaperones, redox potential, presence of post-translationally modifying enzymes as examples, are not ideal there is a chance for misfolding and aggregation resulting in insoluble expression. In *E. coli* such proteins are packed into inclusion bodies, as a mass of high-density insoluble protein. In order to establish conditions that maximise the soluble yield over insoluble expression for a given construct, parameters such as temperature, length of expression, media, cell type and induction method can be varied.

To find the optimal temperature for soluble expression, 0.5 L autoinduction expression cultures were grown in superbroth at 18 °C, 25 °C and at 30 °C, using MH140 *E. coli* cells. The pellet from 1 mL of culture was used to evaluate the level of soluble expression (**Figure 4.2**). A western blot of the soluble

fraction shows that all three temperatures yielded soluble expression, confirmed by the presence of a c. 70 kDa band (**Figure 4.2A**). It also appears that the size of this band is smaller in the 30 °C culture compared to the 18 °C and 25 °C cultures, however this assumes an identical efficiency of lysis and equal cultural densities at the time of harvest for all three samples. In fact the densities at harvest measured at 600 nm were 4.95 AU (18 °C), 2.6 AU (25 °C) and 2.6 AU (30 °C), meaning that the overall yield was the greatest at 18 °C, and the amount of protein expressed per cell may not have been maximal at this temperature.

As discussed above, undesired insoluble expression takes place at the expense of soluble expression. To evaluate the ratio soluble to insoluble expression at 25 °C, the insoluble pellet left over after cell lysis was solubilised with 10% SDS and visualised on a western blot (**Figure 4.2B**).
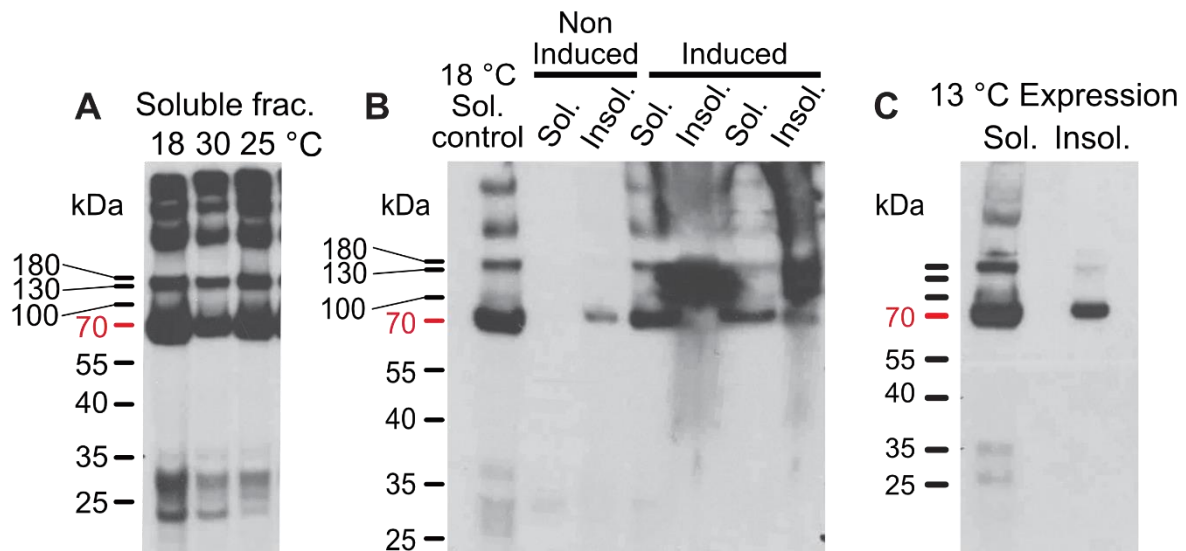
Figure 4.2 ScFv-C5-MBP expression analysis at different temperatures

Anti-His western blot of MBP fused ScFv C5 expression at 18 °C, 30 °C, 25 °C and at 13 °C. (A) shows the result of probing the soluble fractions of expression at different temperature. (B) shows the evaluation of soluble and insoluble expression at 25 °C compared to the amount of soluble expression at 18 °C. (C) shows that the level of insoluble expression is reduced at 13 °C compared to 25 °C.

This shows significant quantities of his-tagged material in the insoluble fraction. In order to reduce this, expression was trialled at 13 °C, for 40 hours. This has achieved successful expression for an EGFR transmembrane construct, which only produced inclusion bodies at higher temperatures (*283*). Reducing the expression temperature to 13 °C significantly lowered the amount of his-tagged protein in the insoluble fraction, while the 70 kDa band in the soluble fraction remained comparable to that at higher temperatures (**Figure 4.2C**).

It can therefore be concluded that while significant levels of soluble expression can be achieved at temperatures up to 30 °C, the amount of insoluble protein can be minimised at little expense of soluble expression by reducing the temperature to 13 °C and extending the length of culture to over 40 hours.

4.3.1 – Purification of the MBP fusion construct

Having established the presence of a his-tagged band of the expected size on a western blot, the 0.5 L cultures expressed at 18 °C and 25 °C were processed for purification. An initial Nickel affinity chromatography step using a 1 mL HisTrap FF column with a load rate of 0.5 mL/min did not show any bound protein to the column upon elution with up to 250 mM of imidazole.

This was unexpected as when designing the MBP fusion expression construct, the available crystallographic data was considered. The Fab format of mAb 17.2 has been crystallised by Pizzaro *et. al*. and this allows us to consider the secondary structure of the C-terminus of the ScFv C5 construct.
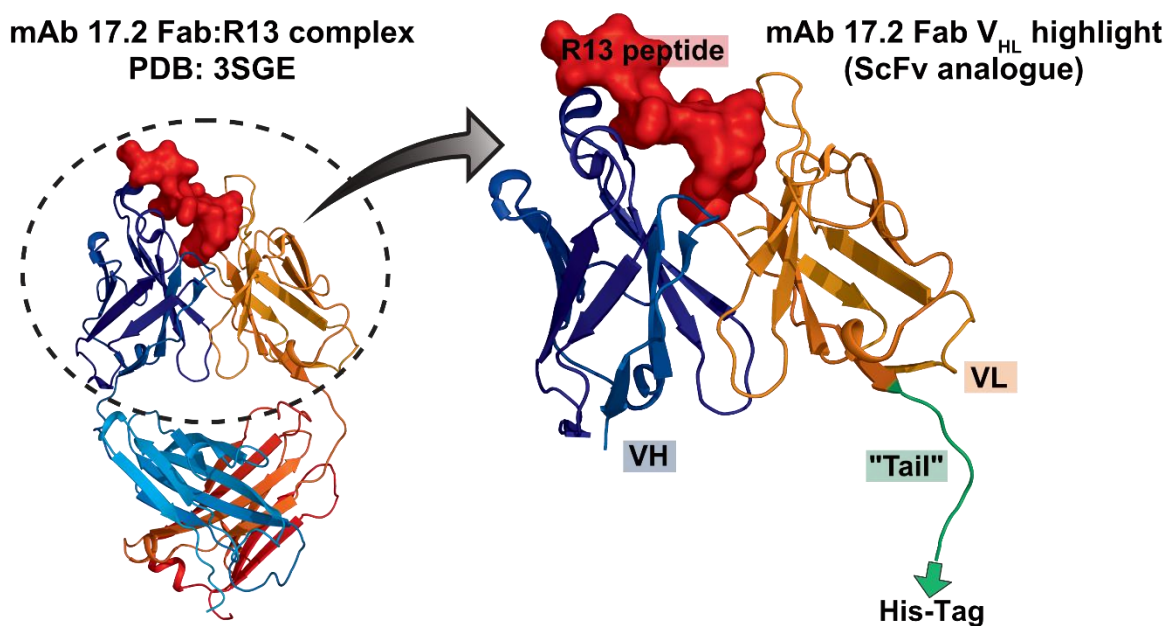
Figure 4.3 ScFv C5 structural model

Structure of Fab 17.2-R13 peptide complex (3SGE), with the VH-VL region in focus (dashed circle). An *in silico* model of the region forming the ScFv region is shown on the right, extracted from 3SGE. The R13 peptide is shown in space fill, the V_H region in blue, V_L in orange and the tail region forming the ScFv C-terminus or the linker to the C_L in the Fab is shown in green.

As shown above, a highlight of the variable regions of Fab 17.2 in complex with the R13 peptide, the variable light (VL) region which connects to the constant light (CL) domain adopts no defined secondary structure in the region of sequence which is still present in our ScFv C5 construct (**Figure 4.3 – ScFv model**). This is shown as a green "tail" on the right. It would therefore make sense to expect an affinity tag sequence placed at this C-terminus to be exposed for purification.

The observation of poor or no binding to the nickel affinity column, in the face of a detectable anti-his signal on a western blot, must mean that the tag is somehow obscured or its binding to the resin sterically hindered in solution. If this obstruction is the result of a reversible process, and that the tag exists in an equilibrium between an obscured and an available form, then if the loading rate onto the packed column happens at a faster rate than the transition between these two states, then a reduction in the flow rate and increase in the on-column residence time may increase the binding efficiency. This could make a nickel affinity step a viable form of purification for this construct.

To get around the limitations posed by the apparent unavailability of the his-tag as a means of purification the maltose binding property of the MBP fusion partner can be exploited for purposes of affinity chromatography. Consequently the HisTrap flow through was reserved and purification was attempted using the MBP tag using a 1 mL capacity MBPTrap column, which consists of packed dextran Sepharose beads, using 20 mM maltose as the eluting agent.
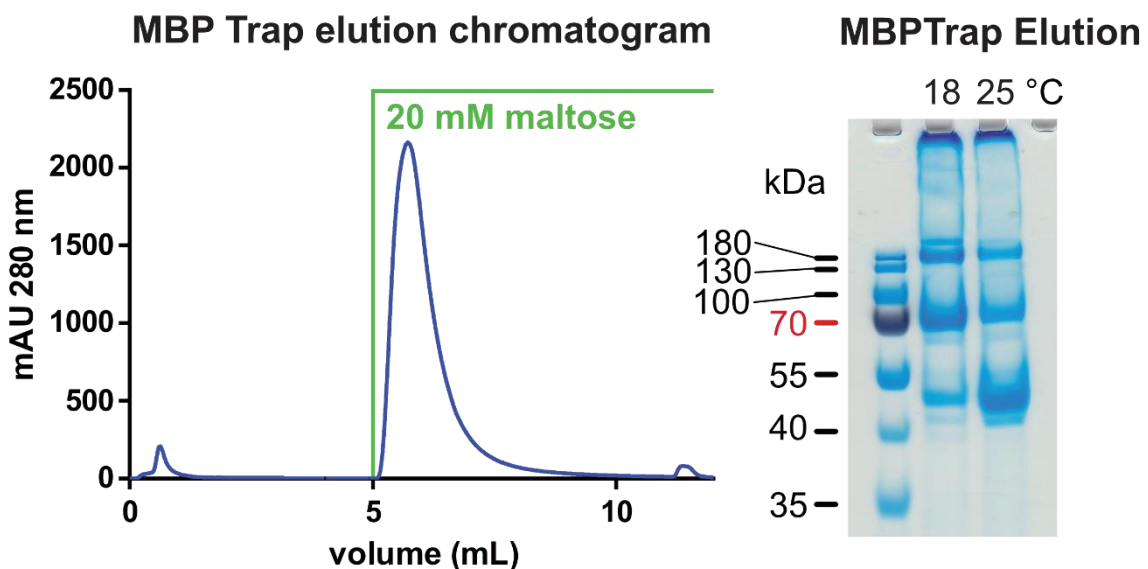
Figure 4.4 MBPTrap purification of the ScFv C5 expressed at 18 °C.
The eluted fractions from expressions ate 25 °C and at 18 °C are shown on an SDS-PAGE gel.

This yielded a significant elution peak for both the 18 °C and the 25 °C culture. Typical of affinity chromatograms using stepwise elution, the elution profile shows a single peak, which contains species of multiple molecular weights (**Figure 4.4**). As mentioned earlier, the fusion construct corresponds to a 71.8 kDa band, which appears present in both purifications. In addition to this band there is a strong band appearing between the 40 and 55 kDa markers, which likely corresponds to MBP with some of the linker sequence, which has a minimal size of 40.2 kDa. This suggests that for some copies, the protein translation terminated prior to the ScFv sequence, producing significant quantities of MBP on its own.

In order to separate these different molecular mass proteins, size exclusion chromatography (SEC) (or gel filtration – GF) was used as a second step purification. The eluted fraction from the MBPTrap column was concentrated and loaded onto a Superdex S200 Increase column. The resulting chromatogram and gel is shown below (**Figure 4.5**).
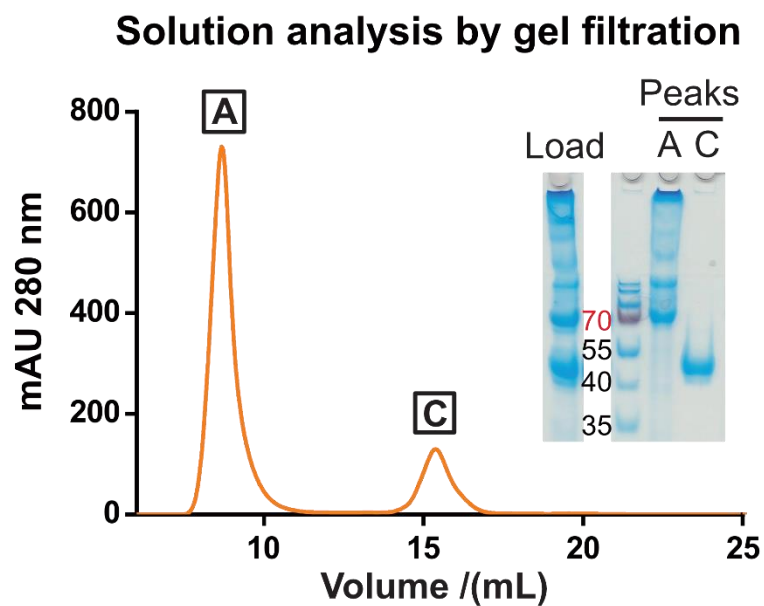
Figure 4.5 Size exclusion chromatography of the MBPTrap purified ScFv-MBP construct. The load fraction as well as peaks A and C are shown on an SDS-PAGE gel.

Peak A elutes at 8.69 mL, which is shortly after the void volume of 8.465 mL for the S200 10/300 Increase column. This corresponds to a molecular weight of 1048 kDa, likely representing large molecular weight aggregates or oligomers of mixed species. The corresponding gel of peak A also indicates a mix of very high molecular weight proteins. Peak C elutes at 15.36 mL, representing a 46.8 kDa size molecule. This likely corresponds to MBP alone with some of the linker sequence between MBP and the ScFv, corresponding to a size of about 43 kDa. The measured difference may arise from

the unknown position of cleavage/truncation, or from the fact that SEC only takes into account the hydrodynamic volume of proteins, not their molecular mass.

Using MBP as an affinity tag exploits a very specific protein-protein interaction, therefore the relative impurity of the resulting eluate is likely due to truncated expression, proteolysis or oligomerisation. In order to purify the intended ScFv construct from this mixture, it had to be cleaved from the MBP fusion partner. The designer construct affords this opportunity either through an enterokinase site or a 3CP site. The latter is considered more specific, cleaving at the LEVLFQGP sequence between the glutamine and glycine residues, therefore 3CP was chosen.

3CP, like TEV is a cysteine protease and it its activity requires a reducing environment (*284*). This can be achieved through the addition of reducing agents such as DTT or TCEP into the cleavage reaction mixture. The mixture eluted from the MBPTrap column was subjected to parallel cleavage reactions, using 3CP and 2 mM DTT, 3CP and 2 mM TCEP and just reducing agents without enzyme. 25 µg of 3CP was used per reaction, which represented a large excess in terms of minimal recommended amounts of commercial PreScission protease (GE) for the given amounts of substrate. The cleavage was monitored by resolving the end product mixture on an SDS-PAGE gel and by gel filtration (**Figure 4.6**).
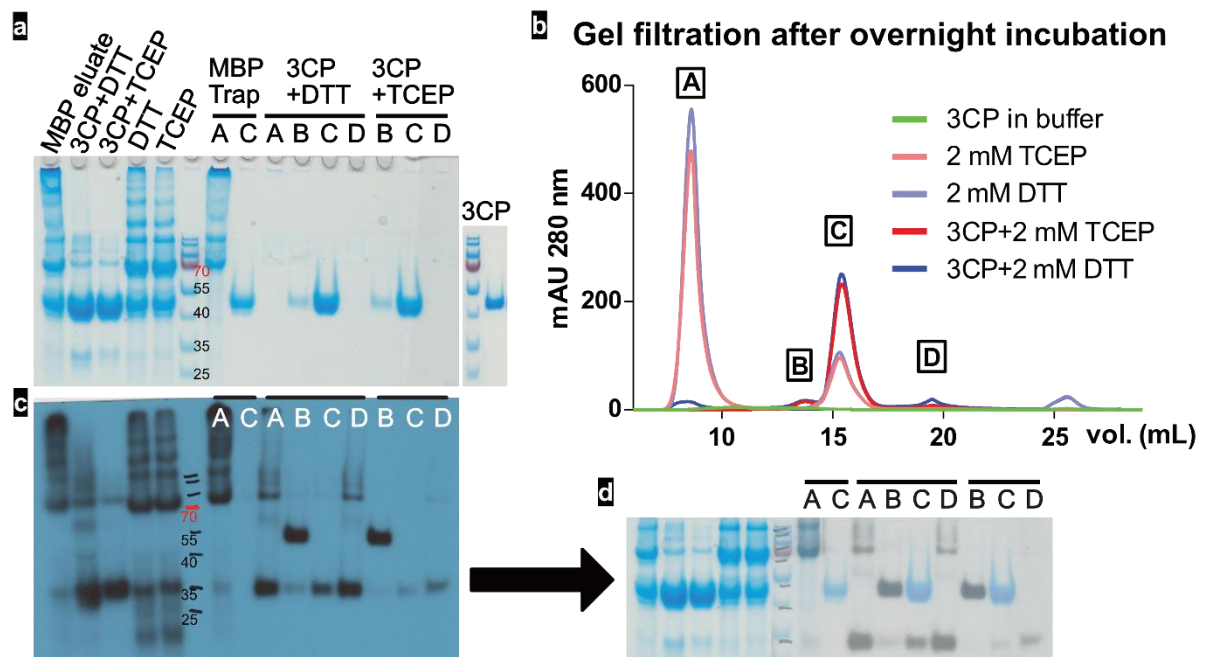
Figure 4.6 Cleavage of the MBP fusion ScFv C5 construct

SDS-PAGE, and western blot of SEC peaks obtained from ScFv-MBP cleavage by 3CP protease. (a) SDS-PAGE gel of the pre- and post-cleavage products as well as of the various SEC peaks shown on the right-hand side chromatogram (b). (c): western blot is of the SDS-PAGE (a). (d) shows a superposition of the SDS-PAGE and western blots for

The SDS-PAGE profile of the cleavage reactions shows that the presence of the enzyme changes the profile of the MBPTrap eluate starting mixture, with the noticeable disappearance of very high molecular weight bands. The mixture resolves to a major 40-55 kDa band corresponding to the 43.68 kDa MBP product and a minor band at 25-35 kDa, likely matching to the 28.13 kDa ScFv product. The result was very similar for both 3CP cleavage reactions with DTT and TCEP. Reactions omitting the protease show no change in the gel profile compared to the starting material, indicating that the reducing agents alone do not cause a significant change in the species present. A his-probe western of the same samples shows no band that corresponds to the presumed MBP product, which serves to verify its identity as that portion of the construct lacks a his-tag. There is however a band under 35 kDa, which is likely the ScFv product, containing the his-tag.

The four reactions were then analysed for their composition using gel filtration, with the aim of separating the cleaved ScFv from the MBP part. This is shown above; the four reactions were concentrated and separated on an S200 10/300 Increase column, alongside an equal amount of 3CP stock to that used in the reactions. This is shown by the green chromatogram, displaying an almost negligible amount of protein in comparison to the other traces. The traces of the DTT and TCEP only mixtures (pink and lilac) are identical to the chromatogram of the MBPTrap eluate trace shown earlier (**Figure 4.5**), with two predominant peaks A and C. These are not shown on the gels above (**Figure 4.6b**) as they were identical to the MBPTrap eluate peaks A and C as shown before. The cleaved end-product once 3CP is present shows a different profile not just on the gel but also on gel filtration (red and blue traces). Peak A has almost completely disappeared for both the 3CP+DTT and 3CP+TCEP reactions, which corresponds well to the abrogation of large molecular weight bands on the gel. Peak C has increased considerably, which matches the large amount of MBP fragment generated. This strong band does not show up on the western, and this is easily visible on the composite gel-western overlay figure on the bottom right. In addition to these changes to peaks A and C, two small peaks, B and D also appear. B shows up as a faint band on the gel and appears as a very strong band on the western. This is very likely to be 3CP, which also has a his-6 tag. The position of 3CP is shown above on the small gel labelled 3CP, aligned with the molecular weight markers of the adjacent gel. This shows that 3CP runs at an identical position as peak B. Band D does not present on the SDS-PAGE gel, but it shows as a 25-35 kDa band on the western. It elutes at 19.21 mL, translating to an estimated molecular weight of 7.7 kDa. To put this into comparison, salts and other small molecules elute at about 22 mL on this column, hence peak D either represents an unfolded state with a non-globular

conformation, which may explain the discrepancy between the size on the gel and western and the size in solution. The same band between 25-35 kDa also appears on the western for peaks A and C, and it matches the band previously noted in the cleavage product mix (on the left hand side of western blot – **Figure 4.6c**). It is likely therefore that this band is the ScFv fragment, however it is completely missing from the gel in the quantities expected based on the amount of MBP fragment present. This was puzzling as it can be seen from the gels and westerns (overlaid composite image) the ScFv fragment is present in peak A and in the cleavage end-mixture, but disappears in observable form during gel filtration.

Reducing agents such as DTT or TCEP are essential for the activity of cysteine proteases like 3CP. This is shown below (**Figure 4.7**), as 3CP in the absence of any reducing agent does not result in cleavage, and the size of peak A remains unchanged. While reducing agents are crucial for protease cleavage, they may adversely affect crucial disulphide bonds in other proteins. Hence it is a possibility that the vanishing act of the ScFv cleavage fragment may be due to the susceptibility of the 2 disulphide bonds present within the molecule in the presence of 2 mM TCEP or DTT. An alternative method of achieving the right conditions for cleavage is using a redox shuffle system, where oxidised and reduced versions of the same molecule are supplied together, creating a tuneable redox environment. An example of such a system is based on the oxidised (GSH) and reduced (GSSG) form of glutathione. Redox shuffle systems can also be used to rearrange disulphide bonds in misfolded proteins. The ratio of the two redox forms must be empirically optimised to achieved the desired conditions, as the redox potential of the system depends on both the concentration and ratio of the two species (*285*, *286*).

A cleavage reaction with an equimolar ration of GSSG/GSH at 0.3 mM produced identical results to cleavage in the absence of reducing agents. A ration of 10:1 GSSG to GSH (3 mM to 0.3 mM) however successfully repeated a chromatogram observed for cleavage with DTT and TCEP as shown earlier. This reduced peak A, and produced a peak C (identical to previous peak Cs) and a significant peak D (**Figure 4.7c**). The corresponding gel shows the emergence of a 25-35 kDa band in the post cleavage reaction, but yet again this band does not appear in either peaks A, C or D (**Figure 4.7b**). The western blot corresponding to the post cleavage band confirms the presence of a his-tagged product, making it highly likely that this is the 28 kDa ScFv C5 fragment. Furthermore, when compared to the western blot lane of the MBPTrap eluate, this band is missing, implying that it was released as the consequence of cleavage. The nature of peak D was uncertain, as it didn't show up as any material on the gel, not even at the very low molecular weight section (cropped). GSSG was the only new component present in larger quantities in comparison to the 1:1 GSSG/GSH reaction. It is not known if GSSG absorbs at 280 nm.

At this point the identity of the 28 kDa band (shown below in dashed square) was verified by mass spectrometry fingerprinting against the expected protein sequence (**Figure 4.7d**).

a **3CP cleavage without reducing agent**

b **1:10 GSSG:GSH Cleavage Results**

c **3CP cleavage with GSH redox system**

— 10:1 GSSG/GSH
— 1:1 GSSG/GSH

d **MS Fingerprinting of candidate band**

```
GPGGSGEVQL EESGGRLVQP KGSLKLSCAA  30
SGFSFNTNAM NWVRQAPGKG LEWVARIRSK  60
INNYSTYYAD SVKDRFTISR DDSQSMLYLQ  90
MNNLKTEDTA MYYCVRGTTY WGQGTLVTVS 120
GGSGGELVMT QSPLTLSVTI GQPASISCKS 150
SQSLLDSDGK TYLNWLLQRP GQSPKRLIYL 180
VSKLDSGVPD RFTGSGSGTD FTLKISRVEA 210
EDLGVYYCWQ GSHFPYTFGG GTKLEIKRAA 240
GQFYLNEAHH HHHHH*               256
```

Figure 4.7 ScF-MBP fusion cleavage with alternative reducing agents

3CP cleavage in the absence of reducing agent and with the glutathione redox shuffle system. (a) and (c) shows corresponding gel filtration chromatograms. SDS-PAGE and anti-his western blots of the redox shuffle cleavage reaction are shown in (b). Mass Spectrometry fingerprinting confirms the presence of cleaved ScFv C5 (d).

The red segments of sequence shown above represent peptides that have been detected from the reference sequence. To estimate the certainty of the match protein scores are given to peptide matches against a database, where a calculated probability threshold value is given to a match which occurs at random. Any scores above this are considered significant at that significance threshold. The protein score for this analysis was 714, with a score of 33 or above considered significant at the 95% confidence interval. The experiment and analysis was carried out by the PNAC facility at the Biochemistry Department.

This verifies for the first time that this 28 kDa band is the ScFv fragment and that it has been expressed and cleaved successfully. It however remained a challenge to isolate it in any purity from the cleaved mixture by the purification strategy of MBPTrap, cleavage and gel filtration.

4.4 – Cell-free expression of a non-MBP fusion ScFv C5 construct

To solve the problem of being unable to separate the cleaved ScFv fragment from the MBP fusion construct, an alternative approach was sought. The ScFv construct was subcloned into a cell-free expression vector (pIVEX2.3d), without a fusion partner. This was done simultaneously to subcloning into a periplasmic expression vector (pET28b) using the same restriction enzyme sites. While the intention was to use periplasmic expression in large scale bacterial cultures, delays in the cloning into the pET28b vector meant that in the meantime the cell-free construct proved a useful tool in trying to obtain ScFv C5 by any means.

A19 *E.coli* extracts (S12) prepared by Henry Chih-Ta Chien and Niclas Frei were used to express the ScFv C5 construct in pIVEX2.3d, as shown below, which is intended to produce a 27.7 kDa protein with a TEV cleavable His-8 tag. Redox shuffling was achieved using 0.3 mM GSSG and 3 mM GSH, in a 1 mL reaction mix.

The reaction mixture was incubated overnight at 30 °C. To try and purify any expressed ScFv, a 50% slurry of nickel-sepharose beads were equilibrated in 20 mM Tris-HCl, 150 mM NaCl and 200 μL beads were added to the cleared cell-free supernatant, with the precipitate also retained and solubilised with SDS. The beads were incubated with the product by end-to-end mixing for 1.5 hours. The flow through was then collected and the beads were washed with the previous buffer yielding a wash fraction. The beads were then washed with buffer supplemented with 250 mM imidazole, yielding elution fractions 1 and 2. These fractions were then analysed using SDS-PAGE and western blot as shown below.
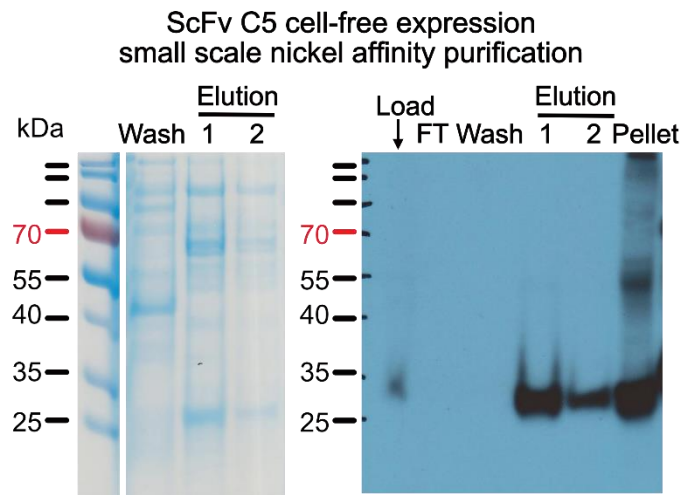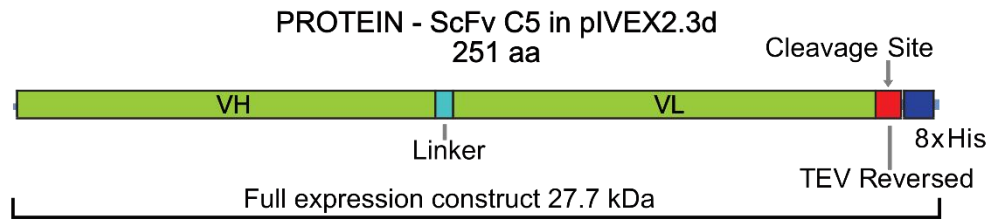
Figure 4.8 Cell free ScFv C5 expression

Cell free expression construct of ScFv C5 (top) and SDS-PAGE and western blot of the ScFv C5 cell free expression reaction (bottom).

The anti-his western blot has shown a band at the expected size in the load fraction, confirming that protein expression has taken place, with a considerable amount in the pellet, representing insoluble expression or crashed out aggregates. This band was even stronger in elution fractions 1 and 2, meaning that for the first time it was possible to bind the ScFv to nickel beads, making it a viable option for purification. It must be noted that the long batch binding mode incubation was likely a contributing factor, as was perhaps the lack of the MBP fusion partner which may have previously taken a role in obscuring the His-tag. The gel also shows a band of the expected size in the elution fractions, indicating that even at this low scale of expression, significant quantities of ScFv can be produced.

This endeavour hence provided two significant observations that encouraged further attempts of expression in large scale *E. coli* cultures. These were the observation of binding to nickel beads, and that the yields, even at this small scale, were non-negligible, building an expectation that with scaled up bacterial expression, the yields may make this protein amenable to structural biology projects, such as NMR experiments.

4.5 – Bacterial expression of a non-MBP fusion ScFv Construct

The successful cloning of the ScFv construct into the pET28b expression vector allowed for the periplasmic expression of the antibody fragment via the PelB export sequence. The pIVEX2.3d vector lacks the necessary gene for the lac repressor (LacI) necessary for lactose induced expression. It only contains a T7 promoter and a ribosome binding site (RBS) upstream of the gene of interest and lacks the lac operator sequence between the T7 promoter and the RBS, which is the site of lac repressor binding. In this arrangement, inserts in the pIVEX2.3d cell-free vector are not under the control of an inducer such as IPTG or lactose, making it an ideal system for cell free expression. The inducible expression permitted by the pET28b vector makes it possible to scale up the bacterial culture and for protein production to commence at the desired time, when temperature, culture density and culture size are right. This setup also benefits from the possibility of employing autoinduction, which can achieve culture densities far superior to those under IPTG induction, where the culture usually only doubles after induction. With autoinduction densities of 10 OD_{600} or more can be achieved.
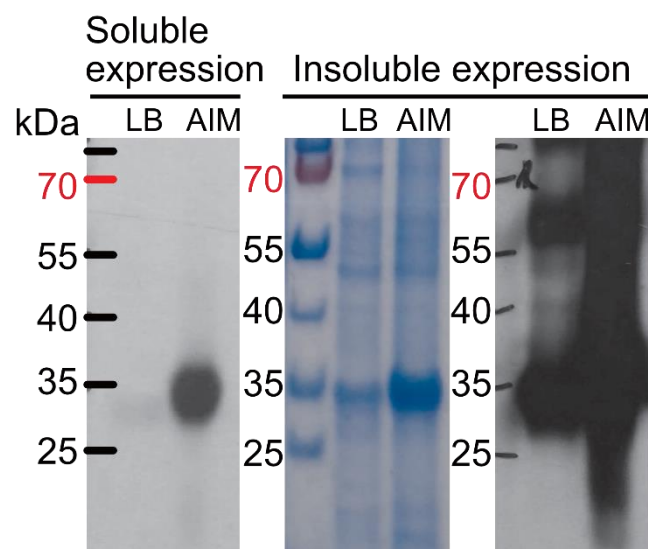
Figure 4.9 Analysis of periplasmic ScFv expression

Expression analysis of the soluble and insoluble fraction of ScFv periplasmic expression in LB and AIM (autoinduction media)

Two 0.5 L cultures were expressed at 25 °C, as a starting condition, in IPTG induced LB media and in autoinduction media (AIM) superbroth (AIM-SB). A 1 mL sample was taken for expression analysis and the soluble and insoluble fractions analysed by SDS-PAGE and western blot (**Figure 4.9**). This showed successful soluble expression in large quantities in AIM-SB, with a significant amount of insoluble

expression. This was also the case for LB media, but with a lower yield. The faint LB band on the western blot of the soluble fraction was the result of a compromise in exposure time, to balance the intensity difference of the two bands. The conclusion from this analysis is that the construct expresses well in this system and that the soluble yield would benefit from a lowered temperature, as observed for the MBP construct.

Having observed that nickel bead binding is possible for the ScFv construct without a fusion partner, a HisTrap chromatography step was tried as first round of purification. A very low flow rate of 0.1 mL/min was used, and the lysate was re-circled for 16 hours (overnight). The column was washed with PBS and eluted with PBS+250 mM imidazole. This achieved a very low purity preparation that was further analysed by gel filtration in PBS buffer. This is shown below (**Figure 4.10**).

Calibrated SEC estimates:
Peak A: 64.43 kDa
Peak B: 31.8 kDa
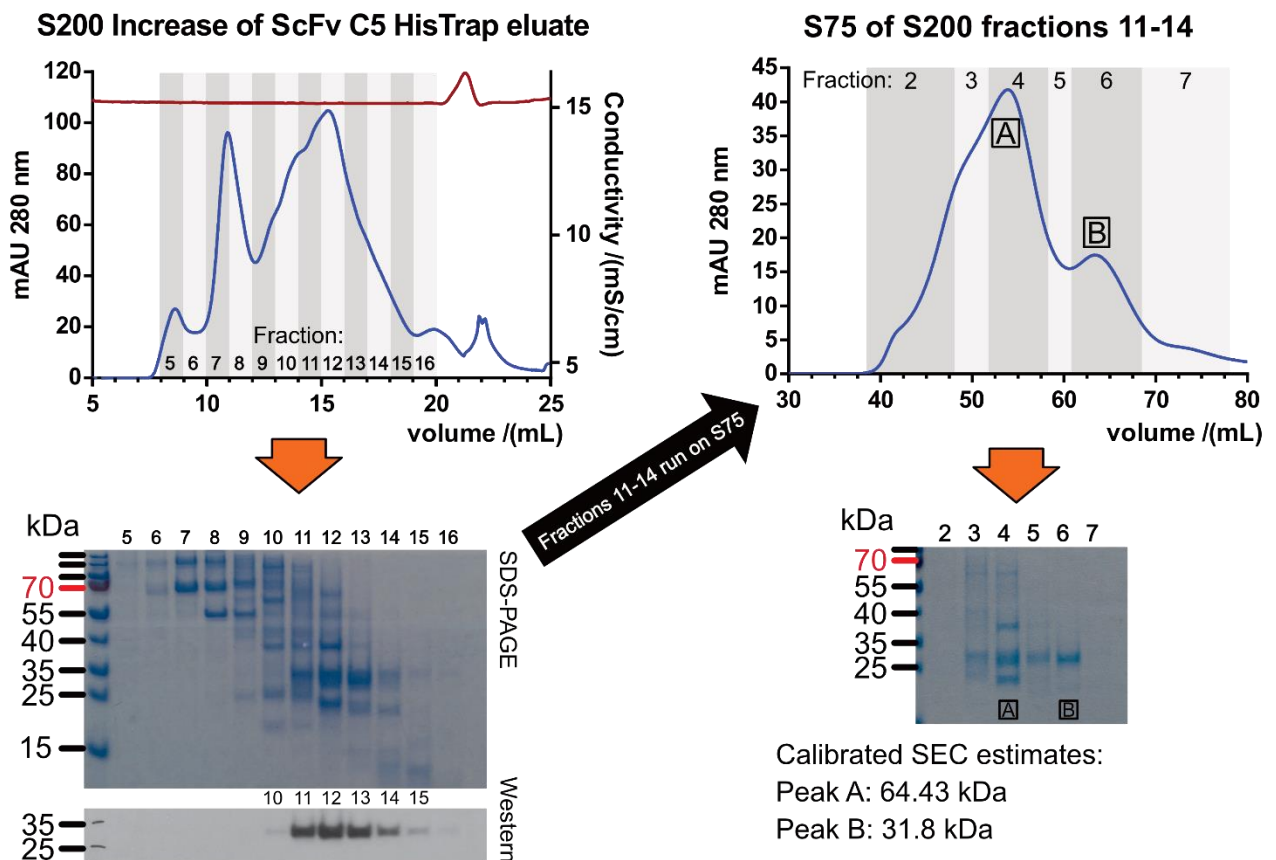
Figure 4.10 Assessment of ScFv C5 oligomerisation state by SEC

SEC assessment of the dimerisation status of the expressed ScFv C5. The elution fraction of ScFv C5 purified by HisTrap columns was concentrated and separated according to size on an S200 Increase column. Selected fractions were further separated using an S75 gel filtration column, revealing two peaks which correspond to the same -band on an SDS-PAGE gel.

Separating the HisTrap combined eluate pool on an S200 Increase column revealed that fractions 11-14 contained a band that matched to the expected 27.9 kDa construct. These fractions also showed a His-tagged band at the expected size.

ScFv C5 is published as a dimeric fragment of the mAb 17.2 antibody. It differs from another fragment version called ScFv B7 in that C5 has a five amino acid sequence linking the VH and VL chains, compared to the 15 amino acid linker for B7 (*230*). In all other respects including their CDR sequence, the two fragments were identical. This difference however resulted in major differences in the structural nature and physiological properties of the two constructs. ScFv B7 appears as a monomer in solution, in comparison to C5, determined by calibrated SEC. C5 was shown to bind the human β_1AR ECL2 peptide H26R with an affinity of 10 μM using surface plasmon resonance. While immunohistochemistry of β_1AR transfected cells showed binding of both C5 and B7, B7 had a significantly lower cAMP stimulating effect than (> 5 fold) than C5, and it was also shown to block the cAMP stimulation in the presence of isoprenaline. Tested on spontaneously beating cardiomyocytes, C5 induced a propranolol negatable 40 beat/min increase, while B7 alone did not cause any significant increase. Furthermore, B7 also prevented an isoprenaline caused increase in beating frequency. This establishes the monomeric form of the same antibody as a negative (with respect to isoprenaline agonism) or neutral (with respect to B7 alone) allosteric modulator, while the dimeric form as a positive allosteric modulator. In this respect, it is a primary objective to achieve a dimeric preparation of ScFv C5, the form which has been shown to achieve specific β_1AR agonism.

To estimate whether the purified ScFv was in a monomeric or the desired dimeric format, the target containing fractions 11-14 from the S200 separation were pooled and separated over a Superdex S75 16/60 pg column. The S200 column is ideal for separating particles between 10 kDa to 600 kDa, while the S75 column is suited for molecules between 3 kDa to 70 kDa. Hence the range defined by the ScFv monomer of 27.9 kDa and dimer of 55.8 kDa is much better suited for separation on an S75 column, making it easier to distinguish between the two species. This has resolved the S200 pooled fractions previously under one broad peak to two major peaks (labelled A and B above – **Figure 4.10**), with elution volumes of 53.91 mL and 63.41 mL, translating to a calibration size of 64.43 kDa and 31.8 kDa respectively. These deviate from the expected sizes by 8.63 kDa and 3.9 kDa. The accuracy of calibrated SEC in estimating molecular mass has been discussed earlier, however it is undeniable that a SEC-MALS measurement would provide more accurate masses. Given that both peaks show the expected 27.9 kDa band on the SDS-PAGE gel, and that in previous publications on purified ScFv C5 the dimeric protein only manifested itself in solution through its solution size based on gel filtration, and that on a gel it appeared at the size of the monomer, it appears that peaks A and B correspond to a dimer and monomer respectively (*230*).

While this analysis suggests this purification method is able to produce the sought after ScFv dimer, another method of purification was tested with the aim of improving material purity. This replaced the second gel filtration step with a light-chain affinity step. Protein L is a bacterial protein isolated from *Peptococcus magnus*, and it is able to bind immunoglobulin kappa light chains (*287*). To this end it has been used for the affinity chromatography of suitable kappa-containing antibodies in an agarose conjugated format. In this case a 1 mL HiTrap Protein L column was used to bind the HisTrap eluate. As shown below (**Figure 4.11**), by SDS-PAGE and his-probe western blow, the use of the Protein L resin resulted in a very high purity preparation, from a very inhomogeneous starting material, demonstrating the specificity of the interaction. This method also achieves the goal of improving the purity compared to the previous nickel affinity-SEC approach.
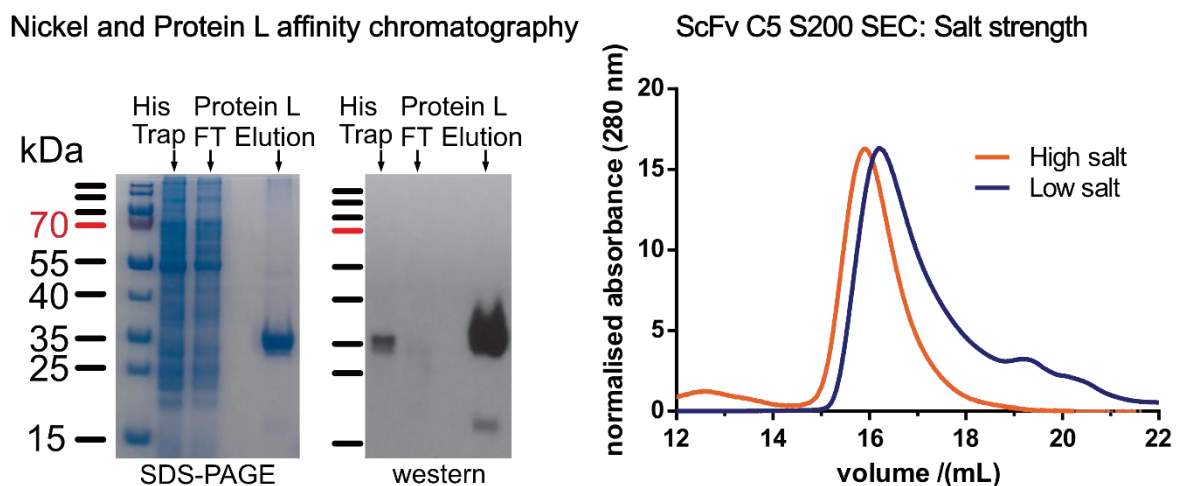
Figure 4.11 Protein L chromatography

Protein L affinity chromatography and SEC analysis of the effect of salt on ScFv C5.

A noteworthy technical consideration concerns the elution of protein from the Protein L column. Following manufacturer's recommendation, the use of 100 mM sodium citrate pH 2.7 only eluted a small fraction (2-5% by peak volume) of the bound material and it was later found to elute in the post-elution water wash. A second attempt of reducing the ionic strength of the elution buffer by using 5 mM NaCitrate achieved complete elution. Similar results were achieved with 100 mM and 5 mM Glycine pH 2.7. This suggests that the interaction of protein L to the ScFv favours high ionic strength conditions, or that the dissociation of the two is not reliant on low pH *per se*, but also on a drop in ionic strength, such that water was able to elute the fragment alone.

The gel filtration of the eluted protein in PBS (with a composition of 10 mM Phosphate pH 7.4, 137 mM NaCl, 2.7 mM KCl) showed a peak at 16.23 mL elution volume, corresponding to a size of 30.19 kDa

(blue trace above – **Figure 4.11**). This likely indicates a monomeric form. Arndt *et. al.* have demonstrated a link between the formation of ScFv dimers and the ionic strength of the buffer (*288*). In order to try and generate dimers from this preparation, the sample was desalted into a high salt buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 50 mM $CaCl_2$). The solution was incubated overnight and ran on the same size exclusion column as the low salt sample, in the new buffer. This produced the orange trace, with a shifted elution volume of 15.92 mL, corresponding to a 35.0 kDa molecule. While the observed shift is in the increased molecular weight direction, it does not quite match the size of a dimer, however as discussed above, an S75 column would provide a much clearer separation between these two peaks. When compared to the HisTrap-SEC approach peaks from the S200 column, both low salt and high salt elution volumes fall under the same peak in fractions 12-13 (**Figure 4.10**), which were later resolved to two peaks on the S75 column.

4.6 – Conclusions on the preparation of purified ScFv C5

In conclusion, two purification approaches should be explored further for the production of dimeric ScFv C5. One option is to employ the HisTrap-SEC approach where the dimeric form of the higher molecular weight peak is confirmed. Alternatively the HisTrap-Protein L option is used, where the purity is superior, but the resulting product appears monomeric, which is not reported to bind the β_1AR. The use of an S75 column (which was not available for the period this work was undertaken or was borrowable for a very limited time) would be highly preferable to judge the dimerisation status of the preparation. The immediate next goal is to produce high purity, confirmed dimers using a combination of the chromatography techniques here explored.

4.7 – Preparation of hECL2 β₁AR chimera

The bystander activation of β_1AR by ScFv C5 was shown through its interactions with the ECL2 H26R peptide, which is a sequence found in the human β_1AR. The figure below depicts a diagrammatic representation of the turkey and human ECL2 loop sequences, with differing residues shown by arrows. As it is unknown if ScFv C5 cross-reacts to the turkey ECL2 loop it is sensible to create a thermostabilised turkey-human-ECL2-chimera.
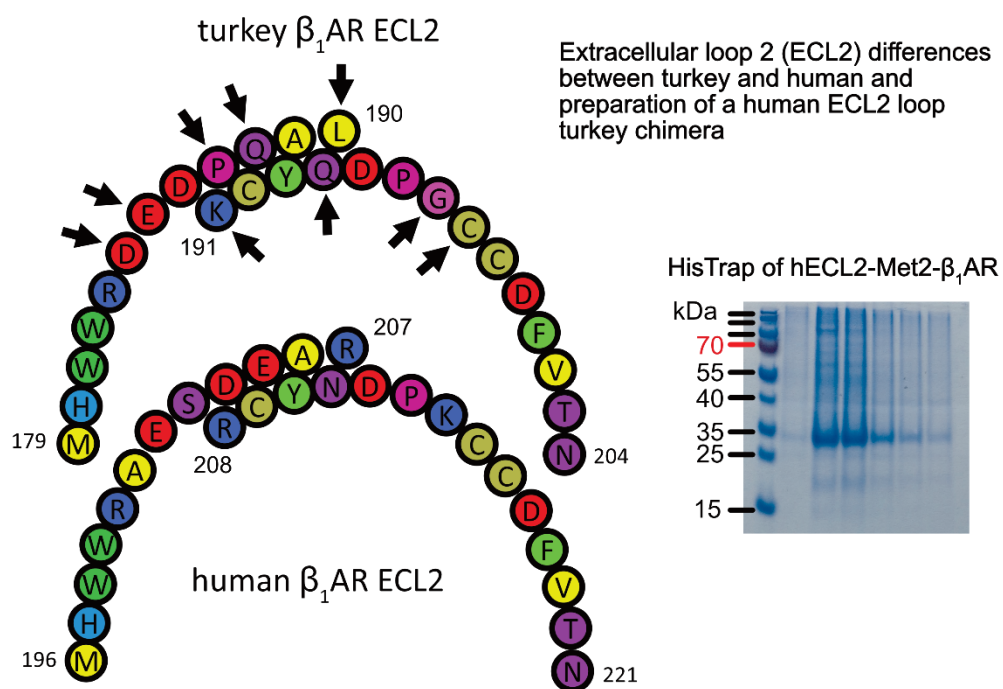
Figure 4.12 Expressing the human ECL2 chimera

Comparison of the human and turkey ECL2 loops and SDS-PAGE of a purified hECL2-turkey-β_1AR construct.

This would be instrumental in using purified receptor for use in biophysical and structural studies of this interaction. To this end, the ECL2 of the tβ_1AR-Met2-Δ5 and tβ_1AR-Met2 constructs was replaced by PCR with the H26R sequence. The created constructs were transfected in Sf9 insect cells for the generation of baculovirus, and a 1 L culture was expressed. The Met2-hECL2 construct was purified by the same method as all other β_1AR constructs in this study, and the SDS-PAGE of the HisTrap eluent shown above confirms the successful expression and purification of the construct.

This makes another component available for the study of receptor agonism by cross-reacting anti-TcP protein antibodies.

4.8 – Future directions

Once the preparation of a dimeric ScFv C5 is confirmed the next step is to evaluate its binding to both the R13 peptide or to the H26R peptide (β_1AR ECL2) in addition to preparations of purified β_1AR in either detergent micelles or lipid nanodiscs. Understanding the quality and nature of this interaction (affinity, on/off rates) will help answer the following questions underpinning the mechanism of chronic Chagas' heart disease.

1. How do ligands modulate receptor agonism by cross-reacting antibodies?
2. What is the molecular mechanism of this agonism?
3. Does activation produce similar molecular signatures as does agonism achieved by ligands and G-proteins, G-protein mimics?
4. Can we stabilise the agonistic conformation for study by structural biology techniques?

The progress outlined in this chapter is far from answering these questions, but it does represent a starting point in terms of preparing the necessary materials needed to address these issues.

Chapter 5 – Methods

5.1 – Transformation of chemically competent *E. coli*

E. coli cells were made competent using either the Inoue buffer method or the rubidium chloride method (*289*, *290*). The former is appropriate for ultracompetent DH5α cells, but it failed to produce long lasting competent cells of any other strain tested. All competent cells were suitable for long term storage and freezing at -80 °C.

Chemically competent or ultracompetent *E. coli* cells were thawed in ice. DNA of appropriate volumes or quantities were added to the cells and incubated for 30 minutes on ice. The cells were then heat shocked at 42 °C for 60 seconds. The cells were then allowed to recover on ice for 2 mins. 750 µL of LB were added to the cells and the culture incubated at 37 °C for 15 to 60 minutes. Appropriate amounts of culture were plated in LB agar enriched with necessary selection.

5.2 – PCR reaction

PCR reactions were all carried out using NEB Q5 polymerase in a 50 µL reaction volume following the manufacturer's protocol. Specific primers and deviating conditions are mentioned at relevant sections in the Results or Methods chapters.

The following tables outline the reaction conditions.

Table 5.1 PCR Reaction mixture

5X Q5 Reaction buffer	10 µL
10 mM dNTP mix	1 µL
10 µM forward primer	2.5 µL
10 µM reverse primer	2.5 µL
Template DNA	1 ng
Q5 polymerase	0.5 µL
GC enhancer (if used)	10 µL
Water	Up to 50 µL total volume

Thermocycler conditions

Table 5.2 PCR conditions

1x	Initial denaturation	98 °C	30 seconds
30 cycles	Denaturation	98 °C	7 seconds
	Annealing	see note below*	20 seconds
	Extension		25 sec/kb
1x	Final extension		2 mins
	Cooled hold		infinite

*Annealing temperatures were calculated using the online NEBTm Calculator and were specific to each primer set.

5.3 – Site Directed Mutagenesis

For assignment purposes, methionine residues were substituted using site directed mutagenesis in a PCR reaction using Phusion DNA Polymerase (ThermoFisher). DpnI digested PCR product was transformed into DH5α E. coli cells along with the transfer plasmid and the resulting plasmid DNA construct was isolated using a commercial Miniprep kit (Qiagen) ready for transfection in insect cells.

5.4 – Baculovirus generation

Baculovirus was generated for expression using FlashBac viral DNA (Oxford Expression Technologies). β_1AR containing pBacPAK8 plasmid was diluted to 100 ng/µl and 1.8 µL, together with 1.8 µl of FlashBac DNA, was added to 1.8 µL of previously dissolved and NaOH neutralised polyethylenimine (linear PEI 25000, Polysciences, 1 mg/ml concentration) and diluted with 360 µL of cell culture media (SF4, Bioconcept). The mixture was incubated at room temperature for 45 min to allow for DNA-polymer complexes to form. The complex mixture was added to 1 mL of mid-log phase Sf9 cells, diluted to 0.5×10^6 cells/mL in serum free SF4 media and incubated at 27 °C shaking for 5 days. On day 5, a small sample was taken to confirm visual signs of infection. The resulting P0 viral stock was diluted to 4 ml with fresh media and incubated for 48 h to generate P1 viral stock. 100 µL of P1 stock was used to infect 50 mL of cells at a density of 1×10^6 cells/mL, incubated at 27 °C, shaking for 48 h, generating a high titre P2 stock for protein expression.

Alternatively some transfection reactions were carried out using Cellfectin II (ThermoFisher). For each transfection reaction 4 µL of Cellfectin II and 50 µL of media were mixed, and separately 0.25 µg of plasmid DNA was mixed with 2.5 µL of FlashBac viral DNA and 50 µL media. The two mixtures were

combined and incubated at room temperature for 15-30 minutes. Following this incubation the transfection mixture was added to cells in a 12 well flat bottom plate, dispensed at a density of 0.5×10^6 cell/mL in 1 mL culture. The plate was not shaken for 3 hours, and the media was changed at this time by replacing it with 1 mL of fresh culture media two times. The generation of virus was identical to the PEI transfection from this point on.

A positive control was used to verify the success of transfection, using an enhanced GFP construct in a pOET1. The vector is compatible with FlashBac through its ORF603 and truncated ORF1629 regions, which are sites of recombination with the viral DNA.
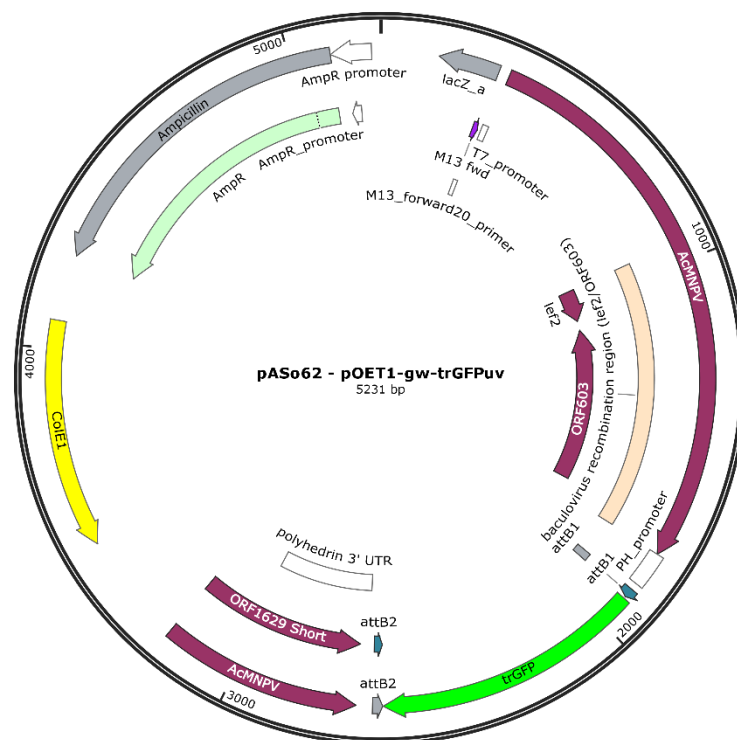
Figure 5.1 GFP control vector for transfection into insect cells

5.5 – β₁AR Protein expression

For β₁AR expression, Sf9 or Sf21 cells, grown in serum free SF4 (Bioconcept) media were centrifuged (500 g, 10 min) and washed with sterile PBS, to reduce the carry-over of unlabelled methionine. The washed cells were diluted to a density of 3×10^6 cells/ml into methionine deficient SF4 media (Bioconcept) or ESF 921 media (Expression systems) at half the intended final culture volume. The culture was then infected with 4 ml/L of high density viral stock (P2), and incubated for 5 h, before supplementing the culture with 250 mg/l of ^{13}C-methyl methionine and diluting to a final density of 1.5×10^6 cells/ml. The initial reduction in culture volume ensured optimal aeration in the initial phase

of the viral infection. Cells were grown at 27 °C for 48-72 hrs and were harvested by centrifugation (3500 g, 15 min), when viability dropped to 80%.

For non-labelled expression the method was identical, except the cells were grown in unlabelled SF4 (Bioconcept) or Insect XPRESS (Lonza) and no methionine supplementation took place.

5.6 – Expression of Nanobodies

Nb80 and Nb6b9 containing plasmids (pET28b(+) and pET26b(+) respectively) were transformed into chemically competent BL21 (DE3) cells and grown in LB media supplemented with the kanamycin at a final concentration of 50 μg/mL. Expression cultures were inoculated with a saturated overnight culture grown at 37 °C, to an OD_{600} density of 0.1 AU and grown to a density of 0.8 AU at 37 °C. Induction was achieved with isopropyl thiogalactopyranoside (IPTG) at a final concentration of 0.5 mM. Expression cultures were grown at 25 °C for 16 h before harvesting by centrifugation (3500 g, 20 min, 4 °C).

Alternatively nanobodies were expressed in Superbroth auto-induction media (Formedium), where the overnight culture was prepared in LB, and 0.5 mL was used to inoculate the expression culture, regardless of overnight OD or culture volume.

The following tables outline the composition of media used:

Table 5.3 Composition of bacterial culture media used

	Final concentration	/(g/L)
LB Media		
Tryptone		10
Yeast Extract		5
NaCl		10
Superbroth Media		
Tryptone		32
Yeast Extract		20
NaCl		5
SOB (Super Optimal Broth)		
Tryptone	0.5% W/V	5
Yeast Extract	2% W/V	20
NaCl	10mM	0.584
KCl	2.5mM	0.186
MgSO4	10mM	1.200
MgCl2	10mM	0.952

Autoinduction Superbroth was either prepared from a commercial powder or made from the following constituents. Base ingredients were dissolved and autoclaved together. Trace elements were sterile filtered and the media was supplemented before use (*291*, *292*).

Table 5.4 Ingredients for autoinduction superbroth media (AIM SB)

Base ingredients	
Component	g/L
Tryptone	35
Yeast Extract	20
$(NH4)_2SO_4$	3.3
KH_2PO_4	6.8
Na_2HPO_4	7.1
Glucose	0.5
α-Lactose	2.0
$MgSO_4$	0.15
Glycerol	5

Table 5.5 Composition of trace element mix for AIM SB

Trace elements		
Ingredient	[Stock]	Added for 500 mL /(mL)
$FeCl_3$	5 mM	0.2 mL
Salt mix	1000x	0.5 mL
$CaCl_2$	100 mM	0.5 mL
NH_4Cl	10%	5 mL (=1g/L)
Antibiotics	1000x	0.5 mL
	$ZnSO_4$	0.1 M
Composition of 1000x salt mix	$MnCl_2$	0.1 M
	H_3BO	0.1 M
	$CuSO_4$	0.1 M

5.7 – Purification of $β_1AR$

The frozen insect cell pellet was thawed with solubilisation buffer (20 mM Tris-HCl pH 8.0, 350 mM NaCl, 3 mM Imidazole, Complete Protease Inhibitor Cocktail (Roche), 2% LMNG) and stirred for 1 h. The solubilised cells were clarified by centrifugation (175,000 g, 45 min) and the soluble fraction was loaded onto a nickel affinity column. The column was washed with equilibration buffer (20 mM Tris-HCl pH 8.0, 350 mM NaCl, 3 mM Imidazole, 0.02% LMNG) and weakly bound impurities were washed out with a 50 mM imidazole wash. The receptor was eluted with the same buffer supplemented with

250 mM imidazole. The final sample was exchanged into 10 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.04% LMNG using a HiPrep Desalting column.

5.8 – Purification of Nanobodies

Frozen cell pellets were thawed in 20 mM Tris-HCl pH 8.0, 150 mM NaCl with Complete Protease Inhibitor Cocktail (Roche) and lysed with an EmulsiFlex C5 homogeniser (Avestin). The solubilised cells were clarified by centrifugation (175,000 g, 15 min) and the soluble fraction was loaded onto a nickel affinity column. Unbound sample was removed with wash buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl), followed by a further wash of the same composition supplemented with 6 mM imidazole. Nanobodies were eluted with wash buffer containing 250 mM imidazole. A further size exclusion chromatography step on a Superdex S200 10/300 Increase column yielded a 90% pure protein preparation in 20 mM Tris-HCl pH 8.0, 150 mM NaCl.

An alternative method of purification involved a nickel affinity chromatography step, followed by buffer-exchange dialysis, cation exchange chromatography and a final desalting step.

Here, following the wash of any unbound material, the HisTrap column was washed with a 15 mM imidazole containing buffer, to remove weakly bound impurities. All further bound material was eluted by equilibration buffer supplemented with 250 mM imidazole. The eluate was dialysed overnight against 2 L of dialysis buffer (50 mM NaAcetate, pH 4.8, 75 mM NaCl) using a 3 kDa cutoff snakeskin membrane. Following dialysis, the solution was cleared of precipitate by centrifugation (4000 rcf, 10 mins) and the supernatant was loaded onto a 1 mL Resource S cation exchange column, pre-equilibrated with dialysis buffer at a rate of 2 mL/min. Subsequent to a column wash, bound species were separated using a liner gradient (60% target, 48 mL duration, 2 mL/min flowrate) using dialysis buffer supplemented with 1M NaCl. Nanobody elution peaked at 350 mM NaCl concentration.

The eluted protein was desalted using a 5 mL HiTrap Desalting column into 20 mM Tris-HCl pH 8, 150 mM NaCl. Nanobody samples were concentrated using a 3 kDa cutoff Amicon concentrator.

5.9 – SDS-PAGE and Western Blots

SDS-PAGE gels were cast according to the following recipe:

Table 5.6 Composition of SDS-PAGE gels

12% Acrylamide Resolving Gel			
	Final concentration	2 Gels	4 gels
Water	-	4.92 mL	8.2 mL
30% Acrylamide/Bis-Acryl 37.5:1	12%	6 mL	10 mL
1.5 M Tris pH 8.8	375 mM	3.75 mL	6.25 mL
10% SDS	0.1%	150 µL	250 µL
10% APS	0.1%	150 µL	250 µL
TEMED (Tetramethylethylenediamine)	-	15 µL	25 µL
TOTAL VOLUME	-	15 mL	25 mL

6% Acrylamide Stacking Gel			
Water	-	2.73 mL	4.75 mL
30% Acrylamide/Bis-Acryl 37.5:1	6%	1 mL	1.8 mL
0.5 M Tris pH 6.8	125 mM	1.25 mL	2.25 mL
10% SDS	0.1%	50 µL	90 µL
10% APS	0.1%	50 µL	90 µL
TEMED (Tetramethylethylenediamine)	-	8 µL	15 µL
TOTAL VOLUME	-	5 mL	9 mL

Samples were prepared with 1x final concentration of LDS loading buffer. Gels were run in NuPAGE MES buffer (Life technologies), or in equivalent custom made MES running buffer (50 mM MES, 50 mM Tris base, 1 mM EDTA, 0.1% SDS) at 200 V. The gels were stained with Instant Blue (Expedeon).

For western blot analysis the resolved gels were transferred onto a 0.45 µm pore size methanol activated PVDF membrane in 1x NuPAGE Transfer buffer (25 mM bicine, 25 mM bis-tris pH 7.2, 1 mM EDTA, 0.05 mM chlorobutanol, 10% methanol) at 30 V for 1 hour. The membrane was then blocked in 4% BSA in TBS-0.05% Tween20 (TBST) for 1 hour and subsequently washed with TBST for 5 minutes, repeated 3 times. The membrane was probed with nickel conjugated HRP in a 1:5000 dilution in TBST as instructed in the SuperSignal West HisProbe Kit (Pierce), for 1 hour. The probed membrane was washed extensively with TBST and incubated with 2 mL SuperSignal West Pico Luminol Enhancer and 2 mL Stable Peroxide for 1 min. Probe binding was visualised on a Fuji RX X-ray film.

5.10 – Dynamic Light Scattering

Purified protein was dialysed into 10 mM Tris-HCl pH 8.0, 0.001% LMNG with varying amount of NaCl, ranging from 50 mM to 350 mM. 70 µL of buffer exchanged sample was placed into a ZEN0040 type

cuvette and measured on a Malvern Zetasizer NanoS instrument. Three repeat measurements were recorded with a 0 second delay.

5.11 – Protein quantification

Proteins quantified by NanoDrop relied on absorbance measurements at 280 nm, owing to the presence of tryptophan, tyrosine, phenylalanine or disulphide bonds all of which have an absorbance maximum at this wavelength. This method allows protein concentrations to be measured without dilution up to 400 mg/mL. A 0.1% protein solution (1 mg/mL) is assumed an absorption of 1.0 AU with a 10 mm path length. For each protein the actual absorption of a 0.1% solution varies depending on the number of absorbing residues present. To this end a correction factor is needed, as 1 AU will likely not equal a 1 mg/mL solution.

$$Absorbance\ 280\ nm\ (0.1\%\ solution)(AU) = {Mw\ (Da)}\Big/{\varepsilon\ (M^{-1}\ cm^{-1})}$$

Multiplying the measured A280 value by the inverse of the obtained 0.1% absorbance value will give the correct protein concentration for the given protein sample in units of mg/mL.

Proteins measured using the Pierce BCA (bicinchoninic acid) assay followed manufacturers protocol using the microplate reader method. Firstly a standard curve is generated through serial dilutions of a 2 mg/mL stock of BSA as described in the manual. Each standard curve point and sample is prepared in triplicate, by mixing 20 µL sample or standard and 200 µL of working reagent. The working reagent is prepared by mixing Reagents A and B in a 50:1 ratio. The samples are mixed and incubated at 37 °C for 30 minutes and the absorbance is measured at 562 nm.

Results were analysed using GraphPad Prism 6.01. Standard curve values were baseline corrected to blank measurements and fitted with a fourth-order polynomial non-linear regression. Unknown sample values were interpolated from this fit. LMNG was found not to be interfering with the assay, while DDM triggered a false positive readout even in the absence of protein.

5.12 - Calibration of size exclusion columns

In order to be able to use gel filtration as a means of estimating molecular mass and oligomeric composition the void volume (V$_0$) and the elution volume of known molecular mass standards had to be established.

The void volume was measured using blue dextran dissolved in 50 mM Tris-HCl pH 8, 150 mM KCl, to a final concentration of 2 mg/mL. Blue dextran has an approximate mass of 2000 kDa, therefore elutes

in the void volume. Calibration was done using the Gel filtration Molecular Weight Calibration Kit (Sigma, #MWGF200).

The following proteins were dissolved in 1 mL of 50 mM Tris-HCl pH 8, 150 mM KCl to the indicated concentrations.

Table 5.7 Calibration markers used for size exclusion chromatography (SEC)

Marker	Molecular Mass /kDa	Concentration /(mg/mL)
Cytochrome c	12.4	2
Carbonic Anhydrase	29	3
Albumin	66	10
Alcohol Dehydrogenase	150	5
β-Amylase	200	4

The molecular masses were plotted against the ratio of measured retention volumes (V_e) to void volume (V_e/V_o). The points were fitted with an exponential curve. The fitting was done in Excel and Graph Pad Prism.

A Superdex S200 Increase and a Superdex S75 16/60 pg columns were calibrated in this way. The resulting standard curves are shown below.
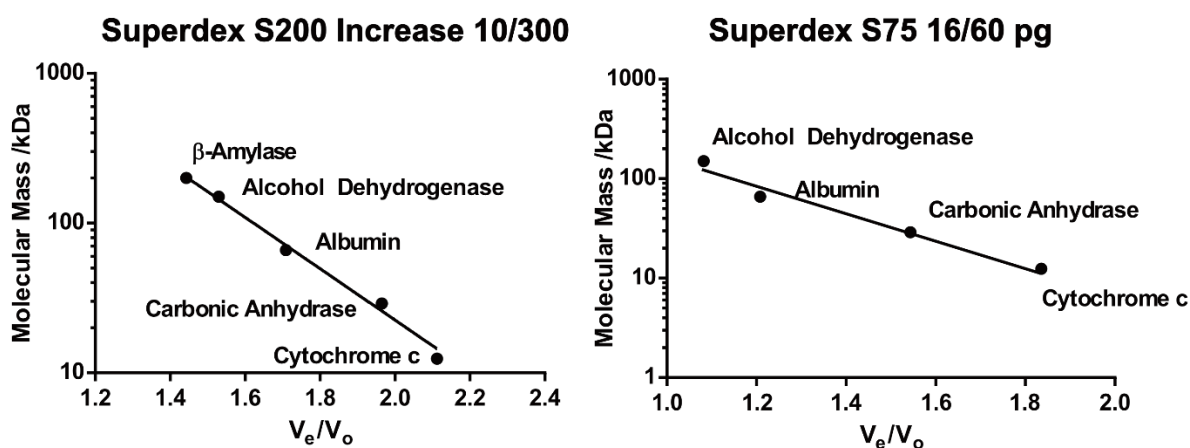
Figure 5.2 SEC calibration curves

These standard curves can be described by the following equations.

Table 5.8 Molecular mass and elution volume calculations for calibrated SEC columns.

	Superdex S200 Increase 10/300	Superdex S75 16/60
Void volume	8.465 mL	41.78 mL
Molecular weight	$Mw = 60244 * e^{-3.946\frac{Ve}{8.465}}$	$Mw = 3540 * e^{-3.105\frac{Ve}{41.78}}$
Elution volume	$Ve = -\frac{8.465}{3.946}\ln\left(\frac{Mw}{60244}\right)$	$Ve = -\frac{41.78}{3.105}\ln\left(\frac{Mw}{3540}\right)$

5.13 - NMR experiments

^{1}H 1D experiments

NMR experiments were recorded on a Bruker Avance AVIII 800 spectrometer (^{1}H 800 MHz) equipped with a 5 mm TXI HCN/z cryoprobe. ^{1}H 1D experiments were recorded with 8 scans and an acquisition time of 0.0512 seconds, recording 1024 data points. Water suppression was achieved with a 3-9-19 180° Watergate sequence.

[^{1}H, ^{13}C] HMQC experiments

Spectra were recorded at 308 K using a SOFAST ^{1}H,^{13}C HMQC experiment with gradient coherence-order selection, using non-uniform sampling (NUS) in the ^{13}C dimension (*293*, *294*). Gradient selection was required to reduce the intense LMNG detergent signals, while selective excitation of methyl groups enabled use of a short recycle delay of 0.5 s. Excitation used a 2.25 ms 120° PC9 pulse and inversion a 1.16 ms, 180° REBURP pulse (*295*, *296*). A 60% Poisson-gap sampling schedule was used (60 complex points from total 100 complex points), with maximum acquisition time (t_{max}) 25 ms and spectral width (^{13}C) of 4000 Hz. The direct (^{1}H) dimension was acquired with 10,000 Hz spectral width, 1024 points and t_{max} = 51.2 ms (*297*). Spectra were recorded with 368 scans, giving an acquisition time of 6.5 h. Where higher sensitivity was required, multiple experiments were recorded and added. Spectra were reconstructed using the iterative hard thresholding (IHT) compressed sensing (CS) implementation in the Cambridge CS package (Mark J. Bostock, unpublished) (*298*). Data was analysed using CCPN Analysis v2 (*299*).

5.14 – ScFv C5 Restriction enzyme cloning

Restriction enzyme cloning was employed to clone the ScFv C5 synthetic construct into the pMAL1-Nb vector using the *Eco*RI and *Hin*dIII restriction sites. 1 µg vector (pMAL1-Nb) and insert (ScFv C5 in

pBSK) was digested in separate 20 μL reactions at 37 °C for 30 mins, using Thermo Fast digest green buffer and ThermoFisher fast digest enzymes. The vector was incubated for a further 30 minutes with alkaline phosphatase to prevent religation. The digested fragments were separated on a 1.1% agarose gel in TAE buffer (40 mM Tris base pH 8.5, 2 mM EDTA, 20 mM acetic acid).The vector produced a 6692 bp fragment, while the C5 insert was 799 bp in size. The relevant fragments were excised from the agarose gel and extracted using an NEB Monarch DNA Gel Extraction Kit, following the manufacturer's protocol. The ligation reaction was carried out using 50 ng linearised vector and 1:1, 3:1 and 5:1 molar ratio of insert to vector in the presence of 1 U of T4 DNA ligase in 1x T4 DNA ligase buffer. The final reaction volume was 20 μL. The ligation reaction was carried out at room temperature for 30 minutes. 5 μL of reaction mixture was transformed into chemically ultracompetent DH5α cells.

Cloning into the cell free vector pIVEX2.3d proceeded along identical lines, with the addition of a prior PCR step to introduce the *Nco*I and *Xho*I restriction sites, alongside extra bases on the overhanging side for enzyme binding. The primers were of the following sequence: XhoRev: cgcataactagctcgagttaatgatggtgatggtggtg, NcoFor: cgatctagaccatggaagttcagctggaagaaag. PCR reaction conditions were followed as described earlier. The PCR product was purified using a NEB Monarch PCR Purification Kit, to remove primers and enzymes. The purified PCR product was used in place of the insert as described above. Thermo fast digest *Nco*I and *Xho*I were used.

5.15 – HRV 3CP protease preparation and cleavage

Human Rhinovirus 3CP protease (referred to as 3CP from hereon) was expressed and purified in house. The expression construct is shown below. The molecular weight is 47.3 kDa with a charge of -7.83 (pI 6.93) at pH 8.0, making it amenable to anion exchange chromatography at this pH.

Human Rhinovirus 3C Protease Expression Construct

MAHHHHHHSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMAIIR
80

His-6 — GST

YIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDF
160

GST

MLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDPGGGLVPRGPNTE
240

GST — thrombin site

FALSLLRKNIMTITTSKGEFTGLGIHDRVCVIPTHAQPGDDVLVNGQKIRVKDKYKLVDPENINLELTVLTLDRNEKFRD
320

HRV 3CP

IRGFISEDLEGVDATLVVHSNNFTNTILEVGPVTMAGLINLSSTPTNRMIRYDYATKTGQCGGVLCATGKIFGIHVGGNG
400

HRV 3CP

RQGFSAQLKKQYFVEKQ*
418

HRV 3CP

Figure 5.3 3C protease construct primary sequence

3CP was expressed in *E. coli* RIL cells, using autoinduction media. Harvested cells were washed and resuspended in PBS and lysed using an Avestin Emulsiflex C5 homogeniser. The supernatant was loaded onto a 5 mL capacity GSTrap FF column and washed with 20 mM Tris-HCl pH 8.0, 50 mM NaCl. The protein was eluted with the same buffer supplemented with 20 mM GSH. The eluate was loaded onto a 1 mL Resource Q anion exchange column and eluted using a linear salt gradient using 20 mM Tris-HCl pH 8.0, 500 mM NaCl. The final preparation was concentrated to 1 mg/mL, and supplemented with 10 mM EDTA and 10% glycerol and stored at -80 °C. No reducing agent was added to the stock.

Cleavage reactions were carried out at 4 °C in the presence of reducing additives (DTT, TCEP, GSSG/GSH), described in more detail where used.

Bibliography

1. R. Fredriksson, M. C. Lagerström, L.-G. Lundin, H. B. Schiöth, The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **63**, 1256–72 (2003).

2. N. Wettschureck, S. Offermanns, Mammalian G proteins and their cell type specific functions. *Physiol. Rev.* **85**, 1159–204 (2005).

3. A. Wise, K. Gearing, S. Rees, Target validation of G-protein coupled receptors. *Drug Discov. Today*. **7**, 235–246 (2002).

4. L. F. Kolakowski, GCRDb: a G-protein-coupled receptor database. *Receptors Channels*. **2**, 1–7 (1994).

5. D. K. Vassilatis *et al.*, The G protein-coupled receptor repertoires of human and mouse. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 4903–8 (2003).

6. H. B. Schiöth, R. Fredriksson, The GRAFS classification system of G-protein coupled receptors in comparative perspective. *Gen. Comp. Endocrinol.* **142**, 94–101 (2005).

7. S. P. H. Alexander *et al.*, The concise guide to PHARMACOLOGY 2013/14: G protein-coupled receptors. *Br. J. Pharmacol.* **170**, 1459–1581 (2013).

8. W. F. Xu *et al.*, Cloning and characterization of human protease-activated receptor 4. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 6642–6 (1998).

9. T. K. Vu, D. T. Hung, V. I. Wheaton, S. R. Coughlin, Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. *Cell*. **64**, 1057–68 (1991).

10. S. Y. Hsu *et al.*, The three subfamilies of leucine-rich repeat-containing G protein-coupled receptors (LGR): identification of LGR6 and LGR7 and the signaling mechanism for LGR7. *Mol. Endocrinol.* **14**, 1257–71 (2000).

11. Q. R. Fan, W. A. Hendrickson, Structure of human follicle-stimulating hormone in complex with its receptor. *Nature*. **433**, 269–77 (2005).

12. M. C. Lagerström, H. B. Schiöth, Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat. Rev. Drug Discov.* **7**, 339–57 (2008).

13. J. J. Onuffer, R. Horuk, Chemokines, chemokine receptors and small-molecule antagonists: recent developments. *Trends Pharmacol. Sci.* **23**, 459–67 (2002).

14. T. Ishihara *et al.*, Molecular cloning and expression of a cDNA encoding the secretin receptor. *EMBO J.* **10**, 1635–41 (1991).

15. B. A. Hofmann *et al.*, Functional and protein chemical characterization of the N-terminal domain of the rat corticotropin-releasing factor receptor 1. *Protein Sci.* **10**, 2050–62 (2001).

16. A. Bazarsuren *et al.*, In vitro folding, functional characterization, and disulfide pattern of the extracellular domain of human GLP-1 receptor. *Biophys. Chem.* **96**, 305–18 (2002).

17. Z. Feng, S. Ma, G. Hu, X.-Q. Xie, Allosteric Binding Site and Activation Mechanism of Class C G-Protein Coupled Receptors: Metabotropic Glutamate Receptor Family. *AAPS J.* **17**, 737–53 (2015).

18.   N. Kunishima *et al.*, Structural basis of glutamate recognition by a dimeric metabotropic glutamate receptor. *Nature*. **407**, 971–7 (2000).

19.   E. Adler *et al.*, A novel family of mammalian taste receptors. *Cell*. **100**, 693–702 (2000).

20.   C. Conte, M. Ebeling, A. Marcuz, P. Nef, P. J. Andres-Barquin, Identification and characterization of human taste receptor genes belonging to the TAS2R family. *Cytogenet. Genome Res.* **98**, 45–53 (2002).

21.   N. Wettschureck, S. Offermanns, Mammalian G proteins and their cell type specific functions. *Physiol. Rev.* **85**, 1159–204 (2005).

22.   G. Milligan, Constitutive activity and inverse agonists of G protein-coupled receptors: a current perspective. *Mol. Pharmacol.* **64**, 1271–6 (2003).

23.   T. Warne, P. C. Edwards, A. G. W. Leslie, C. G. Tate, Crystal structures of a stabilized ??1-adrenoceptor bound to the biased agonists bucindolol and carvedilol. *Structure*. **20**, 841–849 (2012).

24.   D. Wootten, A. Christopoulos, P. M. Sexton, Emerging paradigms in GPCR allostery: implications for drug discovery. *Nat. Rev. Drug Discov.* **12**, 630–44 (2013).

25.   P. J. Flor, F. C. Acher, Orthosteric versus allosteric GPCR activation: the great challenge of group-III mGluRs. *Biochem. Pharmacol.* **84**, 414–24 (2012).

26.   W. Zhu, X. Zeng, M. Zheng, R.-P. Xiao, The enigma of beta2-adrenergic receptor Gi signaling in the heart: the good, the bad, and the ugly. *Circ. Res.* **97**, 507–9 (2005).

27.   M. J. Lohse, C. Hoffmann, Arrestin interactions with G protein-coupled receptors. *Handb. Exp. Pharmacol.* **219**, 15–56 (2014).

28.   K. Watari, M. Nakaya, H. Kurose, Multiple functions of G protein-coupled receptor kinases. *J. Mol. Signal.* **9**, 1 (2014).

29.   C. Krasel *et al.*, Dual role of the beta2-adrenergic receptor C terminus for the binding of beta-arrestin and receptor internalization. *J. Biol. Chem.* **283**, 31840–8 (2008).

30.   A. B. Tobin, G-protein-coupled receptor phosphorylation: where, when and by whom. *Br. J. Pharmacol.* **153 Suppl**, S167-76 (2008).

31.   O. B. Goodman *et al.*, β-Arrestin acts as a clathrin adaptor in endocytosis of the β2-adrenergic receptor. *Nature*. **383**, 447–450 (1996).

32.   S. S. Ferguson *et al.*, Role of beta-arrestin in mediating agonist-promoted G protein-coupled receptor internalization. *Science*. **271**, 363–6 (1996).

33.   S. K. Shenoy, R. J. Lefkowitz, β-Arrestin-mediated receptor trafficking and signal transduction. *Trends Pharmacol. Sci.* **32**, 521–33 (2011).

34.   S. K. Shenoy, R. J. Lefkowitz, Seven-transmembrane receptor signaling through beta-arrestin. *Sci. STKE*. **2005**, cm10 (2005).

35.   S. Sarker, K. Xiao, S. K. Shenoy, A Tale of Two Sites – How ubiquitination of a G protein-coupled receptor is coupled to its lysosomal trafficking from distinct receptor domains. *Commun. Integr. Biol.* **4**, 528–531 (2014).

36.   X. H. Chen, T. K. Harden, R. A. Nicholas, Molecular cloning and characterization of a novel beta-adrenergic receptor. *J. Biol. Chem.* **269**, 24810–9 (1994).

37.   J. G. Granneman, The putative beta4-adrenergic receptor is a novel state of the beta1-adrenergic receptor. *Am. J. Physiol. Endocrinol. Metab.* **280**, E199-202 (2001).

38.   A. R. Kompa, R. J. Summers, Desensitization and resensitization of beta 1- and putative beta 4-adrenoceptor mediated responses occur in parallel in a rat model of cardiac failure. *Br. J. Pharmacol.* **128**, 1399–406 (1999).

39.   G. Wallukat, The beta-adrenergic receptors. *Herz*. **27**, 683–90 (2002).

40.   C. Communal, K. Singh, D. R. Pimentel, W. S. Colucci, Norepinephrine stimulates apoptosis in adult rat ventricular myocytes by activation of the beta-adrenergic pathway. *Circulation*. **98**, 1329–34 (1998).

41.   T. Noma *et al.*, Beta-arrestin-mediated beta1-adrenergic receptor transactivation of the EGFR confers cardioprotection. *J. Clin. Invest.* **117**, 2445–58 (2007).

42.   K. Palczewski *et al.*, Crystal structure of rhodopsin: A G protein-coupled receptor. *Science*. **289**, 739–45 (2000).

43.   S. G. F. Rasmussen *et al.*, Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature*. **450**, 383–387 (2007).

44.   E. Ghosh, P. Kumari, D. Jaiman, A. K. Shukla, Methodological advances: the unsung heroes of the GPCR structural revolution. *Nat. Rev. Mol. Cell Biol.* **16**, 69–81 (2015).

45.   P. W. Day *et al.*, A monoclonal antibody for G protein-coupled receptor crystallography. *Nat. Methods*. **4**, 927–929 (2007).

46.   V. Cherezov, J. Clogston, M. Z. Papiz, M. Caffrey, Room to move: crystallizing membrane proteins in swollen lipidic mesophases. *J. Mol. Biol.* **357**, 1605–18 (2006).

47.   L. V Misquitta *et al.*, Membrane protein crystallization in lipidic mesophases with tailored bilayers. *Structure*. **12**, 2113–24 (2004).

48.   P. Nollert, H. Qiu, M. Caffrey, J. P. Rosenbusch, E. M. Landau, Molecular mechanism for the crystallization of bacteriorhodopsin in lipidic cubic phases. *FEBS Lett.* **504**, 179–86 (2001).

49.   P. S. Chae *et al.*, Maltose-neopentyl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat. Methods*. **7**, 1003–8 (2010).

50.   M. G. Caron, Y. Srinivasan, J. Pitha, K. Kociolek, R. J. Lefkowitz, Affinity chromatography of the beta-adrenergic receptor. *J. Biol. Chem.* **254**, 2923–7 (1979).

51.   V. Cherezov *et al.*, Rastering strategy for screening and centring of microcrystal samples of human membrane proteins with a sub-10 microm size X-ray synchrotron beam. *J. R. Soc. Interface*. **6 Suppl 5**, S587-97 (2009).

52.   H. N. Chapman *et al.*, Femtosecond X-ray protein nanocrystallography. *Nature*. **470**, 73–7 (2011).

53.   W. Liu, A. Ishchenko, V. Cherezov, Preparation of microcrystals in lipidic cubic phase for serial femtosecond crystallography. *Nat. Protoc.* **9**, 2123–34 (2014).

54.   S. G. F. Rasmussen *et al.*, Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature*. **477**, 549–55 (2011).

55.   X. Deupi *et al.*, Stabilized G protein binding site in the structure of constitutively active metarhodopsin-II. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 119–24 (2012).

56. A. C. Kruse *et al.*, Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature*. **504**, 101–6 (2013).

57. Y. Zhang *et al.*, Cryo-EM structure of the activated GLP-1 receptor in complex with a G protein. *Nature*. **546**, 248–253 (2017).

58. J. García-Nafría, Y. Lee, X. Bai, B. Carpenter, C. G. Tate, Cryo-EM structure of the adenosine A2A receptor coupled to an engineered heterotrimeric G protein. *Elife*. **7**, e35946 (2018).

59. C. J. Draper-Joyce *et al.*, Structure of the adenosine-bound human adenosine A1 receptor–Gi complex. *Nature*. **558**, 559–563 (2018).

60. J. Standfuss *et al.*, The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature*. **471**, 656–660 (2011).

61. J. García-Nafría, R. Nehmé, P. C. Edwards, C. G. Tate, Cryo-EM structure of the serotonin 5-HT1B receptor coupled to heterotrimeric Go. *Nature*. **558**, 620–623 (2018).

62. H. W. Choe *et al.*, Crystal structure of metarhodopsin II. *Nature*. **471**, 651–655 (2011).

63. P. Scheerer *et al.*, Crystal structure of opsin in its G-protein-interacting conformation. *Nature*. **455**, 497–502 (2008).

64. S. G. F. Rasmussen *et al.*, Crystal structure of the β2 adrenergic receptor–Gs protein complex. *Nature*. **477**, 549–555 (2011).

65. Y. L. Liang *et al.*, Phase-plate cryo-EM structure of a biased agonistbound human GLP-1 receptor-Gs complex. *Nature*. **555**, 121–125 (2018).

66. Y. L. Liang *et al.*, Phase-plate cryo-EM structure of a class B GPCR-G-protein complex. *Nature*. **546**, 118–123 (2017).

67. A. Koehl *et al.*, Structure of the μ-opioid receptor–Gi protein complex. *Nature*. **558**, 547–552 (2018).

68. B. Carpenter, R. Nehmé, T. Warne, A. G. W. Leslie, C. G. Tate, Structure of the adenosine A2Areceptor bound to an engineered G protein. *Nature*. **536**, 104–107 (2016).

69. T. Warne *et al.*, The structural basis for agonist and partial agonist action on a β(1)-adrenergic receptor. *Nature*. **469**, 241–4 (2011).

70. G. Lebon, P. C. Edwards, A. G. W. Leslie, C. G. Tate, Molecular Determinants of CGS21680 Binding to the Human Adenosine A2A Receptor. *Mol. Pharmacol.* **87**, 907–15 (2015).

71. F. Xu *et al.*, Structure of an agonist-bound human A2A adenosine receptor. *Science*. **332**, 322–7 (2011).

72. B. E. Krumm, J. F. White, P. Shah, R. Grisshammer, Structural prerequisites for G-protein activation by the neurotensin receptor. *Nat. Commun.* **6**, 7895 (2015).

73. J. Zhang *et al.*, Agonist-bound structure of the human P2Y12 receptor. *Nature*. **509**, 119–22 (2014).

74. Y. Kang *et al.*, Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature*. **523**, 561–7 (2015).

75. S. G. F. Rasmussen *et al.*, Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature*. **450**, 383–7 (2007).

76. V. Cherezov *et al.*, High-resolution crystal structure of an engineered human beta2-

adrenergic G protein-coupled receptor. *Science*. **318**, 1258–65 (2007).

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

78. M. P. Bokoch *et al.*, Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. *Nature*. **463**, 108–12 (2010).

79. D. Wacker *et al.*, Conserved binding mode of human beta2 adrenergic receptor inverse agonists and antagonist revealed by X-ray crystallography. *J. Am. Chem. Soc.* **132**, 11443–5 (2010).

80. S. G. F. Rasmussen *et al.*, Structure of a nanobody-stabilized active state of the β(2) adrenoceptor. *Nature*. **469**, 175–80 (2011).

81. D. M. Rosenbaum *et al.*, Structure and function of an irreversible agonist-β(2) adrenoceptor complex. *Nature*. **469**, 236–40 (2011).

82. S. G. F. Rasmussen *et al.*, Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature*. **477**, 549–55 (2011).

83. Y. Zou, W. I. Weis, B. K. Kobilka, N-terminal T4 lysozyme fusion facilitates crystallization of a G protein coupled receptor. *PLoS One*. **7**, e46039 (2012).

84. A. M. Ring *et al.*, Adrenaline-activated structure of β2-adrenoceptor stabilized by an engineered nanobody. *Nature*. **502**, 575–9 (2013).

85. D. Weichert *et al.*, Covalent agonists for studying G protein-coupled receptor activation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 10744–8 (2014).

86. C.-Y. Huang *et al.*, In meso in situ serial X-ray crystallography of soluble and membrane proteins. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **71**, 1238–1256 (2015).

87. D. P. Staus *et al.*, Allosteric nanobodies reveal the dynamic range and diverse mechanisms of G-protein-coupled receptor activation. *Nature*. **535**, 448–452 (2016).

88. P. Ma *et al.*, The cubicon method for concentrating membrane proteins in the cubic mesophase. *Nat. Protoc.* **12**, 1745–1762 (2017).

89. X. Liu *et al.*, Mechanism of intracellular allosteric β 2 AR antagonist revealed by X-ray crystal structure. *Nature*. **548**, 480–484 (2017).

90. J. A. Ballesteros *et al.*, Activation of the beta 2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6. *J. Biol. Chem.* **276**, 29171–7 (2001).

91. M. J. Serrano-Vega, C. G. Tate, Transferability of thermostabilizing mutations between beta-adrenergic receptors. *Mol. Membr. Biol.* **26**, 385–396 (2009).

92. T. Warne, M. J. Serrano-Vega, C. G. Tate, G. F. X. Schertler, Development and crystallization of a minimal thermostabilised G protein-coupled receptor. *Protein Expr. Purif.* **65**, 204–213 (2009).

93. J. L. Miller, C. G. Tate, Engineering an ultra-thermostable β(1)-adrenoceptor. *J. Mol. Biol.* **413**, 628–38 (2011).

94. M. J. Serrano-Vega, F. Magnani, Y. Shibata, C. G. Tate, Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 877–882 (2008).

95. G. S. Balaraman, S. Bhattacharya, N. Vaidehi, Structural insights into conformational stability of wild-type and mutant beta1-adrenergic receptor. *Biophys. J.* **99**, 568–77 (2010).

96. T. Warne *et al.*, Structure of a β1-adrenergic G-protein-coupled receptor. *Nature*. **454** (2008), pp. 486–491.

97. R. Moukhametzianov *et al.*, Two distinct conformations of helix 6 observed in antagonist-bound structures of a beta1-adrenergic receptor. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 8228–32 (2011).

98. J. A. Christopher *et al.*, Biophysical fragment screening of the β1-adrenergic receptor: identification of high affinity arylpiperazine leads using structure-based drug design. *J. Med. Chem.* **56**, 3446–55 (2013).

99. J. L. Miller-Gallacher *et al.*, The 2.1 Å resolution structure of cyanopindolol-bound β1-adrenoceptor identifies an intramembrane Na+ ion that stabilises the ligand-free receptor. *PLoS One*. **9**, e92727 (2014).

100. J. Huang, S. Chen, J. J. Zhang, X.-Y. Huang, Crystal structure of oligomeric β1-adrenergic G protein-coupled receptors in ligand-free basal state. *Nat. Struct. Mol. Biol.* **20**, 419–25 (2013).

101. J. L. Miller-Gallacher *et al.*, The 2.1 Å resolution structure of cyanopindolol-bound β1-adrenoceptor identifies an intramembrane Na+ ion that stabilises the ligand-free receptor. *PLoS One*. **9**, e92727 (2014).

102. T. Sato *et al.*, Pharmacological Analysis and Structure Determination of 7-Methylcyanopindolol-Bound 1-Adrenergic Receptor. *Mol. Pharmacol.* **88**, 1024–1034 (2015).

103. A. G. W. Leslie, T. Warne, C. G. Tate, Ligand occupancy in crystal structure of β 1-adrenergic G protein coupled receptor. *Nat. Struct. Mol. Biol.* **22** (2015), pp. 941–942.

104. E. S. Burstein, T. A. Spalding, M. R. Brann, The second intracellular loop of the m5 muscarinic receptor is the switch which enables G-protein coupling. *J. Biol. Chem.* **273**, 24322–7 (1998).

105. S. O. Smith, Structure and activation of the visual pigment rhodopsin. *Annu. Rev. Biophys.* **39**, 309–28 (2010).

106. K. Palczewski, Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. *Science (80-.).* **289**, 739–745 (2000).

107. R. W. Schoenlein, L. A. Peteanu, R. A. Mathies, C. V Shank, The first step in vision: femtosecond isomerization of rhodopsin. *Science*. **254**, 412–415 (1991).

108. D. A. Baylor, Photoreceptor signals and vision: Proctor lecture. *Investig. Ophthalmol. Vis. Sci.* **28** (1987), pp. 34–49.

109. G. Lebon, T. Warne, C. G. Tate, Agonist-bound structures of G protein-coupled receptors. *Curr. Opin. Struct. Biol.* **22**, 482–90 (2012).

110. G. Lebon *et al.*, Agonist-bound adenosine A2A receptor structures reveal common features of GPCR activation. *Nature*. **474**, 521–5 (2011).

111. A. S. Doré *et al.*, Structure of the adenosine A(2A) receptor in complex with ZM241385 and the xanthines XAC and caffeine. *Structure*. **19**, 1283–93 (2011).

112. D. Milić, D. B. Veprintsev, Large-scale production and protein engineering of G protein-coupled receptors for structural studies. *Front. Pharmacol.* **6**, 66 (2015).

113. A. Gautier, H. R. Mott, M. J. Bostock, J. P. Kirkpatrick, D. Nietlispach, Structure determination

of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy. *Nat. Struct. Mol. Biol.* **17**, 768–74 (2010).

114. S. Reckel *et al.*, Solution NMR structure of proteorhodopsin. *Angew. Chem. Int. Ed. Engl.* **50**, 11942–6 (2011).

115. C. Opitz, S. Isogai, S. Grzesiek, An economic approach to efficient isotope labeling in insect cells using homemade (15)N-, (13)C- and (2)H-labeled yeast extracts. *J. Biomol. NMR* (2015), doi:10.1007/s10858-015-9954-3.

116. A. Sitarska *et al.*, Affordable uniform isotope labeling with (2)H, (13)C and (15)N in insect cells. *J. Biomol. NMR*. **62**, 191–7 (2015).

117. K. Saxena, A. Dutta, J. Klein-Seetharaman, H. Schwalbe, Isotope labeling in insect cells. *Methods Mol. Biol.* **831**, 37–54 (2012).

118. A. D. Gossert, W. Jahnke, Isotope labeling in insect cells. *Adv Exp Med Biol*. **992**, 179–196 (2012).

119. D. Salom *et al.*, Isotopic labeling of mammalian G protein-coupled receptors heterologously expressed in Caenorhabditis elegans. *Anal. Biochem.* **472**, 30–6 (2015).

120. R. Nygaard *et al.*, The dynamic process of β2-adrenergic receptor activation. *Cell*. **152**, 532–542 (2013).

121. Y. Kofuku *et al.*, Efficacy of the β2-adrenergic receptor is determined by conformational equilibrium in the transmembrane region. *Nat. Commun.* **3**, 1045 (2012).

122. J. J. Liu, R. Horst, V. Katritch, R. C. Stevens, K. Wüthrich, Biased signaling pathways in β2-adrenergic receptor characterized by 19F-NMR. *Science*. **335**, 1106–10 (2012).

123. K. Rajewsky, Clonal selection and learning in the antibody system. *Nature*. **381** (1996), pp. 751–758.

124. H. Zhang *et al.*, Structure of the full-length glucagon class B G-protein-coupled receptor. *Nature*. **546**, 259–264 (2017).

125. S. Muyldermans, Nanobodies: Natural Single-Domain Antibodies. *Annu. Rev. Biochem.* **82**, 775–797 (2013).

126. J. Govaert *et al.*, Dual beneficial effect of interloop disulfide bond for single domain antibody fragments. *J. Biol. Chem.* **287**, 1970–1979 (2012).

127. A. S. Solt *et al.*, Insight into partial agonism by observing multiple equilibria for ligand-bound and Gs-mimetic nanobody-bound β1-adrenergic receptor. *Nat. Commun.* **8**, 1795 (2017).

128. M. T. Eddy, T. Didenko, R. C. Stevens, K. Wüthrich, β2-Adrenergic Receptor Conformational Response to Fusion Protein in the Third Intracellular Loop. *Structure*. **24**, 2190–2197 (2016).

129. J. S. Burg *et al.*, Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor. *Science (80-.)*. **347**, 1113–1117 (2015).

130. T. F. Miles *et al.*, Viral GPCR US28 can signal in response to chemokine agonists of nearly unlimited structural degeneracy. *Elife*. **7**, e35850 (2018).

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

132. S. G. F. Rasmussen *et al.*, Structure of a nanobody-stabilized active state of the β(2)

adrenoceptor. *Nature*. **469**, 175–180 (2011).

133. S. G. F. Rasmussen *et al.*, Crystal structure of the β2 adrenergic receptor–Gs protein complex. *Nature*. **477**, 549–555 (2011).

134. A. M. Ring *et al.*, Adrenaline-activated structure of β 2-adrenoceptor stabilized by an engineered nanobody. *Nature*. **502**, 575–579 (2013).

135. M. P. Bokoch *et al.*, Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. *Nature*. **463**, 108–112 (2010).

136. J. García-Nafría, Y. Lee, X. Bai, B. Carpenter, C. G. Tate, Cryo-EM structure of the adenosine A2A receptor coupled to an engineered heterotrimeric G protein. *Elife*. **7**, e35946 (2018).

137. T. Che *et al.*, Structure of the Nanobody-Stabilized Active State of the Kappa Opioid Receptor. *Cell*. **172**, 55-67.e15 (2018).

138. A. C. Kruse *et al.*, Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature*. **504**, 101–106 (2013).

139. A. Ishchenko *et al.*, Structural insights into the extracellular recognition of the human serotonin 2B receptor by an antibody. *Proc. Natl. Acad. Sci.* **114**, 8223–8228 (2017).

140. W. Huang *et al.*, Structural insights into μ-opioid receptor activation. *Nature*. **524**, 315–321 (2015).

141. B. Carpenter, C. G. Tate, Engineering a minimal G Protein to facilitate crystallisation of G protein - coupled receptors in their active conformation Byron Carpenter & Christopher G . Tate MRC Laboratory of Molecular Biology , Cambridge Biomedical Campus , F. *Protein Eng. Des. Sel.* **29**, 583–594 (2016).

142. R. K. Y. Cheng *et al.*, Structural insight into allosteric modulation of protease-activated receptor 2. *Nature*. **545**, 112–115 (2017).

143. J. Li, P. C. Edwards, M. Burghammer, C. Villa, G. F. X. Schertler, Structure of bovine rhodopsin in a trigonal crystal form. *J. Mol. Biol.* **343**, 1409–1438 (2004).

144. J. H. Park, P. Scheerer, K. P. Hofmann, H. W. Choe, O. P. Ernst, Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature*. **454**, 183–187 (2008).

145. G. R. Pathare *et al.*, Crystal structure of the proteasomal deubiquitylation module Rpn8-Rpn11. *Proc. Natl. Acad. Sci.* **111**, 2984–2989 (2014).

146. K. Rostislavleva *et al.*, Structure and flexibility of the endosomal Vps34 complex reveals the basis of its function on membranes. *Science (80-.).* **350**, aac7365 (2015).

147. L. Li *et al.*, Crystal structure of a substrate-engaged SecY protein-translocation channel. *Nature*. **531**, 395–399 (2016).

148. E. R. Geertsma *et al.*, Structure of a prokaryotic fumarate transporter reveals the architecture of the SLC26 family. *Nat. Struct. Mol. Biol.* **22**, 803–808 (2015).

149. G. Hassaine *et al.*, X-ray structure of the mouse serotonin 5-HT3receptor. *Nature*. **512**, 276–281 (2014).

150. I. A. Ehrnstorfer, E. R. Geertsma, E. Pardon, J. Steyaert, R. Dutzler, Crystal structure of a SLC11 (NRAMP) transporter reveals the basis for transition-metal ion transport. *Nat. Struct. Mol. Biol.* **21**, 990–996 (2014).

151. P. Owen, H. R. Kaback, Molecular structure of membrane vesicles from Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.* **75**, 3148–3152 (1978).

152. A. Manglik, B. K. Kobilka, J. Steyaert, Nanobodies to Study G Protein–Coupled Receptor Structure and Function. *Annu. Rev. Pharmacol. Toxicol.* **57**, 19–37 (2017).

153. G. H. Westfield *et al.*, Structural flexibility of the G alpha s alpha-helical domain in the beta2-adrenoceptor Gs complex. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 16086–91 (2011).

154. A. Manglik *et al.*, Structural Insights into the Dynamic Process of β2-Adrenergic Receptor Signaling. *Cell*. **161**, 1101–1111 (2015).

155. J. S. Burg *et al.*, Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor. *Science (80-.).* **347**, 1113–1117 (2015).

156. T. N. Kledal, M. M. Rosenkilde, T. W. Schwartz, Selective recognition of the membrane-bound CX3C chemokine, fractalkine, by the human cytomegalovirus-encoded broad-spectrum receptor US28. *FEBS Lett.* **441**, 209–214 (1998).

157. A. S. Hauser, M. M. Attwood, M. Rask-Andersen, H. B. Schiöth, D. E. Gloriam, Trends in GPCR drug discovery: new agents, targets and indications. *Nat. Rev. Drug Discov.* **16**, 829–842 (2017).

158. Animal Cell Technology Idustrial Platform, Monoclonal Antibodies Approved by the EMA and FDA for Therapeutic Use – ACTIP (2017), (available at http://www.actip.org/products/monoclonal-antibodies-approved-by-the-ema-and-fda-for-therapeutic-use/).

159. Approved antibodies. *Antib. Soc.* (available at https://www.antibodysociety.org/news/approved-antibodies/).

160. J. M. Reichert, C. J. Rosensweig, L. B. Faden, M. C. Dewitz, Monoclonal antibody successes in the clinic. *Nat. Biotechnol.* **23**, 1073–1078 (2005).

161. R. Macarron, Critical review of the role of HTS in drug discovery. *Drug Discov. Today*. **11**, 277–279 (2006).

162. S. P. Alexander *et al.*, THE CONCISE GUIDE TO PHARMACOLOGY 2017/18: G protein-coupled receptors. *Br. J. Pharmacol.* **174**, S17–S129 (2017).

163. N. J. Smith *et al.*, Themed Section: Molecular Pharmacology of G Protein-Coupled Receptors Identifying ligands at orphan GPCRs: current status using structure-based approaches (available at http://onlinelibrary.wiley.com/doi/10.1111/bph.v173.20/).

164. R. Santos *et al.*, A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34 (2017).

165. D. J. Foster, P. J. Conn, Allosteric Modulation of GPCRs: New Insights and Potential Utility for Treatment of Schizophrenia and Other CNS Disorders. *Neuron*. **94**, 431–446 (2017).

166. C. Ullmer *et al.*, Functional monoclonal antibody acts as a biased agonist by inducing internalization of metabotropic glutamate receptor 7. *Br. J. Pharmacol.* **167**, 1448–1466 (2012).

167. E. L. Moore, C. A. Salvatore, Targeting a family B GPCR/RAMP receptor complex: CGRP receptor antagonists and migraine. *Br. J. Pharmacol.* **166**, 66–78 (2012).

168. D. R. Webb, T. M. Handel, A. Kretz-Rommel, R. C. Stevens, Opportunities for functional

selectivity in GPCR antibodies. *Biochem. Pharmacol.* **85**, 147–52 (2013).

169. C. Zöllner *et al.*, Chronic morphine use does not induce peripheral tolerance in a rat model of inflammatory pain. *J. Clin. Invest.* **118**, 1065–1073 (2008).

170. C. Stein, L. J. Lang, Peripheral mechanisms of opioid analgesia. *Curr. Opin. Pharmacol.* **9**, 3–8 (2009).

171. M. Suzuki, C. Kato, A. Kato, Therapeutic antibodies: their mechanisms of action and the pathological findings they induce in toxicity studies. *J. Toxicol. Pathol.* **28**, 133–9 (2015).

172. S. D. Liu *et al.*, Afucosylated Antibodies Increase Activation of Fc RIIIa-Dependent Signaling Components to Intensify Processes Promoting ADCC. *Cancer Immunol. Res.* **3**, 173–183 (2015).

173. I. Zafir-Lavie, Y. Michaeli, Y. Reiter, Novel antibodies as anticancer agents. *Oncogene*. **26**, 3714–3733 (2007).

174. A. T. Kurdi *et al.*, *Mol. Cancer Ther.*, in press, doi:10.1158/1535-7163.MCT-17-0998.

175. J. M. Lambert, C. Q. Morris, Antibody-Drug Conjugates (ADCs) for Personalized Treatment of Solid Tumors: A Review. *Adv. Ther.* **34**, 1015–1035 (2017).

176. M. A. Ayoub *et al.*, Antibodies targeting G protein-coupled receptors: Recent advances and therapeutic challenges. *MAbs*. **9**, 735–741 (2017).

177. A. L. J. Marschall, S. Dübel, Antibodies inside of a cell can change its outside: Can intrabodies provide a new therapeutic paradigm? *Comput. Struct. Biotechnol. J.* **14**, 304–308 (2016).

178. S. M. Shin *et al.*, Antibody targeting intracellular oncogenic Ras mutants exerts anti-tumour effects after systemic administration. *Nat. Commun.* **8**, 15090 (2017).

179. A. Turchick, D. C. Hegan, R. B. Jensen, P. M. Glazer, A cell-penetrating antibody inhibits human RAD51 via direct binding. *Nucleic Acids Res.* **45**, 11782–11799 (2017).

180. M. Che Nordin, S.-Y. Teow, Review of Current Cell-Penetrating Antibody Developments for HIV-1 Therapy. *Molecules*. **23**, 335 (2018).

181. C. J. Hutchings, M. Koglin, F. H. Marshall, Therapeutic antibodies directed at G protein-coupled receptors. *MAbs*. **2**, 594–606 (2010).

182. N. Robertson *et al.*, The properties of thermostabilised G protein-coupled receptors (StaRs) and their use in drug discovery. *Neuropharmacology*. **60**, 36–44 (2011).

183. J. Kean, A. Bortolato, K. Hollenstein, F. H. Marshall, A. Jazayeri, Conformational thermostabilisation of corticotropin releasing factor receptor 1. *Sci. Rep.* **5**, 11954 (2015).

184. D. J. Scott, L. Kummer, P. Egloff, R. A. D. Bathgate, A. Plückthun, Improving the apo-state detergent stability of NTS1 with CHESS for pharmacological and structural studies. *Biochim. Biophys. Acta - Biomembr.* **1838**, 2817–2824 (2014).

185. R. S. Boshuizen *et al.*, A combination of in vitro techniques for efficient discovery of functional monoclonal antibodies against human CXC chemokine receptor-2 (CXCR2). *MAbs*. **6**, 1415–1424 (2014).

186. P. Timmerman, W. C. Puijk, R. H. Meloen, Functional reconstruction and synthetic mimicry of a conformational epitope using CLIPS^{TM} technology. *J. Mol. Recognit.* **20**, 283–299 (2007).

187. P. Giblin *et al.*, Fully human antibodies against the Protease-Activated Receptor-2 (PAR-2)

with anti-inflammatory activity. *Hum. Antibodies*. **20**, 83–94 (2011).

188.   A. Fujimoto *et al.*, Enhancement of antibody responses to native G protein-coupled receptors using E. coli GroEL as a molecular adjuvant in DNA immunization. *J. Immunol. Methods*. **375**, 243–251 (2012).

189.   R. S. Chambers, Antibody production using genetic immunization. *Discov. Med.* **3**, 52–3 (2003).

190.   V. Boivin-Jahns, R. Jahns, "2065 GPCR-autoantibodies in chronic heart failure" (2018), (available at http://www.ncbi.nlm.nih.gov/pubmed/29772546).

191.   T. Miljenović, X. Jia, P. Lavrencic, B. Kobe, M. Mobli, A non-uniform sampling approach enables studies of dilute and unstable proteins. *J. Biomol. NMR*. **68**, 119–127 (2017).

192.   G. Wallukat, H. Prüss, J. Müller, I. Schimke, Functional autoantibodies in patients with different forms of dementia. *PLoS One*. **13**, e0192778 (2018).

193.   O. Cabral-Marques, G. Riemekasten, Functional autoantibodies targeting G protein-coupled receptors in rheumatic diseases. *Nat. Rev. Rheumatol.* **13** (2017), pp. 648–656.

194.   T. Takezako, H. Unal, S. S. Karnik, K. Node, Current topics in angiotensin II type 1 receptor research: Focus on inverse agonism, receptor dimerization and biased agonism. *Pharmacol. Res.* **123** (2017), pp. 40–50.

195.   G. Wallukat *et al.*, Aptamer BC007 for neutralization of pathogenic autoantibodies directed against G-protein coupled receptors: A vision of future treatment of patients with cardiomyopathies and positivity for those autoantibodies. *Atherosclerosis*. **244**, 44–47 (2016).

196.   X. Yu, G. Riemekasten, F. Petersen, Autoantibodies against muscarinic acetylcholine receptor M3 in Sjogren's syndrome and corresponding mouse models. *Front. Biosci. (Landmark Ed.* **23**, 2053–2064 (2018).

197.   G. Wallukat *et al.*, Autoantibodies Directed Against the Endothelin A Receptor in Patients With Benign Prostatic Hyperplasia. *Prostate*. **77**, 458–465 (2016).

198.   E. T. van der Westhuizen, C. Valant, P. M. Sexton, A. Christopoulos, Endogenous Allosteric Modulators of G Protein-Coupled Receptors. *J. Pharmacol. Exp. Ther.* **353**, 246–260 (2015).

199.   N. P. Becker, J. Müller, P. Göttel, G. Wallukat, I. Schimke, Cardiomyopathy — An approach to the autoimmune background. *Autoimmun. Rev.* **16** (2017), pp. 269–286.

200.   A. Jünemann *et al.*, Agonistic autoantibodies to the β2-adrenergic receptor involved in the pathogenesis of open-angle glaucoma. *Front. Immunol.* **9**, 145 (2018).

201.   A. Recke *et al.*, Autoantibodies in Serum of Systemic Scleroderma patients: Peptide-based epitope mapping indicates increased binding to cytoplasmic domains of CXCR3. *Front. Immunol.* **8**, 428 (2018).

202.   J. R. Thyrian *et al.*, Prevalence and Determinants of Agonistic Autoantibodies Against α1-Adrenergic Receptors in Patients Screened Positive for Dementia: Results from the Population-Based DelpHi-Study. *J. Alzheimers. Dis.* **64**, 1091–1097 (2018).

203.   A. Carrea, L. Diambra, Systems Biology Approach to Model the Life Cycle of Trypanosoma cruzi. *PLoS One*. **11**, e0146947 (2016).

204.   P. Bourée, Romana's sign Historical Article. *Our Dermatology Online*. **6**, 248 (2015).

205.   A. Rassi, A. Rassi, J. A. Marin-Neto, Chagas disease. *Lancet*. **375** (2010), pp. 1388–1402.

206. W. C. Van Voorhis, H. Eisen, Fl-160. A surface antigen of Trypanosoma cruzi that mimics mammalian nervous tissue. *J. Exp. Med.* **169**, 641–52 (1989).

207. W. C. Van Voorhis, L. Schlekewy, H. L. Trong, Molecular mimicry by Trypanosoma cruzi: the F1-160 epitope that mimics mammalian nerve can be mapped to a 12-amino acid peptide. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 5993–7 (1991).

208. N. Kerner, P. Li??geard, M. J. Levin, M. Hontebeyrie-Joskowicz, Trypanosoma cruzi: Antibodies to a MAP-like protein in chronic chagas' disease cross-react with mammalian cytoskeleton. *Exp. Parasitol.* **73**, 451–459 (1991).

209. R. S. Tibbetts, T. S. McCormick, E. C. Rowland, S. D. Miller, D. M. Engman, Cardiac antigen-specific autoantibody production is associated with cardiomyopathy in Trypanosoma cruzi-infected mice. *J Immunol*. **152**, 1493–1499 (1994).

210. E. Cunha-Neto *et al.*, Autoimmunity in Chagas disease cardiopathy: biological relevance of a cardiac myosin-specific epitope crossreactive to an immunodominant Trypanosoma cruzi antigen. *Proc. Natl. Acad. Sci. United States Am.* **92**, 3541–3545 (1995).

211. M. J. Levin *et al.*, Identification of major Trypanosoma cruzi antigenic determinants in chronic Chagas' heart disease. *Am. J. Trop. Med. Hyg.* **41**, 530–8 (1989).

212. M. J. Levin, M. Vazquez, D. Kaplan, A. G. Schijman, The Trypanosoma cruzi ribosomal P protein family: Classification and antigenicity. *Parasitol. Today*. **9**, 381–384 (1993).

213. E. A. Mesri *et al.*, Major Trypanosoma cruzi antigenic determinant in Chagas' heart disease shares homology with the systemic lupus erythematosus ribosomal P protein epitope. *J. Clin. Microbiol.* **28**, 1219–1224 (1990).

214. M. J. Ayub *et al.*, Selective blockade of trypanosomatid protein synthesis by a recombinant antibody anti-Trypanosoma cruzi P2β protein. *PLoS One*. **7**, e36233 (2012).

215. H. Yamamoto *et al.*, EF-G and EF4: translocation and back-translocation on the bacterial ribosome. *Nat. Rev. Microbiol.* **12**, 89–100 (2014).

216. P. H. M. Too *et al.*, The C-terminal fragment of the ribosomal P protein complexed to trichosanthin reveals the interaction between the ribosome-inactivating protein and the ribosome. *Nucleic Acids Res.* **37**, 602–610 (2009).

217. A. K. H. Choi, E. C. K. Wong, K. M. Lee, K. B. Wong, Structures of eukaryotic ribosomal stalk proteins and its complex with trichosanthin, and their implications in recruiting ribosome-inactivating proteins to the ribosomes. *Toxins (Basel).* **7** (2015), pp. 638–647.

218. M. J. Levin, M. Vazquez, D. Kaplan, A. G. Schijman, The Trypanosoma cruzi ribosomal P protein family: Classification and antigenicity. *Parasitol. Today*. **9**, 381–384 (1993).

219. K. Petry, H. Eisen, Chagas disease: a model for the study of autoimmune diseases. *Parasitol. Today*. **5**, 111–6 (1989).

220. L. Sterin-Borda, C. Perez Leiros, M. Wald, G. Cremaschi, E. Borda, Antibodies to beta 1 and beta 2 adrenoreceptors in Chagas' disease. *Clin Exp Immunol*. **74**, 349–54. (1988).

221. J. C. Goin, E. Borda, C. Perez Leiros, R. Storino, L. Sterin-Borda, Identification of antibodies with muscarinic cholinergic activity in human Chagas' disease: pathological implications. *J. Auton. Nerv. Syst.* **47**, 45–52 (1994).

222. E. Borda *et al.*, A circulating IgG in Chagas' disease which binds to beta-adrenoceptors of myocardium and modulates their activity. *Clin. Exp. Immunol.* **57**, 679–86 (1984).

223. I. Ferrari *et al.*, Molecular mimicry between the immunodominant ribosomal protein P0 of Trypanosoma crnzi and a functional epitope on the human$β${<}inf{>}1{<}/inf{>}-adrenergic receptor. *J. Exp. Med.* **182**, 59–65 (1995).

224. R. Elies *et al.*, Structural and functional analysis of the B cell epitopes recognized by anti-receptor autoantibodies in patients with Chagas' disease. *J. Immunol.* **157**, 4203–11 (1996).

225. D. Kaplan *et al.*, Antibodies to ribosomal P proteins of Trypanosoma cruzi in Chagas disease possess functional autoreactivity with heart tissue and differ from anti-P autoantibodies in lupus. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 10301–10306 (1997).

226. P. Sepulveda, P. Liegeard, G. Wallukat, M. J. Levin, M. Hontebeyrie, Modulation of cardiocyte functional activity by antibodies against Trypanosoma cruzi ribosomal P2 protein C terminus. *Infect. Immun.* **68**, 5114–5119 (2000).

227. E. Mahler *et al.*, A monoclonal antibody against the immunodominant epitope of the ribosomal P2β protein of Trypanosoma cruzi interacts with the human β1-adrenergic receptor. *Eur. J. Immunol.* **31**, 2210–2216 (2001).

228. E. Mahler, J. Hoebeke, M. J. Levin, Structural and functional complexity of the humoral response against the Trypanosoma cruzi ribosomal P2 beta protein in patients with chronic Chagas' heart disease. *Clin. Exp. Immunol.* **136**, 527–34 (2004).

229. P. Lopez Bergami *et al.*, Immunization with recombinant Trypanosoma cruzi ribosomal P2beta protein induces changes in the electrocardiogram of immunized mice. *FEMS Immunol. Med. Microbiol.* **18**, 75–85 (1997).

230. C. Smulski *et al.*, Structural basis of the cross-reaction between an antibody to the Trypanosoma cruzi ribosomal P2beta protein and the human beta1 adrenergic receptor. *FASEB J.* **20**, 1396–1406 (2006).

231. V. Labovsky, C. R. Smulski, K. Gómez, G. Levy, M. J. Levin, Anti-beta1-adrenergic receptor autoantibodies in patients with chronic Chagas heart disease. *Clin. Exp. Immunol.* **148**, 440–9 (2007).

232. J. C. Pizarro *et al.*, Crystal structure of the complex mAb 17.2 and the C-terminal region of Trypanosoma cruzi P2?? protein: Implications in cross-reactivity. *PLoS Negl. Trop. Dis.* **5**, e1375 (2011).

233. R. Verardi, N. J. Traaseth, L. R. Masterson, V. V Vostrikov, G. Veglia, Isotope labeling for solution and solid-state NMR spectroscopy of membrane proteins. *Adv. Exp. Med. Biol.* **992**, 35–62 (2012).

234. J. G. Baker, R. G. W. Proudman, C. G. Tate, The pharmacological effects of the thermostabilising (m23) mutations and intra and extracellular (β36) deletions essential for crystallisation of the turkey β-adrenoceptor. *Naunyn. Schmiedebergs. Arch. Pharmacol.* **384**, 71–91 (2011).

235. K. Y. Chung *et al.*, Role of detergents in conformational exchange of a G protein-coupled receptor. *J. Biol. Chem.* **287**, 36305–11 (2012).

236. G. S. Rule, T. K. Hitchens, *Fundamentals of Protein NMR Spectroscopy* (Springer Science & Business Media, 2006; https://books.google.com/books?id=pB-hsMnPy60C&pgis=1).

237. G. M. Hautbergue, A. P. Golovanov, Increasing the sensitivity of cryoprobe protein NMR experiments by using the sole low-conductivity arginine glutamate salt. *J. Magn. Reson.* **191**, 335–9 (2008).

238. R. G. Shorr, R. J. Lefkowitz, M. G. Caron, Purification of the beta-adrenergic receptor. Identification of the hormone binding subunit. *J. Biol. Chem.* **256**, 5820–6 (1981).

239. B. Carpenter, R. Nehmé, T. Warne, A. G. W. Leslie, C. G. Tate, Structure of the adenosine A2Areceptor bound to an engineered G protein. *Nature* (2016), doi:10.1038/nature18966.

240. Y. Kang *et al.*, Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature* (2015), doi:10.1038/nature14656.

241. Y. Kofuku *et al.*, Functional dynamics of deuterated β2 -adrenergic receptor in lipid bilayers revealed by NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* **53**, 13376–9 (2014).

242. Y. Kofuku *et al.*, Efficacy of the β_2-adrenergic receptor is determined by conformational equilibrium in the transmembrane region. *Nat. Commun.* **3**, 1045 (2012).

243. R. Sounier *et al.*, Propagation of conformational changes during μ-opioid receptor activation. *Nature* (2015), doi:10.1038/nature14680.

244. S. Isogai *et al.*, Backbone NMR reveals allosteric signal transduction networks in the β1-adrenergic receptor. *Nature*. **314**, 1–17 (2016).

245. J. G. Baker, A full pharmacological analysis of the three turkey β-adrenoceptors and comparison with the human β-adrenoceptors. *PLoS One*. **5**, e15487 (2010).

246. T. Sato *et al.*, Pharmacological Analysis and Structure Determination of 7-Methylcyanopindolol-Bound β1-Adrenergic Receptor. *Mol. Pharmacol.* **88**, 1024–34 (2015).

247. S. Engelhardt, Y. Grimmer, G. H. Fan, M. J. Lohse, Constitutive activity of the human β_1-adrenergic receptor in β_1-receptor transgenic mice. *Mol. Pharmacol.* (2001).

248. S. Wang *et al.*, Solid-state NMR spectroscopy structure determination of a lipid-embedded heptahelical membrane protein. *Nat. Methods*. **10**, 1007–1012 (2013).

249. A. Gautier, H. R. Mott, M. J. Bostock, J. P. Kirkpatrick, D. Nietlispach, Structure determination of the seven-helix transmembrane receptor sensory rhodopsin II by solution NMR spectroscopy. *Nat. Struct. Mol. Biol.* **17**, 768–774 (2010).

250. G. L. Butterfoss *et al.*, Conformational dependence of 13C shielding and coupling constants for methionine methyl groups. *J. Biomol. NMR*. **48**, 31–47 (2010).

251. B. K. Kobilka, X. Deupi, Conformational complexity of G-protein-coupled receptors. *Trends Pharmacol. Sci.* **28**, 397–406 (2007).

252. R. Lape, D. Colquhoun, L. G. Sivilotti, On the nature of partial agonism in the nicotinic receptor superfamily. *Nature* (2008), doi:10.1038/nature07139.

253. M. J. Bostock, A. S. Solt, D. Nietlispach, The role of NMR spectroscopy in mapping the conformational landscape of GPCRs. *Curr. Opin. Struct. Biol.* **57** (2019), pp. 145–156.

254. L. Ye, N. Van Eps, M. Zimmer, O. P. Ernst, R. S. Prosser, Activation of the A_{2A} adenosine G-protein-coupled receptor by conformational selection. *Nature*. **533**, 265–268 (2016).

255. L. Ye *et al.*, Mechanistic insights into allosteric regulation of the A_{2A} adenosine G protein-coupled receptor by physiological cations. *Nat. Commun.* **9**, 1372 (2018).

256. M. T. Eddy *et al.*, Allosteric Coupling of Drug Binding and Intracellular Signaling in the A2A Adenosine Receptor. *Cell*. **172**, 68–80 (2018).

257. M. T. Eddy *et al.*, Extrinsic Tryptophans as NMR Probes of Allosteric Coupling in Membrane

Proteins: Application to the A2A Adenosine Receptor. *J. Am. Chem. Soc.* **140**, 8228–8235 (2018).

258.   L. D. Clark *et al.*, Ligand modulation of sidechain dynamics in a wild-type human GPCR. *Elife*. **6**, 1–27 (2017).

259.   A. S. Solt *et al.*, Insight into partial agonism by observing multiple equilibria for ligand-bound and G_s-mimetic nanobody-bound β_1-adrenergic receptor. *Nat. Commun.* **8**, 1795 (2017).

260.   S. Isogai *et al.*, Backbone NMR reveals allosteric signal transduction networks in the β_1-adrenergic receptor. *Nature*. **530**, 237–241 (2016).

261.   Y. Shiraishi *et al.*, Phosphorylation-induced conformation of β_2-adrenoceptor related to arrestin recruitment revealed by NMR. *Nat. Commun.* **9**, 194 (2018).

262.   R. Nygaard *et al.*, The dynamic process of β_2-adrenergic receptor activation. *Cell*. **152**, 532–542 (2013).

263.   J. J. Liu, R. Horst, V. Katritch, R. C. Stevens, K. Wüthrich, Biased signaling pathways in β_2-adrenergic receptor characterised by ^{19}F-NMR. *Science*. **335**, 1106–1110 (2012).

264.   Y. Kofuku *et al.*, Efficacy of the β_2-adrenergic receptor is determined by conformational equilibrium in the transmembrane region. *Nat. Commun.* **3**, 1045 (2012).

265.   Y. Kofuku *et al.*, Functional dynamics of deuterated β_2-adrenergic receptor in lipid bilayers revealed by NMR spectroscopy. *Angew. Chemie Int. Ed.* **53**, 13376–13379 (2014).

266.   T. H. Kim *et al.*, The role of ligands on the equilibria between functional states of a G protein-coupled receptor. *J. Am. Chem. Soc.* **135**, 9465–74 (2013).

267.   R. Horst, J. J. Liu, R. C. Stevens, K. Wüthrich, β_2 -Adrenergic Receptor Activation by Agonists Studied with ^{19}F NMR Spectroscopy. *Angew. Chemie Int. Ed.* **52**, 10762–10765 (2013).

268.   K. Y. Chung *et al.*, Role of detergents in conformational exchange of a G protein-coupled receptor. *J. Biol. Chem.* **287**, 36305–36311 (2012).

269.   A. Manglik *et al.*, Structural insights into the dynamic process of β_2-adrenergic receptor signaling. *Cell*. **161**, 1101–1111 (2015).

270.   M. Casiraghi *et al.*, Functional Modulation of a G Protein-Coupled Receptor Conformational Landscape in a Lipid Bilayer. *J. Am. Chem. Soc.* **138**, 11170–11175 (2016).

271.   F. Bumbak *et al.*, Optimization and $^{13}CH_3$ methionine labeling of a signaling competent neurotensin receptor 1 variant for NMR studies. *Biochim. Biophys. Acta - Biomembr.* **1860**, 1372–1383 (2018).

272.   R. Sounier *et al.*, Propagation of conformational changes during μ-opioid receptor activation. *Nature*. **524**, 375–378 (2015).

273.   J. Okude *et al.*, Identification of a Conformational Equilibrium That Determines the Efficacy and Functional Selectivity of the μ-Opioid Receptor. *Angew. Chemie - Int. Ed.* **54**, 15771–15776 (2015).

274.   P. Samama, S. Cotecchia, T. Costa, R. J. Lefkowitz, A mutation-induced activated state of the β2-adrenergic receptor. Extending the ternary complex model. *J. Biol. Chem.* (1993).

275.   S. Isogai *et al.*, Backbone NMR reveals allosteric signal transduction networks in the β1-adrenergic receptor. *Nature*. **530**, 237–241 (2016).

276. J. J. Liu, R. Horst, V. Katritch, R. C. Stevens, K. Wüthrich, Biased signaling pathways in β2-adrenergic receptor characterized by 19F-NMR. *Science (80-.).* **335**, 1106–1110 (2012).

277. T. Che *et al.*, Structure of the Nanobody-Stabilized Active State of the Kappa Opioid Receptor. *Cell.* **172**, 55-67.e15 (2018).

278. S. G. F. Rasmussen *et al.*, Structure of a nanobody-stabilized active state of the β2 adrenoceptor. *Nature.* **469**, 175–180 (2011).

279. A. Koehl *et al.*, Structure of the μ-opioid receptor-G i protein complex. *Nature* (2018), doi:10.1038/s41586-018-0219-7.

280. C. Smulski *et al.*, Structural basis of the cross-reaction between an antibody to the Trypanosoma cruzi ribosomal P2beta protein and the human beta1 adrenergic receptor. *FASEB J.* **20**, 1396–1406 (2006).

281. D. P. Humphreys, Production of antibodies and antibody fragments in Escherichia coli and a comparison of their functions, uses and modification. *Curr. Opin. Drug Discov. Devel.* **6**, 188–96 (2003).

282. V. Salema, L. Á. Fernández, High yield purification of nanobodies from the periplasm of E. coli as fusions with the maltose binding protein. *Protein Expr. Purif.* **91**, 42–48 (2013).

283. M. V. Goncharuk *et al.*, Bacterial synthesis, purification, and solubilization of transmembrane segments of ErbB family receptors. *Mol. Biol.* **45**, 823–832 (2011).

284. R. Ullah *et al.*, M. J. van Raaij, Ed., Correction: Activity of the human rhinovirus 3C protease studied in various buffers, additives and detergents solutions for recombinant protein production. *PLoS One.* **11** (2016), p. e0153436.

285. F. Q. Schafer, G. R. Buettner, Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. *Free Radic. Biol. Med.* **30** (2001), pp. 1191–1212.

286. V. Domkin, A. Chabes, Phosphines are ribonucleotide reductase reductants that act via C-terminal cysteines similar to thioredoxins and glutaredoxins. *Sci. Rep.* **4**, 5539 (2014).

287. L. Björck, Protein L. A novel bacterial cell wall protein with affinity for Ig L chains. *J. Immunol.* **140**, 1194–1197 (1988).

288. K. M. Arndt, K. M. Müller, A. Plückthun, Factors influencing the dimer to monomer transition of an antibody single-chain Fv fragment. *Biochemistry.* **37**, 12918–12926 (1998).

289. J. Sambrook, D. W. Russell, *Cold Spring Harb. Protoc.*, in press, doi:10.1101/pdb.prot3944.

290. R. Green, E. J. Rogers, Chemical Transformation of E. coli. *Methods Enzym.* **529**, 329–336 (2013).

291. N. Guthertz, J. Klopp, A. Winterhalter, C. Fernández, A. D. Gossert, Auto-inducing media for uniform isotope labeling of proteins with (15)N, (13)C and (2)H. *J. Biomol. NMR.* **62**, 169–77 (2015).

292. F. W. Studier, Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).

293. P. Schanda, E. Kupĉe, B. Brutscher, SOFAST-HMQC experiments for recording two-dimensional deteronuclear correlation spectra of proteins within a few seconds. *J. Biomol. NMR.* **33**, 199–211 (2005).

155

294.   A. L. Davis, J. Keeler, E. D. Laue, D. Moskau, Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients. *J. Magn. Reson.* **98**, 207–216 (1992).

295.   E. Kupce, R. Freeman, Polychromatic Selective Pulses. *J. Magn. Reson. Ser. A*. **102**, 122–126 (1993).

296.   H. Geen, R. Freeman, Band-selective radiofrequency pulses. *J. Magn. Reson.* **93**, 93–141 (1991).

297.   S. G. Hyberts, K. Takeuchi, G. Wagner, Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data. *J. Am. Chem. Soc.* **132**, 2145–2147 (2010).

298.   M. J. Bostock, D. J. Holland, D. Nietlispach, Compressed sensing reconstruction of undersampled 3D NOESY spectra: application to large membrane proteins. *J. Biomol. NMR*. **54**, 15–32 (2012).

299.   W. F. Vranken *et al.*, The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins Struct. Funct. Genet.* **59**, 687–696 (2005).

Appendix A – Development of turkey β₁AR constructs

P0 – WT-tβ1AR [NP_001290104.1] – 483 AA

MGDGWLPPDCGPHNRSGGGGATAAPTGSRQVSAELLSQQWEAGMSLLMALVVLLIV
AGNVLVIAAIGRTQRLQTLTNLFITSLACADLVMGLLVVPFGATLVVRGTWLWGSFLCEC
WTSLDVLCVTASIETLCVIAIDRYLAITSPFRYQSLMTRARAKVIICTVWAISALVSFLPIMM
HWWRDEDPQALKCYQDPGCCDFVTNRAYAIASSIISFYIPLLIMIFVYLRVYREAKEQIRKI
DRCEGRFYGSQEQPQPPPLPQHQPILGNGRASKRKTSRVMAMREHKALKTLGIIMGVFT
LCWLPFFLVNIVNVFNRDLVPDWLFVFFNWLGYANSAFNPIIYCRSPDFRKAFKRLLCFPR
KADRRLHAGGQPAPLPGGFISTLGSPEHSPGGTWSDCNGGTRGGSESSLEERHSKTSRS
ESKMEREKNILATTRFYCTFLGNGDKAVFCTVLRIVKLFEDATCTCPHTHKLKMKWRFKQ
HQA

- 425E-483A DEL
- C116L
  to increase yield

tβtrunc

P0.1 – tβtrunc (1-424)-C116L

MGDGWLPPDCGPHNRSGGGGATAAPTGSRQVSAELLSQQWEAGMSLLMALVVLLIV
AGNVLVIAAIGRTQRLQTLTNLFITSLACADLVMGLLVVPFGATLVVRGTWLWGSFLCEL
WTSLDVLCVTASIETLCVIAIDRYLAITSPFRYQSLMTRARAKVIICTVWAISALVSFLPIMM
HWWRDEDPQALKCYQDPGCCDFVTNRAYAIASSIISFYIPLLIMIFVYLRVYREAKEQIRKI
DRCEGRFYGSQEQPQPPPLPQHQPILGNGRASKRKTSRVMAMREHKALKTLGIIMGVFT
LCWLPFFLVNIVNVFNRDLVPDWLFVFFNWLGYANSAFNPIIYCRSPDFRKAFKRLLCFPR
KADRRLHAGGQPAPLPGGFISTLGSPEHSPGGTWSDCNGGTRGGSESSLEERHSKTSRS
ESKM

J. G. Baker, R. G. W. Proudman, C. G. Tate, The pharmacological effects…
Naunyn. Schmiedebergs. Arch. Pharmacol. **384**, 71–91 (2011).

- 2G-19G DEL
- 425E-483A DEL
- C-term His-6
- C116L to increase yield

(20-242)

P1 – tβ1AR (20-424)-C116L-His6 – 411AAA

MGATAAPTGSRQVSAELLSQQWEAGMSLLMALVVLLIVAGNVLVIAAIGRTQRLQTLT
NLFITSLACADLVMGLLVVPFGATLVVRGTWLWGSFLCELWTSLDVLCVTASIETLCVIAI
DRYLAITSPFRYQSLMTRARAKVIICTVWAISALVSFLPIMMHWWRDEDPQALKCYQDP
GCCDFVTNRAYAIASSIISFYIPLLIMIFVYLRVYREAKEQIRKIDRCEGRFYGSQEQPQPPPL
PQHQPILGNGRASKRKTSRVMAMREHKALKTLGIIMGVFTLCWLPFFLVNIVNVFNRDL
VPDWLFVFFNWLGYANSAFNPIIYCRSPDFRKAFKRLLCFPRKADRRLHAGGQPAPLPG
GFISTLGSPEHSPGGTWSDCNGGTRGGSESSLEERHSKTSRSESKMHHHHHH

- 22T-33A DEL

T. Warne, J. Chirnside, G. F. X. Schertler, Expression and purification of
truncated, non-glycosylated turkey beta-adrenergic receptors for
crystallization. *Biochim. Biophys. Acta - Biomembr.* **1610**, 133–140 (2003).

β6

P2 – β6 (34-424)-C116L-His6 (=β1AR6, β6, β1AR^ΔNC) – 400 AA

MGAELLSQQWEAGMSLLMALVVLLIVAGNVLVIAAIGRTQRLQTLTNLFITSLACADLV
MGLLVVPFGATLVVRGTWLWGSFLCELWTSLDVLCVTASIETLCVIAIDRYLAITSPFRYQ
SLMTRARAKVIICTVWAISALVSFLPIMMHWWRDEDPQALKCYQDPGCCDFVTNRAYA
IASSIISFYIPLLIMIFVYLRVYREAKEQIRKIDRCEGRFYGSQEQPQPPPLPQHQPILGNGRA
SKRKTSRVMAMREHKALKTLGIIMGVFTLCWLPFFLVNIVNVFNRDLVPDWLFVFFNWL
GYANSAFNPIIYCRSPDFRKAFKRLLCFPRKADRRLHAGGQPAPLPGGFISTLGSPEHSPG
GTWSDCNGGTRGGSESSLEERHSKTSRSESKMHHHHHH

- 368H-424M DEL
- C358A remove palmitoylation site
- m23 mutations:
  - R68S
  - M90V
  - Y227A
  - A282L
  - F327A
  - F338M

T. Warne, J. Chirnside, G. F. X. Schertler, Expression and purification of truncated, non-glycosylated turkey beta-adrenergic receptors for crystallization. *Biochim. Biophys. Acta - Biomembr.* **1610**, 133–140 (2003).

T. Warne, M. J. Serrano-Vega, C. G. Tate, G. F. X. Schertler, Development and crystallization of a minimal thermostabilised G protein-coupled receptor. *Protein Expr. Purif.* **65**, 204–213 (2009).

B. Carpenter, C. G. Tate, Engineering a minimal G Protein to facilitate crystallisation of G protein - coupled receptors in their active conformation *Protein Eng. Des. Sel.* **29**, 583–594 (2016).

β6-m23

P2.1 - β6-m23 (34-424)-(R68S, M90V, Y227A, A282L, F327A, F338M)-C116L-His6 (=β1AR-m23, βAR₃₄₋₄₂₄ in second reference)

- 244C-271R DEL (ICL3)
- 275R-278S DEL (ICL3)

MGAELLSQQWEAGMSLLMALVVLLIVAGNVLVIAAIGSTQRLQTLTNLFITSLACADLVV
GLLVVPFGATLVVRGTWLWGSFLCELWTSLDVLCVTASIETLCVIAIDRYLAITSPFRYQSL
MTRARAKVIICTVVWAISALVSFLPIMMHWWRDEDPQALKCYQDPGCCDFVTNRAYAIA
SSIISFYIPLLIMIFVALRVYREAKEQIRKIDRCEGRFYGSQEQPQPPPLPQHQPILGNGRAS
KRKTSRVMLMREHKALKTLGIIMGVFTLCWLPFFLVNIVNVFNRDLVPDWLFVAFNWLG
YANSAMNPIIYCRSPDFRKAFKRLLCFPRKADRRLHAGGQPAPLPGGFISTLGSPEHSPGG
TWSDCNGGTRGGSESSLEERHSKTSRSESKMHHHHHH

J. L. Miller, C. G. Tate, Engineering an ultra-thermostable β(1)-adrenoceptor. *J. Mol. Biol.* **413**, 628–38 (2011).

M. J. Serrano-Vega, F. Magnani, Y. Shibata, C. G. Tate, Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 877–882 (2008).

β6-JM3

P2.2 - JM3-(34-424)-(D200E, D322K, I129V, Y343L)-C116L-His6 – 400 AA (44.78 kDa)

MGAELLSQQWEAGMSLLMALVVLLIVAGNVLVIAAIGSTQRLQTLTNLFITSLACADLVV
GLLVVPFGATLVVRGTWLWGSFLCELWTSLDVLCVTASVETLCVIAIDRYLAITSPFRYQSL
MTRARAKVIICTVVWAISALVSFLPIMMHWWRDEDPQALKCYQDPGCCEFVTNRAYAIA
SSIISFYIPLLIMIFVALRVYREAKEQIRKIDRCEGRFYGSQEQPQPPPLPQHQPILGNGRAS
KRKTSRVMLMREHKALKTLGIIMGVFTLCWLPFFLVNIVNVFNRDLVPKWLFVAFNWLG
YANSAMNPIILCRSPDFRKAFKRLLCFPRKADRRLHAGGQPAPLPGGFISTLGSPEHSPGG
TWSDCNGGTRGGSESSLEERHSKTSRSESKMHHHHHH

J. L. Miller, C. G. Tate, Engineering an ultra-thermostable β(1)-adrenoceptor. *J. Mol. Biol.* **413**, 628–38 (2011).

β34

P3 - β34 (34-367)-(C116L, C358A) – 343 AA (39.3 kDa)

- 244C-271R DEL (ICL3)
- 275R-278S DEL (ICL3)

MGAELLSQQWEAGMSLLMALVVLLIVAGNVLVIAAIGRTQRLQTLTNLFITSLACADLV
MGLLVVPFGATLVVRGTWLWGSFLCELWTSLDVLCVTASIETLCVIAIDRYLAITSPFRYQ
SLMTRARAKVIICTVVWAISALVSFLPIMMHWWRDEDPQALKCYQDPGCCDFVTNRAYA
IASSIISFYIPLLIMIFVYLRVYREAKEQIRKIDRCEGRFYGSQEQPQPPPLPQHQPILGNGRA
SKRKTSRVMAMREHKALKTLGIIMGVFTLCWLPFFLVNIVNVFNRDLVPDWLFVFFNWL
GYANSAFNPIIYCRSPDFRKAFKRLLAFPRKADRRLHHHHHH

T. Warne, M. J. Serrano-Vega, C. G. Tate, G. F. X. Schertler, Development and crystallization of a minimal thermostabilised G protein-coupled receptor. *Protein Expr. Purif.* **65**, 204–213 (2009).

P4 - β35 (34-367)-(C116L, C358A) – ICL3 Del1

MGAELLSQQWEAGMSLLMALVVLLIVAGNVLVIAAIGRTQRLQTLTNLFITSLACADLV
MGLLVVPFGATLVVRGTWLWGSFLCELWTSLDVLCVTASIETLCVIAIDRYLAITSPFRYQ
SLMTRARAKVIICTVWAISALVSFLPIMMHWWRDEDPQALKCYQDPGCCDFVTNRAYA
IASSIISFYIPLLIMIFVYLRVYREAKEQIRKIDRASKRVMAMREHKALKTLGIIMGVFTLCWL
PFFLVNIVNVFNRDLVPDWLFVFFNWLGYANSAFNPIIYCRSPDFRKAFKRLLAFPRKADR
RLHHHHHH

T. Warne, M. J. Serrano-Vega, C. G. Tate, G. F. X. Schertler, Development and crystallization of a minimal thermostabilised G protein-coupled receptor. *Protein Expr. Purif.* **65**, 204–213 (2009).

- 244C-271R DEL (ICL3)
- 277T-278S DEL (ICL3)

P5 - β36 (34-367)-(C116L, C358A) – ICL3 Del2

MGAELLSQQWEAGMSLLMALVVLLIVAGNVLVIAAIG<u>R</u>TQRLQTLTNLFITSLACADLV
<u>M</u>GLLVVPFGATLVVRGTWLWGSFLCELWTSLDVLCVTASIETLCVIAIDRYLAITSPFRYQ
SLMTRARAKVIICTVWAISALVSFLPIMMHWWRDEDPQALKCYQDPGCCDFVTNRAYA
IASSIISFYIPLLIMIFV<u>Y</u>LRVYREAKEQIRKIDRASKR<u>K</u>RVM<u>A</u>MREHKALKTLGIIMGVFTLC
WLPFFLVNIVNVFNRDLVPDWLFV<u>F</u>FNWLGYANSA<u>F</u>NPIIYCRSPDFRKAFKRLLAFPRK
ADRRLHHHHHH

T. Warne, M. J. Serrano-Vega, C. G. Tate, G. F. X. Schertler, Development and crystallization of a minimal thermostabilised G protein-coupled receptor. *Protein Expr. Purif.* **65**, 204–213 (2009).

- m23 mutations:
  - <u>R68S</u>
  - <u>M90V</u>
  - <u>Y227A</u>
  - <u>A282L</u>
  - <u>F327A</u>
  - <u>F338M</u>

P6 - β36-m23 (34-367)-(R68S, M90V, Y227A, A282L, F327A, F338M)-(C116L, C358A)

MGAELLSQQWEAGMSLLMALVVLLIVAGNVLVIAAIGSTQRLQTLTNLFITSLACADLVV
GLLVVPFGATLVVRGTWLWGSFLCELWTSLDVLCVTASIETLCVIAIDRYLAITSPFRYQSL
MTRARAKVIICTVWAISALVSFLPIMMHWWRDEDPQALKCYQDPGCCDFVTNRAYAIA
SSIISFYIPLLIMIFVALRVYREAKEQIRKIDRASKRKRVMLMREHKALKTLGIIMGVFTLCW
LPFFLVNIVNVFNRDLVPDWLFVAFNWLGYANSAMNPIIYCRSPDFRKAFKRLLAFPRKA
DRRLHHHHHH

T. Warne *et al.*, Structure of a β1-adrenergic G-protein-coupled receptor. *Nature*. **454** (2008), pp. 486–491.

R. Moukhametzianov *et al.*, Two distinct conformations of helix 6 observed in antagonist-bound structures of a beta1-adrenergic receptor. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 8228–32 (2011).

- T277+S278 ADD

P7 - β44-m23 (34-367)-(R68S, M90V, Y227A, A282L, F327A, F338M)-(T277+S278)-(C116L, C358A)

MGAELLSQQWEAGMSLLMALVVLLIVAGNVLVIAAIGSTQRLQTLTNLFITSLACADLVV
GLLVVPFGATLVVRGTWLWGSFLCELWTSLDVLCVTASIETLCVIAIDRYLAITSPFRYQSL
MTRARAKVIICTVWAISALVSFLPIMMHWWRDEDPQALKCYQDPGCCDFVTNRAYAIA
SSIISFYIPLLIMIFVALRVYREAKEQIRKIDRASKRKTSRVMLMREHKALKTLGIIMGVFTLC

WLPFFLVNIVNVFNRDLVPDWLFVAFNWLGYANSAMNPIIYCRSPDFRKAFKRLLAFPR
KADRRLHHHHHH

T. Warne *et al.*, The structural basis for agonist and partial agonist action on a

T. Warne, P. C. Edwards, A. G. W. Leslie, C. G. Tate, Crystal structures of a stabilized β1-adrenoceptor bound to the biased agonists bucindolol and carvedilol. *Structure*. **20**, 841–849 (2012).

- I129V
- D322K
- Y343L

JM50 ➤

P7.1 – β1AR-JM50 (34-367)-(R68S, M90V, Y227A, A282L, F327A, F338M)-(T277+S278)-(I129V, D322K, Y343L)-(C116L, C358A) (=β44-m23+JM50) – 315 AA, (35.85 kDa)

MGAELLSQQWEAGMSLLMALVVLLIVAGNVLVIAAIGSTQRLQTLTNLFITSLACADLVV
GLLVVPFGATLVVRGTWLWGSFLCELWTSLDVLCVTASVETLCVIAIDRYLAITSPFRYQSL
MTRARAKVIICTVVWAISALVSFLPIMMHWWRDEDPQALKCYQDPGCCDFVTNRAYAIA
SSIISFYIPLLIMIFVALRVYREAKEQIRKIDRASKRKTSRVMLMREHKALKTLGIIMGVFTLC
WLPFFLVNIVNVFNRDLVPKWLFVAFNWLGYANSAMNPIILCRSPDFRKAFKRLLAFPRK
ADRRLHHHHHH

J. L. Miller-Gallacher *et al.*, The 2.1 Å resolution structure of cyanopindolol-bound β1-adrenoceptor identifies an intramembrane Na+ ion that stabilises the ligand-free receptor. *PLoS One*. **9**, e92727 (2014).

- I129V
- E130W
- Y343L

β44-TS ➤

7.2 – β44-TS (34-367)-(R68S, M90V, Y227A, A282L, F327A, F338M)-(T277+S278)-(I129V, E130W, Y343L)-(C116L, C358A) (=β44-m23+JM50) – 315 AA, (35.9 kDa)

MGAELLSQQWEAGMSLLMALVVLLIVAGNVLVIAAIGSTQRLQTLTNLFITSLACADLVV
GLLVVPFGATLVVRGTWLWGSFLCELWTSLDVLCVTASVWTLCVIAIDRYLAITSPFRYQS
LMTRARAKVIICTVVWAISALVSFLPIMMHWWRDEDPQALKCYQDPGCCDFVTNRAYAI
ASSIISFYIPLLIMIFVALRVYREAKEQIRKIDRASKRKTSRVMLMREHKALKTLGIIMGVFTL
CWLPFFLVNIVNVFNRDLVPDWLFVAFNWLGYANSAMNPIILCRSPDFRKAFKRLLAFPR
KADRRLHHHHHH

T. Sato *et al.*, Pharmacological Analysis and Structure Determination of 7-Methylcyanopindolol-Bound β1-Adrenergic Receptor. *Mol. Pharmacol.* **88**, 1024–34 (2015).

- I129V
- D322K
- Y343L
- D200E

TS-β1AR ➤

7.3 – TS-β1AR (34-367)-(R68S, M90V, Y227A, A282L, F327A, F338M)-(T277+S278)-(I129V, Y343L, D200E)-(C116L, C358A) (=β44-m23+JM50) – 315 AA, (35.9 kDa)

MGAELLSQQWEAGMSLLMALVVLLIVAGNVLVIAAIGSTQRLQTLTNLFITSLACADLVV
GLLVVPFGATLVVRGTWLWGSFLCELWTSLDVLCVTASVETLCVIAIDRYLAITSPFRYQSL
MTRARAKVIICTVVWAISALVSFLPIMMHWWRDEDPQALKCYQDPGCCEFVTNRAYAIA
SSIISFYIPLLIMIFVALRVYREAKEQIRKIDRASKRKTSRVMLMREHKALKTLGIIMGVFTLC
WLPFFLVNIVNVFNRDLVPKWLFVAFNWLGYANSAMNPIILCRSPDFRKAFKRLLAFPRK
ADRRLHHHHHH

S. Isogai *et al.*, Backbone NMR reveals allosteric signal transduction networks in the β1-adrenergic receptor. *Nature*. **314**, 1–17 (2016).

- Stabilisation removed:
  V90M, A227Y, L282A
- Added:
  E130W, D322K (JM3)

Met2

P8 - Met2 (34-367)-(R68S, F327A, F338M)-(V90M, A227Y, L282A)-(E130W, D322K)-(TS)-(C116L, C358A)-His8 tag – 317 AA (36.33 kDa)

MGAELLSQQWEAGMSLLMALVVLLIVAGNVLVIAAIGSTQRLQTLTNLFITSLACADLV
MGLLVVPFGATLVVRGTWLWGSFLCELWTSLDVLCVTASIWTLCVIAIDRYLAITSPFRYQ
SLMTRARAKVIICTVVWAISALVSFLPIMMHWWRDEDPQALKCYQDPGCCDFVTNRAYA
IASSIISFYIPLLIMIFVYLRVYREAKEQIRKIDRASKRKTSRVMAMREHKALKTLGIIMGVFT
LCWLPFFLVNIVNVFNRDLVPKWLFVAFNWLGYANSAMNPIIYCRSPDFRKAFKRLLAFP
RKADRRLHHHHHHHH

- Methionines mutated:
  M44L, M48L, M179L, M281A, M338A
- K322D

Met2-Δ5

P8.1 - Met2-Δ5 (34-367)-(R68S, F327A)-(V90M, A227Y, L282A)-(E130W, K322D)-(TS)-(C116L, C358A)-(M44L, M48L, M179L, M281A, M338A)-His8 tag – 317 AA (36.33 kDa) (also called as Met3)

MGAELLSQQWEAGLSLLLALVVLLIVAGNVLVIAAIGSTQRLQTLTNLFITSLACADLVMG
LLVVPFGATLVVRGTWLWGSFLCELWTSLDVLCVTASIWTLCVIAIDRYLAITSPFRYQSL
MTRARAKVIICTVVWAISALVSFLPIMLHWWRDEDPQALKCYQDPGCCDFVTNRAYAIAS
SIISFYIPLLIMIFVYLRVYREAKEQIRKIDRASKRKTSRVAAMREHKALKTLGIIMGVFTLCW
LPFFLVNIVNVFNRDLVPDWLFVAFNWLGYANSAANPIIYCRSPDFRKAFKRLLAFPRKA
DRRLHHHHHHHH

Explanation of mutations: m23 set: R68S, F327A; m23 reversed: V90M, A227Y, L282A; β44-TS mutation (Sato *et al.*): E130W (relevance to a β2AR Zn bridge); salt bridge reversal: K322D; ICL3 adjustment: TS; increased expression: C116L; palmitoylation site: C358A; coincidentally stabilising: M338A