

**Some pages of this thesis may have been removed for copyright restrictions.**

If you have discovered material in Aston Research Explorer which is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please read our [Takedown policy](#) and contact the service immediately (openaccess@aston.ac.uk)

# **Characterisation of Styrene-maleic acid-solubilised G protein-coupled receptors**

**Romez Uddin**

Doctor of Philosophy

Aston University

September 2020

©Romez Uddin asserts his moral right to be identified as the author of this thesis. This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright belongs to its author and that no quotation from the thesis and no information derived from it may be published without appropriate permission or acknowledgement.

I would like to dedicate this thesis to my wife, Noorin and my parents, Shaira and Giash.

## **Acknowledgements**

I would like to thank my supervisor, Prof. D.R. Poyner for his guidance and support throughout my project. I would also like to thank Prof. M. Wheatley for his advice and support. I would like to thank my associate supervisor, Prof. R. Bill and her research group for helping me settle into the lab. I'd like to thank Dr J. Simms for demonstrating the photoaffinity cross-linking assay, mutagenesis and for discussing scientific topics. I'd also like to thank Dr S.J. Routledge for demonstrating different assays and her input in optimally culturing *Pichia pastoris*. I also want to give thanks to Dr M. Claire and Dr A. Patel for demonstrating molecular techniques to help me begin my project.

I would also like to thank my collaborators, who worked with me to obtain novel pharmacological data. My photoaffinity cross-linking project was a continuation of Dr J. Simms project to map the topology of a receptor binding site, where Dr J.Simms showed me to perform the cross-linking experiments. Dr. S. Bridden, Dr Joelle Goulding and Rachel Grime were a part of the fluorescence correlation spectroscopy project, as a part of the COMPARE scheme. Dr V. Bavro and Dr. S Gupta analysed my receptor samples to detect water-exposed regions of a protein. Aiman Gullamhussein from Dr A. Rothnie's research group collaborated with me to test the recently emerging polymer termed DIBMA.

I would finally like to give thanks to UCB, Slough for providing me with their phage libraries to perform the phage display experiments.

## Abstract

Detergents have historically been used to solubilise membrane proteins for structural studies and pharmacological research, however detergents can alter the lipid environment surrounding a membrane protein. The styrene-maleic acid (SMA) copolymer has been designed to solubilise membrane proteins from the cell membrane, with the lipid bilayer intact, thus forming styrene-maleic acid lipid particles (SMALP). This would retain the native conformation of the protein and therefore suitable for applications such as drug discovery. In this project, the adenosine 2A receptor ( $A_{2A}R$ ) and the calcitonin gene-related peptide receptor (CGRPR) were solubilised into SMALPs. Various techniques were used to characterise the SMA-solubilised receptors.

Out of the SMA copolymers tested, SMA2000 was chosen as the copolymer to solubilise the GPCRs. The copolymer was also compared with the new diisobutylene-maleic acid (DIBMA) copolymer, which has better resistance to divalent cations than the SMA copolymer. Molecular techniques confirmed the expression of the GPCRs in membranes and after solubilisation into SMALPs. Radioligand binding assays demonstrated that the  $A_{2A}R$  retained its binding capability when solubilised and purified. The binding assay showed that the  $A_{2A}R$  was more stable in SMALPs than DIBMA lipid particles (DIBMALP). Various techniques were used to characterise the  $A_{2A}R$ -SMALP, providing novel properties of GPCRs in SMALP. The x-ray radiolytic footprinting (XRF) was used to detect regions of the GPCR-SMALP which were exposed to hydroxyl modification. The transmembrane domain, and the intracellular surface of the SMA-solubilised  $A_{2A}R$  were exposed to water, demonstrating the SMALPs can successfully be used in XRF. Fluorescence correlation spectroscopy (FCS) was implemented to characterise the pharmacology of a single ligand binding to a single receptor, where the pharmacological profile of the  $A_{2A}R$  was successfully characterised when in a SMALP.

SMALPs were also tested for their applicability in phage display in order to generate GPCR-specific nanobodies against receptors in their native conformation. The M13 phage used in this project were conjugated with a VHH nanobody. The  $A_{2A}R$ -SMALP was immobilised onto ELISA plates for phage binding, where approximately 22% of the total phage were  $A_{2A}R$ -SMALP specific, which was lower than the control Fab protein. Avi-tagged  $A_{2A}R$  and CGRPR constructs were designed to improve the immobilisation of the SMALPs, to yield a higher enrichment of phage, specific to the GPCR of interest.

Finally, the photoaffinity cross-linking assay was implemented in this project, which has potential implications in drug discovery as receptors can be locked into a particular conformation when cross-linked with a ligand. The technique can theoretically be applied to receptors in SMALPs. Residues of the extracellular loops 1 and 3 of the CGRPR were studied. The assay showed residues A199, N200 and N201 of the extracellular loop 1 forming crosslinks with the ligand, when substituted with azidophenylalanine.

Overall, the project demonstrated techniques which are applicable to study SMA-solubilised receptors. Using the various techniques revealed novel properties of the GPCRs in SMALPs. SMALPs were also applied to the drug discovery technique, phage display, with limited success. Techniques were incorporated into this study to improve the applicability of SMALPs in phage display.

Keywords: GPCRs, adenosine 2A receptor, calcitonin gene-related peptide receptor, styrene maleic-acid lipid particles, nanobodies, phage display

<b>Table of Contents</b>	<b>Page</b>
Abbreviations .....	9
Figure contents .....	12
Table contents.....	15
Chapter 1: Introduction.....	16
1.1: Introduction to pharmacology .....	16
1.2: Overview of receptor pharmacology.....	17
1.2.1: Membrane proteins are key drug targets .....	17
1.2.2: The use of structural and computational biology to identify drug targets in receptors.....	18
1.2.3.1: The pharmacodynamics of ligand-receptor interactions .....	21
1.2.3.2: Mode of ligand binding and ligand classification .....	26
1.3: Introduction to G protein-coupled receptors .....	29
1.3.1: Overview of GPCRs.....	29
1.3.2: Family A GPCRs .....	32
1.3.3: Common mechanisms of family A GPCR activation .....	32
1.3.4: The adenosine 2a receptor and its pharmacology .....	34
1.3.5: Family B GPCRs.....	39
1.3.6: Common mechanisms of family B GPCR activation upon peptide binding.....	42
1.3.7: The calcitonin gene-related peptide receptor .....	45
1.3.8: Allosteric modulation of GPCRs .....	47
1.4: The use of the styrene maleic-acid to solubilize receptors.....	50
1.4.1: How detergents are used to solubilise membrane proteins .....	50
1.4.2: The use of styrene-maleic acid copolymers as an alternative to detergents in solubilizing membrane proteins .....	51
1.5: Characterisation of the SMA-solubilised A <sub>2A</sub> R and the CGRPR .....	54
1.6: Project aims.....	56
Chapter 2: Materials and methods.....	57
2.1: Table of materials used in experimental procedures (see table contents).....	57
2.2: Methods.....	66
2.2.1: Plasmid constructs.....	66
2.2.2: Diagnostic restriction digest using DNA electrophoresis .....	73
2.2.3: Competent <i>E.coli</i> generation and transformation.....	74
2.2.4: Expressing the A <sub>2A</sub> R and CD81 in <i>Pichia Pastoris</i> .....	76
2.2.4.1: Creating competent <i>Pichia pastoris</i> cells for transformation .....	77
2.2.4.2: <i>Pichia pastoris</i> cells transformation and growth .....	77
2.2.4.3: Extracting <i>Pichia pastoris</i> membranes .....	78
2.2.5: Mammalian cell culture and transfection .....	79

2.2.6: SMA copolymer preparation and membrane solubilisation.....	79
2.2.7: Protein purification, dialysis and concentration.....	80
2.2.8: Identifying recombinant proteins via SDS-PAGE .....	81
2.2.8.1: Western blotting .....	83
2.2.8.2: Instant Blue (Coomassie) staining of SDS-PAGE gel .....	84
2.2.8.3: Determining the concentration of a receptor-SMALP sample .....	84
2.2.9: Radioligand binding assay .....	85
2.2.9.1: Competitive radioligand binding assay .....	85
2.2.9.2: Calculating ZM241385-bound dghA <sub>2A</sub> R-SMALP recovery .....	86
2.2.10: cAMP detection assay .....	87
2.2.11: Fluorescence correlation spectroscopy .....	89
2.2.12: The biotinylation and immobilization of avi-tagged SMALPs .....	90
2.2.12.1: Biotinylation of receptor-SMALPs .....	92
2.2.12.2: Indirect ELISA to determine receptor-SMALP biotinylation and streptavidin-biotin interaction using pNPP .....	92
2.2.12.3: Biotin quantitation using HABA/avidin .....	93
2.2.13: Nanobody generation using phage display .....	93
2.2.13.1: Generating A <sub>2A</sub> R-specific phage using phage display .....	94
2.2.13.2: Titration of phage stocks .....	96
2.2.13.3: Detecting phage specificity to A <sub>2A</sub> R using phage ELISA .....	96
2.2.14: Photoaffinity cross-linking .....	98
2.2.14.1: Primer design and mutagenesis of the T8-HA-CLR construct using quick-change PCR .....	99
2.2.14.2: Photoaffinity cross-linking of F-CGRP to the mutant CGRPR .....	100
2.2.15: Statistical analysis .....	101
Chapter 3: Solubilisation of recombinant membrane proteins from <i>Pichia pastoris</i> using SMA .....	102
3.1: Introduction .....	102
3.2: Results .....	103
3.2.1: Growth of X33 yeast expressing A <sub>2A</sub> R .....	103
3.2.2: The use of SMA to solubilise A <sub>2A</sub> R receptors from the <i>P.pastoris</i> membranes.....	105
3.2.3: Recovery of binding-capable A <sub>2A</sub> R-A after solubilisation at different time periods .....	108
3.2.4: Overview of Ni <sup>2+</sup> -NTA purification of histidine-tagged proteins .....	110
3.2.4.1: The purification of SMA solubilized membrane proteins.....	111
3.3: Discussion .....	118
Chapter 4: Biophysical characterisation of GPCR-SMALPs .....	123
4.1: Introduction .....	123
4.2: Results .....	128

4.2.1: Assessing the pharmacology of the A <sub>2A</sub> R-SMALPs using radioligand binding .....	128
4.2.1.1: Using competitive radioligand binding to characterise the pharmacology of the A <sub>2A</sub> R-SMALPs .....	128
4.2.1.2: Comparing the pharmacological activity of the A <sub>2A</sub> R solubilized using either SMA or DIBMA copolymers .....	130
4.2.2: Using FCS to study A <sub>2A</sub> R-SMALP pharmacodynamics in real-time .....	133
4.2.3: The use of XRF to identify water-accessible residues of the SMA-solubilised A <sub>2A</sub> R ....	136
3.3: Discussion .....	138
Chapter 5: Using phage display against dghA <sub>2A</sub> R-SMALPs to generate A <sub>2A</sub> R-specific nanobodies	143
5.1: Introduction .....	143
5.2: Results .....	150
5.2.1: Assessing the A <sub>2A</sub> R specific binding of phage, conjugated with the positive control Fab2838 and ScFc2838 .....	150
5.2.2: Discovering A <sub>2A</sub> R-specific, VHH-conjugated phage using phage display .....	152
5.3: Discussion .....	155
Chapter 6: The biotinylation and immobilization of the avi-tagged A <sub>2A</sub> R-A .....	158
6.1: Introduction .....	158
6.2: Results .....	159
6.2.1: Assessing whether the biotinylated A <sub>2A</sub> R-A-SMALP can interact with streptavidin ....	159
6.2.2: The immobilisation of the biotinylated A <sub>2A</sub> R-A-SMALP .....	161
6.3: Discussion .....	164
Chapter 7: The expressing of CGRPR in HEK293T mammalian cells .....	166
7.1: Introduction .....	166
7.2: Results .....	167
7.2.1: Expression and solubilisation of CGRPR-A from HEK293T cell membranes .....	167
7.2.2: cAMP production upon CGRP binding to CGRPR and CGRPR-A .....	169
7.3: Discussion .....	170
Chapter 8: Analysing ligand-receptor binding interactions using photoaffinity cross-linking .....	172
8.1: Introduction .....	172
8.2: Results .....	175
8.2.1: Generation of the human CLR mutants with the amber stop codon .....	175
8.2.2: Using photoaffinity cross-linking to decipher the ligand binding sites of ECLs 1 and 3 of the CGRPR .....	176
8.3: Discussion .....	182
Chapter 9: General discussion and future experimental considerations .....	191
References .....	197

Appendix .....214

## Abbreviations

A1	Adenosine 1 receptor
A <sub>2A</sub> R	Adenosine 2A receptor
A <sub>2A</sub> R-A	Adenosine 2A receptor-avitag
A <sub>2B</sub> R	Adenosine 2B receptor
A <sub>3</sub> R	Adenosine 3 receptor
Ado	Adenosine
ADP	Adenosine diphosphate
AM	Adrenomedullin
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
b2AR	Beta 2 adrenoceptor
Bio	Biotinylated
Bmax	Maximum binding
BMGY	Buffered complex glycerol media
BMMY	Buffered complex methanol media
BRET	Bioluminescence resonance energy transfer
cAMP	Cyclic adenosine 3', 5' monophosphate
CD81	Cluster of differentiation 81
cDNA	Complimentary deoxyribonucleic acid
CGRP	Calcitonin gene-related peptide
CGRPR	Calcitonin gene-related peptide receptor
CGRPR-A	Calcitonin gene-related peptide receptor-avitag
CLR	Calcitonin-like receptor
CLR-A	Calcitonin
cpm	Counts per minute
CREB	cAMP response element-binding protein
CRF	Corticotrophin-releasing factor
CRFR	Corticotropin-releasing hormone receptor
Cryo-EM	Cryo-electron microscopy
CT	Calcitonin
CTR	Calcitonin receptor
DAG	Diacylglycerol
DDM	Dodecyl-β-D-maltopyranoside
DIBMA	diisobutylene-maleic acid
DIBMALP	diisobutylene-maleic acid lipid particle

DMSO	Dimethyl sulfoxide
dNTP	Deoxyribonucleotide triphosphate
EC50	Half maximal effective concentration of agonist
ECD	Extracellular domain
ECL	Extracellular loop
ELISA	Enzyme-linked immunosorbent assay
FCS	Fluorescence correlation spectroscopy
FRET	Fluorescence resonance energy transfer
GDP	Guanosine diphosphate
GLP1R	Glucagon-like peptide 1 receptor
GLR	Glucagon receptor
GPCR	Guanine nucleotide- binding protein coupled receptor
G protein	Guanine nucleotide- binding protein
G <sub>s</sub>	G-protein stimulatory
GTP	Guanosine triphosphate
HA	Haemagglutinin
hr	Hours
HRP	Horse radish peroxidase
IC50	Half maximum inhibition concentration
ICL	Intracellular loop
Ino	Inosine
IP <sub>3</sub>	Inositol triphosphate
L	Litre
LB	Lennox broth
M	Molar
MBP	Myelin basic protein
mg	Milligram
min	Minute
ml	Millilitre
mM	Millimolar
NF-κB	Nuclear factor kappa B
Ni <sup>2+</sup> -NTA	Nickel nitrilotriacetic acid
nM	Nanomolar
NMR	Nuclear magnetic resonance
OD	Optical density
PACAP	Pituitary adenylyl cyclase-activating protein

PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	polymerase chain reaction
PDB	Protein data bank
pH	Negative logarithm of the hydrogen ion concentration
Pi	Inorganic phosphate
PI3K	Phosphoinositide 3-kinase
PIP <sub>2</sub>	Phosphatidylinositol 4,5-bisphosphate
PKA/B/C	Protein kinase A/B/C
pKi	Negative logarithm inhibitory constant
PLC $\beta$	Phospholipase C beta
PLD	Phospholipase D
PTH	Parathyroid hormone
PTHR	Parathyroid hormone receptor
RAMP	Receptor activity modifying protein
RGS	Regulator of G protein signalling
Rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
SEM	Standard error of the mean
SMA	Styrene-maleic acid
SMALP	Styrene-maleic acid lipid particle
TEMED	N,N,N',N'-Tetramethylethylenediamine
TM	Transmembrane
TRIS	Tris(hydroxymethyl)aminoethane
Tween 20	Polyoxyethylene sorbitan monolaurate
WT	Wild type
w/v	Weight/volume
XRF	X-ray radiolytic footprinting
YNB	Yeast nitrogen base extract
YPD	Yeast/peptone/dextrose
YPDS	Yeast/peptone/dextrose/sorbitol
ZM241385	4-(2-[7-amino-2-(2-furyl)-1,2,4-triazolo[1,5-a][1,3,5]triazin-5-ylamino]ethyl)phenol

<b>Figure contents</b>	<b>Page</b>
Figure 1: Forward and reverse ligand binding reaction equation .....	22
Figure 2: Calculating the $K_d$ .....	22
Figure 3: Drug-receptor binding saturation model.....	23
Figure 4: The receptor occupancy equation .....	24
Figure 5: Sigmoidal curve of a binding assay .....	25
Figure 6: Simplified models of ligand binding .....	26
Figure 7: GPCR response upon ligand binding .....	28
Figure 8: Signalling pathways of G proteins.....	30
Figure 9: Conformational changes upon family A GPCR activation.....	33
Figure 10: Crystal structure of the adenosine-bound $A_{2A}R$ .....	34
Figure 11: Adenosine receptor signalling .....	35
Figure 12: Distribution of the $A_{2A}R$ .....	36
Figure 13: Structures of NECA and ZM241395 .....	37
Figure 14: Amino acid sequences of the three human RAMP isoforms .....	40
Figure 15: Superimposed structures of RAMP1 and RAMP2.....	41
Figure 16: Ser286 of the ECL2 of the CGRPR forms a hydrogen bond with His10 of the CGRP.....	43
Figure 17: V-shaped chalice in family B GPCRs .....	44
Figure 18: Crystal structure of the CGRPR .....	47
Figure 19: Sodium ion in the $A_{2A}R$ crystal structure .....	48
Figure 20: Conversion reaction of styrene-maleic anhydride to SMA .....	52
Figure 21: Illustration of SMALPs and detergent-solubilised membrane proteins.....	52
Figure 22: VHH-conjugated M13 bacteriophage.....	55
Figure 23: cDNA sequences of the $A_{2A}R$ constructs .....	67
Figure 24: Plasmid map and illustration of the pPICZ $\alpha$ A and the $A_{2A}R$ constructs .....	68
Figure 25: cDNA sequences of the CLR constructs .....	70
Figure 26: cDNA sequence of the RAMP1 construct.....	71
Figure 27: Plasmid map of the pcDNA3.1+ vector and the CLR and RAMP1 constructs.....	72
Figure 28: Equation to calculate recovery of binding-capable dgh $A_{2A}R$ -SMALPs .....	86
Figure 29: LANCE cAMP assay illustration .....	88
Figure 30: Principle of the HABA/avidin biotin quantitation assay .....	91
Figure 31: Expression of the $A_{2A}R$ constructs in X33 cells .....	104

Figure 32: A <sub>2A</sub> R solubilisation efficiency of SMA copolymers .....	107
Figure 33: Recovery of active A <sub>2A</sub> R-SMALPs after solubilisation .....	108
Figure 34: Illustration of his-tag protein purification using nickel affinity resin .....	110
Figure 35: dghA <sub>2A</sub> R-SMALP purification .....	111
Figure 36: A <sub>2A</sub> R-A-SMALP and CD81-SMALP purification .....	115
Figure 37: FCS confocal detection volume .....	125
Figure 38: X-ray radiolytic footprinting schematic.....	126
Figure 39: Binding profile of the A <sub>2A</sub> R constructs .....	129
Figure 40: Comparison of the DIBMA and SMA copolymer structure .....	131
Figure 41: Binding profile comparing the stability of the DIBMA and SMA-solubilised A <sub>2A</sub> R .....	132
Figure 42: FCS data showing total and non-specific binding of CA200645 to A <sub>2A</sub> R-SMALPs .....	134
Figure 43: Comparing the binding of CA241385 between the A <sub>2A</sub> R SMALP constructs .....	135
Figure 44: Water-exposed regions of the XRF derived structure of the A <sub>2A</sub> R .....	137
Figure 45: ZM241385 interaction with water networks and the A <sub>2A</sub> R .....	142
Figure 46: Fab2838 binding to the A <sub>2A</sub> R and its allosteric effects on receptor pharmacology .....	144
Figure 47: Illustration of the IgG, HCABs and VHH antibodies/nanobody .....	146
Figure 48: Crystal structure of the VHH-5 nanobody .....	147
Figure 49: Position of the VHH gene in <i>Camelidae</i> .....	148
Figure 50: Components of an IgG antibody Fab and ScFv .....	150
Figure 51: Fab2838 and ScFv2838 bound specifically to the SMA-solubilised A <sub>2A</sub> R .....	151
Figure 52: Illustration of the phage display panning strategy .....	153
Figure 53: Phage ELISA plate showing A <sub>2A</sub> R-specific monoclonal phage .....	154
Figure 54: Potential issues of directly immobilising SMALPs.....	156
Figure 55: Avi-tag biotinylation of the A <sub>2A</sub> R-A-SMALP .....	159
Figure 56: Biotin-streptavidin immobilisation theory.....	161
Figure 57: Immobilisation of A <sub>2A</sub> R-A-SMALPs on streptavidin-coated ELISA plates .....	162
Figure 58: Expression of the CGRPR-A in transfected HEK293T cells .....	167
Figure 59: Dose response curves of the CGRPR and the CGRPR-A, using the cAMP accumulation assay .....	169
Figure 60: Incorporation of the amber stop codon into a DNA construct.....	172
Figure 61: Chemical structure of AzF.....	173
Figure 62: Interactions between ECL2 residues of the CLR and CGRP .....	174

Figure 63: hCLR mutants were present in the pCDNA3.1+ vector .....	175
Figure 64: Pharmacodynamic properties of the photoaffinity cross-linking assay .....	177
Figure 65: A199U, N200U, N201U hCGRPR mutants cross-link with F-CGRP .....	179
Figure 66: No cross-linking between ECL3 hCGRPR mutants and F-CGRP .....	181
Figure 67: Crystal structure of the CGRPR:CGRP complex, illustrating potential interactions between the CGRP and the CLR .....	185
Figure 68: ECL3 does not form molecular contacts with CGRP .....	188

<b>Table contents</b>	<b>Page</b>
Table 1: Adenosine receptor ligands .....	38
Table 2: Polymerase chain reaction reagents .....	57
Table 3: Enzymes used for molecular experiments .....	57
Table 4: Reagents for DNA electrophoresis .....	57
Table 5: <i>E.coli</i> culture and transformation solutions .....	58
Table 6: <i>Pichia pastoris</i> culture and electroporation solutions .....	58
Table 7: <i>Pichia pastoris</i> growth and expression solutions .....	59
Table 8: Mammalian cell culture and growth solutions and reagents .....	60
Table 9: Buffers for SMA solubilisation.....	60
Table 10: Buffers and reagents for his-tag protein purification .....	60
Table 11: Buffers and reagents for his-tag protein purification for XRF experiments .....	60
Table 12: Radioligand binding buffers and reagents .....	61
Table 13: SDS-PAGE buffers and reagents .....	61
Table 14: Western blotting buffers .....	62
Table 15: Coomassie staining reagents .....	62
Table 16: Buffers and reagents used for avi-tag protein biotinylation.....	62
Table 17: Reagents for biotin quantitation.....	62
Table 18: ELISA buffers and reagents.....	63
Table 19: Phage display and phage ELISA buffers, reagents and media.....	63
Table 20: Buffers and reagents used for photoaffinity cross-linking.....	64
Table 21: List of antibodies and phage for recombinant protein detection.....	64
Table 22: List of ligands and their properties .....	65
Table 23: SMA copolymers used to solubilise the A <sub>2A</sub> R .....	106

## **Chapter 1: Introduction**

### **1.1: Introduction to pharmacology**

Pharmacology is the study of how drugs affect living systems and how they interact within living systems (Currie, 2018). The drugs can range from pharmaceutically approved drugs, natural and synthetic compounds and exogenous or endogenous compounds, however they all achieve a similar goal of producing a response within a living organism (Currie, 2018). Pharmacology is split into disciplines, such as pharmacodynamics and pharmacokinetics. Recently, pharmacogenomics has been developed as an additional discipline, however all of them are essential in pharmacology (Currie, 2018)

Pharmacology has been practised by scholars since 150 AD, where theories and experiments were utilised to discover drug properties (Vallance and Smart, 2006). Over time, the information gathered from scientists has led to the development of synthetic drugs, using natural compounds to assist with drug design (Pors, 2011). Drug discovery techniques have eventually improved and expanded, leading to a higher rate of drug and receptor discovery, however current drugs in pharmaceuticals and in clinical trials target approximately 3% of human proteins (Ryaboshapkina and Hammar, 2019). Challenges in drug discovery can hinder the process of the production of new drugs, where only 10-14% of proteins encoded by the human genome are deemed 'druggable' (Pors, 2011). Proteins which are difficult to study are often left aside by pharmaceutical companies.

## **1.2: Overview of receptor pharmacology**

### **1.2.1: Membrane proteins are key drug targets**

Existing in living organisms are proteins, which are macromolecules involved in biochemical processes (Chandel *et al.*, 2018). They also have roles to structurally support a cell, such as cytoskeletons, as well as providing mechanical processes to complex organisms, such as muscle tissue (Chandel *et al.*, 2018). In pharmacology, molecules involved in signalling are of key interest, such as receptors and their ligands.

Membrane proteins are proteins which interact with a phospholipid bilayer. Membrane proteins expressed on the cell surface are the main drug targets for pharmaceutical industries, therefore understanding their function and structure is essential for drug development (Yin and Flynn, 2016). Around 30% of the human genome encodes transmembrane proteins and approximately half of the pharmaceutical agents available target membrane proteins to prevent onset of diseases (Dobson *et al.*, 2015). Membrane proteins differ from each other based on their structure and function and include: enzymes, receptors, ion channels and transport proteins (Cournia *et al.*, 2015).

Receptors are commonly transmembrane proteins, which are essential for cells to respond to and adapt to their surrounding environment (Uings and Farrow, 2000). They function by responding to an external stimuli, which docks into a binding site present on the receptor (Currie, 2018). This in turn leads to receptor activation, which then activates intracellular signalling cascades to achieve a response (Uings and Farrow, 2000). Some substances can inactivate a receptor upon binding, reducing intracellular activity associated with the receptor (Currie, 2018). Exogenous chemicals, polypeptides, neurotransmitters, hormones and sugars are potential stimuli which can activate their respective receptor (Currie, 2018, Uings and Farrow, 2000). There is a huge diversity of receptors encoded by the human genome, categorised by their types. There are cytokine receptors, receptor tyrosine kinases, ligand gated ion channels and G protein coupled receptors (GPCRs), where each type is further divided into classes (Uings and Farrow, 2000).

### **1.2.2: The use of structural and computational biology to identify drug targets in receptors.**

Synthetic chemistry, organic chemistry, pharmacogenomics, proteomics and biology have provided an understanding of synthetic drug design against receptors, categorising receptors based on gene and protein sequences and studying their function in living systems. Structural and computational biology have provided knowledge of the molecular mechanisms of protein activity at an atomic resolution, which is essential for designing synthetic compounds against crevices and binding sites of a receptor (Congreve *et al.*, 2005). Structural biology involves the use of techniques such as x-ray crystallography to generate static structures of proteins or the capture of images of a protein in different orientations. In silico methods can be used to generate a hypothetical scenario of receptor conformational changes, based on existing structural data.

For around 50 years ago, structural biology has been used to discover drug targets (Congreve *et al.*, 2005). 3D structures have initially been identified for soluble proteins such as enzymes and globins, whereupon ligands have been developed against haemoglobin to relieve sickling in sickle cell patients (Beddell *et al.*, 1976). There are now over thousands of structures on the protein data bank (PDB) due to the development of structural techniques (PDB [online] last accessed: 20.12.2019). Pharmacological data can also be rationalised by structural data, which is important in pharmacology to understand the molecular connections a receptor makes with a ligand and second messenger-generating proteins (Congreve *et al.*, 2005). Comparison of different receptor structures allows the organisation of receptors into families or superfamilies, which in turn allows the categorisation of drugs (Congreve *et al.*, 2005). Sequence homology can be combined with structural homology of a family of receptors to identify key residues associated with their function (Congreve *et al.*, 2005). This also simplifies the identification of new proteins, which may have homologous sequences and structure to a family of proteins. There are many structural techniques to solve the structure of a protein, however the most commonly used techniques are: X-ray crystallography, cryo-electron microscopy and nuclear magnetic resonance (Chandel *et al.*, 2018; Moraes *et al.*, 2014; Cheng, 2018).

X-ray crystallography, in biology, is used to study proteins and has generated 3-dimensional (3D) structures (Congreve *et al.*, 2005; Cheng, 2018; Moraes *et al.*, 2014). It is currently the most used method

due to its consistency in providing high resolution crystallisation data and over 112,000 protein structures in the protein data bank (PDB) have been solved in this technique (Carpenter *et al.*, 2008; Loll, 2003; Shoemaker and Ando, 2018). Crystal structures can be obtained using purified protein from natural sources, such as the squid rhodopsin from *Todarodes pacificus* (Shimamura *et al.*, 2008). Recombinant proteins expressed in cell expression systems can be solubilised and purified, prior to crystallisation (Congreve *et al.*, 2005). Once a successful crystal is obtained, the structure can be determined (Congreve *et al.*, 2005). There are issues associated with obtaining a structure of a protein in one conformation, such as GPCRs which exist in multiple conformations at its basal state (Gurevich and Gurevich, 2017). The heterogeneity in conformation of certain proteins can lead to differences in different crystal structures of a protein in one sample.

Cryo-electron microscopy (cryo-EM) is a method which has recently gained popularity as an alternative structural method to study proteins, where the technique has improved over time, with current resolutions reaching approximately 2 - 3 Å (Cheng, 2018; Kumar *et al.*, 2020; Zhao *et al.*, 2020; Liang *et al.*, 2018). Cryo-EM can also provide structural information on protein complexes (Chandel *et al.*, 2018). This technique differs from crystallography due to the protein sample being incorporated into a fine layer of glass-like ice in random orientations, instead of being crystallised (Cheng, 2015). Electron microscopy is subsequently used to take images of the protein orientations, which are then averaged and analysed. The images can then be converged to generate a 3D protein model (Cheng, 2015). This gives cryo-EM an advantage over X-ray crystallography as proteins can be analysed in many orientations (Cheng, 2018). Cryo-EM also requires less protein than crystallography. Furthermore, the improvement of algorithms, beam-induced image motions and the use of a direct electron detection camera have providing resolution of protein images which rival x-ray crystallography (Cheng, 2018; Wang and Wang, 2016). Cryo-EM has had success in the generation of structures of large transmembrane proteins and complexes, which were difficult to achieve with x-ray crystallography (Cheng, 2018). Crystal structures dominated cryo-EM prior to 2014 where the first membrane protein structure was obtained in 2013 using cryo-EM (Cao *et al.*, 2013; Cheng, 2018). 5% of total protein database (PDB) structures were based on cryo-EM in 2014, however this increased to around 35% in 2018 (Cheng, 2018). This

demonstrates the rapid growth of cryo-EM as a new structural technique to identify protein structures. A disadvantage of cryo-EM however is the difficulty of capturing images of small proteins. GPCRs were often considered too small to analyse with cryo-EM, however scientists have coupled GPCRs with a G protein or stabilising antibodies in order to provide enough soluble mass for cryo-EM (Zhang *et al.*, 2017; Gao *et al.*, 2019; Tsai *et al.*, 2019). Therefore proteins which are too small for cryo-EM can be complexed with another complementary protein, such as an antibody fragment or an accessory protein, to provide enough soluble mass for cryo-EM (Cheng, 2018).

Structural methods have demonstrated the importance of analysing protein structures, and this comes hand-in-hand with *in silico* modelling programmes. For drug-receptor interactions, using *in silico* programmes can identify structural grooves a ligand can occupy and how these change the conformation of the receptor (Geux, 1999; Congreve *et al.*, 2005).

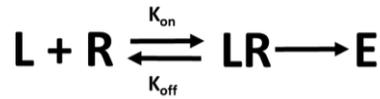
There are challenges when obtaining structures of membrane proteins, in which many issues arise from the solubilisation and preparation of samples as well as the protein itself. Cell expression systems are essential for generating protein structures, where cell systems which expressing proteins in their native conformation at a high yield is considered ideal (Cheng, 2018). Membrane proteins are also insoluble, therefore a detergent is often used to solubilise proteins into micelles. Detergents are particularly important in crystallography where once the protein is solubilised and purified, crystals can be generated using the soluble protein (Moraes *et al.*, 2014). An issue with detergents however is that they can remove some or all of the surrounding lipid bilayer, potentially distorting the proteins conformation (Moraes *et al.*, 2014; Jamshad *et al.*, 2015; Gulamhussein *et al.*, 2020). As detergents have been used frequently in structural research, the protein structures available may have alternations to their natural conformation, affecting the development of synthetic drugs. Methods to incorporate membrane proteins into lipid nanodiscs have been developed to rectify this problem, where styrene-maleic acid has become the latest trend in protein research to solubilise proteins in their native lipid environment (Jamshad *et al.*, 2015). However there are still issues in structural biology with nanodiscs, where capturing atomic structures of protein-nanodiscs provides low resolution data with high noise (Cheng, 2018). The membrane protein itself may be structurally challenging. Large transmembrane proteins often crystallise poorly as the

solubilised material must be folded in its native state and disordered, flexible structure can make the protein difficult to generate a homogenous crystal structure (Carpenter *et al.*, 2008). Thermostabilising mutations provide the protein stability in detergent micelles. The mutations also push a membrane receptor to certain conformational states (Carpenter *et al.*, 2008; Serrano-Vega *et al.*, 2008). A T4 lysozyme has also been used to replace an intracellular loop of GPCRs to stabilise its structure for crystallisation (Carpenter *et al.*, 2008; White *et al.*, 2018; Thorsen *et al.*, 2014). The T4 lysozyme has been used as it is well-folded and soluble and can crystallise under different conditions (Thorsen *et al.*, 2014). Nanobodies have also been used to stabilise GPCRs for crystallography (Warne *et al.*, 2018). Although there are challenges in structural studies, the techniques to resolve protein structures are improving over time.

### **1.2.3.1: The pharmacodynamics of ligand-receptor interactions**

Pharmacodynamics is a major discipline in pharmacology, studying the mechanisms of a drug binding to a receptor. Binding kinetics and drug affinity are important factors to consider when designing a drug and many methods have been developed to identify the potency and affinity of drug-receptor interactions (Chandel *et al.*, 2018). Ligand-receptor interactions are governed by the law of mass action and chemical events such as the formation of non-covalent and covalent bonds (Chandel *et al.*, 2018). Many techniques have been developed to study ligand-protein interactions, such as: structural techniques, radioligand binding assays, UV/Visible absorption and fluorescence spectroscopy and surface plasmon resonance (Chandel *et al.*, 2018).

A receptor and its ligand can form a complex, which can also dissociate back into receptor and ligand. The equilibrium between ligand, receptor and ligand-receptor complex depend upon the reactions which drive complex formation and its dissociation. This equilibrium can be altered by adjusting the concentration of ligand and receptor within an assay (Chandel *et al.*, 2018). The ratio of forward and reverse reactions (or on and off reactions) in **figure 1** determine the affinity of a ligand to its receptor, where at an equilibrium, the reactions occur at the same rate. A ligand with a high affinity for a receptor will push the equilibrium to where the [LR] state predominates.



**Figure 1** | A chemical equilibrium equation showing the forward and reverse ligand binding interactions, based on the law of mass action. The forward and backward binding reactions. (L) is a ligand and (R) is a receptor.  $K_{\text{on}}$  is the rate constant for the forward reaction, where a ligand and receptor form a complex and  $K_{\text{off}}$  is rate constant reverse reaction where the LR complex becomes uncoupled. E is the effect produced upon complex formation. (Upton and Mould, 2014; Chandel *et al.*, 2018).

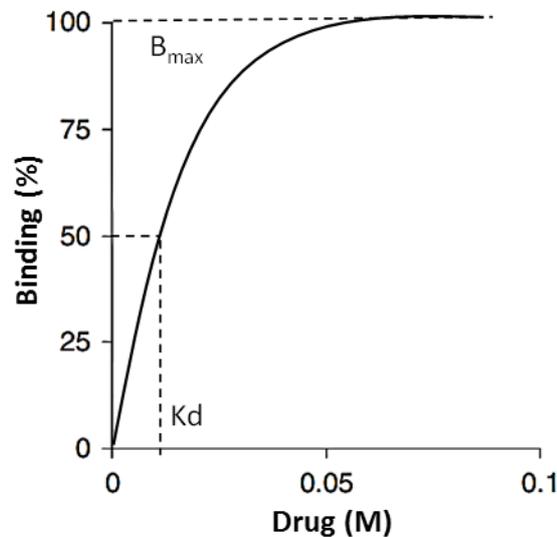
Drug affinity is determined by the dissociation constant (Kd) (Strange, 2008, Upton and Mould, 2014). The Kd was derived from the law of mass action and is now commonly used to compare the affinity of different ligands for the same receptor (Goutelle *et al.*, 2008). Kd is defined as the ratio of the two rate constants,  $k_{\text{off}}$  reaction and  $k_{\text{on}}$ , showing the concentration of a drug which occupies half the receptors (**figure 2**). The Kd can be calculated from the concentration of ligand [L] and receptor [R] divided by the concentration of ligand-receptor complex [LR], all at an equilibrium.

$$\frac{K_{\text{on}}}{K_{\text{off}}} = K_d = \frac{[L][R]}{[LR]}$$

**Figure 2** | The dissociation constant, Kd, is defined by the ratio of the forward and reverse reactions. L] is the concentration of ligand, [R] is the concentration of receptor and [LR] is the concentration of ligand-receptor complexes.  $K_{\text{on}}$  is the forward reaction and  $K_{\text{off}}$  is the reverse reaction (Upton and Mould, 2014; Chandel *et al.*, 2018).

In saturation binding models, when receptor occupancy is plotted against drug concentration, the resulting curve is a hyperbola (Uddin *et al.*, 2018; Hein and Michel, 2005). The graph (**figure 3**), also known as a Rosenthal plot, is used to identify the Bmax of a drug and the Kd, which is the drug concentration which occupies half the receptors (Hein *et al.*, 2005). Radioligand binding saturation

experiments are often used to identify the  $K_d$  (Hulme and Trevethick, 2010). The saturation experiments are done at equilibrium to observe the formation of ligand-receptor complexes whilst increasing the concentration of labelled ligand (Uddin *et al.*, 2018). The  $K_d$  can be calculated through kinetic experiments, which observe the binding of a concentration of ligand at increasing time points (Uddin *et al.*, 2018). This can establish the  $k_{on}$  and  $k_{off}$  rate constants, effectively providing a  $K_d$  value of a drug from their ratio.



**Figure 3** | Drug-receptor binding model demonstrating the relationship between an increasing drug concentration and an increasing drug-receptor complex. In the graph the ligand concentration is plotted on a linear scale on the x axis where a hyperbolic relationship between an increasing agonist concentration and amount bound,  $b$ , is seen. The  $B_{max}$  is the maximum receptor binding achieved by the drug. The  $K_d$  is the concentration of ligand to occupy half the receptors

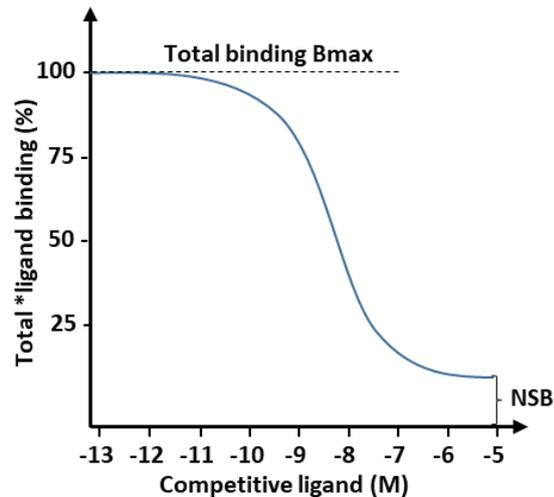
When plotting a saturation curve on a semi-logarithmic scale, the curve becomes sigmoidal. The occupancy equation can be used to describe drug-receptor occupancy, using the  $K_d$  (**figure 4**) (Kenakin, 2016).

$$[LR] = \frac{[L] \times B_{max}}{[L] + K_d}$$

**Figure 4** | [LR] represents the proportion of bound receptors. [L] represents the concentration of free ligand. B<sub>max</sub> is the maximum binding achieved by the ligand. K<sub>d</sub> is the dissociation equilibrium constant. Equation was obtained and modified from (Kenakin, 2016).

Competition binding experiments are useful to assess the affinity of a ligands and to observe their displacement by a competitive ligand. The experiments often involve a radioligand, used at a constant concentration, as well as a non-labelled ligand at increasing, logarithmic concentrations to outcompete the radioligand. The concentration of receptors used in competitive binding experiments are also kept constant (Hulme and Trevethick, 2010). The data obtained can be plotted on a semi-logarithmic scale, where the law of mass action equation can be applied to determine the relationship between receptor occupancy and drug concentration (Kenakin, 2016). **Figure 5** demonstrates how a competitive binding assay is analysed and plotted on a semi-logarithmic scale (Hulme and Trevethick, 2010). When experimenting with two different ligands in a binding curve and one is an inhibitor, the Cheng-Prusoff transformation ( $K_i = IC_{50} / (1 + [L] / K_d)$ ) can be used (Hein and Michel, 2005; Hulme and Trevethick, 2010). The IC<sub>50</sub> value of competitive inhibitors is calculated by finding the concentration of inhibitor which inhibits half the specific binding to the receptors (Burlingham and Widlanski, 2003). The IC<sub>50</sub> value takes into consideration the concentration of the labelled ligand and its K<sub>d</sub> (Uddin *et al.*, 2018), therefore conversion of IC<sub>50</sub> to K<sub>i</sub> via the Cheng-Prusoff equation can determine the affinity of the competitive inhibitor. (Uddin *et al.*, 2018). A successful displacement with the competitive ligand would yield a mass action sigmoidal curve as shown in **figure 5**. The extent of sigmoidicity is determined by the hill coefficient. A homogenous receptor sample would theoretically follow the Hill slope of unity, where a 1:1 relationship between a ligand and a receptor is present (Hein and Michel, 2005; Uddin *et al.*, 2018). If the Hill coefficient is higher than a value of 1, the data could show positive cooperativity, suggesting the sample containing multiple ligand binding sites, where binding of one ligand eases the binding of another ligand (Uddin *et al.*, 2018). A value less than 1 suggests negative cooperativity, where

a ligand may be interacting with more than one receptor type or binding site. In a majority of binding experiments, non-specific binding will be observed. The non-specific binding should be eliminated from the total binding data in order to observe specific binding of a ligand to its receptor (Uddin *et al.*, 2018). The  $\text{LogIC}_{50}$  can be calculated from a specific binding curve by identifying the concentration of ligand required to occupy half the receptors (half the  $B_{\text{max}}$ ).

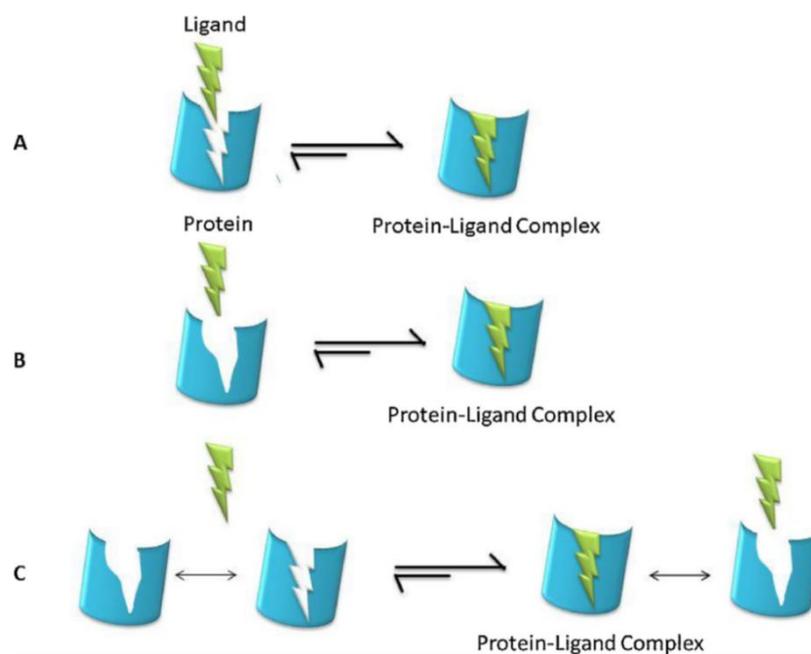


**Figure 5** | Hypothetical graph demonstrating a sigmoidal curve of a competitive binding assay, on a semi-logarithmic scale. \*Ligand is the labelled ligand where its concentration is kept constant for all assay points. The competitive ligand is used to compete against the \*Ligand for the receptor binding site at various concentrations. The total \*ligand binding observed at high concentrations of competitor ligand ( $10^{-6}$ - $10^{-5}$  M) show non-specific binding (Hulme and Trevethick, 2010)

### 1.2.3.2: Mode of ligand binding and ligand classification

Upon a ligand binding a receptor, a signal is generated by the receptor to produce the appropriate response to trigger signalling cascades. The ligand must make contact with the receptors binding site in order to achieve signalling. The understanding of ligand-receptor interactions have been developed over time. The initial model proposed by Emil Fischer in the year 1894 was the lock and key model, suggesting a protein-ligand complex being rigid with perfect complementarity (Chandel *et al.*, 2018). However, there is often an imperfect complementarity between a protein and ligand in structural research (Chandel *et al.*, 2018). The induced-fit model proposed that binding of a ligand to a receptor induces a conformational change to the receptor, as the binding site is flexible (Chandel *et al.*, 2018).

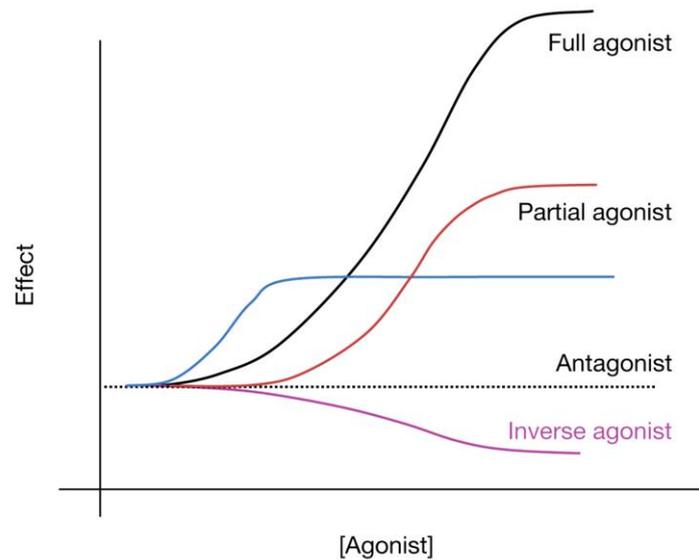
However both the lock and key model and the induced-fit model assumes a receptor undergoing a single, stable conformational change. Many receptors such as GPCRs can experience many conformational changes to reach an active conformation. The conformational selection model presumes that a receptor is in a multiple conformations at equilibrium, where the interaction between a ligand and a receptor pushes the equilibrium of the receptor to a complex formation state, in order for binding to occur (Chandel *et al.*, 2018). Therefore the receptor is not in a rigid and single state at equilibrium. **Figure 6** illustrates the three models which propose different modes of ligand binding (Chandel *et al.*, 2018).



**Figure 6|** An Illustration of the three simplified models of ligand binding. **A)** The lock and key model proposes the idea of perfect complementarity between a ligand and a protein receptor, allowing complex formation. **B)** The induced-fit model demonstrates a ligand altering the binding site of a protein in order to bind. **C)** The conformational selection model demonstrates a protein assuming different conformations at equilibrium. Interaction with a ligand promotes the protein to form an active conformation, allowing the ligand to complex with the protein. Illustration from Chandel *et al.*, (2018).

In the absence of ligands, receptors are in their “quiescent” state, although not all receptors are inactive. GPCRs in their quiescent state can differ from each other in terms of basal activity (Wootten *et al.*, 2018). At equilibrium, GPCRs can exist in different conformations, where the addition of ligands stabilise a conformation of a GPCR (Sato *et al.*, 2016) and so push the equilibrium towards a particular state. As GPCRs can undergo multiple conformations to reach an active or inactive state, there are many types of ligands to stabilise a particular configuration of the receptors. Therefore there is more to ligands than agonists and antagonists. Ligands can alter the response and conformation of a receptor and are classified as: agonists, antagonists, partial agonists and inverse agonists, depending on their effect of a protein (Currie, 2018). Agonists are ligands which activate a receptor. Endogenous ligands can activate a receptor and drugs which are made to activate receptors are often mimics of an endogenous ligand (Currie, 2018). They shift the equilibrium of a receptor to an active conformation, where a signalling response would be activated (Sato *et al.*, 2016). Partial agonists have the characteristics of both an agonist and an antagonist. They alter the conformation of a receptor and produce a response, however the efficacy of partial agonists are not as high as full agonists, therefore a maximum response would not be observed (Currie, 2018. Sato *et al.*, 2016). Reversible, neutral antagonists compete with an agonist for the binding site of a receptor. Once bound, the antagonist prevents agonist binding, however they do not alter the activity of the receptor (Sato *et al.*, 2016. Currie, 2018). The equilibrium of a receptor would not be changed by neutral antagonists, therefore basal activity of the receptor would still occur (Sato *et al.*, 2016). A higher concentration of an agonist can outcompete a reversible antagonist from the receptor binding site. Irreversible antagonists however do not dissociate from the receptor due to their high affinity (Currie, 2018). Inverse agonists have a negative efficacy, where they push the equilibrium of a receptor to an inactive conformation, inhibiting the activity of a receptor (Strange, 2008. Currie, 2018). In a clinical setting, inverse agonists would be used to inactivate the basal signal of a receptor. Mutant receptors with aberrant signalling may lead to disease states, such as hyperthyroidism caused by a mutation of the thyrotropin receptor, leading to constitutive activity (Sato *et al.*, 2016). Inverse agonists are useful to treat constitutively active receptors by lowering their activity; which neutral antagonists would not be able to do (Sato *et al.*, 2016). **Figure 7** demonstrates a simulated signalling response with the addition of different ligand types (Sato *et al.*, 2016).

In the field of GPCRs, there is an emerging class of ligands which have an additional level of pharmacological selectivity (Violin *et al.*, 2014). Biased ligands are of interest to pharmaceutical companies. These ligands can now target specific pathways of a receptor instead of fully activating them, which has benefits such as the limitation of adverse effects that may occur with certain drug types (Violin *et al.*, 2014).

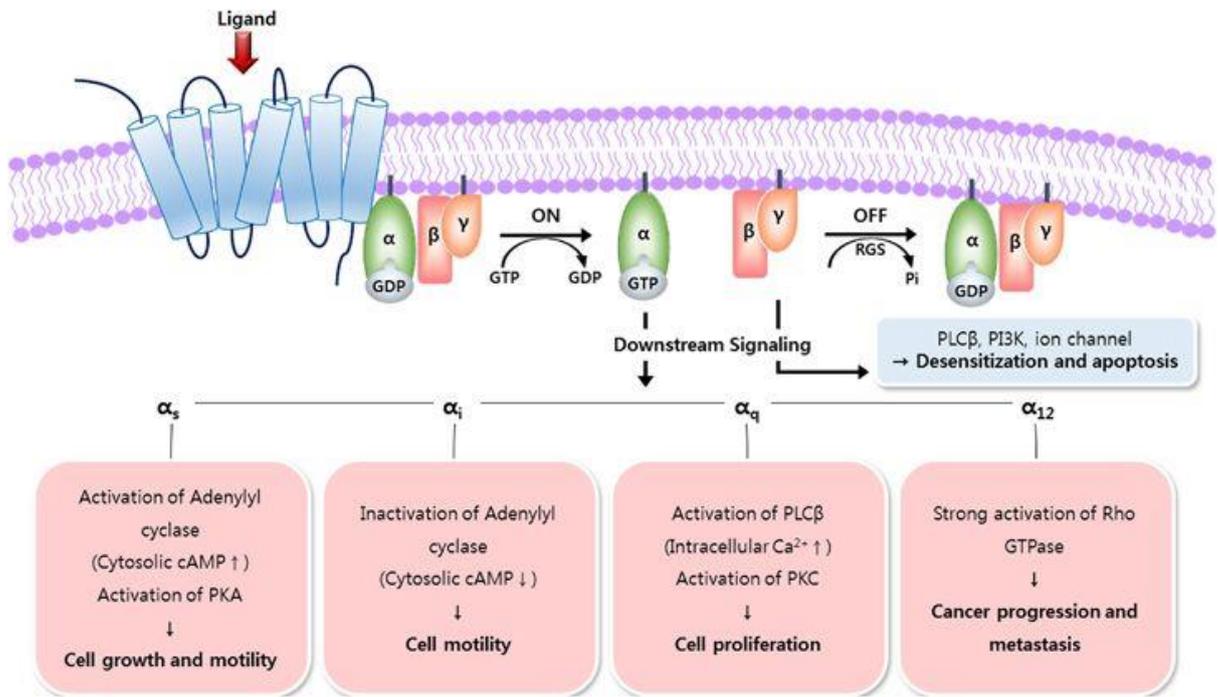


**Figure 7** | Graph demonstrating how GPCR ligands affect GPCR response. The blue line demonstrates basal activity of a receptor without the presence of ligands. A full agonist shifts the receptor to an active conformation, producing a positive response. A partial agonist also produces a positive response upon binding, however the efficacy of GPCR activity is not as high as what a full agonist could achieve. Antagonists competitively inhibit a GPCR, preventing agonist binding, however the equilibrium of the receptor would not be altered. An inverse agonist opposes a full and partial agonist, producing a negative response upon binding. Graph obtained from (Sato *et al.*, 2016).

### **1.3: Introduction into G protein-coupled receptors**

#### **1.3.1: Overview of GPCRs**

GPCRs are the largest protein superfamily and are expressed on the cell surface of many different cell types. They are characterised by their seven  $\alpha$ -helical transmembrane domains, which are arranged in an anti-clockwise bundle (Hu *et al.*, 2017). Over 800 GPCRs have been identified in mammalian organisms, however there are orphan receptors which are yet to be assigned (Sato *et al.*, 2016; Ngo *et al.*, 2017; Wootten *et al.*, 2018). The diversity of GPCRs is augmented by transcription and translational mechanisms, such as: alternative splicing, post-translational modifications and protein-protein interactions (Wootten *et al.*, 2018). The functional purposes of GPCRs are also varied and numerous, where they are involved in physiological responses such as olfactory senses, secretion, neurotransmission, metabolism, cellular growth and differentiation, inflammatory and immunological responses (Hu *et al.*, 2017). The molecules GPCRs interact with vary in their properties and structure, where molecules can be photons, ions, odorants, neurotransmitters, hormones, peptides and small molecule ligands (Wootten *et al.*, 2018). Even though each GPCR has a distinct purpose, the functional response can differ for a single GPCR, which is determined by: the different types of ligands a GPCR interacts with, allosteric modulators that associate with the receptors as and the signalling molecules found in the tissue a GPCR is expressed in (Wootten *et al.*, 2018). Abnormalities and mutations of GPCRs are associated with human diseases such as: cancer, diabetes, cardiac and central nervous system dysfunction (Hu *et al.*, 2017). Due to the growing research in GPCR signalling and pharmacology and the diseases GPCRs are associated with, the receptor superfamily has become a huge target for pharmaceutical companies to manufacture drugs against. Approximately 30-40% of approved pharmaceutical drugs target GPCRs, however only 10% of total GPCRs are deemed 'druggable' (Hu *et al.*, 2017. Wootten *et al.*, 2018). Some GPCRs interact with peptide ligands, which can be over 30 amino acid residues in size, making synthetic drug development difficult (Wootten *et al.*, 2018). This has often led to drug screening failures and halts in drug manufacturing processes.



**Figure 8|** Diagram illustrating the downstream signalling pathways of different G proteins. Upon GPCR activation, the G protein becomes active. The G $\alpha$ - subunit dissociates from the G $\beta\gamma$ -subunits and GDP of the G $\alpha$ - subunit is displaced by GTP. Depending on the type of G $\alpha$ - subunit, different downstream processes are initiated (pink boxes). The G $\beta\gamma$ -subunits are also capable of signalling. Abbreviations: GTP, guanosine triphosphate; GDP, guanosine diphosphate; RGS, regulator of G protein signalling; cAMP, cyclic adenosine monophosphate; PKA, protein kinase A; PKC, protein kinase C; PLC $\beta$ , phospholipase C beta; PI3K, phosphoinositide 3-kinase; Pi, inorganic phosphate. Diagram obtained from (Jo and Jung, 2016).

Agonist binding to GPCRs leads to the activation of downstream signalling processes, such as: cyclic adenosine monophosphate (cAMP) accumulation, calcium (Ca<sup>2+</sup>) release and inositol 1,4,5-trisphosphate activity (Wootten *et al.*, 2018). GPCRs interact with G proteins, which in turn activate effectors to transduce signalling cascades. All GPCRs interact with intracellular, heterotrimeric GTPases known as guanine nucleotide-binding proteins (G proteins) (Jo and Jung, 2016), which are approximately 90 kDa in size as a heterotrimer (Milligan and Kostenis, 2006). The protein consists of an  $\alpha$ ,  $\beta$  and  $\gamma$  subunit. There are 16  $\alpha$  subunits, 5  $\beta$  subunits and 13-14  $\gamma$  subunits in humans, where each  $\alpha$ ,  $\beta$ ,  $\gamma$  subunit can potentially combine with each other to form a certain type of G protein (Wootten *et al.*, 2018. Milligan and Kostenis, 2006). Upon GPCR activation via agonist binding, the coupled G protein becomes active. Guanine triphosphate (GTP) replaces guanosine diphosphate (GDP) residing

within the GTP-binding pocket of the  $\alpha$  subunit, resulting in the  $\alpha$  subunit dissociating from the  $\beta\gamma$  subunits (Wootten *et al.*, 2018. Sato *et al.*, 2016). The  $\alpha$  subunit is capable of signalling independently from the  $\beta\gamma$  subunits as shown above in **figure 8** (Jo and Jung, 2016. Sato *et al.*, 2016. Wootten *et al.*, 2018). GPCRs also signal through the  $\beta$ -arrestin signalling pathway by interacting with  $\beta$ -arrestin. There are two forms of  $\beta$ -arrestin, termed  $\beta$ -arrestin 1 and  $\beta$ -arrestin 2, which can interact with GPCRs (Ma and Pei, 2007) They are around 45 kDa in size, and are typically associated with the ablation of GPCR signalling, as well as functioning as scaffolds for mitogen-activated protein (MAP) kinase activation (Ma and Pei, 2007; Wootten *et al.*, 2018).

There are three superfamilies of GPCRs: family A, family B and family C. Each family differ in the structure of the GPCRs and their conserved sequences as well as the type of ligands which interact with them, varying from small molecules to hormones to large peptides. Family A GPCRs are also termed the ‘rhodopsin-like’ family of GPCRs due to their sequence and structural homology to rhodopsin (Hu *et al.*, 2018). This family holds 80% of the total GPCRs discovered, being the largest family known with over 800 members (Fredriksson and Schioth, 2005; Hu *et al.*, 2018). Their ligands include hormones and neurotransmitters. They have an eighth helix and sometimes a C-terminal cysteine, which is palmitoylated (Hu *et al.*, 2018). Family B consists of approximately 70 receptors and is subdivided into secretin-like receptors, with 15 members, and adhesion receptors (Poyner and Hay, 2012; Hu *et al.*, 2018). The characteristic feature of this family is the large N-terminal extracellular domain, consisting of around 120 residues, stabilised by disulphide bonds (Hu *et al.*, 2018). Family C GPCRs include the metabotropic glutamate family of receptors, GABA receptors, calcium-sensing receptors and taste receptors. Their N-terminal domain can be around 600 residues in size, which can act as a ligand binding site (Hu *et al.*, 2018). Family C GPCRs are also obligate dimers (Wootten *et al.*, 2018).

As my project involves receptors from family A and family B GPCR, these two families will be the main focus of this thesis.

### **1.3.2.: Family A GPCRs**

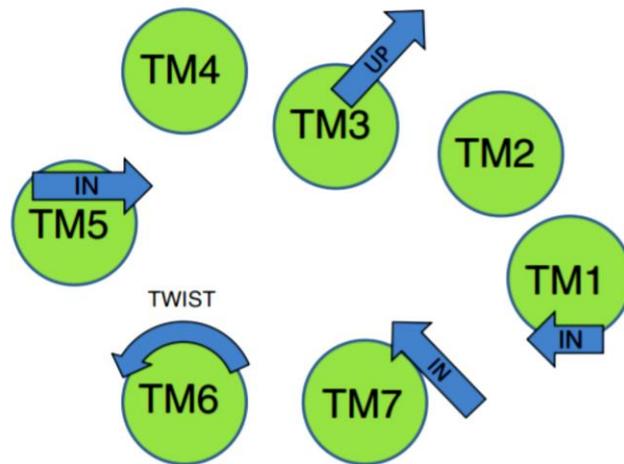
Family A GPCRs are also known as the rhodopsin family of GPCRs due to their homologous sequences to the rhodopsin receptor (the first family A GPCR structurally characterised). Due to the nature of sequences being conserved amongst family A GPCRs, the Ballesteros-Weinstein numbering can be used to label residues based on the transmembrane domain (TMD) on which they are located and how conserved the residue is amongst the family A GPCRs. It consists of two numbers; the first number is the helix the residue is located in and the second number represents the position of the residue, which is relative to the most conserved residue (Isberg *et al.*, 2016). This numbering system is useful to study GPCRs in relation to highly conserved residues of the family A receptors (Isberg *et al.*, 2016). An essential conserved sequence found in family A GPCRs is the E/DRY sequence located between intracellular loop (ICL) 2 and TMD3 and is important in G protein coupling (Rovati *et al.*, 2007; Hanlon and Andrew, 2015). The ionic lock is a feature found in a third of family A GPCRs where the E/DRY residue R3.50 and D/E6.30 form a salt bridge, locking the receptor in an inactive state (Schneider *et al.* 2010). Upon ligand binding, the GPCR changes conformation, breaking the ionic lock thus allowing the receptor to become active (Schneider *et al.* 2010. Tehan *et al.* 2014). Other conserved sequences include GN in TMD1, LXXXD in TMD2, CWXP in TMD5 and the NPXXY motif, which links TMD7 to helix 8 (Urizar *et al.* 2005). X represents variable amino acids.

### **1.3.3: Common mechanisms of family A GPCR activation**

Understanding the concepts of the conformational changes which occur upon GPCR activation is essential for the development of drugs. In this project, the A<sub>2A</sub>R was used as a model to test the suitability of SMALPs in discovering nanobodies which target extracellular epitopes of the receptor.

Upon agonist binding to family A GPCRs, structural changes occur within the transmembrane domains and the intracellular and extracellular loops. A feature important for GPCR activation is the movement of TMD3 and TMD6 upon agonist binding (Venkatakrisnan *et al.* 2013). When superimposing inactive and active conformations of GPCRs, common movements include the movement of TMD5, the upward movement of TMD3, TMD6 rotation and the inward movements of TMD1 and TMD7 (Tehan *et al.*

2014). This is summarised in **figure 9** (Tehan *et al.* 2014). Once the TMs have been rearranged, a binding crevice becomes available for the  $G\alpha$  subunit to bind to via its  $\alpha 5$  helix. This in turn leads to G protein activation (Wheatley *et al.*, 2012).

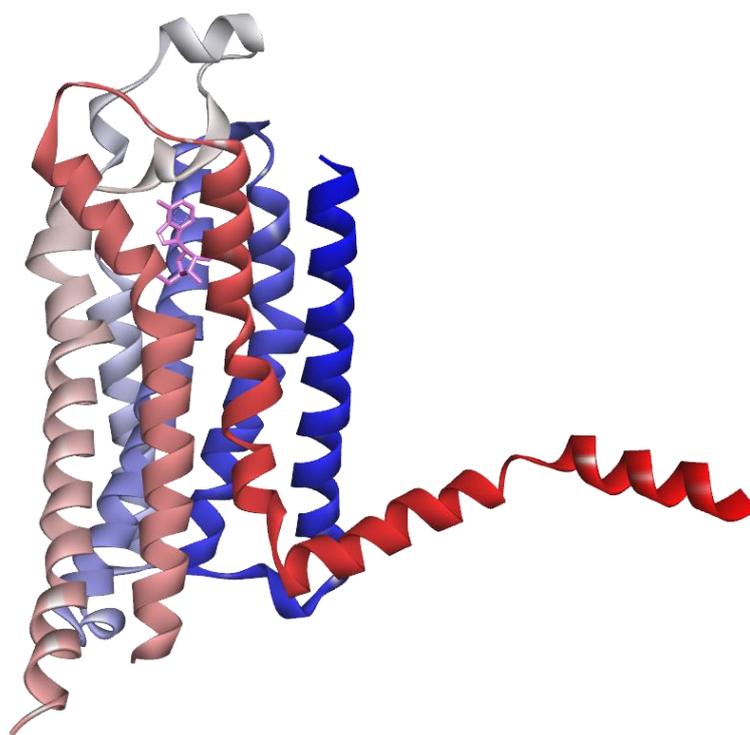


**Figure 9** | Diagram showing basic conformational changes leading to the activation of family A GPCRs. Illustration obtained from (Tehan *et al.* 2014).

Although the TMs are conserved in structure amongst family A GPCRs, the extracellular loops can vary in structure and size (Wheatley *et al.*, 2012). ECL1 and ECL3 are involved in ligand binding, orientation of the TMs and affects the functional response of a GPCR (Wheatley *et al.*, 2012). ECL2 is generally the largest and most diverse extracellular loop within this family of receptors, however there are exceptions such as the melanocortin receptor which has short ECL2 made of a few residues (Wheatley *et al.*, 2012). The ECLs are also involved in conformational rearrangements of a GPCR to transition from an inactive to active conformational state. ECL1 is formed between TM2 and TM3 whilst ECL3 tethers TM6 and TM7 (Wheatley *et al.*, 2012). ECL1 and ECL3 are responsible for ligand binding, orientation of the TMs and affects the functional response of a GPCR.

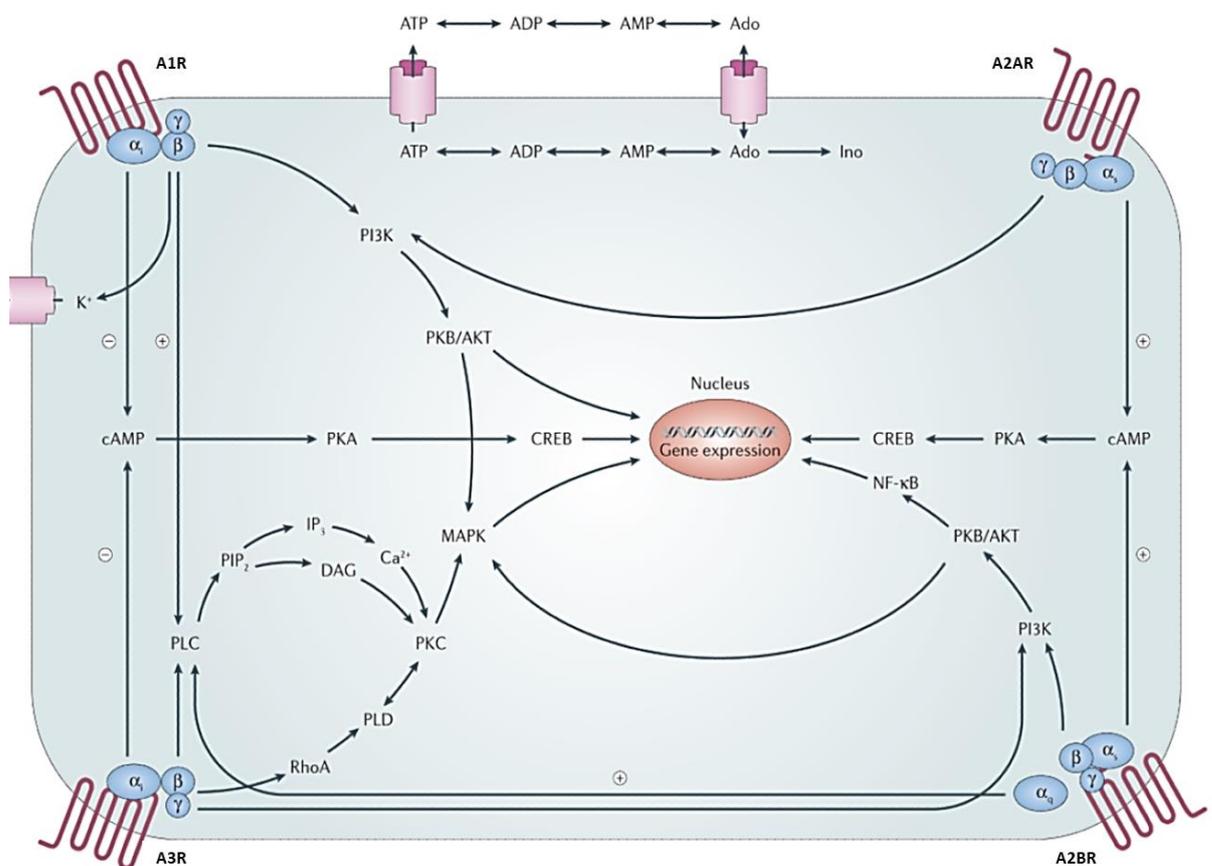
#### 1.3.4. The adenosine 2a receptor and its pharmacology

The adenosine 2a receptor ( $A_{2A}R$ ) is heavily studied amongst the family A GPCRs and has been used for solubilisation, purification and biophysical experiments within this project. Many ligands with different pharmacological effects have been developed against the  $A_{2A}R$ , allowing the characterisation of active, inactive and partially active conformations of the receptor. Furthermore,  $A_{2A}R$  crystal structures are available, with different ligands attached to the receptor as well as different conformations. The structure of the  $A_{2A}R$  complexed with the endogenous ligand, adenosine, is shown in **figure 10**. Within the field of SMALPs, the  $A_{2A}R$  has been well characterised in a SMALP (Jamshad *et al.* 2015). Therefore the  $A_{2A}R$  was a good starting point to characterise SMALPs and to generate phage display nanobodies



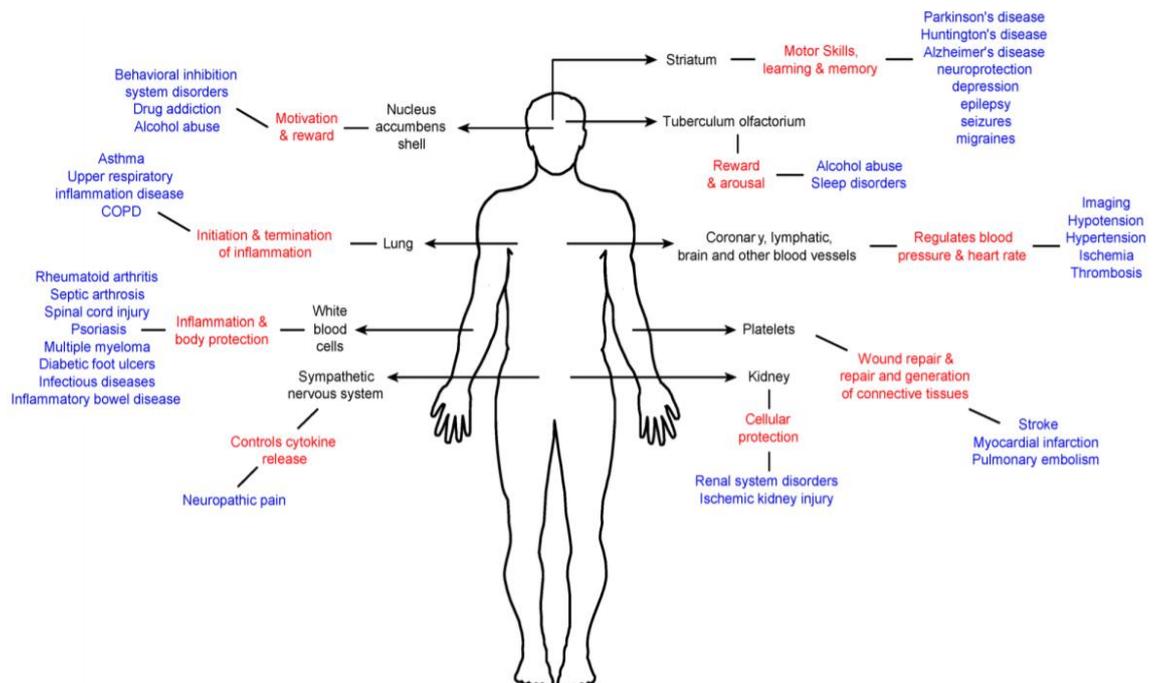
**Figure 10** | Crystal structure of the adenosine-bound  $A_{2A}R$  (PDB ID: 2YD0). The N to C terminus of the  $A_{2A}R$  is coloured from blue to red. The chemical structure of the adenosine agonist is coloured in pink. The structure was analysed using BIOVIA Discovery Studio.

Adenosine receptors are purinergic GPCRs where its natural ligand is adenosine. Adenosine is a potent ligand for adenosine receptors and is essential for cytoprotection, metabolism and cardiovascular regulation (Vecchio *et al.* 2017). Adenosine acts on various cell types including neurones, neutrophils, platelets and smooth muscle cells (Ruiz *et al.* 2013). Common functions of adenosine receptors are: to increase oxygen supply, to condition the cells against ischaemic damage, to initiate angiogenesis and to trigger anti-inflammatory responses (Jacobson and Gao. 2006). Four subtypes of the adenosine receptor have been identified, which are: A<sub>1</sub>, A<sub>2A</sub>, A<sub>2B</sub> and A<sub>3</sub> receptors. Each subtype differ in their pharmacology, tissue distribution and the G protein they couple to (**figure 11**) (Jacobson and Gao. 2006).



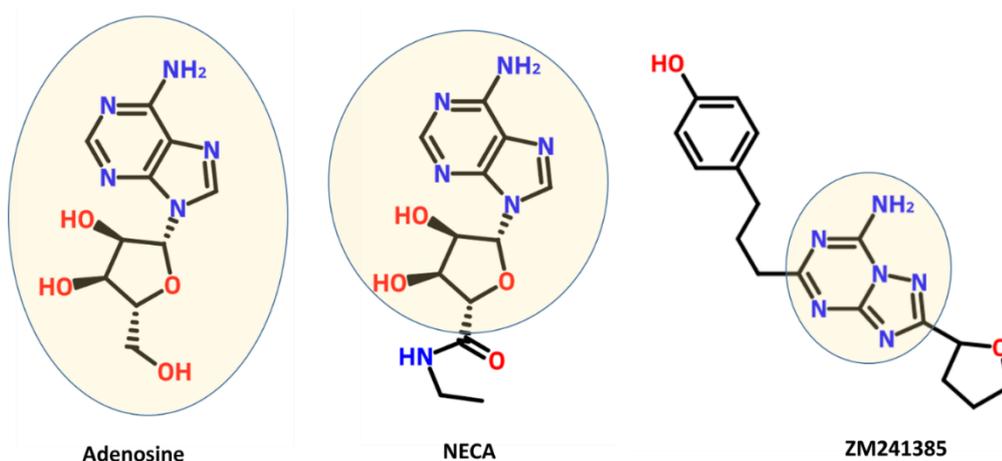
**Figure 11** | Diagram illustrating the four subtypes of adenosine receptors and their signalling pathways. Abbreviations: A<sub>1</sub>R/ A<sub>2A</sub>R/ A<sub>2B</sub>R/ A<sub>3</sub>R, adenosine receptor subtypes; ATP, adenosine triphosphate; ADP, adenosine diphosphate; AMP, adenosine monophosphate; Ado, adenosine; Ino, inosine; α/β/γ, G protein subunits; cAMP, cyclic adenosine monophosphate; PKA, protein kinase A; PKC, protein kinase C; PLCβ, phospholipase C beta; PI3K, phosphoinositide 3-kinase; PLD, phospholipase D; PKB/AKT, protein kinase B; CREB, cAMP response element-binding protein; NF-κB, nuclear factor kappa B; DAG, diacylglycerol; PIP<sub>2</sub>, phosphatidylinositol 4,5-bisphosphate; IP<sub>3</sub>, inositol triphosphate. Diagram taken from (Jacobson and Gao. 2006).

The A<sub>2A</sub>R is involved in many important processes, such as the vasodilation of coronary arteries in order to regulate myocardial blood flow. This in turn increases blood flow in the myocardium however abnormal regulation can lead to hypotension (Ruiz *et al.*, 2013). The regulation of adenosine production is essential for the cardiovascular system as it acts upon the receptor to regulate vasodilation. Within the brain, the secretion of glutamate and dopamine is regulated by the A<sub>2A</sub>R. The A<sub>2A</sub>R can also dimerize with the dopamine D2 receptors to regulate its affinity for dopamine (Ruiz *et al.*, 2013). As the A<sub>2A</sub>R is essential for the regulation of dopamine release and dopamine D2 receptor affinity, targeting the A<sub>2A</sub>R can impact the striatal dopaminergic neurotransmission. Parkinson's disease is associated with the dysfunction of dopaminergic neurotransmission and drugs against the A<sub>2A</sub>R have been studied to see its effect on dopamine D2 receptors. The ligand, SCH58261, has been shown to bind A<sub>2A</sub>R, which leads to the potentiation of dopamine D2 receptors (Ruiz *et al.*, 2013). The A<sub>2A</sub>R is distributed in a wide range of cell types and is associated with numerous bodily functions (**figure 12**) (Ruiz *et al.*, 2013).



**Figure 12** | Diagram showing the distribution of A<sub>2A</sub>R (black text), its function within certain tissues/ cells (red text) and diseases arising from A<sub>2A</sub>R abnormalities within certain tissues/ cells (blue text) (Ruiz *et al.*, 2013).

As the A<sub>2A</sub>R is a heavily studied amongst family A GPCRs, many ligand types have been developed against the receptor. The endogenous agonist is adenosine and synthetic ligands such as 5'-N-ethylcarboxamidoadenosine (NECA) have been made to activate the receptor (Lebon *et al.*, 2011). Structural data have shown that both adenosine and NECA bind to A<sub>2A</sub>R almost identically. 4-(2-[7-amino-2-(2-furyl)-1,2,4-triazolo[1,5-a][1,3,5]triazin-5-ylamino]ethyl)phenol (ZM241385) is a commonly used inverse agonist to study the inactive conformation of A<sub>2A</sub>R. Structural data identified some similarities with the binding of ZM241385 and the agonists. ZM241385 has a triazolotriazine ring structure, similar to the adenine-like ring of the agonists, which binds to the same site in the A<sub>2A</sub>R (Lebon *et al.*, 2011). A structural difference between ZM241385 and the agonists is the triazolotriazine ring of ZM241385 has a furan substituent, whereas the adenine-ring like structure of the agonists contain a ribose substituent (Lebon *et al.*, 2014). The difference in chemical structure between inverse agonist, synthetic agonist and endogenous agonist is shown in **figure 13** (Lebon *et al.*, 2011).



**Figure 13** Structures of NECA and ZM241385, compared with the endogenous agonist, adenosine. Gold circle highlights chemical similarities with adenosine. Structures were adapted from (Lebon *et al.*, 2011).

Many ligands have been created against A<sub>2A</sub>R to pharmacologically characterise the receptor, which are listed in **table 1**.

Ligand	Type	Specificity (receptor)	Affinity (pKi)
Adenosine	Agonist	A <sub>1</sub> , A <sub>2A</sub> , A <sub>2B</sub> , A <sub>3</sub>	7.0, 6.5, 4.8, 6.5
NECA	Agonist	A <sub>1</sub> , A <sub>2A</sub> , A <sub>2B</sub> , A <sub>3</sub>	8.2, 8.7, 6.9, 8.4
ZM241385	Antagonist	A <sub>1</sub> , A <sub>2A</sub> , A <sub>2B</sub> , A <sub>3</sub>	6.6, 9.1, 8.2, 6.1
Xanthine amine congener (XAC)	Antagonist	A <sub>1</sub> , A <sub>2A</sub> , A <sub>2B</sub> , A <sub>3</sub>	7.6, 8.4, 6.9-8.8, 7.0

**Table 1** | The types of ligands which bind to adenosine receptors A<sub>1</sub>, A<sub>2A</sub>, A<sub>2B</sub> and A<sub>3</sub>. pKi is the negative logarithmic Ki value, which is the inhibition constant of a ligand and can be used to measure binding affinity. A higher the pKi value, the higher the affinity of a ligand binding to its receptor. pKi values for the table were obtained from Guide to Pharmacology (Armstrong et al., 2019)

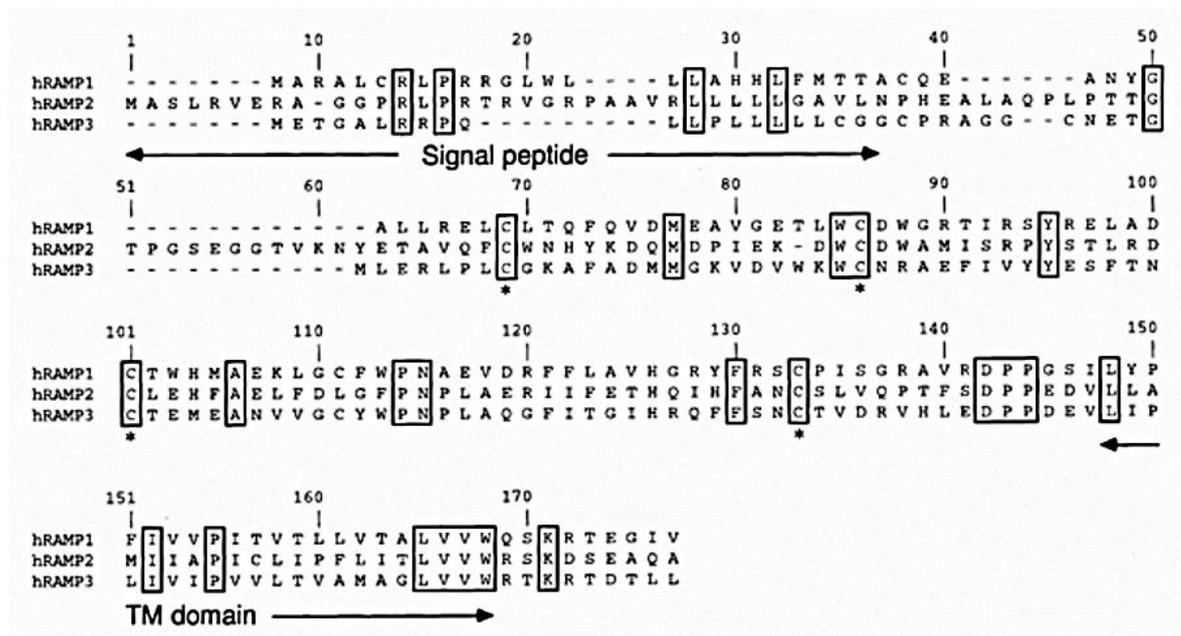
In this project, the A<sub>2A</sub>R was used as a model to assess the efficiency of SMA solubilisation of the receptor, into a SMALP. Various techniques were implemented in this project to assess the pharmacology and biophysical properties of the receptor within a SMALP. Radioligand binding and molecular techniques were implemented to observe the solubilisation and purification of the receptor-SMALP, as well as to test the binding capability of the receptor throughout the solubilisation and purification processes. Various techniques were implemented in this project to study novel properties of the A<sub>2A</sub>R, when incorporated into a SMALP. Such techniques include: fluorescence correlation spectroscopy (FCS), which observed the pharmacodynamics a single ligand and a single receptor-SMALP; and x-ray radiolytic footprinting (XRF), which highlighted the residues of the SMA-solubilised receptor that were exposed to water. The A<sub>2A</sub>R-SMALP was also used in phage display, to test the suitability of SMALPs in drug discovery.

### 1.3.5: Family B GPCRs

The calcitonin gene-related peptide receptor (CGRPR) is a family B secretin-like GPCR and was used in this project to identify residues of the ECL which interact with CGRP using photoaffinity cross-linking. The receptor was also modified with fusion tags and expressed in mammalian cells for SMA solubilisation and phage display.

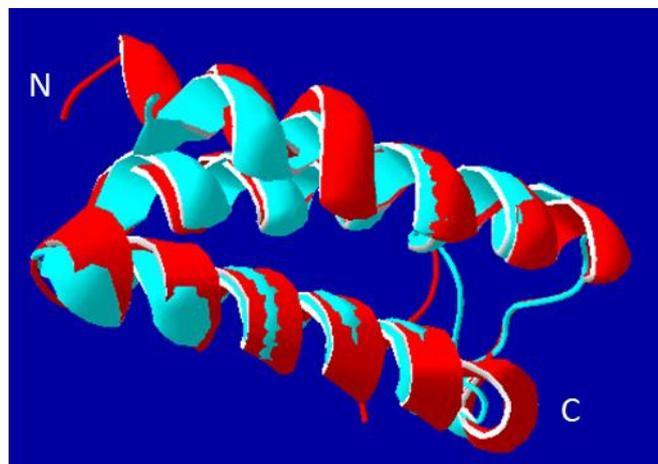
Within the family B GPCRs, there are 15 secretin-like receptors which have been discovered with a sequence homology of 30-50% (Karageorgos *et al.*, 2018. Fredriksson and Schioth, 2005). These receptors differ from the family A counterparts through their elongated extracellular domain (ECD), which is made up of around 150 amino acids (Gingell *et al.*, 2016). Family B ligands are often peptides, which are considerably larger in size when compared with family A ligands and also make more chemical contacts. Manufacturing pharmaceutical drugs against family B GPCRs has been a challenge due to the native ligands being larger than a small molecule drug. Therefore drug design methods involve mimicking the large peptides interacting with a receptor (Karageorgos *et al.*, 2018). Monoclonal antibodies have recently been developed to target family GPCRs, where the antibody, Erenumab, has been made to target the calcitonin gene-related peptide receptor (CGRPR) to prevent migraine (Karageorgos *et al.*, 2018). Allosteric ligands are currently being investigated as a non-peptidic ligand to alter the pharmacology of a family B GPCR, as with the case of the glucagon-like peptide 1 receptor (GLP1R) (Zhao *et al.*, 2020). Family B GPCRs are involved in homeostasis, where they regulate endocrine signalling, the central nervous system and the cardiovascular system. Pathophysiological conditions arise when there are alterations in family B signalling and receptor mutations, which can be life-threatening (Karageorgos *et al.*, 2018). Like family A GPCRs, there are highly conserved residues found amongst family B GPCRs, however the conserved residues differ from motifs seen in family A GPCRs. A numbering scheme called the 'Wootten numbering system' can be applied to family B GPCRs to compare their structures.

Although there are 15 secretin-like receptors within family B GPCRs, many can interact with receptor activity-modifying proteins (RAMPs), which can alter their pharmacology as well as traffic the receptor to the cell surface (Hay and Pioszak, 2016; Gingell *et al.*, 2016). RAMPs are present in the plasma membrane and consist of one transmembrane domain, an ECD made of around 100 amino acids, and an intracellular C terminus (Karageorgos *et al.*, 2018. Booe *et al.*, 2015). Three RAMPs have been discovered in the human genome, termed: RAMP1, RAMP2 and RAMP3. They differ in their amino acid sequence as shown in **figure 14** (McLatchie *et al.*, 1998).



**Figure 14** Amino acid sequence of the three human RAMP isoforms. Boxed residues are conserved amino acids. Amino acids labelled with an asterisk (\*) indicate the conserved cysteine residues. Diagram taken from (McLatchie *et al.*, 1998).

The secondary structure of the ECD of RAMPs is conserved due to the four cysteine residues being present, which can form two disulphide bonds (**figure 15**) (Liang *et al.*, 2018; Kusano *et al.*, 2012; McLatchie *et al.*, 1998). Most of the research on RAMPs has been conducted on the calcitonin receptor (CTR) and the CGRPR. Experiments have demonstrated that RAMPs are essential for the cell surface expression of some family B GPCRs, where the GPCR calcitonin-like receptor (CLR), requires coupling to a RAMP in order to be expressed on the cell surface (Weston *et al.*, 2016). RAMPs can also alter ligand selectivity and pharmacology of the receptor, where CLR/RAMP1 can bind to CGRP with high affinity and is therefore termed the CGRPR (Weston *et al.*, 2016). However when CLR is coupled with RAMP2 or RAMP3, the affinity for CGRP declines, whilst the affinity for adrenomedullin increases (Weston *et al.*, 2016). CTR can also interact with RAMPs, however unlike CLR, it can be expressed by itself. CTR has a high affinity for calcitonin, however once expressed with the RAMPs, its affinity for amylin increases (Karageorgos *et al.*, 2018). RAMPs therefore impact the signalling process of GPCRs as well as ligand selectivity.



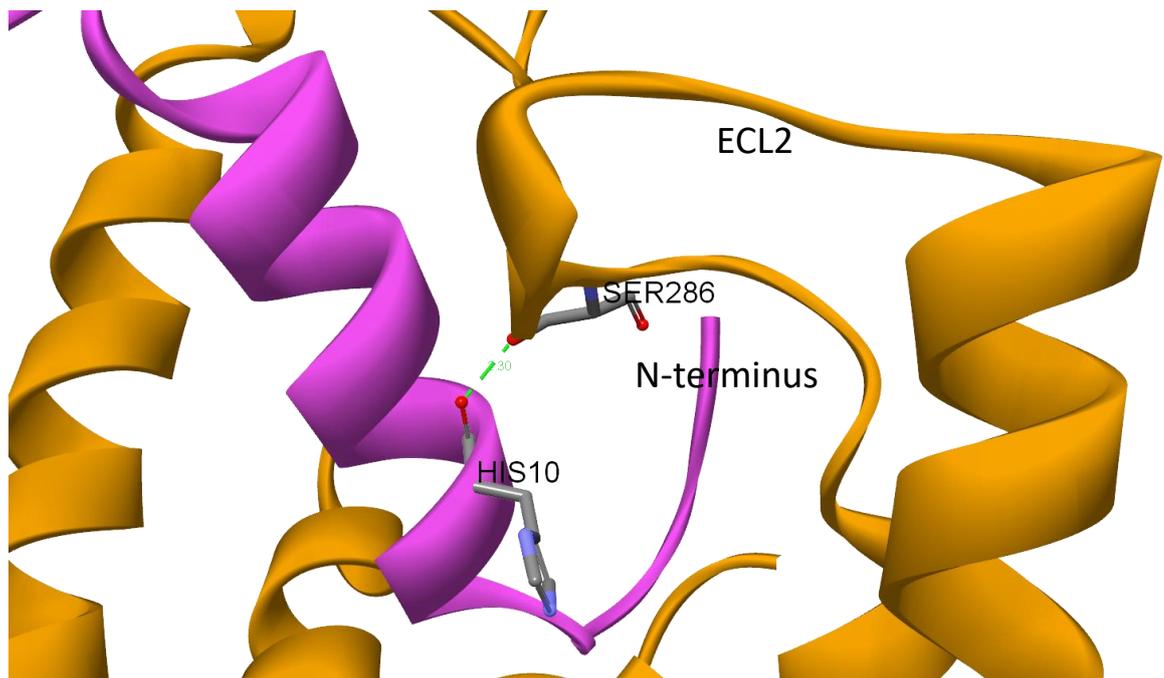
**Figure 15** | Superimposed structures of the ECD of RAMP1 and RAMP2. RAMP1 (PDB ID: 6E3Y) is shown in red and the cysteine residues are shown as red sticks. RAMP2 (PDB ID: 3AQE) is shown in blue and its cysteine residues are shown as blue sticks. N is the N terminus and C is the C terminus. The structures were superimposed by their sequence, using Swiss PDB viewer. PDB files were obtained from (Liang *et al.*, 2018) and (Kusano *et al.*, 2012).

### **1.3.6: Common mechanisms of family B GPCR activation upon peptide binding**

In order to generate conformation specific nanobodies against the CGRPR and to determine ligand binding to the ECL residues of the receptor, understanding the conformational changes occurring upon receptor activation is important, with a main focus on the extracellular region of the receptor.

The family B GPCR ligands are peptides, which make more chemical contacts with a target receptor than family A GPCR ligands. The endogenous peptides are around 25-114 residues in size, which include secretin, glucagon, calcitonin, parathyroid hormone (PTH), CGRP and amylin (Wheatley *et al.*, 2012. Culhane *et al.*, 2015). The N terminus of a peptide interacts with the juxtamembrane domain of a GPCR (which includes the TM bundle and the ECLs), whilst the C terminus makes contact with the ECD (Karageorgos *et al.*, 2018. Culhane *et al.*, 2015; Liang *et al.*, 2018). Deletion studies have shown that removing the termini residues of the  $\alpha$ CGRP reduces the binding affinity to its target receptor, demonstrating that both the N terminus and C terminus of a peptide are capable of making molecular contacts (Watkins *et al.*, 2013). The positioning of a peptide ligand within a receptor has been identified in recent structures, such as the CGRPR and the glucagon receptor (Liang *et al.*, 2018; Zhang *et al.*, 2018).

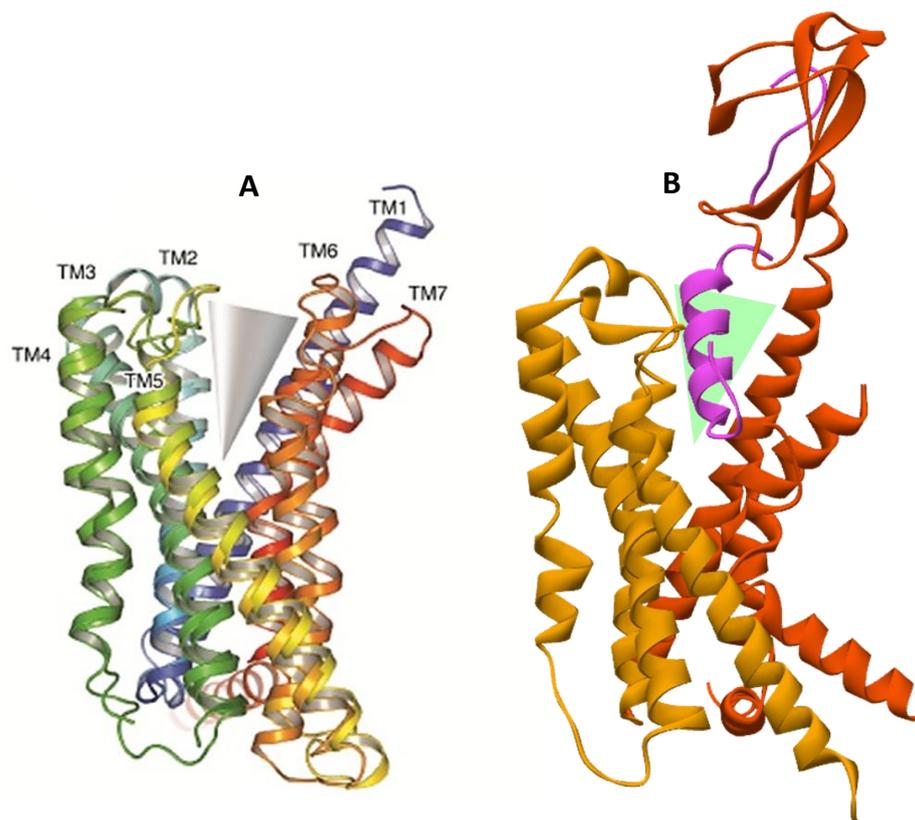
Analysis of different family B GPCR structures have also informed the process of ligand binding. Studies with the glucagon-like peptide 1 (GLP-1) receptor have shown that the hydrophobic motif of the C-terminus of the peptide ligand makes contact with the ECD of the receptor. The ECD reorientates itself, to allow the peptide to interact with the TM bundle. The middle of the peptide interacts with the stalk and ECL1, which has been confirmed in Liang *et al.*, (2018) and the ECL1 photoaffinity cross-linking experiment in **Chapter 8**. The ECL2 of the receptors also makes contact with a peptide ligand, where CRF1 receptor studies have demonstrated that Lys257 of ECL2 is in close proximity to the N-terminal domain of CRF (Karageorgos *et al.*, 2018). The closeness between the peptide and ECL2 can also be observed in the full length cryo-EM structure of the CGRPR, where the loop is orientated towards the N terminus of the CGRP and residue (**figure 16**) (Liang *et al.*, 2018). The N terminus of the peptide also forms interactions within the TM bundle, leading to receptor activation (Hollenstein *et al.*, 2014; Underwood *et al.*, 2010; Zhao *et al.*, 2020).



**Figure 16** | Cryo-EM structure of the CGRPR:CGRP complex (PDB ID: 6E3Y). RAMP1 is not shown in the figure. Ser286 of ECL2 forms a hydrogen bond with the backbone of His10 of the peptide. Bond length is 2.30 Å. Structural analysis conducted using BIOVIA Discovery Studio.

The ECD of family B GPCRs consists of a conserved  $\alpha$ - $\beta$ - $\beta$ - $\alpha$  structural motif (alpha helices and beta sheets). The N-terminus is alpha helical, followed by two anti-parallel  $\beta$ -sheets, which are known as  $\beta$ 1- $\beta$ 2 and  $\beta$ 3- $\beta$ 4, followed by an alpha helical structure (Karageorgos *et al.*, 2018). The two anti-parallel  $\beta$  sheet structure is also termed the sushi domain, which is essential in protein-protein interactions.

Family B GPCRs are configured into an open V-shape formation, where the open end is at the extracellular surface of the GPCRs, as shown with the superimposed structures of the CRF1 receptor and the glucagon receptor **figure 17A** (Bortolato *et al.*, 2014) and the more recent cryo-EM structure of the CGRPR (**figure 17B**) (Liang *et al.*, 2018).



**Figure 17** | **A**) Superimposed structures of the CRF1 receptor (PDB ID: 45KY) and the glucagon receptor (PDB ID: 4L6R). Receptors shown in rainbow coloration to denote sequence polarity from blue to red (N-C terminus). Receptor is viewed from the membrane point-of-view. The TM helices form a characteristic V shape formation. Diagram taken from (Bortolato *et al.*, 2014). **B**) Structure of the CGRP-bound CGRPR, without the RAMP1 (PDB ID: 6E3Y). Structure in red marks TMs 1, 6 and 7 and the ECD of CLR. Structure in orange marks TMs 2, 3, 4 and 5. The CGRP ligand is shown in purple. The green triangle highlights the V-shape chalice. Structure was made using BIOVIA Discovery Studio.

### **1.3.7: The calcitonin gene-related peptide receptor**

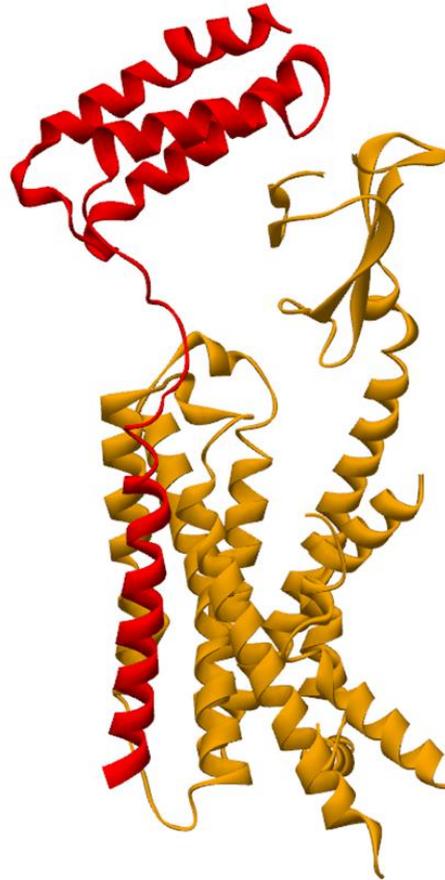
CGRPR is a complex of two proteins, CLR and RAMP1. It is a heavily studied receptor amongst the family B GPCRs and was the first to be identified to interact with a RAMP (McLatchie *et al.*, 1998). The RAMPs are essential for transporting CLR to the cell surface, as well as allosterically altering ligand selectivity to favour the peptide, CGRP (Simms *et al.*, 2019). It was hypothesised that the three RAMPs compete with each other to form a complex with CLR in the Golgi apparatus (Russel *et al.*, 2014). Co-expression experiments have shown that the CLR has a preference for RAMP1 (Buhlmann *et al.*, 1999). Although CLR/RAMP1 has been the consensus complex for CGRP binding, the peptide can also interact with CLR/RAMP2 and CLR/RAMP3 although to a lower extent (Russel *et al.*, 2014). The CGRP receptor is expressed in different tissue types with many roles.

The CGRPR is important for many biological processes and is an important drug target. In the cardiovascular system, CGRP can act as a potent microvascular vasodilator upon binding to CGRPR (Russel *et al.*, 2014). Upon inhibition by antagonists, the vasodilatory effects of CGRP become reduced. Studies have shown that the application of CGRP to the endothelium increases cAMP accumulation through the adenylate cyclase pathway upon G protein activation (Russel *et al.*, 2014). CGRP has been shown to be essential for regulating blood pressure, where pathogenesis of hypertension occurs through reduced CGRP activity, as demonstrated by knockout CGRP mice (Russel *et al.*, 2014). Migraine is a complex condition involving the neurovascular system. It is a disease affecting 8% of males and 18% of females (Russel *et al.*, 2014). These symptoms include: a painful headache, high photosensitivity, sensitivity to sound, vomiting, nausea and fatigue (Russel *et al.*, 2014). Studies have shown that the intravenous administration of CGRP in patients elevates the onset of migraine (Lassen *et al.*, 2008). To combat the effects of migraine, many antagonists have been made to prevent CGRP binding to CGRPR in the neurovascular system. Antagonists include: Olcegepant and telcegepant, which bind the extracellular domain of the CLR:RAMP1, preventing important contacts being made between the CGRP and its receptor (Russel *et al.*, 2014; Charles and Pozo-Rosich, 2019; ter Haar *et al.*, 2010; Warfvinge *et al.*, 2019). Efforts have been made to target the receptor to prevent the onset of migraine, including the generation of CGRPR antagonists, olcegepant and telcegepant. Unfortunately hepatotoxicity was

associated with their administration (Charles and Pozo-Rosich, 2019; Hay *et al.*, 2017; Akhtar, 2019). Due to the difficulties associated with generating small molecule antagonists against the CGRPR, an antibody-based alternative approach is being investigated to target the receptor (Charles and Pozo-Rosich, 2019). Galcanezumab, fremanezumad, erenumab and eptinezumab are antibodies which have entered clinical trials to prevent onset of migraine, with erenumab targeting the receptor and the rest targeting the peptide ligand (Akhtar, 2019). Out of the four monoclonal antibodies, eptinezumab demonstrated the best results in reducing migraine. The antibodies were deemed generally safe in the clinical trials, with a slight risk of adverse effects (Akhtar, 2019). Using phage display to generate CGRPR-specific nanobodies will therefore lead to the discovery of safer, therapeutic drugs to relieve migraine.

The peptide for CGRPR exists as two isoforms, termed  $\alpha$ CGRP and  $\beta$ CGRP. In chromosome 11, two different genes are transcribed and translated into the two isoforms which are 37 amino acids in size (Russel *et al.*, 2014; Kee *et al.*, 2018). Human  $\alpha$ CGRP and  $\beta$ CGRP have 90% sequence similarity, where three amino acids differ between each other (Russel *et al.*, 2014).  $\alpha$ CGRP is mainly found in the central and peripheral nervous system, whilst the  $\beta$  isoform is involved in enteric transmission (Russel *et al.*, 2014; Kee *et al.*, 2018). The isoforms have similar roles to each other however the  $\alpha$ CGRP has been most heavily studied. Both isoforms consist of four domains.

As the CGRPR is composed of the CLR and RAMP1, it is fundamental to understand the molecular contacts made between the two proteins as well as the peptide, The recent cryo-EM structure provides valuable insight on how family B GPCRs function as well a structure of the CLR, in complex with RAMP1, which will be useful for drug discovery (**figure 18**) (Liang *et al.*, 2018). The TM of RAMP1 resides between TMs 3-5 of CLR, which forms an interface for RAMP1 (Liang *et al.*, 2018). A majority of the TM interactions are formed between the TM5 of CLR and the TM of RAMP1 (Liang *et al.*, 2018), Mutagenesis of the RAMP1 interface of CLR reduces binding of CGRP (Liang *et al.*, 2018). Residues within ECL2 of CLR were also shown to make contact with RAMP1.



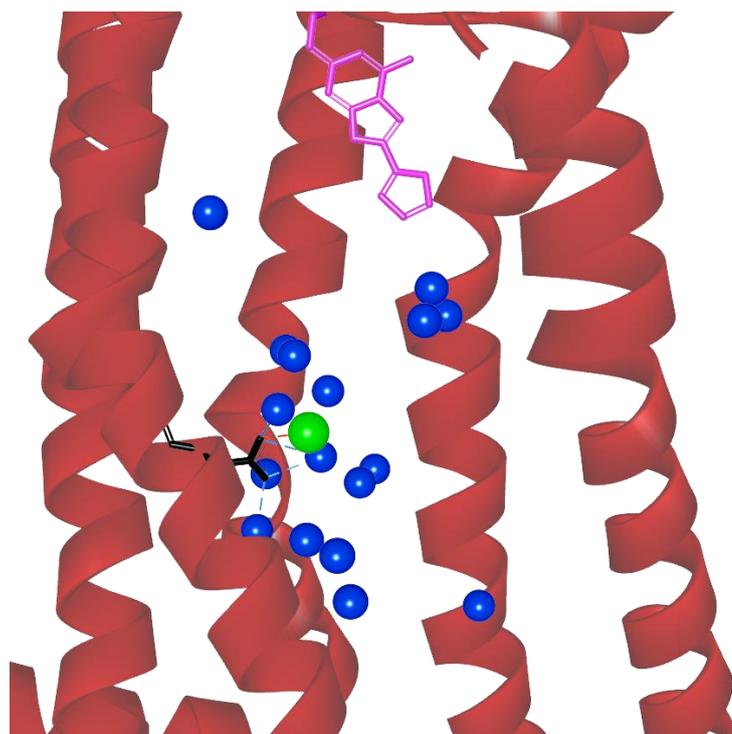
**Figure 18|** Crystal structure of the CGRPR (PDB ID: 6E3Y). The CLR is shown in orange and RAMP1 is shown in red. Structure was made using BIOVIA Discovery Studio.

### 1.3.8. Allosteric modulation of GPCRs

GPCRs contain allosteric binding sites as well as the orthosteric ligand binding site. Orthosteric ligand binding sites are sites where orthosteric ligands bind and are often exploited by pharmaceutical companies for drug development (Jakubik *et al.*, 2019). An example of an orthosteric ligand would be adenosine and CGRP binding to the orthosteric site in the  $A_{2A}R$  and the CGRPR respectively. An allosteric site is positioned away from where endogenous orthosteric ligand binds to, where molecules binding to the allosteric site are termed allosteric ligands or modulators (Digby *et al.*, 2010; Jakubik *et al.*, 2019). The addition of multiple binding sites of a receptor makes their pharmacology more complex. RAMPs, G proteins and arrestin, which have been mentioned in previous sections, are examples of endogenous allosteric modulators which affect GPCR pharmacology. However there are many additional protein and non-protein components which can alter a receptors behaviour. These can either

be positive allosteric modulators which influence the receptor to be in an active state, or a negative allosteric modulators which push a receptor into an inactive conformation. Neutral allosteric modulators do not influence receptor pharmacology however they may compete with other allosteric modulators (Wold *et al.*, 2019). As this project involves the use of SMALPs to solubilise GPCRs with the surrounding lipid bilayer, which can act as an allosteric modulator, understanding allosteric effects of GPCRs was fundamental to this project.

In the A<sub>2A</sub>R, a sodium ion binding pocket was discovered within the ZM241385-bound A<sub>2A</sub>R structure (**figure 19**) (PDB ID 4EIY, Liu *et al.*, 2012). A sodium ion binding region was detected around the highly conserved Asp52<sup>2.50</sup> (Massink *et al.*, 2015). There are four residues in the TM bundle which form a water network with water molecules. The residues are Ser91<sup>3.39</sup>, Trp246<sup>6.48</sup>, Asn280<sup>7.45</sup>, and Asn284<sup>7.49</sup> (Massink *et al.*, 2015). The water network coordinates the sodium ion into the allosteric site of A<sub>2A</sub>R. Upon transitioning to an active state, the sites which interact with structural waters change (Massink *et al.*, 2015). This indicates that water molecules as well as ions can act as allosteric modulators, interacting with GPCRs and stabilising their conformation. From molecular dynamic simulations and wet lab experiments, increasing the sodium ion concentration promotes the receptor to transition into an inactive state (Massink *et al.*, 2015).



**Figure 19** | Crystal structure of the ZM241385-bound A<sub>2A</sub>R (PDB ID: 4EIY) with the sodium ion present within the TM bundle of the receptor. The sodium ion is shown as a green sphere. The blue spheres represent water molecules. The red dashed line marks an electrostatic interaction between Na<sup>+</sup> and Asp52<sup>2.50</sup>. The blue dashed lines indicate hydrogen bonds formed between the water molecules and Asp52<sup>2.50</sup>. The pink chemical structure is the ligand, ZM241385. The black stick chemical structure is the highly conserved Asp52<sup>2.50</sup>. The red ribbons represent the A<sub>2A</sub>R. The structure was analysed using BIOVIA Discovery Studio.

Lipids also impact the pharmacology of receptors. The use of detergents to solubilise proteins often removes the allosteric effects of a native lipid on its target protein. The phospholipids within the lipid bilayer have shown to affect GPCR ligand binding, stability of receptor conformation and receptor dimerization (Song *et al.*, 2019). Radioligand binding experiments with  $\beta$ 2 adrenergic receptor have shown the impact lipids can have on a receptor. 1,2-Dioleoyl-sn-glycero-3-phosphoglycerol (DOPG) is a negatively charged lipid which stabilised the active conformation of the receptor. Replacing surrounding lipids with 1,2-Dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE), which has no net charge, pushes the conformation of the receptor to an inactive state, as shown by a reduced displacement of [<sup>3</sup>H]-DHAP by isoproterenol (Dawaliby *et al.*, 2016). Cholesterol is also an important allosteric modulator that binds to GPCRs such as the A<sub>2A</sub>R. Seven cholesterol binding sites were discovered in the hydrophobic regions of the TM bundle (Song *et al.*, 2019). Experiments involving the depletion of membrane cholesterol using methyl- $\beta$ -cyclodextrin reduced downstream accumulation of cAMP through adenylyl cyclase (McGraw *et al.*, 2019). Both lipids and cholesterol therefore have an impact on GPCR pharmacology, and should therefore be considered when generating membrane protein structures and drugs.

Drugs which allosterically modulate GPCRs are a recent trend, as recent allosteric drugs have been approved by the Food and Drug Administration (FDA) (Wold *et al.*, 2019). Allosteric binding sites between GPCRs show less homology, which can be exploited by pharmacologist to design subtype specific drugs (Wold *et al.*, 2019).

## **1.4: The use of the styrene-maleic acid copolymer to solubilise receptors**

### **1.4.1: How detergents are used to solubilise membrane proteins**

As mentioned previously, in order to generate crystal structures of a membrane protein, solubilisation of the protein must occur in order for the protein to be successfully purified and crystallised. However, many classes of detergents with different properties have been used for solubilisation. Detergents are amphipathic molecules with a hydrophobic tail and a hydrophilic headgroup (Moraes *et al.*, 2014). Due to the amphipathic nature of detergents, they are expected to disrupt the lipid bilayer without degrading the protein. Detergents can form a micellar structure around the protein and mimic the lipid bilayer in an attempt to retain protein structure, rendering the protein soluble (Moraes *et al.*, 2014). When the protein-detergent micelle is in solution, the polar headgroups of the detergent can form hydrogen bonds with water molecules, whilst the hydrophobic tails complex together to form the micellar structure (Moraes *et al.*, 2014). The micelles have a weight of less than 100 kDa, depending on the type of detergent used (Seddon *et al.*, 2004). There are different classes of detergents to choose from, each with different properties.

Ionic detergents contain an anionic/ cationic headgroup. They are capable of disrupting the hydrophobicity of a protein, causing its distortion, therefore this class of detergent is not generally applicable to solubilise proteins in a biologically relevant conformation (Moraes *et al.*, 2014). Ionic detergents can also be applied when less harsh methods of solubilisation do not work with a particular protein.

Non-ionic detergents consist of a hydrophilic headgroup and therefore possess no net charge (Moraes *et al.*, 2014). They are ubiquitously used to solubilise and purify membrane proteins for characterisation and crystallisation for structural studies. Compared to ionic detergents, non-ionic detergents are mild due to their properties in disrupting protein-lipid and lipid-lipid interactions without affecting protein-protein interactions (Moraes *et al.*, 2014). This class of detergents also come in different sizes, where short length varieties are often used for solubilisation and crystallography, whilst larger variants are used for reconstitution studies (Moraes *et al.*, 2014). Another advantage of using non-ionic detergents

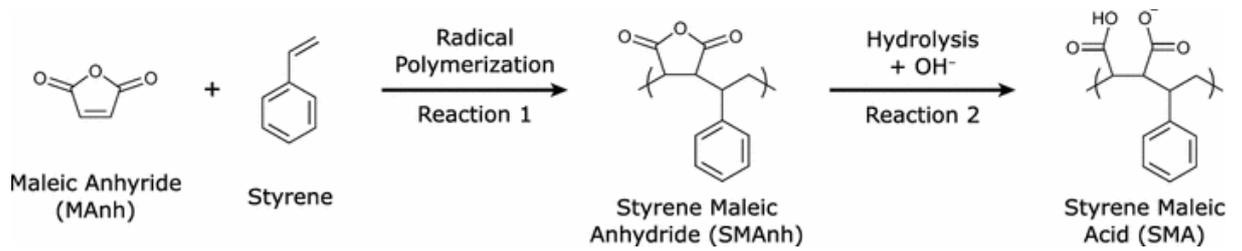
is that UV measurements are not skewed after solubilisation and purification, with exception to Triton due to its impurities (Moraes *et al.*, 2014). Maltosides and glucosides belong to the non-ionic class of detergents such as n-Dodecyl  $\beta$ -D-maltoside (DDM), which is a mild and non-denaturing detergent (Moraes *et al.*, 2014). Zwitterionic detergents are milder than ionic detergents, containing ionic and non-ionic properties (Moraes *et al.*, 2014). They have been used in crystallisation and NMR studies.

#### **1.4.2: The use of styrene-maleic acid copolymers as an alternative to detergents in solubilising membrane proteins**

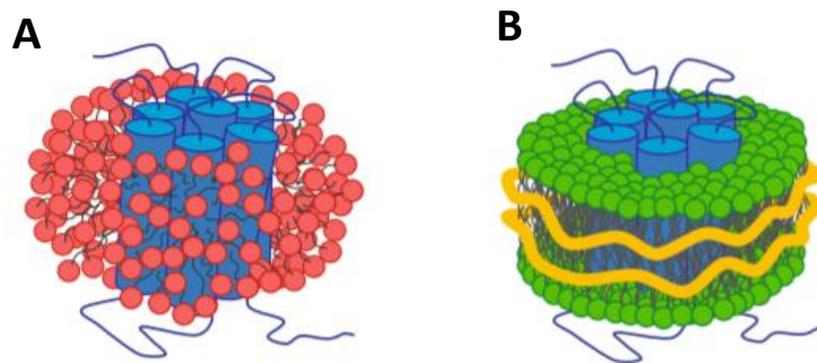
Progress in resolving structures of membrane proteins has been hindered due to the use of detergents to isolate membrane proteins. Detergents are capable of removing some or all of the surrounding lipid bilayer, potentially disrupting the folding of the protein, such as the lipophilic membrane-spanning protein domains (Jamshad *et al.* 2015; Lee *et al.*, 2016). Crystal structures of receptors therefore may not reflect a native conformation of a protein. Detergents are often made fresh for each time it is being used and finding a suitable detergent for a particular protein can be tedious to identify (Moraes *et al.*, 2014). The limitations associated with some detergents can be alleviated by using SMA copolymers.

In the year 2000, an organic polymer was discovered, which can solubilise proteins with the lipid bilayer intact (Knowles *et al.* 2009). The copolymer is styrene-maleic acid (SMA) which can form SMALPs when applied to a membrane. SMA is formed upon the hydrolysis of the styrene-maleic anhydride copolymer, which is formed through the copolymerization of styrene and maleic anhydride, as shown in **figure 20** (Dorr *et al.* 2016). Within life sciences, SMA copolymers can be used to treat cancer by encapsulating anti-cancer drugs for drug delivery (Tsukigawa *et al.*, 2015; Xian *et al.*, 2015). Tighe and Tonge discovered that the SMA copolymer can interact with phospholipids and solubilise lipid bilayers (Dorr *et al.* 2016). SMA can form a ringed structure on a membrane, which can be released from the membrane to form a discoidal structure containing lipid and protein (**figure 21**) (Dorr *et al.* 2016). Different styrene to maleic acid ratios in SMA copolymers can lead to different disc sizes when applied to the membrane, where SMA2000P (the most commonly used SMA copolymer) can form a ring shape that is 10nm in diameter (Dorr *et al.* 2016). The styrene component of SMA can be altered to increase the ratio, where SMA2000P consists of a 2 styrene: 1 maleic acid ratio (Dorr *et al.*, 2016). SMALPs are

therefore a useful tool to solubilise receptors and to keep them in a native conformation surrounded by a phospholipid bilayer.



**Figure 20** | Diagram showing the conversion reaction of styrene maleic-anhydride to SMA in a 1:1 ratio. Reaction 1 demonstrates the polymerization of styrene and maleic anhydride to form styrene maleic-anhydride. Reaction 2 demonstrates the conversion of the anhydride to an acid. Diagram taken from (Dorr *et al.*, 2016).



**Figure 21** | A) Illustration showing a membrane protein solubilised using detergents (red). The lipid has been stripped away by the detergents. B) Illustration showing a membrane protein solubilised using SMA. The protein is surrounded by a lipid bilayer (green head groups). The SMA (yellow) forms a ring-like structure around the lipid bilayer and protein to form a SMALP. Diagram taken from (Dorr *et al.* 2016).

The A<sub>2A</sub>R was previously expressed in human embryonic kidney cells (HEK293T) and *Pichia pastoris*, where SMA2000 was used to solubilise the receptor into SMALPs (Jamshad *et al.*, 2015). Different aspects of the A<sub>2A</sub>R-SMALP was tested, such as the thermostability of the copolymer, where the SMALPs were more stable than DDM micelles as the temperature increased (Jamshad *et al.*, 2015). Circular dichroism (CD) spectroscopy has also shown that the stability of the A<sub>2A</sub>R helix content was intact at 65°C (Jamshad *et al.*, 2015) demonstrating the better stability of the A<sub>2A</sub>R within a SMALP, than DDM. Radioligand binding experiments have also confirmed that the receptor was functional within a SMALP when the temperature was pushed to 65°C, albeit some loss of binding (Jamshad *et al.*, 2015). SMALPs were also proven resilient over time when compared with DDM, where binding experiments with the A<sub>2A</sub>R showed no loss of binding after 16 days of incubation at 37°C (Jamshad *et al.*, 2015). DDM-solubilised A<sub>2A</sub>R however lost binding after 2 days of incubation at 37°C. From the data gathered by Jamshad *et al.* (2015), SMALPs showed superiority over the commonly used detergent, DDM, in keeping the protein stable and functional over a long period of time.

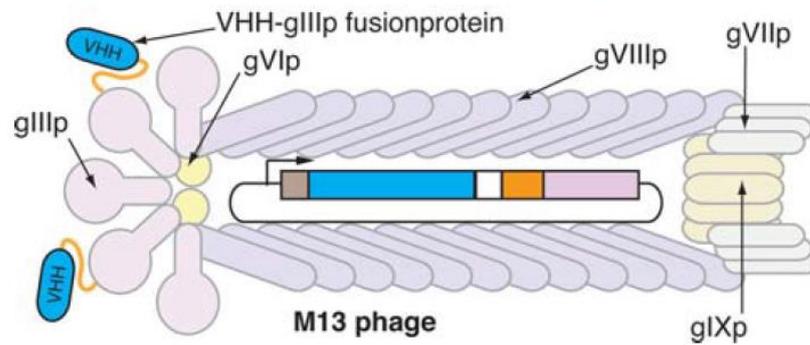
As SMALP research is a new field in membrane protein biology, the understanding of how SMA interacts with the lipid bilayer is somewhat limited. In 2018, molecular dynamic simulations and analysis have suggested how SMA interacts with the membranes to solubilise proteins (Xue *et al.*, 2018). SMA copolymers in solution formed a disordered configuration, where aggregation could occur through the hydrophobic regions of the copolymer (Xue *et al.*, 2018). Once the copolymer was applied to a lipid environment, the hydrophobic styrene embeds itself between the acyl tails of the lipid bilayer at a 20 ns time point (Xue *et al.*, 2018). After 400 ns, the copolymer fully immersed itself into the lipid bilayer, where electrostatic interactions between the SMA carboxyl group and phospholipid head groups were mediated by sodium ions (Xue *et al.*, 2018). The structure of SMA changed into a stretched conformation. The copolymer, once absorbed into the lipid, caused membrane bending, lipid tilting, lipid extraction and water infiltration, according to the simulations (Xue *et al.*, 2018). The stability of the lipid bilayer became disrupted, where water infiltration enlarged the pores of the lipid bilayer (Xue *et al.*, 2018). The SMA copolymers formed disc shapes around the lipid, between the phospholipid head groups of the bilayer to form SMALPs through self-assembly of the copolymers (Xue *et al.*, 2018).

### 1.5: Characterisation of the SMA-solubilised A<sub>2A</sub>R and CGRPR.

As SMALPs have been shown to retain a GPCR in its native, functional conformation with the lipid bilayer intact, assessing further characteristics of the receptor-SMALP would provide new information on the pharmacology of the A<sub>2A</sub>R (Jamshad *et al.*, 2015). In this project, different techniques were used to assess the pharmacology and properties of the A<sub>2A</sub>R and the CGRPR within a SMALP.

Radioligand binding assays were used to compare the ligand binding affinity of ZM241385 to the A<sub>2A</sub>R in membranes and in SMALPs, to observe whether the GPCR retains its pharmacology after solubilisation (Uddin *et al.*, 2018; Jamshad *et al.*, 2015). Once the binding capability of the receptor was confirmed, additional techniques were used to identify novel features of the SMA-solubilised receptor. Fluorescence correlation spectroscopy applied the confocal microscope to observe the pharmacodynamics of a single fluorescent ligand and a single receptor-SMALP (Briddon *et al.*, 2018). The technique was used to observe the pharmacodynamics of the A<sub>2A</sub>R-SMALP, as well as the lipophilicity of ligands, to see how they non-specifically interact with the SMALP. Water-exposed residues of the A<sub>2A</sub>R-SMALP were analysed using X-ray radiolytic footprinting. The technique involved the irradiation of water molecules, to observe which regions of the SMA-solubilised A<sub>2A</sub>R interacts with water, which can be compared with the distribution of water networks in existing crystal structures (Gupta *et al.*, 2016; Gupta, 2019). Changes in water exposure of the A<sub>2A</sub>R can be observed by applying agonists and antagonists, which will inform how water networks may change when a GPCR transitions between an active and an inactive state, as observed with rhodopsin (Orban *et al.*, 2010).

The A<sub>2A</sub>R-SMALP was also used as a drug discovery target, to test whether phage display would be applicable in a SMALP context, in order to develop A<sub>2A</sub>R-specific nanobodies. Phage display applies phage technology to express nanobodies onto the coat protein during phage assembly (Hoogenboom, 2005). The phage used for nanobody generation were conjugated with a nanobody to the coat protein of the phage (**figure 22**) (Wesolowski *et al.*, 2009). A naïve phage library, was used to isolate the phage specific to the A<sub>2A</sub>R, which was in its native confirmation in a SMALP.



**Figure 22|** Illustration of an M13 bacteriophage, with a VHH nanobody fused to the gIIIp coat protein. The genome encoding the VHH nanobody can be digested and cloned into a phagemid vector, from the *Camelidae* genome. Within the phagemid, the VHH gene was fused to the gIIIp coat protein. The phagemid can then be transfected into *E.coli*, alongside a helper phage to assemble VHH-conjugated phage. Diagram taken from (Wesolowski *et al.*, 2009).

Photoaffinity cross-linking was also incorporated into this project to identify the ECL residues of the CGRPR, which interact with the fluorescein-conjugated  $\alpha$ CGRP. The experiments were conducted on detergent-solubilised receptors, however the technique could be transferrable to SMA-solubilised receptors. The technique is also beneficial in locking an SMA-solubilised GPCR in an active or an inactive state, which could be applied in drug discovery techniques such as phage display, to discover conformation and state-specific nanobodies.

## 1.6: Project aims

As SMALPs are becoming increasingly popular in membrane protein research, the characteristics and pharmacology of GPCRs in SMALPs must be understood, to ensure their binding capability and functionality is retained after solubilisation. Previous research with the A<sub>2A</sub>R has demonstrated successful SMA solubilisation, however further experimental techniques can be implemented to understand the properties and pharmacology of the receptor in a SMALP context (Jamshad *et al.*, 2015). In this project, the first aim was to characterise the A<sub>2A</sub>R and the CGRPR in SMALPs, using a range of techniques, such as x-ray radiolytic footprinting, fluorescence correlation spectroscopy, cyclic AMP accumulation and radioligand binding. This is to ensure the receptors are binding capable and to discover properties, such as water-exposed residues of an SMA-solubilised receptor. The second objective was to assess the applicability of using SMALPs as a drug discovery tool, using the phage display technique. SMALPs were used as a platform for phage display nanobody discovery due to their ability to retain the native conformation of a GPCR, therefore allowing nanobody-conjugated phage to bind conformational epitopes of the receptor within the SMALP. (Wheatley *et al.*, 2016; Jamshad *et al.*, 2015). Finally, the photoaffinity cross-linking assay was used to covalently lock the CGRPR into an active state in order to discover nanobodies which are conformationally selective against the active GPCR and to study whether the technique is applicable with SMALP GPCRs.

## Chapter 2: Materials and methods

### 2.1: Table of materials used in experimental procedures

**Table 2: Reagents for quick-change polymerase chain reaction (PCR)**

Enzyme/ Reagent	
PfuUltra II Hotstart PCR master mix	

**Table 3: Restriction enzymes and buffers used for molecular biology experiments**

Buffers/ enzymes	Company
NotI	
EcoRI	
PmeI	
DpnI	
NEBuffer 3.1	
Cutsmart buffer	

**Table 4: Reagents for DNA electrophoresis**

Solutions/ reagents	Components
1 X TRIS-acetate-EDTA (TAE buffer)	20 ml 50 X TAE, pH 8.3 (from Merck) 980 ml dH <sub>2</sub> O
Hi-Res Standard agarose (from AGTC bioproducts).	
RedSafe nucleic acid staining solution- (from Intron)	
6 X Bromphenol blue loading dye	4 g sucrose (Sigma Aldrich) 25 mg bromophenol blue powder (Sigma Aldrich) Dissolved in dH <sub>2</sub> O until the final volume of 10 ml was reached.
Gene ruler DNA ladder- 1 kb (from Thermofisher Scientific)	

**Table 5: Solutions and reagents for competent *E. coli* cell generation, transformation and growth.**

<b>Solutions/ reagents</b>	<b>Components</b>
Transformation buffer 1 (pH 5.8)	6 g RbCl (from Sigma-Aldrich) 5 g MnCl <sub>2</sub> ·4H <sub>2</sub> O (from Sigma-Aldrich) 15 ml 1 M Potassium acetate (pH 7.5) (from Sigma-Aldrich) 0.75 g CaCl <sub>2</sub> ·2H <sub>2</sub> O (from Sigma-Aldrich) 75 ml glycerol (from Fisher Bioreagents) Dissolved in 500 ml dH <sub>2</sub> O
Transformation buffer 2 (pH 6.8)	10 ml 0.5 M 3-(N-morpholino)propanesulfonic acid (MOPS) (pH 6.8) (from Sigma-Aldrich) 0.6 g RbCl 5.5 g CaCl <sub>2</sub> ·2H <sub>2</sub> O 75 ml glycerol Dissolved in 500 ml dH <sub>2</sub> O
Lennox Broth (LB) media	20 g LB dissolved in 1 L dH <sub>2</sub> O
LB agar	5 g LB (from Fisher Bioreagents) 5 g agar (from Becton, Dickinson) Dissolved in 250 ml dH <sub>2</sub> O 50 µg/ml ampicillin (from Sigma-Aldrich)

**Table 6: Solutions and reagents for competent *Pichia pastoris* generation and electroporation**

<b>Solution/ reagents</b>	<b>Components</b>
Yeast extract, peptone, dextrose (YPD) media	10 g yeast extract (from Becton, Dickinson) 20 g peptone (from Becton, Dickinson) Dissolved in 900 ml dH <sub>2</sub> O After sterilising, 100 ml of 10 X dextrose (from Sigma-Aldrich) was added.
BEDS solution	10 mM bicine- NaOH, pH 8.3 (from Sigma-Aldrich) 3% (v/v) ethylene glycol (from Sigma-Aldrich) 5% (v/v) dimethyl sulfoxide (DMSO) (from Thermofisher Scientific) 1 M sorbitol (from Sigma-Aldrich)
DTT solution	1 M dithiothreitol (DTT) (from Sigma-Aldrich)
YPD, sorbitol (YPDS) agar plate	5 g yeast extract 10 g peptone 91.1 g sorbitol 10 g agar Zeocin (0-1000 µg/ml) (from Invitrogen)

**Table 7: Solutions and media for *Pichia pastoris* growth and expression**

<b>Buffer/ reagent</b>	<b>Components</b>
0.02 % Biotin	20 mg Biotin (from Supelco) Dissolved in 100 ml dH <sub>2</sub> O Filter sterilised and stored at 4°C
20 % Dextrose	200 g D-glucose Dissolved in 1 L dH <sub>2</sub> O Autoclaved and stored at room temperature (RT)
5 % Methanol	50 ml methanol (from Fisher Scientific) Up to 1 L in dH <sub>2</sub> O Filter sterilised and stored at 4°C
10 % Glycerol	100 ml glycerol Dissolved in 1 L in dH <sub>2</sub> O
10 X Yeast nitrogen base extract (YNB)	With ammonium sulphate Without amino acids Dissolve YNB (according to suppliers instructions) in dH <sub>2</sub> O to 500 ml to obtain 10 X YNB Filter sterilised and stored at 4°C
1 M Potassium phosphate buffer (pH6)	132 ml 1M K <sub>2</sub> PO <sub>4</sub> (pH6) 868 ml 1M KH <sub>2</sub> PO <sub>4</sub> (pH6) Autoclaved and stored at RT
Yeast extract peptone dextrose (YPD) agar plate + zeocin	1 % yeast extract 2 % peptone 2 % dextrose 2 % agar 100 µg zeocin (from Thermofisher) Dissolved in dH <sub>2</sub> O Filter sterilised and stored at 4°C
Buffered glycerol-complex medium (BMGY)	1 % yeast extract 2 % peptone 100 mM potassium phosphate (pH6) 1.34 % YNB 4 x 10 <sup>-5</sup> % biotin 1 % glycerol Dissolved in dH <sub>2</sub> O Autoclaved and stored at 4°C
Buffered methanol-complex medium (BMMY)	1 % yeast extract 2 % peptone 100 mM potassium phosphate (pH6) 1.34 % YNB 4 x 10 <sup>-5</sup> % biotin 0.5 % methanol dissolved in dH <sub>2</sub> O Autoclaved and stored at 4°C
Breaking buffer (pH8)	50 mM sodium phosphate (pH8) (from Sigma-Aldrich) 2 mM ethylenediaminetetraacetic acid (EDTA) (from Sigma-Aldrich) 100 mM NaCl (from Sigma-Aldrich) 5 % glycerol 2 tablets of protease inhibitor cocktail (Roche)

**Table 8: Solutions and reagents for mammalian cell culture and growth.**

<b>Solutions/ reagents</b>	<b>Components</b>
Dulbecco's modified eagle media (DMEM), Glutamax, high glucose and pyruvate (from Thermofisher scientific)	Standard DMEM components 25 mM glucose 1 mM pyruvate
Fetal bovine serum (FBS), Gibco (from Thermofisher scientific)	
Opti-MEM, reduced DMEM, Gibco (from Thermofisher scientific)	
Lipofectamine 3000 reagent (from Thermofisher scientific)	P3000 Lipofectamine 3000
Poly-D-lysine hydrobromide (from Sigma-Aldrich)	Lyophilized poly-D-lysine reconstituted in sterilised dH <sub>2</sub> O to 50 mg/ml
100 X Penicillin Streptomycin (from Generon)	

**Table 9: Buffers for SMA solubilisation**

<b>Buffer</b>	<b>Components</b>
Buffer A	50 mM HEPES 150 mM NaCl 10 % Glycerol

**Table 10: Buffers and reagents for general deca-his-tagged protein purification**

<b>Buffers/reagents</b>	<b>Components</b>
Nickel-nitrilotriacetic acid agarose (Ni <sup>2+</sup> -NTA) (from Qiagen)	Stored in 100 % ethanol (from Fisher Scientific)
Wash buffer (pH8)	50 mM TRIS (from Sigma-Aldrich) 300 mM NaCl 10-20 mM imidazole (from Sigma-Aldrich)
Elution buffer (pH8)	50 mM TRIS 300 mM NaCl 250 mM imidazole (but can vary depending on experiment)
Dialysis buffer (pH8)	50 mM HEPES 150 mM NaCl

**Table 11: Buffers and reagents for deca-his-tagged protein purification for X-ray radiolytic footprinting.**

<b>Buffers/reagents</b>	<b>Components</b>
Nickel-nitrilotriacetic acid agarose (Ni <sup>2+</sup> -NTA) from Qiagen	Stored in 100 % ethanol
Wash buffer (pH8)	50 mM NaH <sub>2</sub> PO <sub>4</sub> (from Sigma-Aldrich) 300 mM NaCl 20 mM imidazole
Elution buffer (pH8)	50 mM NaH <sub>2</sub> PO <sub>4</sub> 300 mM NaCl 250 mM imidazole
Dialysis buffer (pH8)	20 mM NaH <sub>2</sub> PO <sub>4</sub> 50 mM NaCl

**Table 12: Radioligand binding buffers and reagents**

<b>Buffers/reagents/ligands</b>	<b>Components</b>
Binding buffer	50 mM HEPES 150 mM NaCl 5 % Glycerol
DMSO (Dimethyl sulfoxide)	
Ultima gold LLT scintillation cocktail (from Perkin Elmer)	

**Table 13: Buffers and reagents for sodium dodecyl sulfate (SDS)- polyacrylamine gel electrophoresis (PAGE)**

<b>Buffers/reagents</b>	<b>Components</b>
12 % Separating gel	For 1 gel: 1.9 ml 30 % Protogel polyacrylamide (from National Diagnostics) 2.2 ml dH <sub>2</sub> O 1.5 ml 1.5 M TRIS-HCL (pH8.8) (from Sigma-Aldrich) 60 µl 10% Sodium dodecyl sulfate (SDS) (from Sigma-Aldrich) 20µl 20% Ammonium persulphate (APS) (from Sigma-Aldrich) 4.5 µl TEMED
4 % Stacking gel	For 1 gel: 300 µl 30 % Polyacrylamide 600 µl dH <sub>2</sub> O 1.5 ml 0.5 M TRIS-HCL (pH6.8) 25 µl 10 % SDS 10 µl 20 % APS 2.5 µl TEMED
Running buffer	100 ml TRIS/glycine/SDS mixed with 900ml dH <sub>2</sub> O
5 X Laemmli sampling buffer	4 ml 1.5 M TRIS-HCL pH 6.8 10 ml glycerol 5 ml β-mercaptoethanol (from Sigma Aldrech) 2 g SDS 1 ml 1% bromophenol blue (from Sigma Aldrech)
His ladder protein standard (from ProtoMetrics)	
PageRuler Prestained protein ladder, 10-180 kDa (from Thermofisher scientific)	

**Table 14: Buffers for western blotting**

<b>Buffers/antibodies</b>	<b>Components</b>
Transfer buffer	200 ml 100 % Methanol 100 ml TRIS/glycine (from Geneflow) Mixed with 700 ml dH <sub>2</sub> O
Blocking buffer	2.5 g non-fat milk powder (from BIORAD) dissolved in 50 ml PBS
Electrophoresis buffer	100 ml 10 X TRIS/glycine/SDS (from Geneflow) 900 ml dH <sub>2</sub> O
Phosphate buffer saline (PBS) + Tween	1 X PBS dissolved in 1 L dH <sub>2</sub> O 0.2 % Tween-20 (from Sigma-Aldrich)
EZ-ECL Solutions A and B (from Geneflow)	

**Table 15: Reagents for Coomassie staining**

<b>Reagents</b>	<b>Components</b>
Instant Blue protein stain (from Expedeon)	

**Table 16: Buffers and reagents used for biotinylation of avi-tagged proteins**

<b>Buffers/ reagents</b>	<b>Components</b>
Biotin ligase reaction buffer	2.5 mM MgCl (from Sigma-Aldrich) 10 mM Adenosine triphosphate (ATP) Bir500 biotin ligase - volume depends on protein concentration. (from Avidity) 0.3 mM D-Biotin (from Sigma-Aldrich) 1X PBS Dissolved in 500 µl dH <sub>2</sub> O
Bir500 biotin ligase (from Avidity)	

**Table 17: Reagents used for biotin quantitation**

<b>Reagent</b>	<b>Components</b>
HABA/avidin reagent (from Sigma Aldrich)	Lyophilised HABA/avidin reagent was dissolved in 10 ml deionized water, yielding a solution at pH 7.3 with components: 0.3 mM HABA 0.45 mg/ml avidin 0.3 M NaCl 0.01 M HEPES 0.01 M MgCl <sub>2</sub> 0.02% sodium azide (used as a preservative)

**Table 18: Buffers and reagents used for indirect enzyme-linked immunosorbent assays (ELISA).**

<b>Buffers/ reagents</b>	<b>Components</b>
Dilution buffer	1 X PBS (pH 7.4) 0.15 µg/ml Bovine serum albumin (BSA) (from Sigma-Aldrich) Dissolved in dH <sub>2</sub> O
Blocking buffer	1% BSA 1% non-fat milk powder 1 X PBS Dissolved in dH <sub>2</sub> O
Streptavidin-alkaline phosphatase solution	40 µg/ml streptavidin-alkaline phosphatase (from Sigma-Aldrich) 1 X PBS Dissolved in dH <sub>2</sub> O
Streptavidin solution	1 mg/ml streptavidin (from Sigma-Aldrich) Dissolved in dH <sub>2</sub> O
p-nitrophenol phosphate (pNPP) solution (from Bio-Rad)	
SIGMAFAST o-Phenylenediamine dihydrochloride (OPD) tablets (from Sigma Aldrich)	

**Table 19: Buffers, reagents and media for phage display and phage ELISA**

<b>Buffers/reagents/media</b>	<b>Components</b>
Naïve llama VHH-expressing phage library	IP protected by UCB
Wash buffer	PBS (pH7.4) + 0.1 % Tween
Blocking buffer	PBS (pH7.4) + 3 % Bovine serum albumin (BSA)
Milk buffer	PBS (pH7.4) + 3 % BSA + 3 % skimmed milk
Elution buffer	0.1 M HCl
Neutralization buffer	1 M TRIS (pH8)
PEG solution	20 % Polyethylene glycol + 2.5 M NaCl
2TY media	100 µg/ml carbenicillin 50 µg/ml Kanamycin 1 % Glucose
LB agar plates	LB agar (manufactured by UCB) 100 µg/ml carbenicillin 1 % Glucose

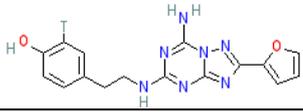
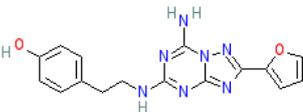
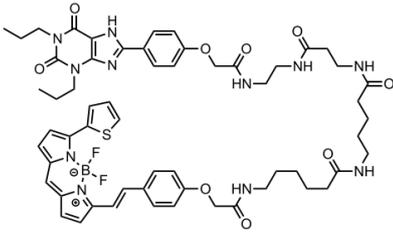
**Table 20: Buffers and reagents used for photoaffinity cross-linking**

<b>Buffers/ reagents</b>	<b>Components/ properties</b>
DDM solubilisation buffer	1% (w/v) dodecyl-beta-maltoside (from Anatrace) 10 mM TRIS-HCL (pH6.8) 50 mM NaCl 1 mM CaCl <sub>2</sub> 0.1 mM <i>phenylmethylsulfonyl fluoride (PMSF)</i> Roche protease inhibitor cocktail (1 tablet/ 10 ml).
Binding buffer (pH7.4)	1 X Hanks balanced salt solution (HBSS) (from Life technologies) 0.1% BSA

**Table 21: List of antibodies and phage for recombinant protein detection**

<b>Antibody</b>	<b>Company</b>	<b>Product code</b>	<b>Epitope</b>
Anti-6xHis mouse IgG	Clontech	631212	HHHHHH
Horseradish peroxidase (HRP)-linked horse anti-mouse IgG	Cell Signaling Technology	7076	Mouse IgG
Anti-FLAG rabbit IgG	Neo Biotech	NB-22-8776	DYKDDDDK
Anti-avi-tag mouse IgG	Neo Biotech	NB-22-15720	GLNDIFEAQKIEWHE
Anti-HA mouse IgG	Cell Biolabs	AKR-006	YPYDVPDYA
HRP-linked goat anti-rabbit IgG	Southernbio	4030-05	Rabbit IgG
HRP-conjugated anti-His mouse IgG	N/A (Given by UCB)	N/A	HHHHHH
Anti-phage (M13) mouse IgG	N/A (Given by UCB)	N/A	M13 bacteriophage
Fab2838-conjugated M13 phage	Made in UCB	N/A	A <sub>2A</sub> R
ScFv2838-conjugated M13 phage	Made in UCB	N/A	A <sub>2A</sub> R
Helper (control) M13 phage	Made in UCB	N/A	Non-specific

**Table 22: List of ligands and their properties**

Ligand	Company	Target receptor	Properties
[ <sup>3</sup> H]ZM241385	American Radiolabelled Chemicals inc.	A <sub>2A</sub> R	Stock is in 50 % ethanol Stock concentration is 1 mCi/ml  Structure: 
ZM241385	TOCRIS Bioscience	A <sub>2A</sub> R	Made stock concentration in DMSO at 30 mM.  Structure: 
CA200645	CellAura Technologies	A <sub>2A</sub> R	Structure (taken from Stoddart <i>et al.</i> (2012)) 
α-CGRP	Sigma Aldrich	CGRPR	Made stock concentration in DMSO at 10 <sup>-3</sup> M.  Peptide sequence: NH <sub>2</sub> - ACDTATCVTHRLAGLLSRSGGVVKNNF VPTNVGSKAF- COOH
α-CGRP 5(6)-carboxyfluorescein	University of Auckland, New Zealand	CGRPR	Made stock concentration in DMSO at 10 <sup>-3</sup> M. Molecular weight: 4163 g/mol Batch: PWHPEP7-193-1  Peptide sequence: NH <sub>2</sub> -ACDTATCVTHRLAG-K(fluoro)- LSRSGGVVKNNFVPTNVGSKAF- NH <sub>2</sub> (disulphide).

## 2.2: Methods

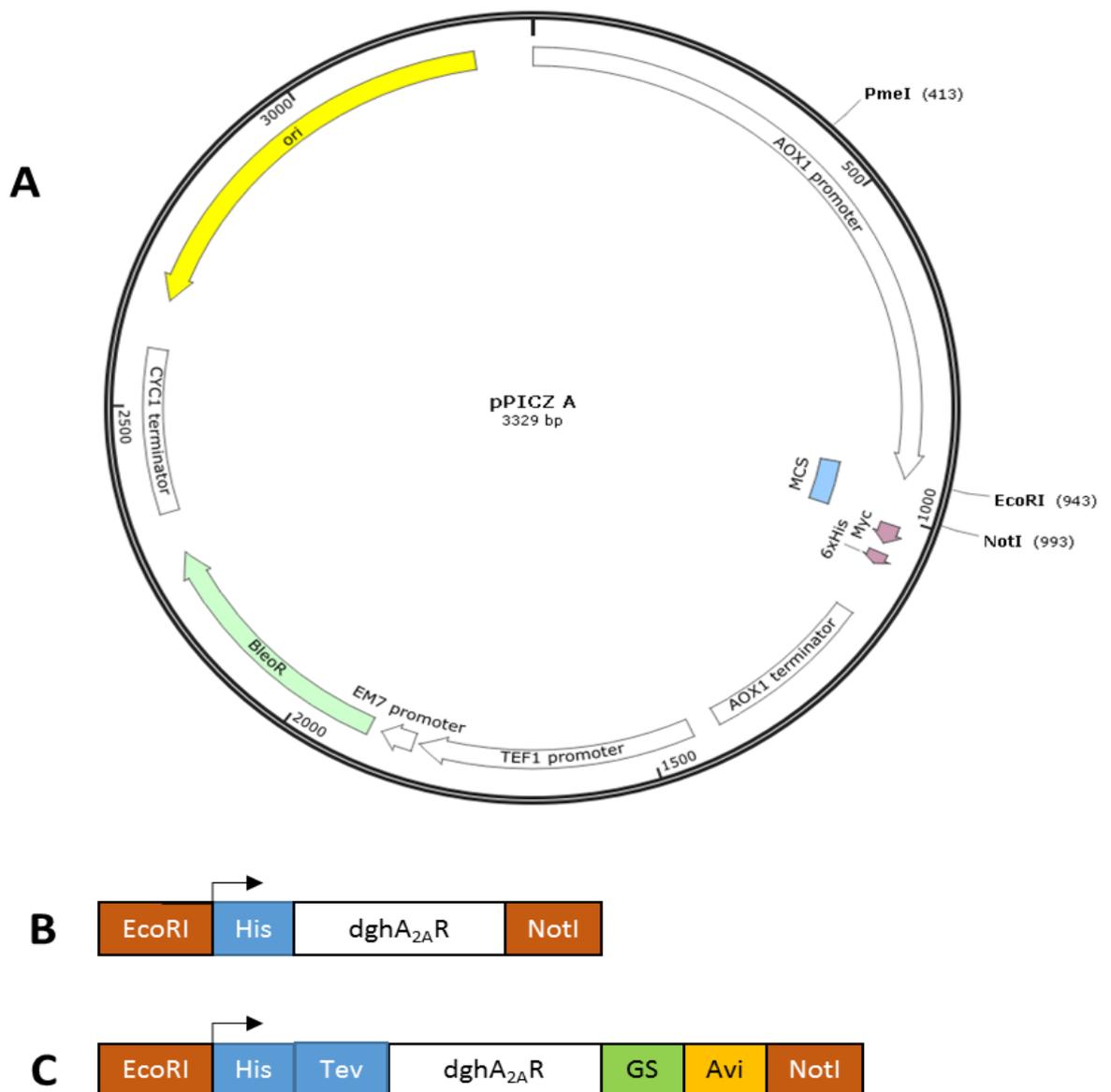
### 2.2.1: Plasmid constructs

The deglycosylated human adenosine 2a receptor (dghA<sub>2A</sub>R) gene construct was obtained from Fraser *et al.* 2006. Expression of dghA<sub>2A</sub>R was done similarly to Fraser (2006). A mutation was induced to replace Asparagine154 to Glutamine to prevent hyperglycosylation, without affecting receptor pharmacology, when expressing the construct in *Pichia pastoris* (Jamshad *et al.* 2015). The construct was ligated into pGEM-T Easy Vector (Promega) using EcoRI and NotI restriction enzymes and was amplified in *Escherichia coli* (DH5-alpha). The construct was excised out of pGEM-T Easy Vector using the same restriction enzymes and was inserted into an upstream 10x His tag (Fraser 2006), using the primer 'CGAGAAAAGAGAGGGCTGAAGCTCATCATCATCATCATCATCATCATCATCATGAATTCATGCCATCATGGGC'. The pPICZ $\alpha$  A-dghA<sub>2A</sub>R construct was expressed in X33 *Pichia pastoris*. pPICZ $\alpha$  A consists of an alcohol oxidase (AOX1) promoter region which is induced by methanol for gene expression. A stop codon was incorporated downstream the hA<sub>2A</sub>R construct to prevent incorporation of downstream c-myc and polyhistidine. A modified version of the dghA<sub>2A</sub>R, containing a C terminal linker glycine-serine (GS) and an avi-tag was made, which was termed dghA<sub>2A</sub>R-A. The cDNA sequence was obtained from Fraser (2006) and was modified in the software 'Gene designer' and was translated in 'Expasy translate'. 'Gene designer' was also used to confirm the sequence being in the correct frame upon expression. The sequence was sent to Genscript for codon optimization. The gene construct was ligated into pPICZ $\alpha$  A, using EcoRI and NotI restriction sites, by Genscript and was sequenced by the company. Human CD81 tetraspanin was also expressed in *Pichia pastoris*. The protein construct was the same construct used in Jamshad *et al.* (2008) and was provided by Hoor Ayub. Human CD81 was ligated into the pPICZB vector, which also allowed protein expression via AOX1 transduction. The CD81 construct was expressed in X33. CD81 was used as a negative control for this PhD project.

**A** **GAATTC**ATGCCCATCATGGGCTCCTCGGTGTACATCACGGTGGAGCTGGCCATTGCTGTGCTGGCCAT  
 CCTGGGCAATGTGCTGGTGTGCTGGGCCGTGTGGCTCAACAGCAACCTGCAGAACCTCACCAACTACT  
 TTGTGGTGTCACTGGCGGCGGCCGACATCGCAGTGGGTGTGCTCGCCATCCCCTTTGCCATCACCATC  
 AGCACCGGGTTCTGCGCTGCCTGCCACGGCTGCCTCTTCATTGCTTCCGTCCTCGTCCCTGCTCAGCA  
 GAGCTCCATCTTCAGTCTCCTGGCCATCGCCATTGACCGCTACATTGCCATCCGCATCCCGCTCCGGT  
 ACAATGGCTTGGTGACCGGCACGAGGGGCTAAGGGCATCATTGCCATCTGCTGGGTGCTGTCGTTTGGC  
 ATCGGCCTGACTCCCATGCTAGGTTGGAACAACCTGCGGTGACCCAAAGGAGGGCAAGCAGCACTCCA  
 GGGCTGCGGGGAGGGCCAAGTGGCCTGTCTCTTTGAGGATGTGGTCCCCATGAACTACATGGTGTACT  
 TCAACTTCTTTGCCTGTGTGCTGGTGGCCCTGCTGCTCATGCTGGGTGTCTATTTGCGGATCTTCCTG  
 GCGGCGGACGACAGCTGAAGCAGATGGAGAGCCAGCCTCTGCCGGGGAGCGGGCACGGTCCACACT  
 GCAGAAGGAGGTCCATGCTGCCAAGTCACTGGCCATCATTGTGGGGCTCTTTGCCCTCTGCTGGCTGC  
 CCCTACACATCATCAACTGCTTCACTTTCTTCTGCCCCGACTGCAGCCACGCCCTCTCTGGCTCATG  
 TACCTGGCCATCGTCTCTCCCACACCAATTCGGTTGTGAATCCCTTCATCTACGCTACCGTATCCG  
 CGAGTTCCGCCAGACCTTCCGCAAGATCATTGCGAGCCACGTCTGAGGCAGCAAGAACCTTCAAGG  
 CAGCTGGCACCAAGTGGCCGGTCTTGGCAGCTCATGGCAGTGACGGAGAGCAGGTGAGCCTCCGTCTC  
 AACGGCCACCCGCCAGGAGTGTGGGCCAACGGCAGTGTCCCCACCCTGAGCGGAGGCCAATGGCTA  
 CGCCCTGGGGCTGGTGTGAGTGGAGGGAGTGGCCAAAGAGTCCCAGGGGAACACGGGCCTCCAGACGTGG  
 AGCTCCTTAGCCATGAGCTCAAGGGAGTGTGCCAGAGCCCCCTGGCCTAGATGACCCCTGGCCAG  
 GATGGAGCAGGAGTGTCTGA**CGGGCCGC**

**B** **GAATTC**ATG**CACCATCATCACCACCATCACCATCAC****GAGAACCTGTACTTTTCAGTCT**CCAATCATGGG  
 CTCTCCGTGTATATTACTGTTGAACTGGCCATCGCTGTTTTGGCCATTCTGGGTAATGTGTTGGTTT  
 GCTGGGCTGTGTGGTTGAACTCTAATCTGCAGAACGTTACCAATTACTTCTGTGGTTTCCCTGGCTGCC  
 GCTGATATTGCTGTGGGCGTTTTGGCCATTCCCTTTGCTATTACAATCTCTACTGGCTTCTGTGCCGC  
 TTGCCATGGTTGTCTGTTTTATCGCCTGTTTTCGTGTGGTTCTGACCCAGTCTTCCATTTTCTCTCTGT  
 TGGCCATTGCTATCGACAGATACATTGCTATCAGAATCCCACTGAGATATAACGGCTTGGTTACTGGT  
 ACCAGAGCCAAGGGTATTATCGCTATCTGCTGGGTGCTGTCTTTTCGCTATTGGATTGACACCAATGCT  
 GGGCTGGAACAATGTGGTCAGCCTAAGGAAGGTAAACAGCATTCCCAGGGCTGCGGAGAGGGACAGG  
 TTGCCTGTCTGTTTGAAGATGTGGTTCCCATGAACTACATGGTGTATTTCAACTTTTTTCGCTTGCCTG  
 TTGGTTCCACTGTTGCTGATGTGGGAGTGTACCTGAGAATCTTTTTGGCCGCTAGAAGACAGCTGAA  
 ACAGATGGAGTCTCAGCCTTTGCCCGGAGAAAGAGCCAGATCCACACTGCAGAAGGAAGTGCACGCCG  
 CTAAATCTTTGGCCATTATCGTTGGCTTGTGTGCTCTGTGCTGGTTGCCTCTGCACATCATCAACTGT  
 TTCATTTCTTTTGTCCCGACTGTTCCCATGCCCATTTGTGGCTGATGTATCTGGCTATTGTTTTGTCT  
 TCACACCAACTCCGTGGTTAACCCATTTCATCTACGCTACAGAATTAGAGAGTTTTAGACAGACATTCA  
 GAAAGATTATCAGATCTCATGTGCTGAGACAGCAGGAACCTTTCAAAGCCGCTGGTACTTCCGCTAGA  
 GTGTTGGCCGCTCACGGATCTGATGGCGAGCAGGTTTCCCTGAGACTGAACGGACACCCACCTGGCGT  
 GTGGGCCAATGGATCTGCTCCACATCCTGAGAGAAGACCTAACGGTTATGCCTTGGGACTGGTTTCTG  
 GTGGATCCGCTCAGGAATCTCAGGGTAATACCGGACTGCCCGACGTGGAGTTGCTGTCCCATGAATTG  
 AAGGGAGTTTGTCCAGACCCCGAGCCCTGGATGACCCCTGGCTCAGGATGGTGTGCTGGTGTGTCT**GG**  
**CGGTGGAGGCTCCGGTGGAGGCGGTTCTGGAGGCGGTGGATCCGGTCTGAACGACATCTTCGAAGCTC**  
**AGAAAAATTGAGTGGCACGAA**TGA**CGGGCCGC**

**Figure 23** | **A**) cDNA sequence of the deglycosylated human A<sub>2A</sub>R (dghA<sub>2A</sub>R) construct, made by Fraser (2006). Text highlighted in red are the EcoRI and NotI restriction sites. The cDNA was sequenced by Dr Sarah Routledge prior to use. **B**) cDNA sequence of the deglycosylated, avi-tagged human A<sub>2A</sub>R (A<sub>2A</sub>R-A) construct. Text highlighted in red are the EcoRI and NotI restriction sites. Text highlighted in blue represents the decahistidine-tag sequence. Text highlighted in purple represents the TEV cleavage sequence. Text highlighted in green represents the 15 amino acid linker (GS) sequence. Text highlighted in yellow represents the avi-tag sequence. Sequence was confirmed by Genscript.



**Figure 24| A)** Plasmid map of the pPICZ $\alpha$  A vector. EcoRI and NotI were the restriction sites used to ligate the dghA<sub>2A</sub>R and A<sub>2A</sub>R-A cDNA sequences into the vector. The PmeI restriction site was used to linearise the vector for genome integration into *Pichia pastoris*. Zeocin resistance is encoded by the BleoR sequence of the vector. Snapgene viewer was used to illustrate the pPICZ $\alpha$  A vector. **B** and **C)** Illustrations of the A<sub>2A</sub>R constructs which were ligated into the vector backbone. EcoRI and NotI were the restriction site added to the 5' and 3' positions of the cDNA construct. His represents the deca-his-tag. GS is the 15 amino acid linker sequence composed of glycine and serine. Avi is the avi-tag. Panel B illustrates the Fraser dghA<sub>2A</sub>R construct. Panel C illustrates the A<sub>2A</sub>R-A construct.

The pcDNA3.1- T8-HA-CLR construct was provided by Dr. S. M. Foord (from Glaxo-Smith Kline, Stevenage UK). The vector pcDNA3.1- was modified to containing the T8 signal peptide, taken from the T-cell surface glycoprotein CD8 alpha chain (T8) signal peptide sequence, and the hemagglutinin (HA)-tag sequence (Wootten *et al.*, 2013). The CALCRL DNA sequence, encoding the CLR, was ligated into the vector using EcoRI and NotI restriction enzymes (Wootten *et al.*, 2013). The signal sequence of the CALCRL was removed prior to ligation in order to incorporate the upstream T8 signal peptide sequence and the HA-tag. This was to increase surface expression of the CGRPR. The T8-HA-CLR construct was used for mutagenesis experiments and was sequenced by Eurofins Genomics, using primers targeting the T7 and BGH sequences of the vector backbone. A modified version of the T8-HA-CLR, containing a C terminal strep II-tag and an avi-tag was made, which was termed CLR-A. The wild type CALCRL gene sequence was modified in the software 'Gene designer' and was translated in 'Expasy translate'. 'Gene designer' was also used to confirm the sequence being in the correct frame upon expression. The sequence was sent to Genscript for codon optimization. The gene construct was ligated into pcDNA3.1+, using EcoRI and NotI restriction sites, by Genscript and was sequenced within the company. The pcDNA3.1+-RAMP1 contained the RAMP1 cDNA sequence with an upstream N terminal FLAG-tag located within the ECD of the RAMP1.

**A**

CGGCCGCGTCGACGGAATTGCGCCACCATG**GCCTTACCAGTGACCGCCTTGCTCCTGCCGCTAGCCTT**  
**GCTGCTCCACGCCGCCAGGCCGATTACGCG**TCT**TACCCGTATGACGTCCCAGATTACGCAT**CGCTGG  
GAGGCCCTTCACTCGAGGGATCCGCA**GAA**TTAGAAGAGAGTCCTGAGGACTCAATTCAGTTGGGAGTT  
ACTAGAAAATAAAATCATGACAGCTCAATATGAATGTTACCAAAGATTATGCAAGACCCCATTCACA  
AGCAGAAGGCGTTTACTGCAACAGAACCCTGGGATGGATGGCTCTGCTGGAACGATGTTGCAGCAGGAA  
CTGAATCAATGCAGCTCTGCCCTGATTACTTTTTCAGGACTTTGATCCATCAGAAAAAGTTACAAAGATC  
TGTGACCAAGATGGAAACTGGTTTAGACATCCAGCAAGCAACAGAACATGGACAAATTATACCCAGTG  
TAATGTTAACACCCACGAGAAAAGTGAAGACTGCACTAAATTTGTTTTACCTGACCATAATTGGACACG  
GATTGTCTATTGCATCACTGCTTATCTCGCTTGGCATATTTCTTTTATTTCAAGAGCCTAAGTTGCCAA  
AGGATTACCTTACACAAAAATCTGTTCTTCTCATTTGTTTGTAACTCTGTTGTAACAATCATTACCT  
CACTGCAGTGGCCAACAACCAGGCCTTAGTAGCCACAAATCCTGTTAGTTGCAAAGTGTCCCAGTTCA  
TTCATCTTTTACCTGATGGGCTGTAATTACTTTTGGATGCTCTGTGAAGGCATTTACCTACACACACTC  
ATTGTGGTGGCCGTGTTTGCAGAGAAGCAACATTTAATGTGGTATTATTTTCTTGGCTGGGGATTTCC  
ACTGATTCCTGCTTGTATACATGCCATTGCTAGAAGCTTATATTACAATGACAATTGCTGGATCAGTT  
CTGATACCCATCTCCTCTACATTATCCATGGCCCAATTTGTGCTGCTTTACTGGTGAATCTTTTTTTC  
TTGTTAAATATTGTACGCGTTCTCATACCAAGTTAAAAGTTACACACCAAGCGGAATCCAATCTGTA  
CATGAAAGCTGTGAGAGCTACTCTTATCTTGGTGCCATGCTTGGCATGAAATTTGTGCTGATTCCAT  
GGCAGCTGAAGGAAAGATTGCAGAGGAGGTATATGACTACATCATGCATCCTTATGCATTCCAG  
GGTCTTTTGGTCTCTACCATTTTCTGCTTCTTTAATGGAGAGGTTCAAGCAATTTCTGAGAAGAACTG  
GAATCAATACAAAATCCAATTTGGAAACAGCTTTTCCAACCTCAGAAGCTCTTCGTAGTGCCTTTACA  
CAGTGTCAACAATCAGTGATGGTCCAGGTTATAGTCATGACTGTCTTAGTGAACACTTAAATGAAAA  
AGCATCCATGATATTGAAAATGTTCTCTTAAAACCAGAAAATTTATATAATTGAGAATTC

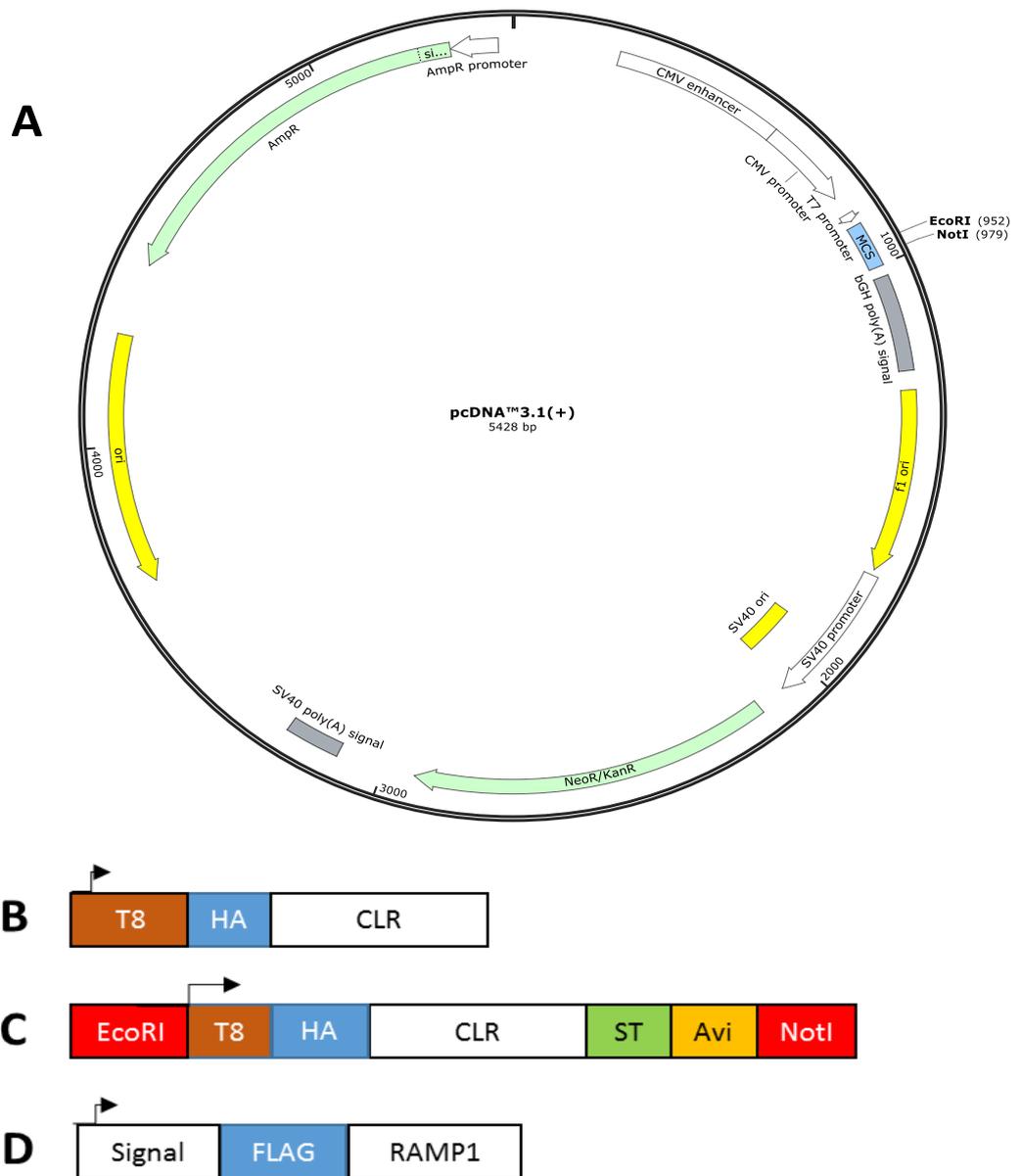
**B**

**GAATTC**AGACCCCGAAGACGGAACTGCGCTACAATG**GCTCTGCCTGTGACTGCCCTGCTGCTGCCCTT**  
**GGCTCTGCTGCTGCACGCTGCTAGGCCAGACTACGCC**TCT**TACCCCTTATGACGTGCCAGATTATGCAA**  
GCCTGGGAGGACCTTCCCTGGAGGGATCTGCCGAGCTGGAGGAGAGCCCAGAGGACTCCATCCAGCTG  
GGCGTGACCAGGAACAAGATCATGACAGCCCAGTACGAGTGCTATCAGAAGATCATGCAGGATCCCAT  
CCAGCAGGCCGAGGGCGTGTACTGCAACCGCACATGGGACGGCTGGCTGTGCTGGAATGATGTGGCAG  
CAGGAACCGAGTCCATGCAGCTGTGCCAGATTATTTCCAGGACTTTGATCCCTCTGAGAAGGTGACA  
AAGATCTGTGACCAGGATGGCAACTGGTTTCAGGCACCCTGCCTCTAACCGCACCTGGACAAAATTACAC  
CCAGTGTAACTGAAATACACACGAGAAGGTGAAGACCGCCCTGAATCTGTTTTATCTGACAATCATCG  
GACACGGCCTGAGCATCGCATCCCTGCTGATCAGCCTGGGCATCTTCTTTTACTTCAAGTCTCTGAGC  
TGCCAGAGGATCACCCCTGCACAAGAACCTGTTCTTCAGCTTCGTGTGCAATTTCTGTGGTGACCATCAT  
CCACCTGACAGCCGTGGCCAACAATCAGGCCCTGGTGGCCACAAACCCAGTGTCTTGCAAGGTGAGCC  
AGTTCATCCACCTGTACCTGATGGGCTGCAATTATTTTTGGATGCTGTGCGAGGGCATCTACCTGCAC  
ACCCTGATCGTGGTGGCCGTGTTTCGCCGAGAAGCAGCACCTGATGTGGTACTATTTTCTGGGCTGGGG  
CTTTCCCTGATCCCTGCATGCATCCACGCAATCGCACGGTCCCTGTACTATAACGACAATTGTTGGA  
TCAGCTCCGATAACCCACCTGCTGTACATCATCCACGGACCAATCTGTGCCGCCCTGCTGGTGAACCTG  
TTCTTTCTGCTGAATATCGTGAGAGTGCTGATCACCAGCTGAAGGTGACACACCAGGCCGAGAGCAA  
CCTGTATATGAAGGCCGTGCGGGCCACACTGATCCTGGTGCCTCTGCTGGGCATCGAGTTTCGTGCTGA  
TCCCATGGCGGCCCGAGGGCAAGATCGCAGAGGAGGTGTACGACTATATCATGCACATCCTGATGCAC  
TTCCAGGGCCTGCTGGTGTCCACCATCTTTTGTCTTTAATGGCGAGGTGCAGGCCATCCTGCGGAG  
AACTGGAATCAGTACAAGATCCAGTTCGGCAACTCCTTTTCTAATAGCGAGGCCCTGCGGTCCGCT  
CTTACACCGTGTCTACAATCAGCGACGGCCCCGGCTATAGCCACGATTGTCTTCCGAGCACCTGAAC  
GGCAAGAGCATCCACGATATCGAGAATGTGCTGCTGAAGCCCAGAACCTGTATAAT**TGGTCCCACCC**  
**ACAGTTTGAGAAA****GGGCTGAATGATATTTTTGAAGCACAGAAGATTGAATGGCACGAATAA****GCGGCCG**

**Figure 25** | cDNA sequence of **A**) T8-HA-CLR and **B**) CLR-A. Sequences highlighted in red indicate the EcoRI and NotI restriction sites. Sequences highlighted in green indicate the 24 amino acid cleavable T8 signal peptide. The T8 signal peptide replaced the endogenous signal peptide of the CLR. Sequences highlighted in pink indicate the hemagglutinin (HA) tag sequence. The sequence highlighted in blue represents the strep II tag sequence. The sequence in yellow is the avi-tag sequence. Sequences were confirmed by Genscript and Eurofins.

GCACGGATGCGTAACTTAAGCTTGGTACCGAGCTCGGATCCACCATGCTCGGGCCCTGTGCCGCCT  
CCCGCGGCGCGGCCTCTGGCTGCTCCTGGCCCATCACCTCTTCATGGACTACAAGGACGACGACGACA  
AGACCACTGCCTGCCAGGAGGCTAACTACGGTGCCCTCCTCCGGGAGCTCTGCCTCACCCAGTTCCAG  
GTAGACATGGAGGCCGTCGGGGAGACGCTGTGGTGTGACTGGGGCAGGACCATCAGGAGCTACAGGGA  
GCTGGCCGACTGCACCTGGCACATGGCGGAGAAGCTGGGCTGCTTCTGGCCCAATGCAGAGGTGGACA  
GGTTCTTCTGGCAGTGCATGGCCGCTACTTCAGGAGCTGCCCCATCTCAGGCAGGGCCGTGCGGGAC  
CCGCCCGGCAGCATCCTCTACCCCTTCATCGTGGTCCCCATCACGGTGACCCCTGCTGGTGACGGCACT  
GGTGGTCTGGCAGAGCAAGCGCACTGAGGGCATTGTGTAGGCGGGCGGAT

**Figure 26|** The cDNA sequence of the FLAG-tagged RAMP1. The FLAG-tag sequence is highlighted in yellow. The start codon is highlighted in green and the stop codon is highlighted in red. Sequence was confirmed by Eurofins Genomes, using primers T7 and BGH.



**Figure 27** | **A**) Plasmid map of the pcDNA3.1+ vector. EcoRI and NotI were the restriction sites used to ligate the CLR-A sequence into the vector. Ampicillin resistance is encoded by the AmpR sequence of the vector. Snapgene viewer was used to illustrate the pcDNA3.1+ vector. **B-D**) Illustrations of the constructs which were ligated into the vector backbone. T8 represents the signal peptide. HA represents the HA-tag. Avi is the avi-tag. ST is the streptavidin II tag. FLAG represents the FLAG-tag Panel B illustrates the T8-HA-CLR construct. Panel C illustrates the CLR-A construct. Panel D represents the FLAG-RAMP1 construct.

## 2.2.2: Diagnostic restriction digest using DNA electrophoresis

### Purpose of the methodology

To confirm the presence of a recombinant gene construct within a plasmid vector, a diagnostic digest can be performed to separate the gene construct from the vector using two restriction sites. The digested DNA can be separated on an agarose gel, based on the size of the DNA fragments and the electrical charge applied to the gel. The ribose-phosphate backbone of DNA gives the molecule a net negative charge, allowing the DNA fragments to migrate down the agarose gel, towards the positively charged electrode (Lee *et al.*, 2012). As the DNA mass to charge ratio is the same amongst different sized fragments, the digested DNA can be separated based on its size, where smaller fragments move faster than larger fragments (Lee *et al.*, 2012). The RedSafe nucleic acid staining solution can be applied to the gel, which binds to the DNA similarly to ethidium bromide, in order to detect the DNA fragments using excitation at 514 nm. Ethidium bromide can insert itself between adjacent nucleotide base pairs of a DNA strand, therefore migrating with the DNA fragments for visualisation. The RedSafe variant has a lower mutation rate than ethidium bromide, according to the manual supplemented by the kit.

### Procedure

A restriction digest was set up using: 1 µl of the restriction enzymes in **table 3**, 2 µl NEB buffer 3.1, 500 ng of the DNA construct. The digestion reaction was topped up to 20 µl using dH<sub>2</sub>O. The digestion was incubated at 37°C for 1 hour. After the reaction, 4 µl of the digested DNA was mixed with 12.5 µl 1 X TAE buffer and 3.5 µl 6 X bromophenol blue loading dye.

A 1% agarose gel was made by dissolving 0.5 g of Hi-res Standard Agarose in 50 ml 1 X TAE buffer. The solution was heated in a conical flask until all the agarose particles were dissolved. The agarose gel was cooled down to approximately 50°C. 2.5 µl of RedSafe was pipetted into the agarose solution, which was subsequently poured into a gel casting tray. A comb was inserted into the gel to create wells for loading the digested DNA. Once the gel was solidified and the comb was removed, the gel was placed into a benchtop electrophoresis tank. 1 X TAE buffer was used to fill the tank, submerging the agarose gel. The digested DNA: loading dye mixture and 10 µl of the DNA ladder was pipetted into the

wells of the agarose gel, which was ran at 80 volts and 60 amps until the dye reached the end of the gel. The gel was visualised in the Syngene G:Box, using UV light to highlight the positioning of the digested bands in relation to the size of the DNA ladder.

### **2.2.3: Competent *E.coli* generation and transformation.**

#### **Purpose of methodology**

In order to expand clones of DNA constructs, *E.coli* was used to generate a sufficient amount of DNA for studies involving cell transfection. *E.coli* can also be transformed with PCR products to increase the yield of DNA as well as to ligate the linear PCR products into a circular structure via DNA ligase (Adam *et al.*, 1999; Sriskanda and Shuman, 2001). Another useful feature of *E.coli* is their ability to methylate chromosomal and plasmid DNA (Marinus and Lobner-Olesen, 2014). This feature is useful for PCR-based experiments, where methylated template DNA, obtained from transformed *E.coli*, can be digested using DpnI, leaving the unmethylated PCR products intact (Li and Mullins, 2002).

Prior to transforming *E.coli* with DNA, the cells must be made competent. In this project, chemically competent *E.coli* cells were generated using a rubidium and calcium based method, similarly to Green and Rogers (2002) protocol. Both rubidium and calcium have been shown to enhance the transformation efficiency of bacteria due to their physiochemical properties affecting the membrane of *E.coli*, as well as the temperature imbalances during the process (Asif *et al.*, 2017). Once made competent, the cells were transformed with DNA constructs, where heat shocking the bacteria in a calcium rich environment can make the bacterial membrane transiently permeable for DNA uptake (Rahimzadeh *et al.*, (2016).

## Procedure

*E.coli* strain, XL Gold, was streaked onto an LB agar plate without antibiotics. The plate was left to incubate at 37°C overnight. A single colony of *E.coli* was grown overnight in a starter culture of 5 ml, in LB broth, at 37°C in a shaking incubator, shaking at 200 rpm. 1 ml of the starter culture was transferred to a flask containing 30 ml of LB at the same conditions overnight. 8 ml of the culture was transferred to a 2 L flask containing 200 ml LB and was incubated at 37°C, shaking at 200 rpm until an OD<sub>550</sub> of 0.6 was achieved. The culture was decanted into 50 ml Falcon tubes, which was chilled at 4°C for 15 minutes. The tubes were spun down at 3000 g for 20 minutes at 4°C, in a pre-cooled centrifuge. The supernatant was drained from the cell pellet, which was resuspended in transformation buffer 1 (16 mls per 50 ml Falcon tube). The cell pellet was resuspended by repeated pipetting. The tubes were incubated for 1 hour, followed by centrifugation at the same settings. The transformation buffer 1 supernatant was removed from the cell pellet. All the cell pellets were resuspended in transformation buffer 2, where the total volume of the buffer and cell pellet was 6 ml. 50 µl aliquots were made using the cell-transformation buffer 2 mixture. The aliquots were stored in -80°C until they were ready for transformation.

For transformation, the competent cell aliquots were thawed on ice. 1pg-100ng of plasmid DNA was mixed with a competent cell aliquot via repetitive pipetting. The tubes were left on ice for 30 minutes. The DNA-cell mixture was then heatshocked at 42°C for 45 seconds. The tubes were then placed on ice for 2 minutes. 1 ml of LB was added to the tubes, which were incubated at 37°C on a roller for 1 hour. The transformed bacteria were plated out onto LB agar with ampicillin to select for transformants. The transformants were grown in LB media with ampicillin at 37°C shaking at 200 rpm. The cells were then lysed for DNA extraction by following the miniprep (from Sigma Aldrich) and maxiprep (from Origene) instructions.

#### **2.2.4: Expressing the A<sub>2A</sub>R and CD81 in *Pichia Pastoris***

##### **Purpose of methodology**

To generate a high yield of recombinant protein for downstream experiments, the *P.pastoris* expression system was used. Prior to introducing exogenous DNA to *P.pastoris*, the yeast cells were made electroporation-competent by weakening the cell wall, using the methodology derived from Lin-Cereghino *et al.* (2005). The pPICZ $\alpha$ A vector containing the gene construct was also linearized at the PmeI restriction site in order for the exogenous plasmid to integrate into the genome of the yeast cells (Vogl *et al.*, 2018). Multiple cassettes of the gene construct can be integrated into *P.pastoris* to increase the yield of protein upon yeast expression in methylotrophic conditions (Vogl *et al.*, 2018). Once the exogenous plasmid has been linearised, electroporation was used to quickly insert and integrate the DNA into the yeast genome.

In order to express the protein in the yeast expression system, the cells were subjected to media containing methanol in order to induce the expression of the protein via induction of the AOX1 promoter (Chang *et al.*, 2018). After growing the cells to an adequate amount, with the recombinant protein expressed, the cells were homogenized to obtain the membrane containing the membrane proteins. The cell wall of the yeast must first be broken in order to harvest the membranes, requiring the C3 homogenizer to apply a high pressure to the cells (Routledge *et al.*, 2016). Homogenization using the C3 homogenizer has a cell wall breaking efficiency of over 90% when applying 30,000 psi of pressure (Bawa *et al.*, 2014). The membranes can subsequently be separated from cell debris through a series of centrifugation steps.

#### **2.2.4.1: Creating competent *Pichia pastoris* cells for transformation**

A 20 ml starter culture was prepared in YPD media in a baffled flask, with one colony of a *Pichia pastoris* strain, plated on YPD agar without antibiotics. The starter culture was grown overnight at 30°C in a shaking incubator, rotating at 220 rpm. 3 ml of the starter culture was incubated into 500 ml of YPD media, which was incubated at 30°C at the same rotation speed until an OD<sub>600</sub> of 0.8-1 was achieved. The culture was centrifuged at 3000 g for 10 minutes; the supernatant was removed and the cell pellet was resuspended in 100 ml of YPD + 0.02 M HEPES. 2.5 ml of 1 M DTT (filter sterilised) was applied to the culture slowly. The culture was then incubated at 30°C at 220 rpm. 400 ml of autoclaved dH<sub>2</sub>O (kept on ice) was added to the culture until 500 ml of total volume was reached. The cell culture was then centrifuged for 10 minutes at 3000 g, 4°C. The supernatant was removed and 500 ml of ice cold dH<sub>2</sub>O was used to resuspend the cell pellet. The cells were centrifuged again at the same speed, temperature, time and the supernatant was removed and replaced with 250 ml of ice cold dH<sub>2</sub>O. The culture was centrifuged again at the same conditions, however the supernatant was replaced with 20 ml of 2 M sorbitol. The sorbitol-yeast culture was centrifuged at 3000 g, 10 minutes at 4°C. The supernatant was replaced with 1 ml of sorbitol. The cell pellet was also resuspended in solution and aliquoted into sterile 1.5 ml Eppendorf tubes. The tubes were stored in -80°C.

#### **2.2.4.2: *Pichia pastoris* transformation and growth**

pPICZα A protein constructs were linearized using PmeI. The reaction was done in multiple 1.5 ml Eppendorf tubes to achieve 10000 ng of linearized plasmids for transformation. The reaction conditions for each tube was: 20 µl cutsmart buffer, 4 µl PmeI, 30 µl DNA (4000 ng of DNA), dH<sub>2</sub>O. The reaction was incubated at 37°C for one hour. The linearized plasmid was then purified using the DNA mini-spin purification kit (from Qiagen). The plasmid was then transformed into competent X33 wild type *Pichia pastoris*. 100 ng of the purified, linearized plasmid construct were mixed with 40 µl of competent yeast cells on ice. The mixture was transferred to a 0.4 mm electroporation cuvette and were left to incubate on ice for 5 minutes. The cell-DNA mixture was electroporated at 1800 v and 1 ml of cold 0.5 M sorbitol. The cell-DNA mixture was left to incubate for 1 hour at 30°C. The cells were then plated in YPD +

sorbitol plates with increasing concentrations of zeocin to select colonies with a high copy number of the plasmid.

A single colony of transformed *Pichia pastoris*, expressing the A<sub>2A</sub>R construct, was incubated in a baffled flask containing 50ml of BMGY overnight at 30°C and shaking at 220 rpm within an incubator. After overnight culture, the yeast were pelleted and reincubated in BMMY, which contained methanol to induce hA<sub>2A</sub>R expression. The incubation was at 30°C in a shaking incubator set at 220 rpm. The yeast were grown in BMMY for 48 hours where at 24 hours, methanol was added to the culture at a final concentration of 2%. Once the 48 hour growth was complete, the culture was decanted into 500ml centrifuge tubes. The *Pichia* were pelleted down and were decanted into sterile plastic bags, which were then stored in -80°C.

#### **2.2.4.3: Extracting *Pichia pastoris* membranes**

In order to extract the membrane, containing the A<sub>2A</sub>R, the cell wall had to be broken. The Avestin EmulsiFlex C3 homogenizer was used to break the cell wall using high pressure. Prior to operating the C3 homogenizer, the extracted yeast pellet was thawed. An equal weight to volume ratio of breaking buffer was added to the yeast pellet and Roche protease inhibitor cocktail tablets were applied to the sample to prevent A<sub>2A</sub>R degradation. The sample was stored and mixed in 4°C for 30 minutes. More breaking buffer was applied to the sample if the sample remained viscous. Meanwhile, the C3 was washed and primed using 70% ethanol, 0.5M sodium chloride and breaking buffer. The sample was then cracked in the C3 homogenizer for 20 minutes. The oxygen tank was set to 80 psi and the air regulator was set to 40 psi. Around 22000-25000 bar pressure was used to crack the yeast cell walls. After 20 minutes, the samples were decanted into centrifuge tubes.

The cracked *Pichia pastoris* was spun down at 5000 RPM (or 3000 g) for 20 minutes using a JA 25.50 rotor. The supernatant was decanted into Beckman coulter centrifuge tubes and were spun down at 100,000 g for 1 hour at 4°C in rotor type Ti70. The membrane was pelleted at the bottom of the tube and the supernatant was removed. The pellet was homogenized with buffer A to yield a final concentration of 40 mg/ml. The reconstituted pellet was kept in -80°C.

### **2.2.5: Mammalian cell culture and transfection**

#### **Procedure**

Human embryonic kidney (HEK) 293T cells were stored in liquid nitrogen in 90% DMEM, 10% DMSO in 500 µl aliquots. An aliquot was thawed at room temperature and was inserted into a T-25 flask with 5 ml DMEM, 10% FBS, 1 X penicillin streptomycin (complete DMEM). Once the cells were 80-90% confluent, normally taking two days to reach high confluency, the cells were passaged into new flasks at a 1:5 or a 1:10 dilution, depending on when the cells were being used for an assay. The instructions supplied in the ThermoFisher Lipofectamine 3000 kit were followed to transiently transfect the cells.

### **2.2.6: SMA copolymer preparation and membrane solubilisation**

#### **Procedure**

The SMA copolymer was prepared by hydrolysing styrene-maleic anhydride under refluxing conditions, as described in Lee *et al.* (2016). The SMA buffer was prepared using 5% (w/v) SMA, dissolved in Buffer A.

The homogenized yeast membranes were solubilised using the 5% SMA buffer at a 1:1 ratio, where the final concentration of the membrane was 40 mg/ml and the final concentration of the SMA copolymer was 2.5% (w/v). The sample was vortexed and inverted until the polymer was dissolved. The sample was mixing on a shaker at room temperature for 2 hours for SMALP solubilisation. Once the solubilisation process was complete, the sample was sonicated for three 15 second pulses. The sample was then spun down at 100,000 g for 45 minutes to pellet the membrane. The SMALP proteins remain in the supernatant due to their solubility and were separated from the un-solubilised membrane pellet.

The HEK293T cells expressing the CGRPR were homogenized using the Polytron tissue-grinder homogenizer. The cell suspension was placed on ice and homogenized at 15 second bursts, 6 times with a 10 second rest interval. After homogenization, the cell suspension was centrifuged at 100,000 g to pellet the cell membrane and the supernatant was removed. The membrane was resuspended in Buffer A, with the concentration of the membrane being 80 mg/ml. The membrane solution was solubilised

using the 5% SMA buffer at a 1:1 ratio, where the final concentration of the membrane was 40 mg/ml and the final concentration of the SMA copolymer was 2.5% (w/v). The membrane was solubilised for 2 hours at room temperature on a shaker. The sample was centrifuged at 100,000 g for 20 minutes. The SMALP supernatant was separated from the un-solubilised membrane pellet.

## **2.2.7: Protein purification, dialysis and concentration**

### **Purpose of methodology**

In order to separate the non-specific, SMA-solubilised membrane proteins from the A<sub>2A</sub>R and CD81, the histidine-tags fused to the recombinant membrane proteins were isolated using nickel-nitrilotriacetic acid agarose (Ni<sup>2+</sup>-NTA) resin. The nickel ions are captured onto the chelating nitrotriacetic acid agarose beads (Spriestersback *et al.*, 2015). The Ni<sup>2+</sup>-NTA can be applied to the SMALP samples in order to bind the deca-histidine-tagged protein SMALPs to the resin. The thin membrane of the benchtop columns can capture the protein-bound Ni<sup>2+</sup>-NTA whilst allowing the unbound and washed material to flow through the column. Although Ni<sup>2+</sup>-NTA purification is a popular purification strategy, non-specific protein binding can occur on the resin itself; however this can be eliminated by including low concentrations (10-50 mM) of imidazole in the washing buffers (Bornhorst and Falke, 2000). After washing the resin, the specific SMALPs can be eluted by increasing the concentration of imidazole above 100 mM, where imidazole competes against the his-tag to bind the nickel ions (Bornhorst and Falke, 2000).

### **Procedure**

Ni<sup>2+</sup>-NTA was used to purify the A<sub>2A</sub>R through the decahis-tag, fused to the receptor. Ni<sup>2+</sup>-NTA (from Qiagen) was supplemented in ethanol at a 1:1 ratio. Ni<sup>2+</sup>-NTA was spun down at 1000 g for 1 minute and the supernatant was aspirated off. Washing buffer was applied to the Ni<sup>2+</sup>-NTA in order to remove residual ethanol and to prime the gel beads for A<sub>2A</sub>R binding. The Ni<sup>2+</sup>-NTA was centrifuged at 1000 g for 1 minute and the washing buffer was removed. This was repeated three times. After washing, the Ni<sup>2+</sup>-NTA was applied to the yeast membrane SMALP sample, which was left to incubate at 4C on a plate shaker overnight.

The next day, benchtop purification columns were equilibrated using 2 column volumes of wash buffer and the flow-through was discarded. The Ni<sup>2+</sup>-NTA SMALP sample was poured into the purification column. The flow-through was collected and labelled as unbound. 20 column volumes of washing buffer was used to wash the Ni<sup>2+</sup>-NTA SMALP to remove non-specific Ni<sup>2+</sup>-NTA protein interaction. The flow-through was collected for analysis. 1 ml of elution buffer was used 10 times to elute the A<sub>2a</sub>R-SMALP from the nickel resin. Elutions were collected in ten 1.5ml Eppendorf tubes or one universal 25ml tube. The A<sub>2a</sub>R-SMALP was then centrifuged in vivaspin columns, molecular weight cut off (MWCO) 30000 (from GE healthcare life sciences) at 2500 RPM for 20 minutes to remove the imidazole from the sample. Dialysis buffer was applied to the A<sub>2a</sub>R-SMALP and the sample was centrifuged again at the same speed. This was repeated three times in order to remove the imidazole. After the third dialysis round, the sample was spun down to obtain the appropriate concentration of protein.

For X-ray radiolytic footprinting, the TRIS and HEPES based buffer systems were replaced by a sodium phosphate based buffer system. The dialysis buffer contained a low salt concentration in order for a successful protein irradiation and mass spectrometry.

### **2.2.8: Identifying recombinant proteins via SDS-PAGE**

#### **Purpose of methodology**

SDS-PAGE was used to detect proteins based on their molecular weight and the tags fused to them. The initial step was to make the polyacrylamide gel. The 30% Protogel acrylamide was stocked in liquid form prior to casting the gels. When preparing the stacking and separating gels, the polyacrylamide was polymerised when the redox initiators, TEMED and APS were applied to the gel mixtures (Abu-Thabit, 2017). APS initiates the polymerization reaction of acrylamide and bis-acrylamide whilst TEMED lowers the activation energy of APS decomposition for the reaction to occur at room temperature (Abu-Thabit, 2017). As the migration of proteins through the polyacrylamide gel requires an electric field, the net charge of the protein molecules must be uniform and of the same charge if they were to be separated based on size only. SDS is therefore an essential component of SDS-PAGE, where it carries a highly negative charge and can coat the proteins (Jensen, 2012). SDS also acts as an anionic detergent,

denaturing the protein tertiary structure, rendering the proteins linear (Jensen, 2012). The resulting linearised, negatively charged proteins can travel in a cathode-to-anode direction. As proteins differ in their size, they can be separated by the speed of their migration (Jensen, 2012).

Once the electrophoresis was completed, the gel can be subjected to western blotting to identify the A<sub>2A</sub>R and CD81 in the membrane and SMALP samples. The migrated proteins can be transferred onto an activated membrane made of polyvinylidene difluoride (PVDF). A primary antibody can be used to target the fusion tags of the protein and a second antibody, such as the HRP-linked secondary antibodies, can be used to target the primary antibody. The HRP enzyme can be used to catalyse the oxidation of luminol into the chemiluminescent product, which can be detected by a low-intensity light emission (Yang *et al.*, 2015) The blocking buffer is essential in western blotting to remove non-specific binding sites, which may be present on the membrane (Jensen, 2012).

InstantBlue coomassie staining can be used to detect the purity of the purified SMALP samples. The gels, post electrophoresis, were stained in InstantBlue staining solution in order to detect all the proteins in a sample. The coomassie dyes, R-250 and G-250, interact with the proteins in a polyacrylamide gel, however the exact interactions between the dye and protein is not fully elucidated. It is thought that the dyes interact with the proteins through van der Waals and hydrophobic interactions (Chevalier, 2010). Once the gel has been stained, the gel can be washed and the bands are visible without the requirement of a light source. A disadvantage of coomassie staining is its low sensitivity when compared with western blotting and silver staining (Chevalier, 2010). However the stained samples can be recovered for mass spectrometry analysis (Chevalier, 2010).

The 12% separating and 4% stacking gels were made according to table 14. The 12% separating gel was loaded into a 1.0 mm gel cassette until  $\frac{1}{4}$  away from the top. 1 ml of 100% isopropanol was added to the top of the separating gel to help settle the gel. The gel was left to set for 20 minutes. After 20 minutes, isopropanol was washed off the gel using dH<sub>2</sub>O. The 4 % stacking gel was loaded into the cassette. A 1.0 mm comb was inserted into the cassette to form the wells. The gel was left to set for 15 minutes. Meanwhile, 5  $\mu$ l of 5X Laemmli sampling buffer was added to 15  $\mu$ l of protein samples. After gel was set, the cassette was inserted into buffer tank. Running buffer was poured into the chamber of the cassette until full. Running buffer was also poured around the cassette. The comb was removed from the cassette and samples were loaded into the wells, as well as the his-ladder and protein ladder (10-180 kDa). 120V was used to run samples down the gel for approximately 1 hour.

### **2.2.8.1: Western blotting**

#### **Procedure**

Whilst samples were running down an SDS-PAGE gel, PVDF membrane (0.2mm) were dipped in 100% methanol and were then washed in transfer buffer for 30 minutes. After 1 hour of running samples, the gel was placed on top of four filter cards and a fiber pad. The PVDF membrane was placed on top of the gel and four more filter cards and another fiber pad was placed on top of the membrane to form a sandwich. The sandwich was placed in a transferring cassette, which was placed in a buffer tank. Transfer buffer was poured into the tank until full. 100 V was used to transfer proteins and ladders from the gel to the membrane for 1 hour. After 1 hour, membrane was placed in 2.5% blocking buffer overnight at 4°C. The next day, the milk buffer was removed from the membrane. Anti-his mouse IgG antibody was pipetted into fresh 2.5% blocking buffer at a 1/5000 dilution. The antibody-blocking buffer solution was poured onto the membrane, which was then placed on a shaking incubator for 1 hour at room temperature. After an hour, the antibody solution was removed from the membrane. The membrane was washed in PBS + 0.2% Tween for three minutes, three times. An anti-mouse HRP-linked goat IgG antibody was prepared in blocking buffer at 1/5000 dilution. After the final wash, PBS + Tween was removed from the membrane. The HRP-linked goat anti-mouse antibody milk solution was poured onto the membrane, which was then placed on a shaking incubator for 1 hour at room temperature. The

membrane was washed with PBS + 0.2% Tween three minutes three times. EZ-ECL solutions A and B were mixed at a 1:1 ratio and was poured onto the membrane. The membrane was shaken on a shaker for 10 minutes. Membranes were visualised in Syngene G:Box.

#### **2.2.8.2.: Instant Blue (Coomassie) staining of SDS-PAGE gel.**

##### **Procedure**

After samples ran down the SDS-PAGE gel, 20 ml of Instant Blue staining solution was pipetted onto the gel. The gel was shaken on shaker for 30 minutes. Residual Instant Blue staining solution was washed off with dH<sub>2</sub>O. The gel was visualised in Syngene G:Box.

#### **2.2.8.3: Determining the concentration of a receptor-SMALP sample**

##### **Procedure**

The concentration of a purified protein-SMALP sample was determined by electrophoresis of increasing concentrations of bovine serum albumin (BSA) (0.125-1.25 mg/ml) and 15 ul of the protein-SMALP. 5 ul of lamelli sampling buffer was added to the BSA concentrations and the protein-SMALP sample to achieve a final volume of 20 ul. The sample and the BSA concentrations was loaded into the SDS-PAGE gel, which ran at 120 v for approximately one hour. The gel was then stained with Instant Blue coomassie stain for 30 minutes. Once bands were visible, an image was taken, which was uploaded onto the software 'Image J'. The density of the BSA concentrations were analysed and a linear line of best fit was illustrated in excel.

## **2.2.9: Radioligand binding assay**

### **2.2.9.1: Competitive radioligand binding**

#### **Procedure**

Various concentrations of ZM214385 were made up in DMSO. The concentration range was between 0.001  $\mu$ M to 1 mM. 100% DMSO was also used as a control to identify maximum A<sub>2A</sub>R binding when using radioligand. The radioligand [<sup>3</sup>H]ZM214385 was made up at a concentration of 100 nM. Prior to the ligand binding reaction, adenosine deaminase was added to the A<sub>2A</sub>R sample to remove adenosine from the A<sub>2A</sub>R binding site.

For membrane radioligand binding, 5  $\mu$ l of different concentrations of ZM214385 and DMSO were pipetted into respective assay tubes. 5  $\mu$ l of 100 nM [<sup>3</sup>H]ZM214385 was pipetted into all assay tubes. 500  $\mu$ l of membrane (at 1mg/ml) containing A<sub>2A</sub>R were pipetted into each tube. The tubes were vortexed and were left to incubate at 30°C for 30 minutes in order for ligands to bind A<sub>2A</sub>R. After 30 minutes, the assay tubes were spun down at 14000RPM on a benchtop centrifuge for 7 minutes to pellet the membrane. The supernatant was washed away with water to remove unbound ligands. Assay tubes were left to air dry for 1 hour. After drying, 100  $\mu$ l of solouene was added to each assay tube to solubilise the membrane pellet overnight. The next day, assay tubes were vigorously vortexed to dislodge and homogenize the pellet. 1 ml scintillant was added to each assay tube and the solution was shaken. Assay tubes were analysed in a scintillation counter.

For soluble protein samples; experiments were done in a similar way to membrane proteins. The difference being the use of P-30 desalting columns (MWCO 30000) from Biorad to separate the unbound ligands from the protein. Columns were washed three times in dH<sub>2</sub>O and equilibrated once with binding buffer. 50-75  $\mu$ l of soluble protein sample was loaded on to the desalting columns and were spun at 1000 g for 4 minutes to remove unbound ligands. 1 ml scintillant was added to each assay tube and the solution was homogenized. Assay tubes were analysed in HP scintillation counter.

The Cheng-Prusoff equation:  $K_i = IC_{50} / 1 + ([R] / K_dR)$  was used to calculate the  $pK_i$  of ZM241385 (Cheng, 2001).  $K_i$  is the dissociation equilibrium constant of an inhibitor/antagonist, which was converted to a negative logarithm ( $pK_i$ ). The  $\log IC_{50}$  was obtained by identifying the log concentration (M) of ZM241385 to occupy half of the receptors (half of  $B_{max}$ ). [R] represents the concentration of the radioligand and  $K_dR$  represents the  $K_d$  affinity of the radioligand. 0.34 nM was the  $K_d$  of ZM241385, when binding to the  $dghA_{2A}R$  construct expressed in *Pichia pastoris* (Fraser, 2006).

### 2.2.9.2: Calculating the recovery of ZM241385-binding $dghA_{2A}R$ -SMALPs

A single point concentration radioligand binding assay was used to assess the binding in membranes and SMALPs, using 1 nM [ $^3H$ ]ZM241385 and either 10  $\mu$ M ZM241385 or DMSO, to identify total and non-specific binding. The binding methodology for membranes and SMALPs is the same as **chapter 2.2.9.1**. To calculate the recovery of ZM241385-binding  $dghA_{2A}R$ -SMALPs, the specific binding of the SMALPs was compared to the specific binding of the receptor expressed in yeast membranes. The concentration of membranes used for SMA-solubilisation or radioligand binding were kept the same, in order to assess the proportion of active receptors recovered after solubilisation. The equation in **figure 28** was used to calculate recovery.

$$\frac{\text{SMALP SB}}{\text{Membrane SB}} \times 100 = \text{Recovery (\%)}$$

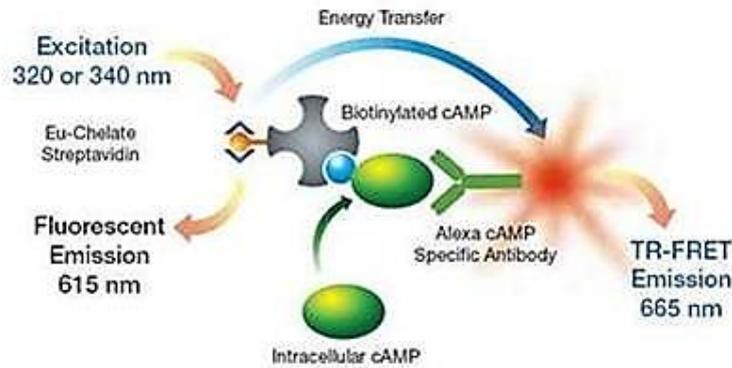
**Figure 28**| Equation to calculate the recovery of ZM241385-binding  $dghA_{2A}R$ -SMALPs after membrane solubilisation. SB stands for specific binding.

### 2.2.10: cAMP detection assay

#### Purpose of methodology

The LANCE cAMP detection kit (from Perkin Elmer) was used to determine whether the CGRPR and the CGRPR-A produced cAMP upon CGRP stimulation. This would demonstrate whether the receptors, transfected in the HEK193T cells, were functional. The kit utilises a FRET-based approach in determining the level the cAMP accumulation (Perkin Elmer [online], last accessed 23.11.2019). Alexa Fluor 647-conjugated anti-cAMP antibodies are occupied by either the intracellular cAMP produced by receptor stimulation or by the competitive Europium-labelled cAMP (Eu-cAMP) tracer complex, as shown in **figure 29** (Perkin Elmer [online], last accessed 23.11.2019). The Eu-cAMP complex is composed of the biotin-cAMP and Europium-labelled streptavidin, which form a tight complex. When the Eu-cAMP complex is in contact with the anti-cAMP fluorescent antibody, a signal will be detected. However when the intracellular cAMP levels out-compete the Eu-cAMP complex, the signal will decrease (Perkin Elmer [online], last accessed 23.11.2019). For Gs-coupled GPCRs like the CGRPR, agonist stimulation will lead to a decrease in the fluorescent signal due to intracellular cAMP occupying the antibody binding site, instead of the Eu-cAMP complex (Liang *et al.*, 2018; Perkin Elmer [online], last accessed 23.11.2019). This would suggest that upon agonist stimulation, the Gs-coupled GPCR activates adenylyl cyclase, resulting in cAMP accumulation.

## LANCE cAMP assay principle



**Figure 29** | An illustration demonstrating the principle of the LANCE cAMP detection assay. When the Eu-cAMP complex, composed of biotin-cAMP and Europium-streptavidin, occupied the Alexa Fluor 647-conjugated antibody, a fluorescent signal will be obtained upon light excitation. The Europium-labelled cAMP tracer complex can be excited at 320 or 340 nm, resulting in an energy transfer from the europium, to the Alexa-conjugated antibody. The fluorescence emission can then be detected from the antibody, using a 665 nm emission filter. Intracellular cAMP can compete against the Eu-cAMP complex for the antibody binding site, where no energy transfer occurs between the intracellular cAMP and the antibody, resulting in no fluorescence emission. Diagram taken from (Perkin Elmer [online], last accessed 23.11.2019)

### Procedure

The methodology was derived from the LANCE cAMP detection kit and Weston *et al.* (2015). The assay kit was used to analyse cAMP production in T8-HA-CLR and CLR-A transfected cells. HEK293T cells were transiently transfected in six-well cell culture plates, using RAMP1 and either T8-HA-CLR or CLR-A. Control cells transfected with empty vector pcDNA3.1+ was prepared as a control. The lipofectamine 3000 kit (from Thermofisher) was used to transfect the cells. Two days post-transfection, the LANCE cAMP 384 kit was used to detect cAMP accumulation in the cells. A standard curve was generated by following the instructions provided by the kit.  $\alpha$ CGRP was the ligand used in the assay at different concentrations. Forskolin was used to assess the maximum production of cAMP from the cells of each condition, acting as a positive control and to normalize the data. The white 384-well Optiplate was used to analyse cAMP accumulation, using the Mithras plate reader with a 340 nm excitation filter and a 665 nm emission filter.

### **2.2.11: Fluorescence correlation spectroscopy**

#### **Purpose of methodology**

FCS is a developing technique to study ligand-receptor interactions at a single molecule level and has recently been used to study pharmacodynamics in SMALPs (Kilpatrick and Hill, 2016). Traditional binding techniques, such as competitive radioligand binding, often study a population of receptors within a sample, where receptors can exist in multiple conformations and can oligomerize, impacting the pharmacological profile of the receptor. (Kilpatrick and Hill, 2016). FCS avoids the issue of studying a population of receptors by observing the pharmacology of a single ligand binding to a single receptor. The technique is also useful to study the pharmacology of receptors in a SMALP context, such as looking at the lipophilicity of a ligand and its non-specific interaction with the lipid bilayer.

The procedure was in the optimization stage whilst studying CA200645 binding to the A<sub>2A</sub>R-SMALP, therefore the methodology will include the priming of the confocal microscope and related parameters to study the diffusion of SMALPs. The confocal microscope was placed on a shock absorbent desk and the room temperature was controlled, to prevent movement of the microscope and a change in temperature influencing the assay. The microscope was aligned and calibrated using the Cyanine-5 dye, which emitted a similar fluorescent wavelength as the CA200645 ligand. After the microscope was optimised for single molecule detection, the SMALPs were loaded into the coverslip, with the addition of the ligands. The data was recorded in real-time to observe the changes in the displacement of the CA200645 fluorescent ligand by ZM241385.

#### **Procedure**

Prior to FCS analysis, the room was set at a constant temperature of 25°C. The confocal microscope was calibrated prior to experimentation in order to align the machine to detect a specific detection volume. Stocks of the Cyanine-5 (Cy5) (from Amersham Biosciences) dye were made at 10 mM, (for calibration) and 500 nM (for alignment) concentrations, using 1 X HBSS buffer. The 633 laser of the ZEISS LSM510 Confocor2 confocal microscope was used to analyse the Cy5 fluorophore and the fluorescent ligand, CA200645. A 500 nM concentration of the Cy5 dye was made up using the 10 mM

stock solution with 1 X HBSS, in a final volume of 500  $\mu$ l and 200  $\mu$ l of the diluted dye was pipetted onto a coverslip, within a Nunc Lab-Tek 8-well chambered cover-glass (from Thermo Scientific, UK). The coverslip was 150 microns thick, so the bottom and the top of the coverslip was detected using the reflection of the laser beam path and the confocal volume was positions 200  $\mu$ m above the coverslip surface. A 10 nM concentration of the Cy5 dye was made using the 10 mM stock concentration, in 1 X HBSS in a total volume of 500  $\mu$ l. 200  $\mu$ l of the 10 nM Cy5 was pipetted onto a coverslip within a glass chamber and was used to adjust the parameters for FCS. The beam path was finely adjusted until 30000-40000 photons was reached, representing counts per molecule. The structural translation parameter was checked to see if the value was 7. Once the microscopy was calibrated and adjusted, 180  $\mu$ l of the SMALP samples were pipetted onto a coverslip within a glass chamber. 20  $\mu$ l of the 500 mM stock CA200645 was pipetted onto the SMALP samples, which was then measured using the adjusted parameters for FCS. The cold ligand was added at an equal final concentration to observe displacement of CA200645. The Zeiss AIM 4.2 software was used to analyse the data.

### **2.2.12: The biotinylation and immobilisation of avi-tagged SMALPs**

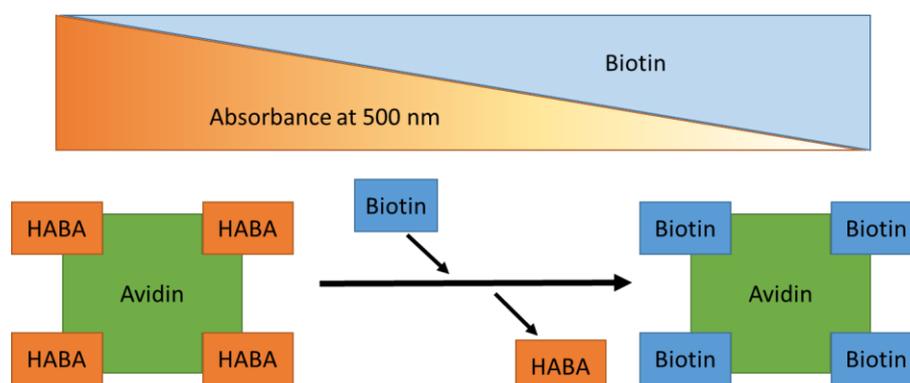
#### **Purpose of methodology**

In order to produce antibody-based drugs using phage display on SMA-solubilised GPCRs, the SMALPs must be immobilised onto an ELISA plate to behave as a target antigen. NUNC maxisorp plates are routinely used in UCB to generate phage libraries against target antigens. The plates consist of a polystyrene surface, which is highly charged, having a high affinity for polar and hydrophilic proteins (ThermoFisher Scientific [online] last accessed: 25.11.2019). SMALPs were directly immobilised onto the NUNC maxisorp plates for phage display, however the immobilisation of the SMALPs was inefficient. In order to solve this problem, avi-tagged GPCR-SMALPs were made. Biotin ligase was used to biotinylated the avi-tag of the GPCR-SMALPs, in order to immobilise them onto streptavidin-coated NUNC maxisorp plates (Li and Sousa, 2012). The BirA500 biotinylation kit from Avidity was used for avi-tag biotinylation, however the buffers provided by the kit were replaced by buffers which were suitable for GPCR-SMALP biotinylation, without causing SMA precipitation. The Biomix B component consisted of 100 mM ATP, 100 mM MgOAc and 500  $\mu$ M D-Biotin. Solutions of each

individual component of Biomix B were prepared so the magnesium concentration can be reduced and the ATP and Biotin content can be increased.

Once the samples were biotinylated, ELISA assays were conducted to determine whether the biotinylated A<sub>2A</sub>R-A-SMALP can interact with streptavidin and to see whether the biotinylation of the A<sub>2A</sub>R-A-SMALP affected its immobilisation onto streptavidin-coated NUNC maxisorp plates. For the pNPP assay, the alkaline phosphatase-conjugated streptavidin catalyses the colourless pNPP into p-nitrophenol, which produces a yellow colour, and a phosphate (Behera *et al.*, 2017). For the OPD assays, the protein-SMALPs were immobilised onto streptavidin-coated ELISA plates, with the primary and secondary antibodies targeting the protein. The HRP component of the secondary antibodies oxidised the OPD, in the presence of H<sub>2</sub>O<sub>2</sub> (Cao *et al.*, 2017). The oxidised OPD produced a detectable yellow colour.

The HABA/avidin biotin quantitation assay was used to determine how many biotin molecules were present per protein molecule. HABA and avidin form a complex, which forms an orange colour. In the presence of biotin, the HABA molecules dissociate from avidin, which forms a complex with biotin. The displacement of HABA reduces the absorbance at 500 nm and produces a yellow colour (**figure 30**). The lower the absorbance, the more biotin molecules are present to outcompete HABA for avidin binding (Gonzalez *et al.*, 2016).



**Figure 30** Principle of the HABA/avidin biotin quantitation assay. The HABA component of the HABA/avidin complex is displaced by biotin, reducing the absorbance of the solution at 500 nm. The reduction in absorbance correlates with the presence of biotin in the sample.

### **2.2.12.1: Biotinylation of receptor-SMALPs**

#### **Procedure**

The Bir500 biotinylation kit was from Avidity. The instructions for biotinylation were performed with slight alterations to the biotinylation reaction buffer, as shown in **table 16**. Once the purified SMALP sample was pipetted into the biotinylation reaction buffer, the sample was left to incubate at 4°C, overnight.

### **2.2.12.2: Indirect ELISA to determine receptor-SMALP biotinylation and streptavidin-biotin interaction using pNPP.**

#### **Procedure**

For streptavidin binding detection, A<sub>2A</sub>R-A-SMALP-biotin, myelin basic protein-biotin (MBP) and myelin basic protein-non-biotinylation (MBP-non-biotinylation) were diluted in dilution buffer (PBS + 0.15 µl/ml BSA) to achieve various concentrations. MBP was loaded onto an NUNC maxisorb 96-well plate, with an amount of 1-45 ng loaded into the wells. 45 ng of non-biotinylated MBP was loaded into empty wells and served as a negative control. 5-200 ng of A<sub>2A</sub>R-A-SMALP was loaded into empty wells. The total volume in each well was 50 µl. The proteins were absorbed into the plate for 1 hour with gentle rotational shaking. Plates were subsequently washed four times with 200 µl of PBST (PBS + 0.05% tween). The plate was then incubated with 200 µl of blocking buffer (PBS + 40 µg/ml BSA) at room temperature. The plate was then washed four times with PBST. 50 µl of streptavidin-alkaline phosphatase (40 µg/ml) was pipetted into each well with sample and two empty wells as a background control. The plate was left to incubate at room temperature for 1 hour with gentle shaking. The plate was again washed four times with PBST. The plate was then washed twice with TBS (10 mM Tris, pH 7.5, 150 mM NaCl). 200 µl of p-nitrophenyl phosphate (pNPP) developing solution was applied to each well and incubated for 15 minutes at room temperature. Absorption was monitored at 405 nm.

### **2.2.12.3: Biotin quantitation using HABA/avidin.**

#### **Procedure**

The A<sub>2A</sub>R-A-SMALP, CD81-SMALP, MBP-biotin and MBP were diluted to 100 µg/ml, prior to biotin quantitation. 1 ml of HABA/avidin solution was pipetted into a cuvette and was read at A<sub>500</sub>. 100 µl of each protein sample was pipetted into cuvettes containing 900 µl of the HABA/avidin solution and was mixed via inversion. The cuvettes were read at A<sub>500</sub>. The HABA calculator by Thermofisher Scientific was used to calculate the ratio of biotin per mole of protein, using the absorbance readings of HABA/avidin and HABA/avidin/protein samples.

### **2.2.13: Nanobody generation using phage display**

#### **Purpose of methodology**

The A<sub>2A</sub>R-SMALP was used as a drug discovery platform to identify A<sub>2A</sub>R-specific monoclonal nanobodies using phage display. The filamentous M13 bacteriophage, containing the nanobody genes in its phagemid, was used to generate A<sub>2A</sub>R-specific nanobodies (Ledsgaard *et al.*, 2018). The general procedure involved the identification of target-specific phage by incubating the phage with the A<sub>2A</sub>R-SMALPs, eluting the phage and expanding monoclonal phage from colonies of bacteria. An ELISA test was used to identify the phage specific to the A<sub>2A</sub>R.

For the phage display selection biopanning rounds, the A<sub>2A</sub>R-SMALPs and the Fab positive control were immobilised onto the NUNC maxisorp ELISA plates. The M13 phage library expressing the naïve VHH were incubated with 1,2-dimyristoyl-*sn*-glycero-3-phosphorylcholine (DMPC)-SMALPs, which was used for negative pre-selection to isolate the lipid and SMA binding phage. The phage-DMPC-SMALP mixture was added to the immobilised A<sub>2A</sub>R-SMALP in order for the VHH expressed on the phage to bind the epitopes of the A<sub>2A</sub>R. The plate was washed to remove non-specific phage from the assay (Ledsgaard *et al.*, 2018). The acidic elution buffer, containing 0.1 M HCL, was used to elute the phage from the A<sub>2A</sub>R-SMALP by breaking the interaction between the target-specific phage and the A<sub>2A</sub>R (Wu *et al.*, 2016; Ledsgaard *et al.*, 2018). For phage expansion, the eluted phage were transfected into phage-

competent TG1 cells *E.coli*, which can tolerate multiple freeze-thaw cycles (Carmen and Jermutus, 2002). Once the phagemid has been inserted into the TG1 *E.coli*, the helper phage were transfected into the bacteria. The helper phage contained the complete M13 genome which was essential for phage assembly, capsid production, chromosome replication and budding (Ledsgaard *et al.*, 2018). Once the A<sub>2A</sub>R-specific phage was expanded, PEG (a non-denaturing water-soluble polymer), was used to precipitate the phage (Castro-Mejia *et al.*, 2015; Ingham, 1984). The process was repeated two more times to remove the non-specific and background binding phage from the A<sub>2A</sub>R-specific phage (Ledsgaard *et al.*, 2018).

### **2.2.13.1: Generating A<sub>2A</sub>R-specific phage using phage display**

#### **Procedure**

To generate the A<sub>2A</sub>R specific phage, three rounds of biopanning was used to select the phage, as shown in **figure 52** in **chapter 5**. In day 1, 50 µl of approximately 5 µg/ml of target protein (A<sub>2A</sub>R) was pipetted into 8 wells of a NUNC Maxisorb ELISA plate. Fab fragments at approximately 5 µg/ml were also pipetted into another 8 wells of the ELISA plate to serve as a positive control. The plate was incubated overnight at 4°C. In day 2, the supernatant was removed from the immobilised target proteins on the ELISA plate. 200 µl of blocking buffer was added to each well containing target proteins. The ELISA plate was incubated for 1 hour at room temperature. A naïve llama VHH expressing bacteriophage (containing ~10<sup>12</sup>phages) was removed from -80°C and left to thaw at room temperature. The phage stock was mixed with DMPC-SMALPs at a 3 (phage):1 (DMPC-SMALPs) ratio. The DMPC-SMALP served as negative selection to remove SMALP-binding phage. The phage-DMPC-SMALP mixture was diluted in dried milk buffer at a 1:1 ratio. 10 ml of TG1 bacteria was prepared as a starter culture in 2TY media, which was incubated at 37°C in a shaking incubator. After 1 hour incubation, the ELISA plate was washed 4 times with wash buffer. 50 µl of the phage mixture was pipetted onto each well containing target protein and the plate was incubated for another hour at room temperature on a shaker. Once the TG1 starter culture reached an OD<sub>600</sub> of 0.4, 2 ml of the culture was transferred to a flask containing 50ml 2TY media, which was then incubated at 37°C in a shaking incubator. The ELISA plate was washed 4 times in wash buffer. The plate was then dried by tapping it onto tissue paper. 50µl of elution

buffer was added to each well containing target solution. The plate was left to incubate for 10 minutes at room temperature. The supernatant from each well containing a target protein was transferred to 400  $\mu$ l of the neutralisation buffer. Once the TG1 culture reached an OD<sub>600</sub> of 0.4-0.6, the neutralised phage was transferred to 10ml of the TG1 culture. The culture was incubated at 37°C without shaking. 200  $\mu$ l of culture was used for Phage titering. The rest of the culture was spun down for 10 min at 3500 RPM. The supernatant was removed from the bacterial pellet. The pellet was spread on a 15cm plate containing 2YT carbenicillin + 1% glucose agar plates. The plate was incubated at 30°C. On day 3, 5 ml of 2YT + 15% glycerol was added to the 15 cm plate. Cells were scraped from the plate and transferred to cryovials. The vials were dipped in liquid nitrogen and stored at -80°C, which concluded the first round of biopanning.

The second and third rounds of biopanning was similar to the first round. The differences were: the phage library being more specific to the immobilised target proteins, the increase in wash cycles from 4 to 12 to 20 and the use of helper phage, incubated with the stock phage mixture to help form the A<sub>2A</sub>R specific phage when extracted from TG1 bacteria. At the start of the second and third rounds, an aliquot of bacteria + phage was thawed and diluted in 11 ml of 2TY + 200  $\mu$ l carbenicillin + 1% glucose until an OD<sub>600</sub> of 1 was achieved. The culture was then grown at 37°C in a shaking incubator until the culture reached an OD<sub>600</sub> of 0.4. 100  $\mu$ l of helper phage (containing  $\sim 10^{12}$  phages) was added to the culture to rescue the bacterial clones. The culture was incubated for 60 minutes at 37°C without shaking, followed by spinning the culture for 10 minutes at 3500RPM to pellet the bacteria. The pellet was resuspended in 10 ml 2TY + 200  $\mu$ l carbenicillin + 200  $\mu$ l kanamycin (both antibiotics from 1000X stock). The culture was grown in a shaking incubator overnight at 30°C. The next day, the cells were centrifuged for 15 minutes at 4700RPM. The supernatant was transferred to a 50 ml Falcon tube and 2.5ml of PEG solution was applied to the supernatant, which was left on ice with shaking at 15 minute intervals. The supernatant was centrifuged at 4700 rpm for 20 minutes at 4°C to pellet the precipitated phage.

### **2.2.13.2: Phage tittering**

During the phage display protocol, 200 µl of TG1-phage culture was used for phage tittering. The 200 µl of culture was pipetted into two wells of a 96-well plate. A serial dilution using 180 µl of 2TY media and 20 µl of culture was created in the empty well, with a total of 6 dilutions. The dilutions were done in duplicate. 10 µl of the dilutions were spotted onto an LB agar plate containing carbenicillin + 1% glucose. The spots were left to dry and grow overnight at 30°C. The next day, colonies were counted to identify the efficiency of TG1 transfection with A<sub>2A</sub>R-binding phage.

### **2.2.13.3: Detecting phage specificity to A<sub>2A</sub>R using phage ELISA**

After the third phage display biopanning round, spots of bacteria were plated for phage tittering. A single colony from the plate was streaked onto another LB agar plate (carbenicillin + 1% glucose), which was incubated overnight at 37°C.

The next day, 1 ml of 2TY (carbenicillin + 1% glucose) was pipetted into each well of a 96 well 2 ml plate. Single colonies were picked from the LB plate and placed into wells of the 96 well 2ml plate. Each well therefore contained a single clone of TG1 bacteria, harbouring monoclonal phage. A breathable foam lid was placed onto the 96 well 2ml plate, which was then stored in a 37°C shaking incubator for 4 hours. 100 µl of culture from each well was pipetted into a new 96 well 2 ml plate. A 1/10 dilution of helper phage was prepared in 2TY media (carbenicillin + 1% glucose). 50 µl of diluted helper phage was pipetted into each well containing 100 µl culture. The plate was incubated at 37°C for 1 hour, followed by spinning the plate at 3500RPM for 7 minutes. The supernatant was removed from each well of the plate and 100 µl of 2TY media (carbenicillin + 1% glucose) was pipetted into each well. The 96 well 2ml plate was placed on a shaker at room temperature at high speed to resuspend cell pellet. 800 µl of 2TY media (carbenicillin + 1% glucose) was applied to each well. The plate was incubated at 37°C shaking overnight. A NUNC Maxisorb 96 well ELISA plate was prepared with 50 µl of either target solution (A<sub>2A</sub>R-SMALP or Fab fragment), DMPC-SMALP and PBS for control. The ELISA plate was left to incubate at 4°C overnight.

The next day, the 96 well 2 ml plate which was left to incubate overnight was spun down at 4700RPM for 7 minutes. The supernatant from wells containing the immobilised A<sub>2A</sub>R-SMALP, DMPC-SMALP and control was replaced with 200 µl blocking buffer. The ELISA plate was placed on a shaker at room temperature for 1 hour. After the 2 ml plates were centrifuged, 200 µl of the supernatant containing the phage were placed in new 96 well 2 ml ELISA plates and 200 µl of milk buffer was applied to each well. Plates containing the phage were placed on a shaker for 1 hour at room temperature. The ELISA plates were washed four times with 200 µl of wash buffer. 50 µl of phage solution from the 2 ml plates were pipetted to wells containing either A<sub>2A</sub>R-SMALP, DMPC-SMALP or control wells. For the positive control, A<sub>2A</sub>R-SMALP was replaced by Fab and phage against Fab was pipetted instead of Phage against A<sub>2A</sub>R. The ELISA plate was placed on a plate shaker for 1 hour at room temperature. 1/5000 dilution of anti-phage antibody was prepared in blocking buffer. The ELISA plate was washed four times with 200 µl of wash buffer and was tapped to dry. 50µl of diluted anti-phage antibody was pipetted into each well of the ELISA plate, which was then placed on a plate shaker for 1 hour at room temperature. The supernatant was aspirated and the plate was washed again with 200 µl of wash buffer, four times. 100 µl of 1 step ultra XMB-ELISA substrate was applied to each well and the plate was read in a plate reader.

#### **2.2.14: Photoaffinity cross-linking F-CGRP to CGRPR ECL mutants**

##### **Purpose of methodology**

The photoaffinity cross-linking assay was used to map the residues of the extracellular loops of the CGRPR which interact with the CGRP. The residues of ECL1 and ECL3 were mutated to AzF, which is a photo-crosslinkable unnatural amino acid. The fluorescent CGRP was used to bind the mutant receptors to detect cross-linking after UV exposure.

Quick-change PCR was used to generate the AzF CLR mutants. The pCDNA3.1- -T8-HA-CLR construct was used as a template to generate the mutants. The primers were complementary to the template, with exception of the site of mutation, which was replaced by the amber stop codon. The Pfu Hotstart master mix contained all the components necessary for the PCR reaction, including: PfuUltra II Fusion HS DNA polymerase, PCR buffer, magnesium and dNTPs. The quick-change PCR protocol was used to synthesize the mutants. DpnI digested the methylated template DNA, so the mutant PCR products remain.

The mutant CLR constructs were co-transfected with the RAMP1 construct to express mutant CGRPR in HEK293T cells. The F-CGRP was cross-linked to the receptors using UV light. The cells were spun down and washed three times to remove the F-CGRP which has not formed cross-links with the receptors. DDM was then used to solubilise the receptors from the cell membrane. The solubilised sample was analysed to assess the residues with which the fluorescent ligand formed cross-links.

### **2.2.14.1: Primer design and mutagenesis of the T8-HA-CLR construct using quick-change PCR**

#### **Procedure**

The FASTA DNA sequence of the CLR construct was obtained from Uniprot. The CALCRL gene encodes the CLR and was used to create mutations. Primers were 35 bases long and a forward and reverse primer sequence was generated (**Appendix**). A single amber stop codon (tag) mutation was inserted into the primer, whilst the rest of the sequence complements the CALCRL DNA. The website 'Netprimer' was used to calculate the melting temperature ( $t_m$ ) of the primer, where the average  $t_m$  was between 75-85°C. The site 'Expasy translate' tool was used to confirm the amino acid sequence of the primers. The web tool 'reverse complement' was used to generate the reverse primer, using the forward primer as a template. Primers were ordered from 'Eurofins genomics'. The lyophilised primers were reconstituted to 100 pmol/ $\mu$ l. Complementary forward and reverse primers were mixed and diluted with dH<sub>2</sub>O, where each primer was at 10 pmol/ $\mu$ l.

Quick-change PCR was used to create CLR mutants with the amber stop codon. The reaction conditions were: 25  $\mu$ l pfu ultra II hotstart mastermix, 2  $\mu$ l 10 pmol/ $\mu$ l primer mix, 1  $\mu$ l pCDNA3.1-CLR template (at 30 ng/ $\mu$ l) and 22  $\mu$ l dH<sub>2</sub>O. The PCR tube containing the reaction ingredients were placed in a PCR machine. The reaction was proceeded with the instructions: 1) 95°C for 30 seconds (melting), 2) 95°C for 30 seconds (melting), 3) 55°C for 1 minute (annealing), 4) 68°C for 7 minutes (extension), 5) 4°C until ready to use. Steps 2-4 were repeated for 16 cycles. Once the reaction was complete, 1  $\mu$ l *dpnI* was used to digest methylated template DNA. 1  $\mu$ l of the PCR product was used for bacterial transformation to confirm *dpnI* digestion (no colonies on an ampicillin LB agar plate) and presence of mutant constructs (colonies on ampicillin LB agar plate). The transformed bacteria were grown and lysed to obtain a high concentration of plasmid construct, which were then sequenced by Eurofins Genomics using primers: T7 and BGH.

### 2.2.14.2: Photoaffinity cross-linking of F-CGRP to the mutant CGRPR

#### Procedure

HEK293T cells were seeded into six-well plates (from NUNC) at  $8 \times 10^5$  cells per well and were incubated at 37°C, 5% CO<sub>2</sub> overnight. The cells were approximately 70% confluent the next day. The lipofectamine 3000 transfection kit was used to transfect the HEK293T cells according to the instructions providing by the kit. The DNA mixture used to transfect the cells consisted of: 0.5 µg pCDNA3.1+- RAMP1, 0.5 µg pCDNA3.1-CLR, pSVBpUC-Yam (the suppressor tRNA, recognising the 'tag' amber stop codon), pcDNA3.1+-RS-V1 (the aminoacyl-tRNA synthetase for AzF incorporation). The total volume of the transfection mixture was 200 µl, which was left to incubate for 15 minutes. 800 µl of opti-MEM was pipetted into the transfection mixture to achieve a total volume of 1 ml.

Once the transfection was completed, p-azido-L-phenylalanine (AzF) was incorporated into the receptor structure, as shown in the schematic in **Chapter 8, figure 60** (Nodling *et al.*, 2019). The complete DMEM was removed from the adherent HEK293T cells and the 1 ml transfection mixture was added to the wells. AzF was prepared in a 25 mM stock. Initially, the lyophilised AzF was dissolved in 0.2 M NaOH, which was added to DMEM, 1 X penicillin streptomycin. Droplets of 20% (v/v) HCL, dH<sub>2</sub>O was added to the soluble AzF until pH 7.4 was reached. The 25 mM AzF was diluted in DMEM, 20% FBS, 1 X penicillin streptomycin, to achieve a 1 mM AzF concentration. 1 ml of 1 mM AzF was added to each well of the six-well plates containing the transfected HEK293T cells, which were subsequently left overnight at 37°C, 5% CO<sub>2</sub> overnight.  $10^{-8}$  M  $\alpha$ CGRP-carboxyfluorescein (F-CGRP) was prepared by diluting the stock F-CGRP ( $10^{-5}$  M) in binding buffer. The supernatant from the transfected HEK293T cells was removed from the wells, where each well was washed using 1 X HBSS, followed by aspiration of the washing solution. 1 ml of  $10^{-5}$  M F-CGRP was added to each well and the cells were incubated at 37°C, 5% CO<sub>2</sub> for 15 minutes. The cells were then placed on ice and a scraper was used to resuspend the cells in the binding buffer. 450 µl of cells were placed in 1.5 ml tubes, wrapped in aluminium foil and were termed the control conditions. 450 µl of cells were also placed in tubes, however they were exposed to ultraviolet (UV) light (254 nm wavelength) for crosslinking. The control

and assay tubes were placed in an ice slurry bath and were exposed to UV light for 30 minutes. The tubes were then centrifuged at 3000 g for 10 minutes to pellet the cells, where the supernatant was aspirated. The cell pellet was resuspended in PBS via vortexing, and were centrifuged using the same conditions. This was repeated three times. In the final centrifugation process, the supernatant was drained so the cell pellet only remained. The pellet was solubilised using 100 µl DDM solubilisation buffer. The cells were left on ice for 1 hour. The solubilised cells were centrifuged at 15000 g for 15 minutes at 4°C. The supernatant was pipetted into black ELISA plates. The Mithras plate reader was used to read fluorescein absorbance, using a 485 nm excitation filter and a 535 nm emission filter.

### **2.2.15. Statistical analysis**

The analysis of variance (ANOVA) tests were used to compare the means of three or more groups of data (Kim., 2017). The one-way ANOVA test compares the means of three or more independent groups with only one independent variable (Mishra et al., 2019). The two-way ANOVA test is an extension of the one-way ANOVA test, in which an additional independent variable is being analysed (Mishra et al., 2019). Once a significant difference is observed by the ANOVA tests, a post-hoc test is used to compare the means. The Tukey's multiple comparisons test compares the mean of each group with every other mean (Graphpad [online] last accessed: 19.03.2020). The Dunnett's multiple comparison test compares a control mean with every other mean (Graphpad [online] last accessed: 19.03.2020). The Sidak's multiple comparisons test functions similarly to the Tukey's test albeit more powerful in making independent comparisons and was used alongside the two-way ANOVA test (Graphpad [online] last accessed: 19.03.2020).

Non-linear regression was used to analyse competitive radioligand binding data. Regression analysis is the correlation between a dependent (Y) variable and independent variables. Linear regression compares the two variables with the  $y = mx + c$  equation, normally generating a straight line. Non-linear regression assumes that the dependent variable is random, with the line being curved and is useful to determine the  $B_{max}$  and  $K_d$  (Motulsky and Neubig, 2002).

## Chapter 3: Solubilisation of recombinant membrane proteins from *Pichia pastoris* using SMA.

### 3.1: Introduction

Yeast cell expression systems have been favoured by scientists due to the low cost of growing cells, which can be grown in large cultures and can generate a high yield of recombinant protein (Byrne, 2015). The methylotrophic yeast species, *Pichia pastoris*, can use methanol as a carbon source and contains eukaryotic post-translational modification machinery, making it a useful alternative to mammalian cells (Byrne, 2015). The alcohol oxidase (AOX1) promoter present in the genome of *Pichia* can be induced by methanol, in absence of the inhibitory glucose and glycerol (Byrne, 2015). The induction process can be taken advantage of, where plasmid-protein constructs containing the AOX1 promoter can be integrated into the genome of *Pichia*. The pPICZ series of plasmids are often used as a vector for recombinant protein expression, as the plasmids contain the AOX1 promoter sequence upstream of the multiple cloning site. *Pichia pastoris* has successfully expressed over 100 GPCRs and can yield double the amount of receptors when compared with *E. coli*, making the yeast strain an effective cell expression system for GPCRs. The deglycosylated A<sub>2A</sub>R construct has been successfully expressed in the wild type *Pichia* strain, X33, retaining its pharmacological characteristics and has been studied in SMALPs (Fraser, 2006; Jamshad *et al.*, 2015).

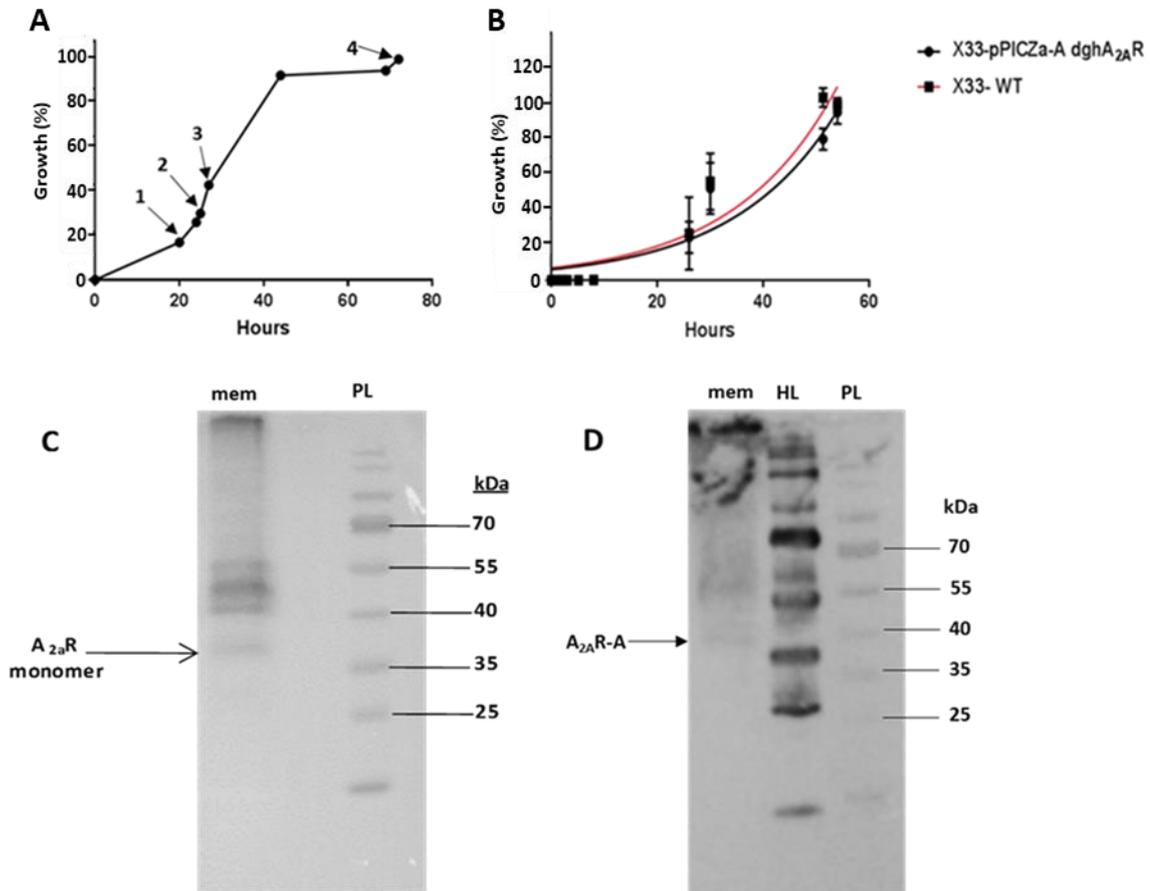
The SMA copolymer has been used to solubilise the deglycosylated human A<sub>2A</sub>R and the A<sub>2A</sub>R-A, where previous research with the A<sub>2A</sub>R has demonstrated successful solubilisation of binding-capable receptors (Jamshad *et al.*, 2015; Wheatley *et al.*, 2016). The SMA copolymer has been studied instead of detergents due to its ability to solubilise receptors in a nanodisc, with the native lipid bilayer intact (Dorr *et al.*, 2016). This is important for downstream applications, particular experiments involving drug discovery. Receptors need to be in a stable and functional conformation, which lipids have been proven to stabilise, in order to discover drugs that bind an active target (Neale *et al.*, 2015). Recent molecular dynamic simulations and mass spectrometry analysis have shown phospholipid interaction sites on the A<sub>2A</sub>R, which stabilise particular states of the receptor (Song *et al.*, 2019; Yen *et al.*, 2018).

Replacing the lipids with detergent micelles will alter the allostereism and structural stability the lipids provide to the receptors and can potentially distort the GPCR structure (Wheatley *et al.*, 2016; Dorr *et al.*, 2016; Song *et al.*, 2019). Within this chapter, the expression, solubilisation and purification of the A<sub>2A</sub>R-SMALP was confirmed, using molecular techniques. The confirmation of the presence of the receptor within a SMALP is essential for downstream applications.

## **3.2: Results**

### **3.2.1: Growth of X33 yeast expressing A<sub>2A</sub>R.**

The growth of the X33 *Pichia pastoris* strain was analysed using a spectrophotometer, measuring light absorbance, to assess the growth rate of the yeast in glycerol and methanol conditions. Yeast expressing the recombinant A<sub>2A</sub>R construct was compared with X33, containing no recombinant gene construct, in order to assess differences between their growth rates.



**Figure 31| X33 *Pichia pastoris* cells expressed the dghA<sub>2A</sub>R and the A<sub>2A</sub>R-A pPICZαA constructs.** **A)** A line graph illustrating the percentage growth of dghA<sub>2A</sub>R-expressing X33 cells over time in media (10% yeast extract, 20% peptone and 0.02% biotin). 8 Samples were taken from the culture for OD analysis. Arrow 1 represents the start of the glycerol feed (BMGY), arrow 2 indicates the stoppage of glycerol feed arrow 3 represents start of methanol feed (BMMY) and arrow 4 indicates the point of harvest. The OD<sub>595</sub> values for the samples were converted to percentages, with the highest growth point being 100%. **B)** Exponential growth curve comparing the growth of X33-pPICZα-A dghA<sub>2A</sub>R cells and wild type X33 cells. The cells were cultured in BMGY for 54 hours at 30°C in baffled flasks. The OD<sub>595</sub> readings for each sample was converted to a percentage value, relative to the maximum growth point (100%) of the X33 yeast strain. Experiment was repeated three times. Error bars represent the standard error of the mean. **C** and **D)** Western blot showing the expression of: **C)** the dghA<sub>2A</sub>R and **D)** the A<sub>2A</sub>R-A, in X33 cells. 15 μl of an 80 mg/ml dghA<sub>2A</sub>R-expressing X33 membrane was mixed with 5 μl LSB and was loaded into a well. Anti-his mouse IgG bound the his-tag of the dghA<sub>2A</sub>R. HRP-linked goat anti-mouse IgG bound the anti-his mouse IgG. EZ-ECL was the chemoluminescent substrate for the HRP-linked goat anti-mouse IgG. ‘mem’ was the 80 mg/ml sample of X33-dghA<sub>2A</sub>R membranes, ‘PL’ was the Pageruler protein ladder and ‘HL’ was the histidine ladder. The blots are representative of three different experiments.

X33 cells were transfected with the pPICZ $\alpha$ -A vector containing the dghA<sub>2A</sub>R construct, using electroporation at 1800V. Cells were grown in a bioreactor at 30°C for 72 hours (figure 31A). Cells were initially growing slowly between 0 and 20 hours. Upon glycerol induction, the cells grew rapidly as they used glycerol as a carbon source. Cells were still growing rapidly after stopping the glycerol feed for one hour. At 24 hours, cells were fed with methanol to induce the expression of dghA<sub>2A</sub>R, where methanol activates the AOX1 promoter upstream the dghA<sub>2A</sub>R construct. The growth rate decreased slightly, when compared to the glycerol feed, however cells were still growing rapidly. At around 42 hours onwards, the cells stopped growing (figure 31A). The exponential growth rate of X33-pPICZ $\alpha$ -A dghA<sub>2A</sub>R cells was compared with X33 wild type cells, which were grown in BMGY at 30°C in baffled flasks. Growth rate of X33-pPICZ $\alpha$ -A dghA<sub>2A</sub>R was slightly slower than X33 wild type cells however both cells types continued to grow (figure 31B). Western blotting was used to confirm the presence of the dghA<sub>2A</sub>R in X33-pPICZ $\alpha$ -A dghA<sub>2A</sub>R cell membranes (figure 33C) and the A<sub>2A</sub>R-A in X33 pPICZ $\alpha$ -A A<sub>2A</sub>R-A cell membranes (figure 31D). For both membrane receptor samples, a band was located between 35-40 kDa, which was close to the 34 kDa band observed in Fraser *et al.* (2006). However multiple bands were detected above 40 kDa. The his molecular ladder was used to confirm the viability of the primary and secondary antibodies used for the western blot.

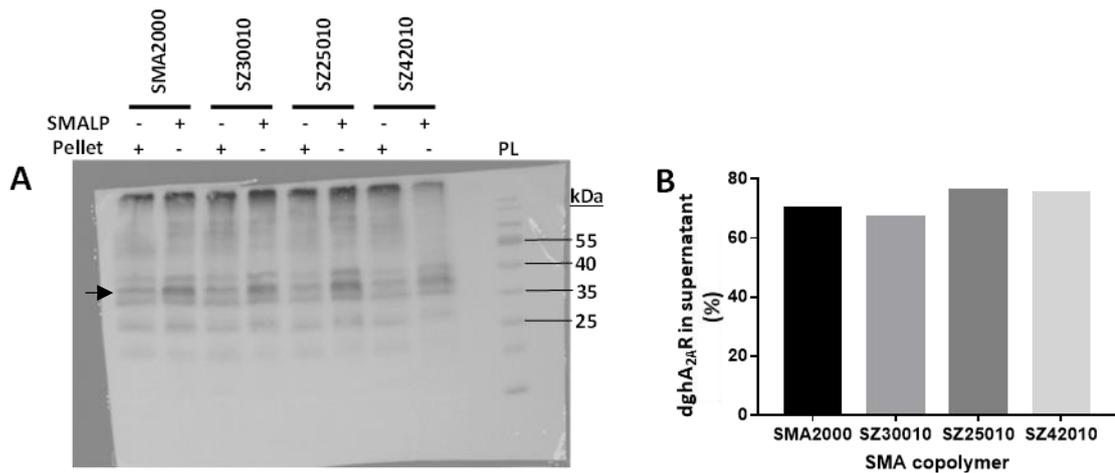
### 3.2.2: The use of SMA to solubilise A<sub>2A</sub>R receptors from the *P. pastoris* membranes

Different SMA copolymers were tested to select the copolymer which solubilised the most receptors from an 80 mg/ml sample of membranes expressing A<sub>2A</sub>R. The SMA copolymers differ from each other through their styrene to maleic acid ratio, which affects the diameter of the SMALP as shown in **table 23**. The four copolymers were selected due to their availability at the time. SMA2000 was previously used with the A<sub>2A</sub>R and has demonstrated successful solubilisation of the receptor as well as retaining ligand-receptor binding (Jamshad *et al.*, 2015). SMA2000 is made of a 2:1 ratio of styrene to maleic acid and forms a 10 nm diameter nanodisc upon membrane solubilisation (Stroud *et al.*, 2018). From literature, Xiran SZ30010 achieved similar membrane protein solubilisation, yield of purified protein and protein thermostability to SMA2000 (Gulamhussein *et al.*, 2019). The Xiran SZ25010 copolymer demonstrated successful solubilisation, although the purified SMALPs showed less stability than

SMA2000 and SZ30010, according to previous research (Gulamhussein *et al.*, 2019). SZ25010 consists of a 3:1 styrene to maleic acid ratio, which can cause a decrease in the strength of the acidic component of the SMA, slightly increasing the pH tolerance of the copolymer (Scheidelaar *et al.*, 2016). The Xiran SZ42010 contains a higher percentage of maleic acid, 42%, than SMA2000, which has a maleic acid content of 33% (Stroud *et al.*, 2018; Morrison *et al.*, 2016). The solubilisation efficiency of the Xiran SZ42010 copolymer was not as efficient as SMA2000, which may be due to the higher acidity content, preventing lipid insertion of the copolymer (Morrison *et al.*, 2016).

<b>SMA copolymer</b>	<b>Mw (g/M)</b>	<b>Disc diameter (nm)</b>	<b>Styrene: maleic acid ratio</b>
SMA2000 (from Cray Valley)	7500	10	2:1
Xiran SZ30010 (from Polyscope)	6500/ 10000	N/K	2.3:1
Xiran SZ25010 (from Polyscope)	10000	N/K	3:1
Xiran SZ42010 (from Polyscope)	10000	N/K	N/K

**Table 23** | The SMA copolymers used to solubilise A<sub>2A</sub>R receptors from yeast membranes and their properties (Stroud *et al.*, 2018; Gulamhussein *et al.*, 2019). N/K means ‘not known’.

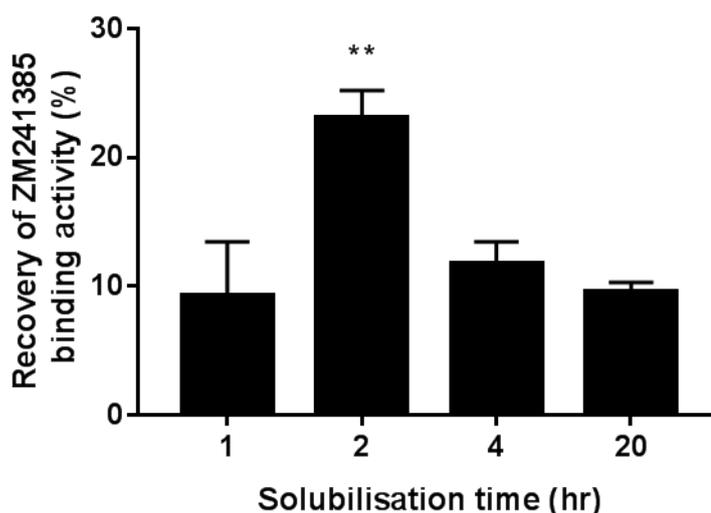


**Figure 32| The solubilisation efficiency of different SMA copolymers used to solubilise dghA<sub>2A</sub>R-expressing yeast membranes.** 40 mg/ml X33 membranes expressing the dghA<sub>2A</sub>R was solubilised for 20 hours at room temperature, using 2.5% (w/v) of SMA copolymers: SMA2000, SZ30010, SZ25010, SZ42010. The resulting SMALPs were separated from the unsolubilised pellet by centrifugation at 100,000 g. The supernatant was extracted and the pellet was suspended in buffer A at the same volume of SMA buffer used for solubilisation. **A)** Western blot showing the density of bands of SMALPs and the resuspended membrane pellet. 15 µl of the SMALP and pellet samples were mixed with 5 µl LSB and were loaded into an SDS-PAGE gel for electrophoresis and western blotting. Anti-his mouse IgG bound the his-tag of the dghA<sub>2A</sub>R. HRP-linked goat anti-mouse IgG bound the anti-his mouse IgG. EZ-ECL was the chemiluminescent substrate for the HRP-linked goat anti-mouse IgG. The arrow points to the position of the dghA<sub>2A</sub>R. The blot represents one experiment. **B)** Bar chart illustrating solubilisation efficiency of each copolymer by using densitometry analysis of panel A. The band between the 35-40 kDa region, representing the dghA<sub>2A</sub>R, was compared between all the copolymers. ImageJ was used to obtain the arbitrary density values of each band. The percentage of dghA<sub>2A</sub>R in the supernatant reflects the density of the bands in panel A. Graph represents one experiment.

Multiple bands were detected for each copolymer condition, however a band between 35-40 kDa was the densest, representing the dghA<sub>2A</sub>R (**figure 32A**). For the SMALP conditions for each copolymer, the bands were denser than the bands in the pellet condition. The programme, ImageJ, was used to detect the density values of each dghA<sub>2A</sub>R band. A densitometry bar chart was made using the data in **figure 32A**, where the solubilisation efficiency of each polymer was similar (**figure 32B**).

### 3.2.3: Recovery of binding-capable A<sub>2A</sub>R-A after solubilisation at different time periods

The solubilisation times for the best recovery of binding-capable receptors was investigated, using the SMA2000 copolymer. A single point, saturation radioligand binding assay was used to detect binding of the ligand, [<sup>3</sup>H]ZM241385, to the A<sub>2A</sub>R-A after solubilisation. **Figure 33** shows the time point where the most pharmacologically active receptors could be obtained from the yeast membrane.

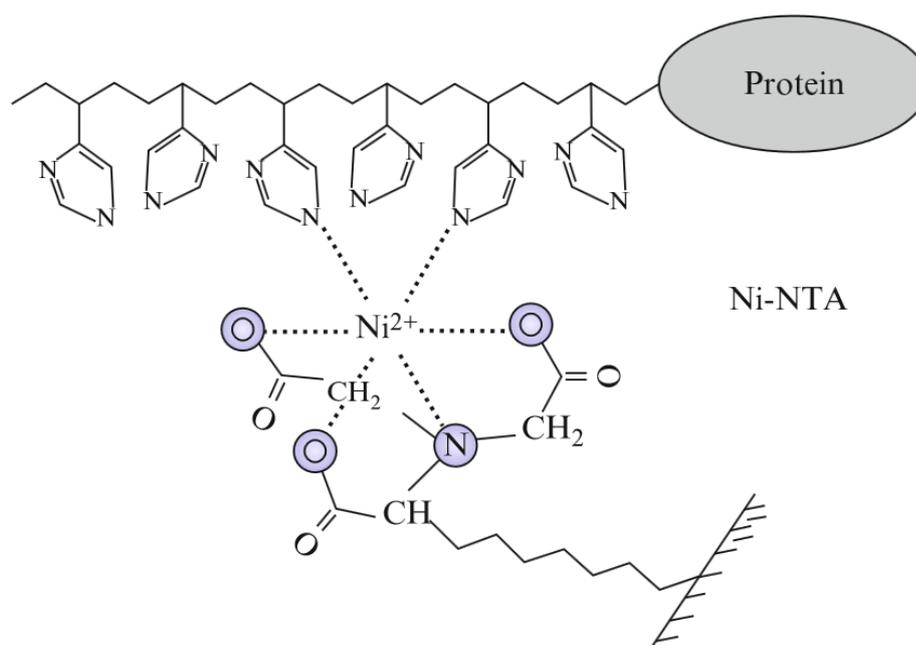


**Figure 33| Solubilising membranes for 2 hours using SMA2000 yielded the most pharmacologically stable A<sub>2A</sub>R-A-SMALPs.** Histogram displaying the percentage of binding-capable A<sub>2A</sub>R-A-SMALPs recovered from the A<sub>2A</sub>R-expressing X33 membranes after SMA solubilisation. 1 nM of the radioligand, [<sup>3</sup>H]ZM241385 and 10<sup>-5</sup> M of ZM241385 were used to calculate specific binding. The percentage values were obtained by dividing the cpm of the A<sub>2A</sub>R-A-SMALPs by the cpm observed from the A<sub>2A</sub>R-A-membranes prior to solubilisation, followed by converting the values to percentage (**chapter 2.2.9.2**). The geometric mean of three independent experiments was calculated from the binding data. The error bars represent the standard error of the mean. The one-way ANOVA statistical test, followed by Tukey's multiple comparison test was used to compare the differences between the solubilisation times (\*\*, p < 0.01).

X33 membranes expressing A<sub>2A</sub>R-A were solubilised for 1 hour, 2 hours, 4 hours and 20 hours, using SMA2000. The receptors in the membrane and the SMALPs were pharmacologically tested, using the radioligand binding assay. [<sup>3</sup>H]ZM241385 was used as the radioligand. To assess total binding of the radioligand, the membrane and SMALP samples were incubated with the radioligand and DMSO. To assess the non-specific binding of the membranes and SMALPs, the samples were incubated with the radioligand and the cold ligand, ZM241385. The specific binding values of the SMALPs were converted into a percentage value, based on the proportion of binding-capable receptors solubilised from the 40 mg/ml membrane sample, using the methodology is **chapter 2.2.9.2**. The percentage recovery was compared between the solubilisation time periods, in order to identify the most efficient solubilisation time length (**figure 33**). The 2 hour solubilisation time demonstrated a significantly higher percentage recovery of active receptors between the four conditions (**figure 33**). The percentage recovery of solubilisation for 2 hours was two-fold higher than the 1 hour, 4 hours and 20 hours solubilisation times. There was no significant different in the recovery of active receptors between the 1 hour, 4 hours and 20 hours solubilisation time, where the percentage recovery was around 10% (**figure 33**). The receptor occupancy equation:  $[L] / ([L] + K_d)$  was used to assess whether the experiment was conducted in saturating conditions. The K<sub>d</sub> of [<sup>3</sup>H]ZM241385 is 0.34 nM, which is 75% occupancy of the receptors.

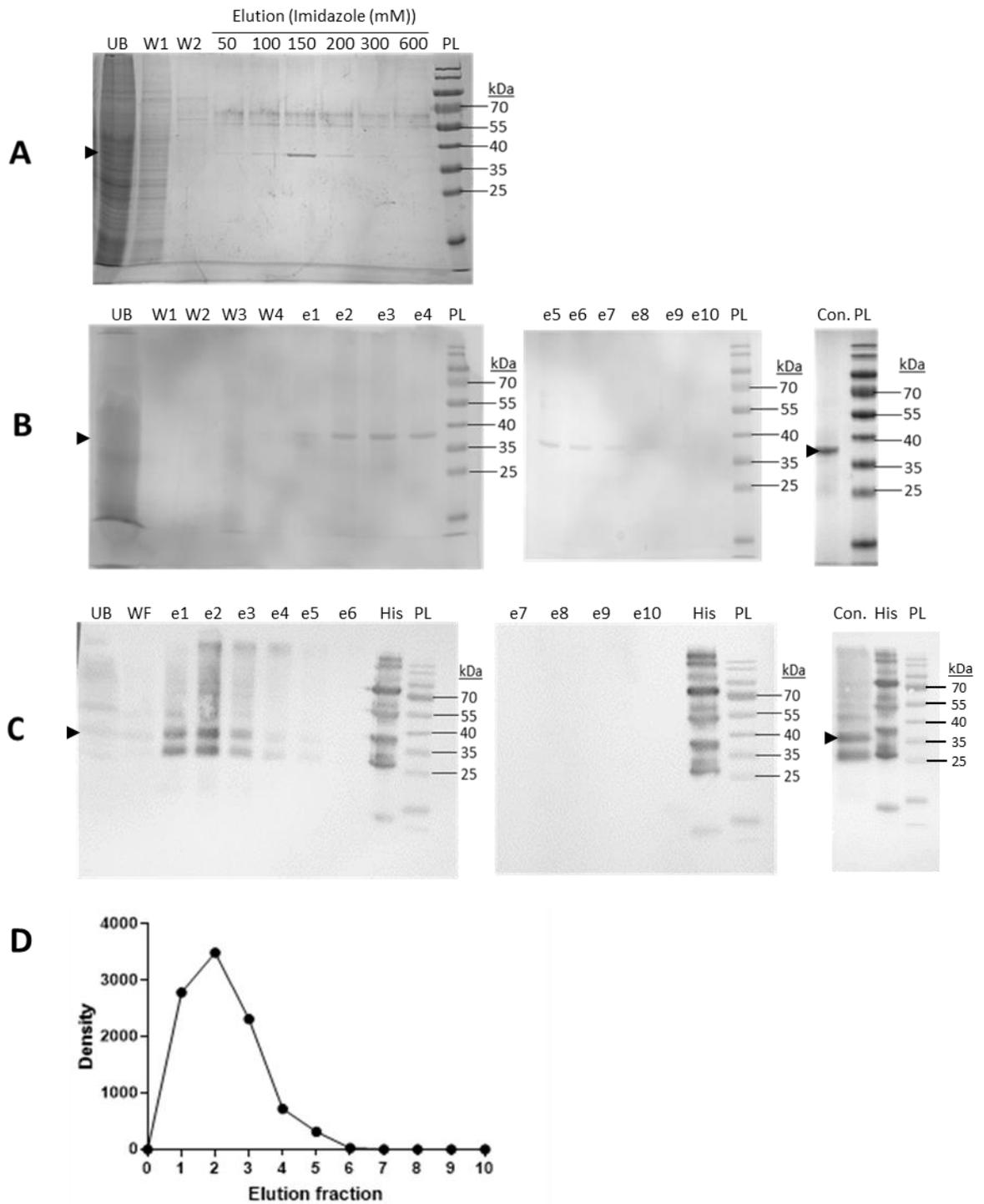
### 3.2.4. Overview of Ni<sup>2+</sup>-NTA purification of histidine-tagged proteins

The purification of the his-tagged receptors in SMALPs was achieved using the Ni<sup>2+</sup>-NTA affinity resin (from Qiagen). Histidine tags have been a popular choice for protein purification as the hexahistidine tag is 0.84 kDa in size, is uncharged at pH 7.4, it does not interfere with the folding of a protein and has low immunogenicity (Spriestersbach *et al.*, 2015). The purification procedure was done through binding the protein-SMALPs to the Ni<sup>2+</sup>-NTA, followed by a series of washes to remove non-specifically bound protein and elution of target SMALPs. Ni<sup>2+</sup>-NTA resins have been widely used to purify his-tagged proteins (Spriestersbach *et al.*, 2015). NTA consists of 4 metal-chelating sites, which stably bind Ni<sup>2+</sup> (Spriestersbach *et al.*, 2015). The his-tag sequence of a fusion protein can bind the Ni<sup>2+</sup> portion of Ni<sup>2+</sup>-NTA with high affinity (**figure 34**) (Block *et al.*, 2009).



**Figure 34** | Illustration demonstrating the interaction between a his-tagged protein and the Ni<sup>2+</sup>-NTA resin. Diagram taken from (Block *et al.*, 2009).

### 3.2.4.1. The purification of SMA solubilised membrane proteins; dghA2aR, A2aR-A and CD81



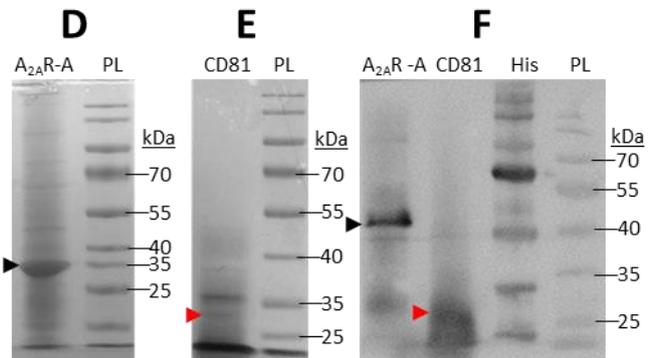
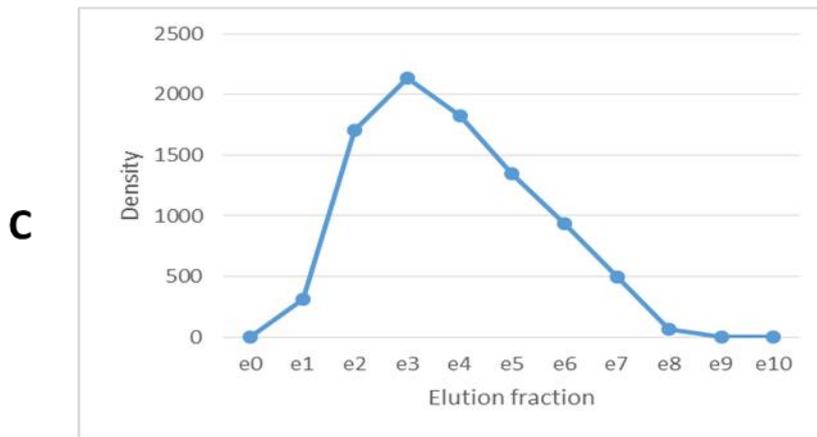
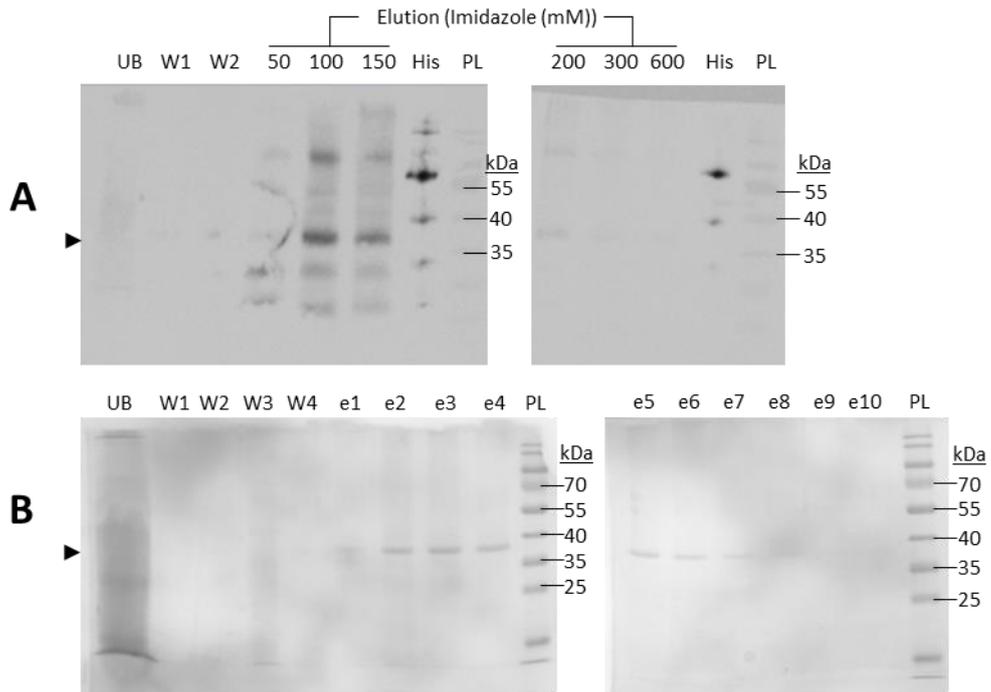
**Figure 35| dghA<sub>2A</sub>R-SMALP purification from contaminating SMALPs.** 40 mg/ml of the X33-dghA<sub>2A</sub>R cell membranes was solubilised for 20 hours using 2.5% (w/v) SMA2000, at room temperature. Membranes were spun down at 100,000 g to separate the membranes from SMALPs. The A<sub>2A</sub>R-SMALPs were purified using nickel affinity chromatography **A)** Instant Blue staining of an SDS-PAGE gel showing the elution of purified dghA<sub>2A</sub>R-SMALPs, using different imidazole concentrations (50-600 mM) in elution buffer. Unbound flow through (UB) and wash flow throughs (W1 and W2) were analysed. 2 column volumes of wash buffer, containing 20 mM imidazole, were used to remove non-specific protein. PL represents the protein ladder. Blot is representative of one experiment. **B)** Instant Blue staining of an SDS-PAGE gel showing the elution of purified dghA<sub>2A</sub>R-SMALPs. Unbound flowthrough (UB) was analysed. 20 column volumes of wash buffer (~300 ml), containing 20 mM imidazole, was used to remove non-specific protein. Each wash (W1, W2, W3 and W4) represents 5 column volumes of washes. 1ml elution buffer containing 250mM imidazole was used 10 times to elute the A<sub>2A</sub>R-SMALP (e1-e10). Conc. represents the purified, concentrated and dialysed dghA<sub>2A</sub>R-SMALP sample, using the 10 collated elution fractions, which were concentrated and dialysed to remove imidazole in Vivaspin 20 centrifugal concentrators. Gel is representative of two experiments. **C)** Western blot demonstrating the presence of dghA<sub>2A</sub>R-SMALPs in the 10, 1 ml elution fractions (e1-e10) and after concentration and dialysis of the pooled elution fractions. The Elution buffer contained 250mM imidazole. Anti-his mouse IgG was used to label the his-tag, fused to the dghA<sub>2A</sub>R. HRP-linked goat anti-mouse IgG was used to label the anti-his mouse IgG. EZ-ECL solutions A and B were mixed together at 1:1 ratio and was used as the substrate for the HRP attached to the goat antibody. UB was the unbound flowthrough, WF was the wash flowthrough, His was the molecular histidine marker and PL was the protein ladder. Blot is representative of two independent experiments. The arrowheads on panels **A-C** indicate the position of the dghA<sub>2A</sub>R. **D)** A densitometry graph based on panel C, illustrating the elution peak when fractions e1-e10 were used to elute the target A<sub>2A</sub>R-SMALP. ImageJ was used to obtain the arbitrary density values of the bands at the 35-40 kDa region of panel C.

The nickel affinity resin was used to purify the dghA<sub>2A</sub>R-SMALPs from contaminants in the SMALP sample (**figure 35**). 40mg/ml of dghA<sub>2A</sub>R-expressing X33 yeast membranes were solubilised using 2.5% (w/v) SMA2000. Solubilisation was done for 18 hours followed by sonication. Membranes were separated from SMALPs by centrifuging at 100,000 g. 1 ml of nickel resin (without ethanol) was used per 10 ml of SMALP sample. SMALPs were incubated with the nickel-NTA resin for 24 hours at 4°C. The 10 ml of nickel resin-SMALP sample was loaded into one benchtop chromatography column to remove all contaminating protein during chromatography as well as to elute the dghA<sub>2A</sub>R-SMALP, with the flow rate being determined by gravity. Initially, the imidazole concentration was optimised to elute as much of the SMALPs from the resin (**figure 35A**). The flow through, containing unbound protein, was obtained and analysed using SDS-PAGE and Instant Blue Coomassie staining, revealing multiple bands which could not be distinguished in the SDS-PAGE gel (UB) (**figure 35A**). The wash buffer, containing 20mM imidazole, was used to wash away any non-specific proteins which may have bound to the columns. The flow through was collected and analysed after each wash cycle. From the first wash, multiple bands were detected in the SDS-PAGE, which decreased after the second wash. After washing, the elution buffer containing imidazole was used to break the interaction between the his-tagged receptor from the nickel resin. Elution buffers with different imidazole concentrations were used to elute the dghA<sub>2A</sub>R-SMALP, where the 150mM imidazole elution buffer eluted the highest amount of receptor-SMALPs (**figure 35A**).

Once, the optimal imidazole concentration was identified, the yield of eluted SMALPs was analysed from each elution cycle, where an increase in wash cycles was used to remove all the non-specific protein bound to the column during purification (**figure 35B**). Again, multiple bands which could not be distinguished were detected in the unbound flow through lane of the coomassie stained SDS-PAGE gel. No clear bands were detected in the wash lanes, W1-W4. Once the resin was washed, the dghA<sub>2A</sub>R-SMALP was eluted, with the flow rate determined by gravity. The elution buffer containing 250mM imidazole was used to elute the dghA<sub>2A</sub>R-SMALP where 1 ml of elution buffer was pipetted onto the resin, ten times where ten fractions were collected and analysed (e1-e10). From the first elution, a band could be detected at the 35-40 kDa region. From the second to fifth elution, the bands reached their

highest density. After the fifth elution the density of the bands decreased. No bands were detected in elution fractions e8-e10 (**figure 35B**). After collating the elutions, the pooled sample was concentrated and dialysed to replace the elution buffer with the dialysis buffer containing no imidazole. The coomassie stained gel showed one dense band at the 35-40 kDa region (**figure 35B**).

The samples used for the coomassie gel in **figure 35B** were ran through an SDS-PAGE gel for western blotting (**figure 35C**). Western blotting detected the presence of the his-tagged dghA<sub>2A</sub>R. A faint band at the 35-40 kDa region was detected in the unbound and wash lanes of the blot. However the bands became dense at the start of eluting the dghA<sub>2A</sub>R-SMALP (**figure 35C**). The densest band was detected at the second elution lane (e2). The density of the band at 35-40 kDa declined after the second elution (**figure 35C**). No bands were visible after elution 5 (e6-e10). There was also a band located underneath the 35 kDa marker which was not present in the coomassie stained SDS-PAGE gels (**figure 35C**). The band representing the dghA<sub>2A</sub>R was still present after western blotting when the elution fractions were collated, dialysed and concentrated (**figure 35C**). A his-ladder was used to confirm whether the anti-his mouse IgG can detect polyhistidine.



**Figure 36| A<sub>2A</sub>R-A-SMALP and CD81-SMALP purification from contaminating SMALPs.** 40 mg/ml of the X33-A<sub>2A</sub>R-A and X33-CD81 cell membranes was solubilised for 2 hours at room temperature, using 2.5% (w/v) SMA2000. The membranes were spun down at 100,000 g to separate them from the SMALPs. The A<sub>2A</sub>R-A-SMALPs and CD81-SMALPs were purified using nickel affinity chromatography **A)** Western blot, showing the elution of purified A<sub>2A</sub>R-A-SMALPs, using different imidazole concentrations (50-600 mM) in elution buffer. Unbound flow through (UB) and wash flow throughs (W1 and W2) were analysed. 20 column volumes of wash buffer, containing 10 mM imidazole, were used to remove non-specific protein, where W1 is a mixture of wash cycles 1-5 and W2 is a mixture of wash cycles 6-10. PL represents the protein ladder. His represents the molecular his ladder. Anti-his mouse IgG was used to label the his-tagged A<sub>2A</sub>R-A. HRP-linked goat anti-mouse IgG was used to label the anti-his mouse IgG. EZ-ECL solutions A and B were mixed together at 1:1 ratio and was used as the substrate for the HRP attached to the goat antibody. Blot is representative of one experiment. **B)** Instant Blue coomassie staining of an SDS-PAGE gel showing the elution of purified A<sub>2A</sub>R-SMALPs. Unbound flowthrough (UB) was analysed. 20 column volumes of wash buffer (~300 ml), containing 10 mM imidazole, was used to remove non-specific protein. Each wash (W1, W2, W3 and W4) represents 5 column volumes of washes. 1ml elution buffer containing 250mM imidazole was used 10 times to elute the A<sub>2A</sub>R-A-SMALP (e1-e10). PL represents the protein ladder. Gel is representative of one experiment. **C)** A densitometry graph based on panel B, illustrating the elution peak when fractions e1-e10 were used to elute the target A<sub>2A</sub>R-A-SMALP. ImageJ was used to analyse the density of the bands at the 35-40 kDa region of panel C. **D and E)** Instant Blue coomassie staining of SDS-PAGE gels, illustrating the presence of the purified A<sub>2A</sub>R-A-SMALP and CD81-SMALP, which were concentrated and dialysed using Vivaspin 20 centrifugal concentrators. PL represents the protein ladder. Gels are representative of three independent experiments. **F)** Western blot comparing the presence of purified, concentrated and dialysed A<sub>2A</sub>R-A-SMALPs and CD81-SMALPs. Anti-his mouse IgG was used to label the his-tag, fused to the A<sub>2A</sub>R-A and CD81. HRP-linked goat anti-mouse IgG was used to label the anti-his mouse IgG. EZ-ECL solutions A and B were mixed together at 1:1 ratio and was used as the substrate for the HRP attached to the goat antibody. His was the molecular histidine marker and PL was the protein ladder. Blot is representative of three independent experiments. For all gels, the black arrowhead indicates the band representing the A<sub>2A</sub>R-A and the red arrowhead indicates the band representing the CD81.

The nickel affinity resin was used to purify the A<sub>2A</sub>R-A-SMALP and the CD81-SMALP from contaminants in the SMALP samples (**figure 36**). CD81 was SMA-solubilised, purified and analysed as a negative control to confirm the expression of the A<sub>2A</sub>R constructs in A<sub>2A</sub>R-transformed yeast. 40mg/ml of X33 yeast membrane expressing either A<sub>2A</sub>R-A or CD81 were solubilised using 2.5% (w/v) SMA2000. Solubilisation was done for 2 hours followed and the membranes were separated from SMALPs by centrifugation at 100,000 g. 1 ml of nickel resin (without ethanol) was used per 10 ml of SMALP sample. SMALPs were incubated with the Ni<sup>2+</sup>-NTA resin for 24 hours at 4°C. The 10 ml of nickel resin-SMALP sample to wash and elute the protein-of-interest, with the flow rate being determined by gravity. 20 column volumes of wash buffer was used to wash the resin, where non-

specific proteins were removed from the resin, which can be seen as an undistinguishable smear in W3 (**figure 36B**). Western blot analysis showed no band representing the A<sub>2A</sub>R-A (35-40 kDa) being present in the first ten wash cycles (**figure 36A**), therefore only the non-specific protein was eluted from the column during the wash cycles. Initially, different concentrations of imidazole (50-600 mM) were used to identify the best concentration to elute the A<sub>2A</sub>R-A-SMALP from the resin. The 100 mM imidazole elution buffer demonstrated the highest yield of A<sub>2A</sub>R-A-SMALP from the nickel resin column, followed by the 150 mM concentration elution buffer, as shown by the density of the bands representing the A<sub>2A</sub>R-A (**figure 36A**). The bands became fainter as higher concentrations of imidazole were applied to the column after a 100 mM concentration. 250 mM imidazole elution buffer was used as the standard elution buffer for further purification processes. After washing the resin, 10 ml of elution buffer was used to elute the A<sub>2A</sub>R-A-SMALP, which was collected in 1 ml fractions (e1-e10), as seen in the coomassie gel (**figure 36B**) and its complementary densitometry graph (**figure 36C**). The peak elution was in fraction 3, with nothing seen after fraction 8 (**figure 36B and C**). No additional bands, besides the band representing the A<sub>2A</sub>R-A, could be seen in the coomassie gel (**figure 36B**).

After purifying the A<sub>2A</sub>R-A-SMALP and collating the 10 elution fractions, the sample was concentrated and dialysed using the dialysis buffer. This was also done with the CD81-SMALPs, which have been purified the same way as the A<sub>2A</sub>R-A-SMALP, using the same buffer system. The SMALP samples were loaded into an SDS-PAGE gel for electrophoresis. The gel was stained with Instant Blue to detect the purified A<sub>2A</sub>R-A and CD81 bands (**figure 36D and 36E**). A dense band was located between 35-40 kDa, which indicated the presence of the A<sub>2A</sub>R-A as it was similar to the size of the dghA<sub>2A</sub>R (Fraser *et al.*, 2006). There were five additional bands present in the purified sample, which were not present in other purification procedures involving the dghA<sub>2A</sub>R-SMALP (**figure 36B and 36D**). The CD81-SMALP sample showed a dense band at the 35 kDa region, where literature has demonstrated a tagged CD81 band positioned at around 27 kDa (Clark *et al.*, 2001). There was a fainter band (indicated by the red arrowhead) below the dense band, which is positioned closer to the 27 kDa region of the gel (**figure 36E**). Western blotting was used to confirm the presence of the membrane proteins using their his-tags (**figure 36F**). A dense band was detected at the 40 kDa marker, which suggests the presence of the

GPCR (**figure 36B**) similarly to the indicated band in the coomassie gel (**figure 36A**). There was a blurred band in the A<sub>2A</sub>R-A-SMALP sample between the 25-35 kDa region, which could be degraded receptors present within the sample. The band representing the A<sub>2A</sub>R-A was not present in the purified CD81-SMALP sample and vice versa (**figure 36F**). Western blotting detected the band representing the CD81 to be positioned around the 27 kDa region, similarly to the arrowhead-indicated band in the coomassie gel (**figure 36E and 36F**).

### 3.3: Discussion

The dghA<sub>2A</sub>R and the A<sub>2A</sub>R-A construct was deglycosylated at residue N154 via mutagenesis to Q154. The deca-his-tag was fused to the N-terminus of the A<sub>2A</sub>Rs by using the primer in **chapter 2.2.1**. The dghA<sub>2A</sub>R was kindly provided by Fraser *et al.* (2006), and its pharmacology was not affected upon deglycosylation (Fraser *et al.*, 2006; Jamshad *et al.*, 2015). The purpose of the deglycosylation in this project was to prevent full occupancy of glycosylation when the construct was expressed in *P.pastoris* (Jamshad *et al.* 2015). In human cells, the extent of glycosylation of the A<sub>2A</sub>R is not as high as when the receptor is expressed in *Pichia pastoris*. The N154 glycosylation site is positioned the middle of the ECL2, which may hinder the accessibility of a nanobody-expressing phage from binding the receptor (Fraser *et al.*, 2006). The construct was ligated into a pPICZ $\alpha$ -A vector, where the AOX1 sequence encodes an alcohol oxidase for methanol metabolism (Kupcsulik and Sevelia, 2004), which was upstream the A<sub>2A</sub>R construct. Methanol is a toxic carbon source for yeast and without alcohol oxidase, methanol would not be metabolised which can harm the cell. Upon methanol induction, alcohol oxidase is activated, which leads to the downstream transcription and translation of the A<sub>2A</sub>R constructs. **Figure 31A** demonstrates that upon starvation of glycerol and the application of methanol, the dghA<sub>2A</sub>R-expressing yeast continued to grow rapidly until the 45 hour time point, where the yeast culture reached the lagging phase. This shows that the dghA<sub>2A</sub>R-expressing yeast is able to survive and use methanol as an alternative carbon source, which activates AOX1, leading to dghA<sub>2A</sub>R expression. When comparing the exponential growth rate of dghA<sub>2A</sub>R-expressing yeast and wild type yeast, the growth rate was almost identical, indicating that the transformation of the X33 yeast strain did not have an impact its growth. Western blot analysis against the his-tag of the A<sub>2A</sub>R constructs demonstrated that dghA<sub>2A</sub>R was present

in the membranes of the dghA<sub>2A</sub>R-expressing *P.pastoris* after the cell wall was broken through C3 homogenization and membrane pelleting (**figure 31C and 31D**).

After the expression of the A<sub>2A</sub>R constructs were confirmed in the X33 *Pichia pastoris*, the homogenized yeast membranes expressing dghA<sub>2A</sub>R were solubilised for 20 hours at room temperature, similarly to Jamshad *et al.* (2015). Different SMA polymers with different properties were used to solubilise the GPCR from the membranes (**figure 32A**). The preliminary data compared the proportion of soluble receptors with the total receptors expressed in the membrane, where all polymers showed a similar solubilisation efficiency of around 70% (**figure 32B**). The similarity in solubilisation efficiency between SMA2000 and SZ25010 was also observed with non-GPCRs in Morrison *et al.* (2016), however SZ42010 was less efficient than SMA2000 and SZ25010 when solubilising non-GPCRs. As the data in **figure 32** was preliminary, SMA2000 was used for subsequent experiments as it is the most popular choice amongst the polymers for GPCR solubilisation (Jamshad *et al.*, 2015; Wheatley *et al.*, 2016; Morrison *et al.*, 2016). Western blot analysis of the SMA-solubilised dghA<sub>2A</sub>R demonstrated a high solubilisation efficiency of the polymers, however the amount of pharmacologically active receptors within the SMALPs was much lower (**figure 32**).

A solubilisation time course was conducted on the A<sub>2A</sub>R-A-expressing membranes for 1-20 hours, to measure the amount of binding-capable receptors solubilised for different time periods (**figure 33**). For the A<sub>2A</sub>R-A construct, a 2 hour solubilisation period showed the highest recovery of ZM241385-binding receptors, whilst the rest of the time periods demonstrated a 2-fold lesser solubilisation efficiency of binding-capable receptors (**figure 33**). The optimum solubilisation time period for the A<sub>2A</sub>R-A differed to the methodology presented in Jamshad *et al.* (2015), where the solubilisation period for the dghA<sub>2A</sub>R was 20 hours at room temperature. As the 2 hour solubilisation time at room temperature showed the highest recovery, subsequent solubilisation experiments with the A<sub>2A</sub>R-A were done with SMA2000 for 2 hours. The highest receptor recovery was around 22% of ZM241385-binding A<sub>2A</sub>R-A, which was lower than the total amount of receptors solubilised in blot **figure 32B**. The recovery percentage differs from existing data with non-GPCRs, where solubilisation efficiencies of 70-90% were observed (Swainsbury *et al.*, 2017; Hardy *et al.*, 2019). Light scattering and western blotting methods were used

to assess solubilisation efficiency, which was similar to the 70-80% efficiency in **figure 32B**. However the techniques do not assess the amount of binding-capable receptors present in the SMALP sample. This demonstrated that the gel-based techniques do not take into account the amount of conformationally stable receptors in the SMALP sample, where binding assays proved to be a superior technique to demonstrate this. ZM241385 is highly specific to the adenosine receptors and can only bind the receptor if it was in its correct conformation.

Once the dghA<sub>2A</sub>R constructs were solubilised into SMALPs, the Ni<sup>2+</sup>-NTA affinity resin was used to purify the receptor SMALPs via the deca-his-tags which were fused to the N terminus of the receptors. Initially, various concentrations of imidazole in elution buffer was used to identify whether the his-tagged dghA<sub>2A</sub>R-SMALP can be eluted from the nickel affinity column. Coomassie staining experiments have identified the 150mM imidazole eluted the most, but not all, of the dghA<sub>2A</sub>R-SMALPs (**figure 35A**). From then on, 250mM of imidazole was chosen to elute all the dghA<sub>2A</sub>R-SMALP from the resin, as this concentration was used previously with success when purifying the dghA<sub>2A</sub>R (Jamshad *et al.*, 2015). The 250mM imidazole elution buffer was used elute the dghA<sub>2A</sub>R-SMALP from the nickel resin, where 1 ml washes with elution buffer ten times eluted all of the A<sub>2A</sub>R-SMALP from the nickel resin in the benchtop column (**figure 35B**). Furthermore, wash cycles with the wash buffer was increased to 20 cycles to ensure all contaminating protein was removed from the sample. From each elution fraction in the coomassie stain, only one band was detected at 35-40 kDa, which is the size of A<sub>2A</sub>R (**figure 35B**). This showed that the purification method eluted all the dghA<sub>2A</sub>R-SMALPs from the nickel resin and the receptor-SMALP sample was pure. The elution fractions were collated and concentrated using vivaspin 20 centrifugal concentrators and the sample was still present as shown by coomassie staining (**figure 35B**). Furthermore, collation of the samples revealed one distinct band at the 35-40 kDa region, showing that the sample was pure. Dialysis was also used on the same day as purification and concentration to remove all the imidazole from the sample to prevent any degradative effects to the receptor. His-tag western blotting also showed that the dghA<sub>2A</sub>R-SMALP was present during the purification procedure and that all the dghA<sub>2A</sub>R-SMALP was eluted during the purification (**figure 35C**). However, the coomassie gels detected the presence of protein up to elution fraction 8 (e8), whereas

the western blot did not detect bands between elution fractions e5-e10 (**figure 35B-35C**). As western blotting is a more sensitive technique than coomassie staining, bands should theoretically be present in elution fractions e5-e8 in **figure 35C**. Possible reasons as to why the bands were not detected in the blot could be due to faults with the anti-His tag antibody, or the coomassie gel may have detected a non-specific protein. The western blot also detected a band between 25-35 kDa, which was not seen in the coomassie stained gels (**figures 35B and 35C**).

A similar purification optimization procedure was also performed on the A<sub>2A</sub>R-A-SMALPs, where the first step was to identify whether the receptor-SMALPs can be eluted from the nickel resin. The elution buffer containing 100 mM imidazole eluted the highest amount of A<sub>2A</sub>R-A-SMALPs from the chromatography column (**figure 36A**) where the rest of the SMALPs were eluted by using increasing concentrations of imidazole. For further purification procedures, elution fractions containing 250 mM imidazole was used for receptor elution. The 20 wash cycles, with the wash buffer containing 10 mM imidazole, removed non-specific protein-SMALPs from the column, without interfering with the captured A<sub>2A</sub>R-A-SMALP (**figure 36B**). An elution peak was created based on the density of the elution fractions of **figure 36B**, where all the A<sub>2A</sub>R-A-SMALPs were eluted from the column (**figure 36C**). The same purification steps done with the A<sub>2A</sub>R-A-SMALPs were also conducted with CD81-SMALPs, which was used as a negative control. The CD81 negative control was used to confirm that the A<sub>2A</sub>R constructs were only expressed in yeast that were transformed with either the dghA<sub>2A</sub>R gene construct, or the A<sub>2A</sub>R-A construct. From the coomassie stained gels (**figures 36D and 36E**), a dense band representing the A<sub>2A</sub>R-A was present between the 35-40 kDa marker, however there were 5 unknown bands present in the concentrated sample. The bands have not been characterised and were not always present in independent purification procedures with A<sub>2A</sub>R-A-SMALPs. The A<sub>2A</sub>R-A band was not present in the CD81-SMALP sample, indicating that the receptor eluted in **figure 36D** could be the receptor of interest. The band representing the A<sub>2A</sub>R-A differed in its molecular weight by approximately 5 kDa, before and after purification (**figure 36D and 36F**). This could be differences in the migration of the protein during each SDS-PAGE experiment or an issue with the anti-His antibody. The SMA copolymer could also have an affect on SDS-PAGE, as it can be separated and possibly co-migrate with

the protein sample (Korotych et al., 2019). The CD81 sample also showed high purity, which was important for downstream experiments when used as a negative control. Western blotting against the his-tag also showed that the bands between the A<sub>2A</sub>R-A-SMALP and CD81-SMALP samples were distinct from each other. As there were a few additional bands in all the purified SMALP samples (**figure 35 and figure 36**), an extra purification step would be beneficial to generate highly pure receptor-SMALPs. The AKTA gel filtration system could be used to further purify the protein-SMALPs based on their molecular weight. Any additional bands after the gel filtration step can be analysed using mass spectrometry.

Both the dghA<sub>2A</sub>R and the A<sub>2A</sub>R-A were successfully expressed in X33 *Pichia pastoris*, as shown by the bands in the western blot, positioned at approximately 37 kDa. This was similar to the bands representing the dghA<sub>2A</sub>R in Jamshad *et al.* (2015), which was positioned at the 37 kDa marker. The positioning of the A<sub>2A</sub>R bands observed in **figures 35 and 36** and the dghA<sub>2A</sub>R band in Jamshad *et al.* (2015) are slightly higher than the band observed in Fraser *et al.* (2006). The difference may be due to the different speed of receptor migration through the polyacrylamide gel. The tetraspanin CD81 has also been successfully expressed in X33 cells, as a band between 25-35 kDa was detected in both the coomassie stained gel and the western blot (**figures 36E and 36F**), which was concordant with the bands observed by anti-CD81 antibodies from various companies and literature (Clark *et al.*, 2001).

## Chapter 4: Biophysical characterisation of GPCR-SMALPs

### 4.1: Introduction

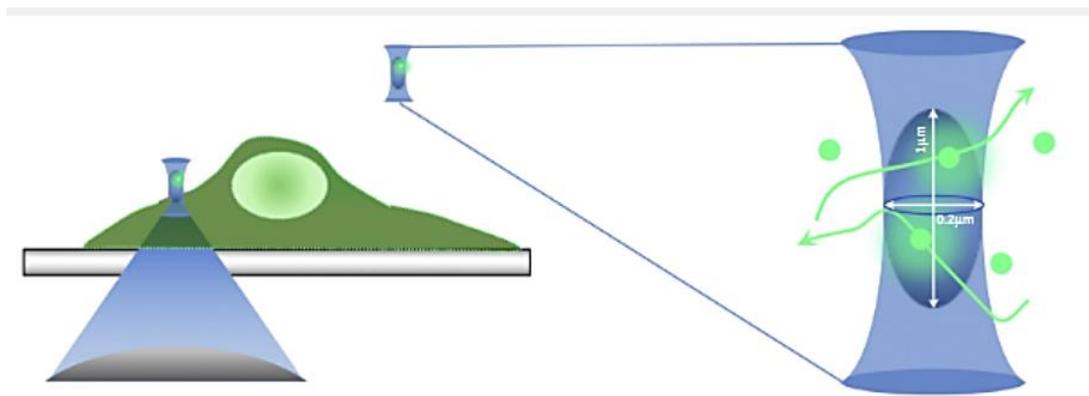
To establish the pharmacology of the SMA-solubilised A<sub>2A</sub>R, different experimental procedures have been implemented into this project to study how the receptor interacts with ligands within a SMALP. Applications have been used to study new properties of the A<sub>2A</sub>R within SMALPs, using techniques such as FCS and XRF.

The radioligand binding assay has been used to observe the displacement of the radioligand, [<sup>3</sup>H]ZM241385, by the competitive ligand, ZM241385, from the SMALP solubilised A<sub>2A</sub>R. GPCRs need to be in their native conformation, which can be stabilised in SMALP (Dorr *et al.*, 2016; Wheatley *et al.*, 2016). Initially radioligand binding was used in this chapter to study the binding capabilities of the A<sub>2A</sub>R in the cell membrane, in SMALPs and after purification. The affinity of [<sup>3</sup>H]- ZM241385 was measured. This was to ensure that the functionality of the receptor was preserved throughout the membrane solubilisation and purification processes (Uddin *et al.*, 2018). The binding capability of the dghA<sub>2A</sub>R in SMALPs was compared with the diisobutylene-maleic acid (DIBMA) copolymer, which is an emerging polymer counteracting the limitations of the SMA copolymers, as described in **chapter 4.2.1.2**.

The use of fluorescent ligands is becoming widespread amongst pharmacology, replacing radioligands as a safer alternative. Fluorescent ligands can be used to measure pharmacological activity in real-time and can be used in imaging studies with the use of fluorescence microscopy. The diversity of the applications of fluorescent ligands is larger than radioligands, which are measured in a scintillation counter (Stoddart *et al.*, 2016; Stoddart *et al.*, 2015). Fluorescent ligands can be used in procedures such as FRET, BRET, flow cytometry and FCS (Stoddart *et al.*, 2016; Stoddart *et al.*, 2015). FCS has previously been used to study ligand-GPCR interactions in live cells and membrane microdomains, however there are currently no publications demonstrating the use of FCS to study GPCR-SMALPs (Bridson *et al.*, 2018). In this procedure, the dghA<sub>2A</sub>R-SMALPs and A<sub>2A</sub>R-A-SMALPs were used. The fluorescent ligand, N-(2-aminoethyl)-2-[4-(2,3,6,7-tetrahydro-2,6-dioxo-1,3-dipropyl-1H-purin-8-

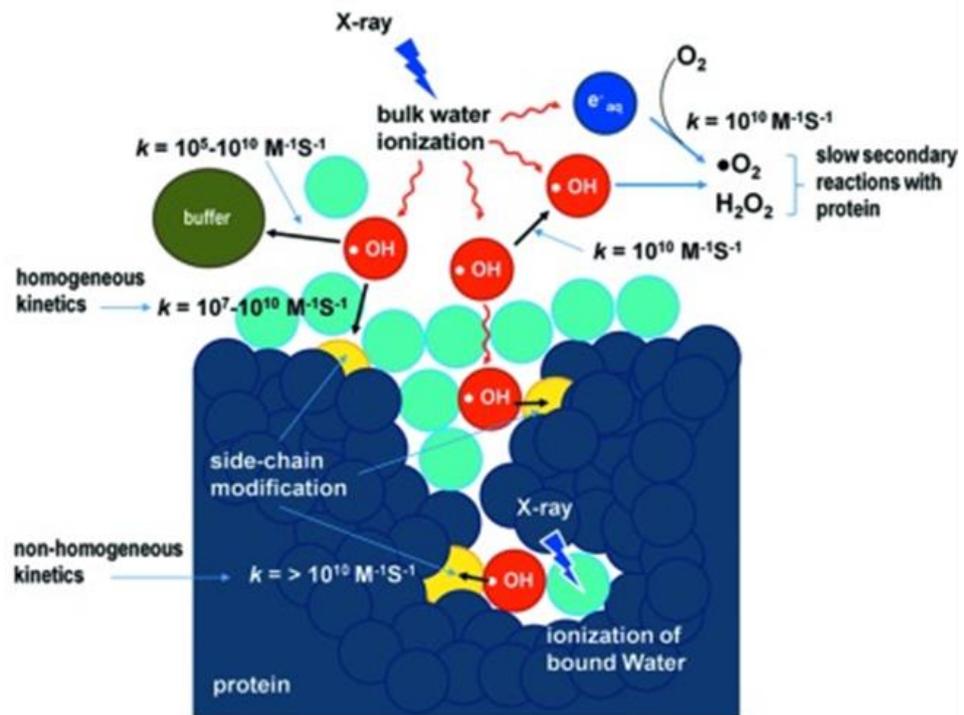
yl)phenoxy]-acetamide (CA200645) and the competitive ligand, ZM241385, was used to assess the pharmacology of the dghA<sub>2A</sub>R-SMALP and the A<sub>2A</sub>R-A-SMALP. CA200645 is a xanthine amine congener (XAC)-based ligand, with the fluorophore BODIPY 630/650 attached via a β-alanine linker (Stoddart *et al.*, 2015). CA200645 binds non-specificly to the adenosine receptor subtypes and antagonises them (Stoddart *et al.*, 2016; Stoddart *et al.*, 2015).

FCS was used to study the pharmacodynamics of ligand binding within a SMALP. It utilises confocal microscopy to identify ligand-receptor interactions. The technique can monitor the spatial and temporal differences in the intensity of fluorescent molecules as they move pass a given detection volume (Briddon *et al.*, 2018). FCS can measure the speed of the molecules, moving through a given space, which is described using the diffusion coefficient (D). The technique also provides information on the brightness of fluorescent molecules as well as the concentration of the molecules. Fluorescent species can be observed through a pinhole within a confocal microscope to generate a detection volume at approximately 100 nm by 200 nm (Briddon *et al.*, 2018). Fluctuations of fluorescence data can be observed as fluorescent molecules move through the detection volume (**figure 37**) (Briddon *et al.*, 2018). When a fluorescent ligand forms a complex with a receptor-SMALP, the complex will have a slower diffusion time when compared with a free fluorescent ligand (Briddon *et al.*, 2018).



**Figure 37** | Illustration of a confocal detection volume used in FCS. A point of interest is selected on a live cell or membrane, which is the confocal detection volume, used for data collection. The confocal lens produce a rugby-shaped detection volume, where fluorescent species can diffuse through. Illustration obtained from (Bridson *et al.*, 2018).

The structural water molecules of the dghA<sub>2A</sub>R-SMALP was also assessed, to see whether the rearrangement of water networks can be observed in the presence of different ligands, under SMALP conditions. In X-ray radiolytic footprinting (XRF), the combination of the synchrotron and mass spectrometry can identify protein sequences which are exposed to modification. The radiolysis of water molecules to form hydroxyl radicals provides an effective labelling technique to identify protein residues exposed to water (**figure 38**) (Gupta *et al.*, 2016). The hydroxyl radicals interact with particular side chains of proteins, acting as a detectable label for mass spectrometry and altering the mass of the protein (Orban *et al.*, 2010). Mass spectrometry is then used to highlight the residues of the receptor sequence which are exposed to hydroxyl modification. Rhodopsin is an example of a GPCR which has undergone radiolytic footprinting to identify a water network. Residue Phe116 of rhodopsin had a high modification rate, suggesting its exposure to water molecules (Orban *et al.*, 2010). In crystallography studies, a water molecule was present near residue Phe116 (Okada *et al.*, 2004). Upon activation of rhodopsin, the modification rate of Phe116 decreased, whereas modification rate of residue Met86 increased (Okada *et al.*, 2004). This indicated that an inactive to active conformational transition of rhodopsin changed the exposure of certain residues to water. This demonstrated the usefulness of X-ray radiolytic footprinting to identify water networks, protein dynamics and structural changes upon receptor activation.



**Figure 38** | Schematic, illustrating the radiolysis of water molecules into hydroxyl radicals to label protein residues. Upon x-ray irradiation, hydroxyl radicals ( $\cdot\text{OH}$ ) and hydrogen peroxide ( $\text{H}_2\text{O}_2$ ) are formed from the water molecules. The hydroxyl radicals, (red) interact with each other, buffer molecules (green) and interact with protein side chains (yellow). The hydroxyl radicals can also interact with electrons (blue) to form superoxide radicals. The kinetics are shown as ‘ $k$ ’ values, representing the rate of various interactions. A short irradiation beam time is favoured as this will lead to fewer superoxide radicals and  $\cdot\text{OH}\cdot\text{OH}$  recombination, which can scavenge the hydroxyl radicals, reducing protein labelling. Diagram taken from (Gupta et al., 2016).

Studying SMALPs using X-ray radiolytic footprinting will provide knowledge on how receptor dynamics may change, with the native lipid bilayer also being a factor to receptor conformation (Wheatley *et al.*, 2016; Gupta *et al.*, 2016; Gupta, 2019). The exclusion of the lipid bilayer in detergent solubilised receptors may provide misleading data on protein dynamics, especially when lipids and cholesterol are allosteric modulators of certain membrane proteins (Tschammer, 2016). Identifying water networks within the dghA<sub>2A</sub>R will provide useful information on the crevices of the receptor, which may give an understanding of how conformational changes occur and how complementary proteins, such as a G protein, may bind. Another important element of XRF is that the changes of protein conformation and solvent accessibility can be compared between active and inactive A<sub>2A</sub>R, in presence of different ligands. This may illustrate how different drugs change the shape of the receptor and which water networks may break or form.

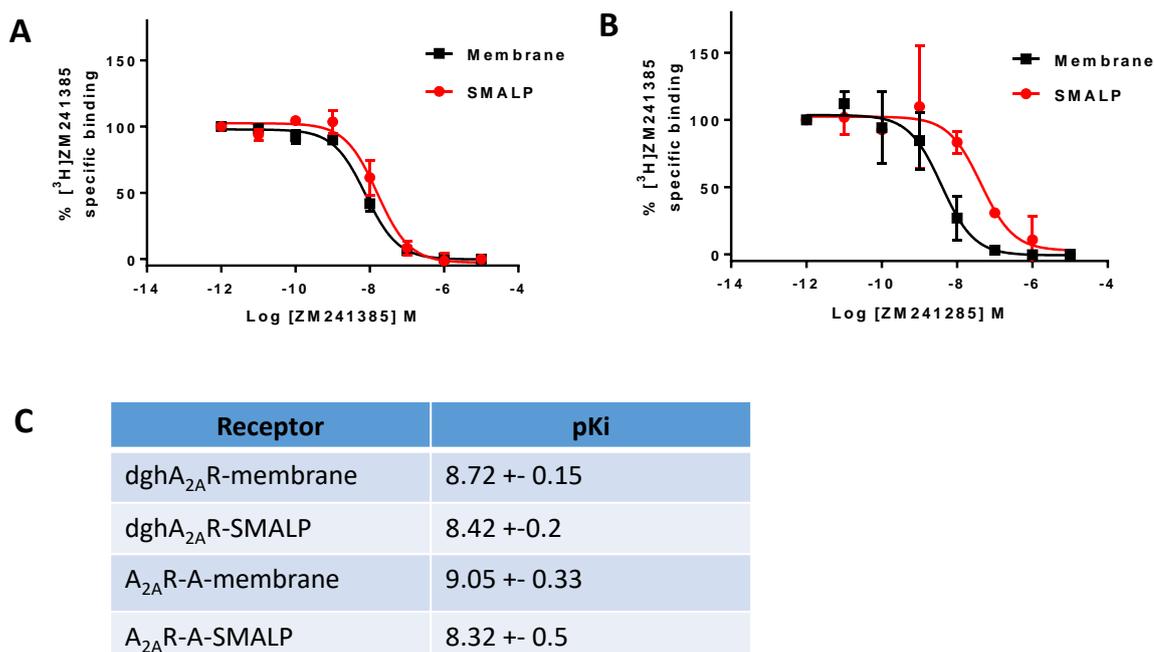
As SMALPs retain the lipid bilayer around the solubilised membrane protein, the impact of the lipid bilayer affects the conformation of the protein, as seen in GPCRs such as the  $\beta_2$ AR (Dawaliby *et al.*, 2016), where phospholipids have altered the pharmacology of the receptor. The conformation of GPCRs are also affected by water molecules, which can stabilise an inactive or active conformation of a GPCR (Okada *et al.*, 2014). As there are no XRF publications on SMALP proteins, discovering the water-exposed residues of the dghA<sub>2A</sub>R-SMALP can provide new insights on the conformational dynamics of the receptor, which may be influenced by water. Furthermore, comparing the water-exposed regions of receptors in SMALPs and DDM micelles could provide a valuable insight on how the dynamics of the receptor may be altered after solubilisation.

## **4.2: Results**

### **4.2.1: Assessing the pharmacology of the A<sub>2A</sub>R-SMALPs using radioligand binding**

#### **4.2.1.1: Using competitive radioligand binding to characterise the pharmacology of the A<sub>2A</sub>R-SMALPs.**

The binding profile of the inverse agonist, ZM241385 was compared between the A<sub>2A</sub>R-expressing membranes and purified A<sub>2A</sub>R-SMALPs. This was done to establish whether its affinity was the same for the receptors expressed in the yeast membranes and after purification of the SMALPs. The comparison of the ligand binding affinity was also made between the dghA<sub>2A</sub>R construct and the A<sub>2A</sub>R-A, to ensure that any modifications made to the receptor did not have a major impact on the A<sub>2A</sub>R-A binding capability.

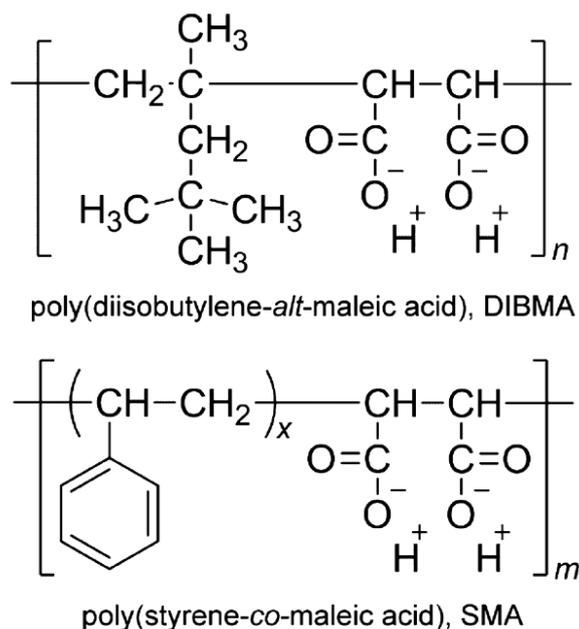


**Figure 39| Competitive radioligand binding profile of the A<sub>2A</sub>R constructs in *Pichia pastoris* membranes and purified SMALPs.** 1 nM [<sup>3</sup>H]ZM241385 was used as the radioligand and 10<sup>-11</sup> – 10<sup>-5</sup> M ZM241385 was used to compete against the radioligand. The pseudo-concentration, 10<sup>-12</sup> M, was 100% DMSO, instead of ZM241385. **A)** The binding curves of the A<sub>2A</sub>R in membranes and purified SMALPs. Graph represents four independent experiments for the A<sub>2A</sub>R, expressed in membranes and three independent experiments for the purified A<sub>2A</sub>R-SMALPs. **B)** The binding curves of the A<sub>2A</sub>R-A in membranes and purified SMALPs. Graph represents three independent experiments for the membrane and purified SMALP conditions. **C)** Table showing the average pKi values for membrane and purified SMALPs containing either the A<sub>2A</sub>R or the A<sub>2A</sub>R-A. The pKi values were calculated using the Cheng-Prusoff equation. The pKi values of the individual repeats from panel A and B were used to compile the data table. A one-way ANOVA test, followed by the Tukey's multiple comparisons test (Graphpad Prism 7) was used to compare the means of the pKi values with each other, which showed no significant difference (p > 0.05).

The A<sub>2A</sub>R expressed in membranes and as purified SMALPs showed a clear mass action displacement curve, where [<sup>3</sup>H]ZM241385 was displaced by increasing concentrations of the cold ligand, ZM241385 (**figure 39A**). The hill slope was of unity and there was no shift in the slope before and after solubilisation. There was also no significant difference in the binding affinity (pKi) of the ligand to the A<sub>2A</sub>R expressed in membranes and solubilised in SMALPs (**figures 39A and 39C**). The same pattern was observed for the membrane A<sub>2A</sub>R-A (**figure 39B**), where there was no difference in ligand binding affinity prior solubilisation and after solubilisation with SMA2000 (**figures 39B and 39C**).

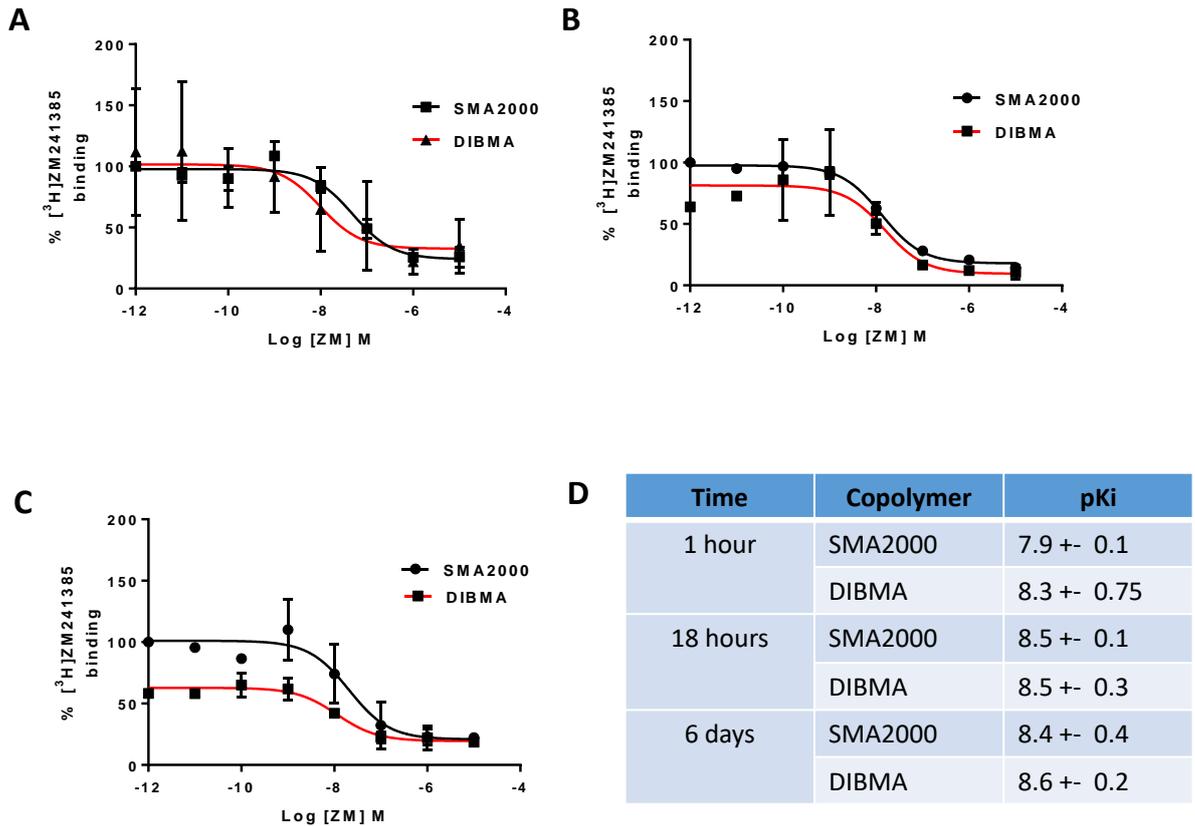
#### **4.2.1.2: Comparing the pharmacological activity of the A<sub>2A</sub>R solubilised using either SMA or DIBMA copolymers**

A novel copolymer is being testing on a variety of membrane proteins, named DIBMA (**figure 40**) (Oluwole *et al.*, 2017). The copolymer is synthesised similarly to SMA. The DIBMA copolymer is theorised to be advantageous over SMALPs, which can alter lipid dynamics (Oluwole *et al.*, 2017). The DIBMA copolymer does not drastically change the lipid-acyl chain order, retaining the phase behaviour of the lipids, similarly to a lipid vesicle preparation (Oluwole *et al.*, 2017). Divalent cations such as Mg<sup>2+</sup> and Ca<sup>2+</sup> are tolerated by DIBMA, unlike SMALPs, where applying 10 mM of Mg<sup>2+</sup> and Ca<sup>2+</sup> into the solubilisation buffer speeds up lipid solubilisation by the DIBMA copolymer (Danielczak *et al.*, 2019). Exceeding 4 mM concentrations of Mg<sup>2+</sup> in buffers containing protein-SMALPs, solubilised using a 2:1 copolymer will cause protein precipitation (Pollock *et al.*, 2018). The tolerance to divalent cations would be advantageous to experiments performed in receptor biotinylation, where the biotin ligase enzyme optimally functions with concentrations of 10 mM magnesium present in the reaction mixture **Chapter 6**. The DIBMA copolymer has also shown similar solubilisation efficiency when compared to SMA (Oluwole *et al.*, 2017). The current literature has demonstrated the advantages of replacing SMA with DIBMA for membrane protein solubilisation, however as research into DIBMA is new, there is a lack of information in protein conformational stability and functionality in DIBMALPs (Danielczak *et al.*, 2019; Oluwole *et al.*, 2017; Oluwole *et al.*, 2017).



**Figure 40** | The structure of DIBMA and SMA copolymer. x represents the number of styrenes in SMA, 1 = (1:1), 2 = (2:1), 3 = (2:1) styrene to maleic acid ratio. Structure taken from (Oluwole *et al.*, 2017).

The dghA<sub>2A</sub>R was solubilised from yeast membranes using either the SMA2000 copolymer or the DIBMA (tradename Sokalan CP9, BASF, Ludwigshafen, Germany) copolymer. The binding affinity and receptor stability post-solubilisation was measured using the competitive radioligand binding assay. The experiments were done in collaboration with Aiman Gulamhussein (Aston University, Birmingham, UK) and is published in Gulamhussein *et al.*, (2020). I have conducted one repeat and Aiman Gulamhussein has repeated the experiment two more times with the same conditions. Data analysis was done by myself and Aiman Gulamhussein.



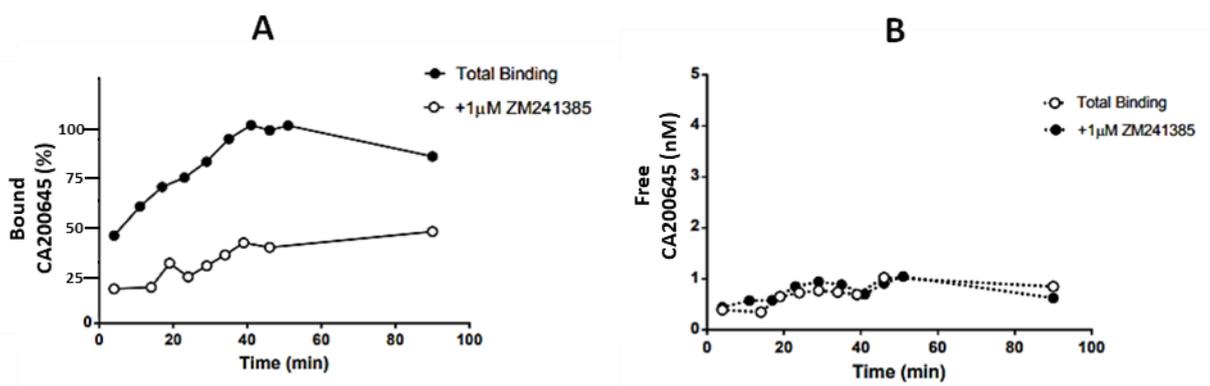
**Figure 41| Competitive radioligand binding assay demonstrating the stability of the dghA<sub>2A</sub>R after solubilisation, using the copolymers SMA2000 and DIBMA.** The dghA<sub>2A</sub>R was solubilised from X33 membranes for 1 hour. 1 nM [<sup>3</sup>H]ZM241385 was used as the radioligand and 10<sup>-11</sup> – 10<sup>-5</sup> M ZM241385 was used to compete against the radioligand. The pseudo-concentration, 10<sup>-12</sup> M, was 100% DMSO, instead of ZM241385. ZM241385 binding to the dghA<sub>2A</sub>R was measured **A**) 1 hour **B**) 18 hours **C**) 6 days post-solubilisation. The graphs represent three independent experiments. **D**) Table showing the pK<sub>i</sub> binding affinity of ZM241385 to the dghA<sub>2A</sub>R, solubilised using either SMA2000 or DIBMA. The affinity was measured 1 hour, 18 hours or 6 days post-solubilisation. The table represents the three independent experiments, shown in graphs A, B and C. The two-way ANOVA test, followed by Sidak's multiple comparisons test was used to compare the pK<sub>i</sub> of SMA2000 and DIBMA for each time post-solubilisation. There was no significant difference in the pK<sub>i</sub> for each copolymer used for solubilisation (p > 0.05).

A mass action sigmoidal curve was observed for both the SMA2000 and DIBMA copolymers, which indicate the displacement of the radioligand by ZM241385 (**figures 41A-C**). The stability of the dghA<sub>2A</sub>R was measured over time when using either SMA2000 or DIBMA copolymers. The percentage of active dghA<sub>2A</sub>R-DIBMALP was compared to the dghA<sub>2A</sub>R-SMALP. After 6 days post-solubilisation, around 50% of the Bmax was lost for the dghA<sub>2A</sub>R-DIBMALPs when compared with the SMALPs (**figures 41A and 41C**). The binding affinity was stable over time after solubilisation, as there was no significant difference in pKi between the SMA and DIBMA copolymers (**figure 41D**).

#### **4.2.2: Using fluorescence correlation spectroscopy to study A<sub>2A</sub>R-SMALP pharmacodynamics in real-time.**

The FCS project was a part of the Centre of Membrane Proteins and Receptors (COMPARE) scheme. The production of the A<sub>2A</sub>R-A-SMALPs was done by myself. The initial experiments presented in this thesis was conducted by myself, alongside Dr. Steve Bridden (University of Nottingham, UK) and Dr. Joelle Goulding (University of Nottingham, UK). Further optimisation experiments were conducted by Rachael Grime (University of Birmingham, UK), Dr Steve Briddon and Dr Joelle Goulding, using the yeast cells prepared by myself. Rachael Grime, Dr Steve Briddon and Dr Joelle Goulding analysed the FCS data, using the equations in the recently published paper by Grime *et al.*, (2020). The work was in collaboration between Aston University, University of Birmingham and University of Nottingham. My contribution to the project is in Grime *et al.*, (2020).

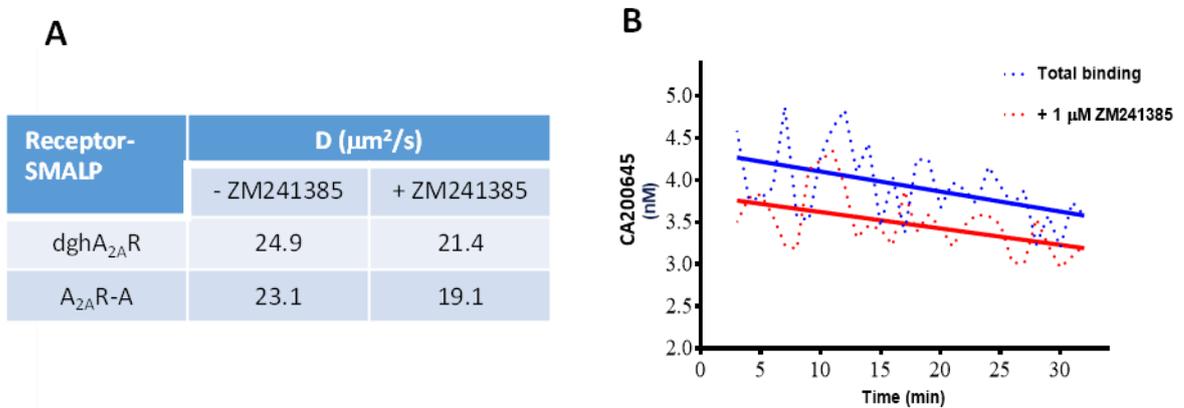
As FCS has mainly been conducted on live cells, the use of this technique has not been done on SMA solubilised GPCRs. Therefore the initial experiments involved the optimisation of FCS in order to detect ligand-binding in a SMALP format, using a fluorescent ligand. The experiments presented in **figure 42** and **figure 43** were conducted to find a suitable concentration to displace the fluorescent ligand with a displacer ligand, in order to measure pharmacodynamics in real-time.



**Figure 42| Fluorescence correlation spectroscopy data showing the displacement of the fluorescent ligand, CA200645, by ZM241385.** The dghA<sub>2A</sub>R-SMALP samples were split into two conditions. The first condition contained 100 nM CA200645 (total binding) , whilst the second condition contained 100 nM CA200645 and 1 μM ZM241385, in order to displace the fluorescent ligand and measure non-specific binding of CA200645. Both conditions contained 5 μg/ml of dghA<sub>2A</sub>R-SMALPs. A 633 nm photon laser was used to excite and detect the fluorescent CA200645. **A)** The percentage of the fluorescent CA200645-bound SMALPs over time, with or without ZM241386 present. The detection volume was gated around the dwell time of the dghA<sub>2A</sub>R-SMALP:CA200645 complex, which was around 625 μs. **B)** The concentration of the unbound, free CA200645 over time, with or without ZM241385 pre-incubation. The detection volume was gated around the dwell time of the CA200645, which was around 68 μs.

The initial FCS experiments were executed to optimise the confocal microscope and the binding conditions of CA200645 and its displacement by ZM241385. **Figure 42A** demonstrated the kinetics of total and non-specific binding. The detection volume was filtered to measure the CA200645, bound to the dghA<sub>2A</sub>R-SMALP, which had a dwell time of 625 μs. The percentage of CA200645:dghA<sub>2A</sub>R-SMALP complexes increased overtime, with a declination after 60 minutes of measurement. Pre-incubation with ZM241385 revealed non-specific binding, where the percentage of CA200645:dghA<sub>2A</sub>R-SMALP complexes was lower than the total binding condition. **Figure 42B** measured the concentration of free CA200645 overtime, with or without ZM241385 pre-incubation. The detection volume was gated around the dwell time of the free fluorescent ligand, which was moving at an average of 68 μs within the detection volume. For the total binding and non-specific binding conditions, the concentration of CA200645 was similar overtime. This differed to the displacement

observed between total binding and non-specific binding conditions of the CA200645:dghA<sub>2A</sub>R-SMALP, when the microscope was detecting slower moving particles (**figure 42A**)



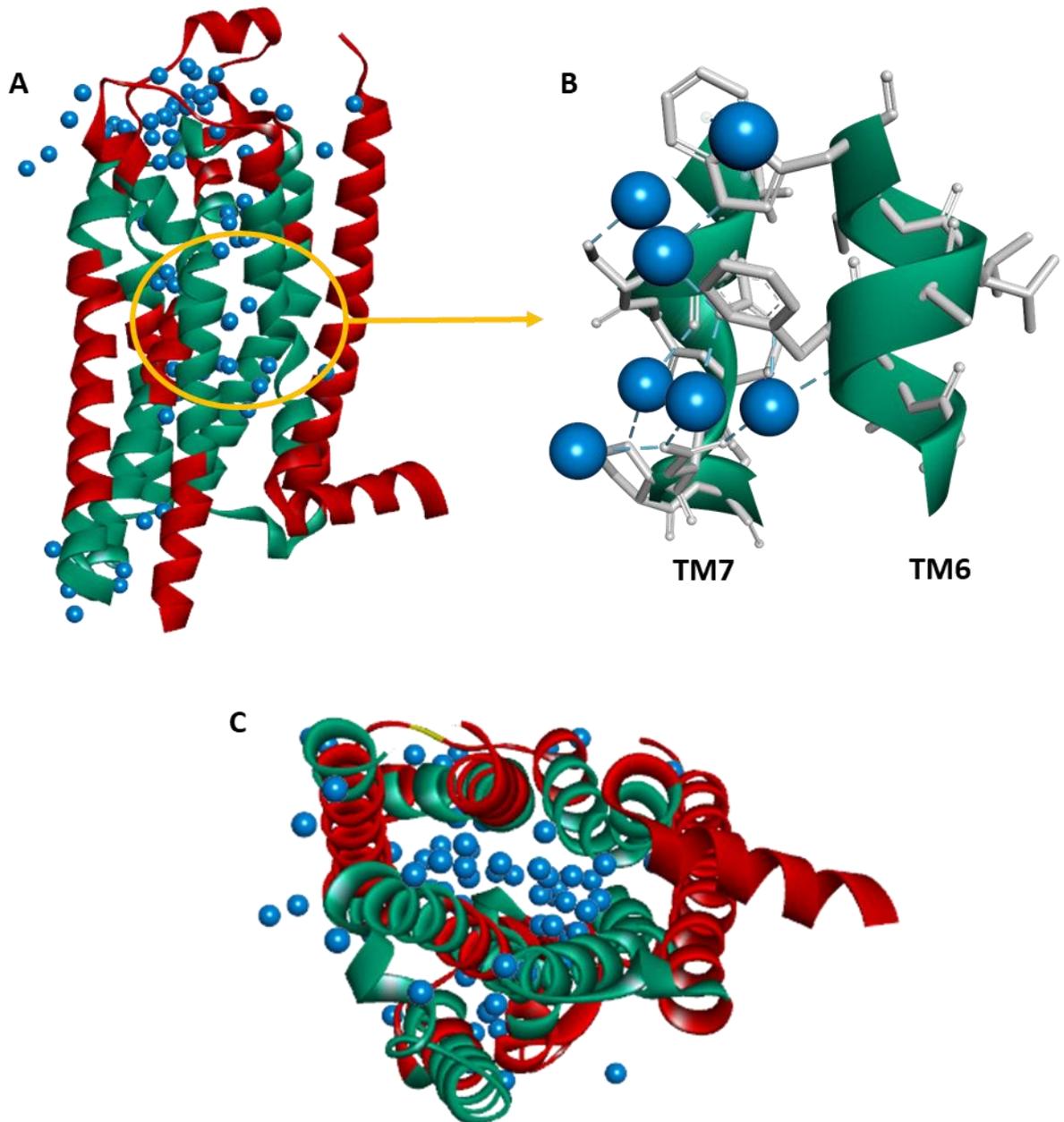
**Figure 43| Comparing the binding of CA200645 between A<sub>2A</sub>R-SMALPs and A<sub>2A</sub>R-A-SMALPs.** 5 μg/ml of the A<sub>2A</sub>R-SMALPs and the A<sub>2A</sub>R-A-SMALPs were pipetted into a glass chamber. Both samples were pre-incubated with 100 nM CA200645. The ZM241385 conditions contained both 100 nM CA200645 and ZM241385. A 633 nm photon laser was used to excite and detect the fluorescent CA200645. The detection volume was filtered to detect fluorescent particles moving at 500 ns **A**) Table showing the diffusion coefficient (D) of CA200645, with (+) and without (-) 100 nM ZM241385 pre-incubation. **B**) The concentration of the fluorescent CA200645-bound SMALPs over time, with or without ZM241386 incubation. The detection volume was gated around the dwell time of the dghA<sub>2A</sub>R-SMALP:CA200645 complex, which was around 625 μs. A line of best fit was generated by analysing the data using linear regression in Graphpad Prizm 7.

The diffusion coefficient (D) of the CA200645-bound A<sub>2A</sub>R-SMALP and the A<sub>2A</sub>R-A-SMALP was similar (**figure 43A**). When ZM241385 was pre-incubated with the SMALPs, D decreased by 5 μm<sup>2</sup>/s for the A<sub>2A</sub>R-A and dghA<sub>2A</sub>R SMALPs. **Figure 43B** demonstrated the concentration of CA200645:A<sub>2A</sub>R-A-SMALP complexes formed overtime, after the reaction reached an equilibrium. The graph shows the specific binding of the fluorescent ligand to the avi-tagged receptor.

#### **4.2.3: The use of X-ray radiolytic footprinting to identify water-accessible residues of the SMA solubilised A<sub>2A</sub>R.**

The buffer system during the purification of the A<sub>2A</sub>R-SMALP was changed into a sodium phosphate based system instead of using HEPES and TRIS. This was to prevent the interruption of hydroxyl radicals modifying the water-exposed residues. TRIS, HEPES and DMSO can scavenge the hydroxyl radicals, reducing the dosage required for residue oxidization, leading to inaccurate modifications and skewed data. After the receptor was purified, a soluble protein radioligand binding assay was used to determine whether the A<sub>2A</sub>R-SMALP was pharmacologically stable in a sodium phosphate buffer system with a lowered salt concentration.

The purified A<sub>2A</sub>R-SMALP, was concentrated to 1 mg/ml. The sample was sent to the ALS Berkeley institute (California, USA) for the irradiation of the sample using x-ray light. Dr Sayan Gupta (Lawrence Berkeley National Laboratory, California, USA) conducted the synchrotron experiments. After fragmentation of the receptor into peptides, Dr Vassiliy Bavro (University of Essex, UK) used mass spectrometry to analyse the oxidised residues of the protein sequence.



**Figure 44| Structure of the A<sub>2A</sub>R, illustrating the regions of the receptor which were exposed to hydroxyl modification by XRF.** Structure of the A<sub>2A</sub>R and surrounding crystallised water molecules (blue spheres) (PDB ID: 5OLG) at a 1.86 Å resolution. The water-exposed sequences detected by XRF are shown in green. The unmodified sequences are shown in red. **A)** Side view of the crystal structure. **B)** Molecular analysis of the orange circle (panel A), showing intermolecular hydrogen bonds (blue dashed lines) between the hydroxyl-modified, water-exposed amino acid residues of TM6 and TM7 and the crystallised water molecules. The residues are shown as grey sticks. **C)** Intracellular view of the crystal structure. BIOVIA Discovery Studio was used to analyse the structure.

The structure in **figure 44** was made using the mass spectrometry data, was analysed using Matrix Science Mascot. The mass spectrometry data covered 71% of the sequence, with peptide sequence modified by oxidation. The hydroxyl-modified sequences were predominantly found in TMs 2, 3, 4, 6, 7 and the TM5-ECL2 interface. The intracellular surface of the receptor was also exposed to irradiation, as seen with the intracellular loops (ICL) 1 and 2. ICL3 was missing from the structure therefore it is unknown whether this region of the sequence was modified. The ECLs were not modified, according to the Mascot data and a majority of the ECL2 was not detected by mass spectrometry, which could be due to peptide degradation during peptide lysis. Water molecules in the crystal structure are in line with the hydroxyl-modified residues within the TM bundle (**figure 44A and 44C**). A majority of the water molecules are present at the extracellular surface and the ligand binding site of the receptor (**figure 44A**). The structure in **figure 44B** demonstrates potential interactions between the crystallised water molecules and the residues of the water-exposed regions of TM6 and TM7.

### **4.3: Discussion of chapter**

After the dghA<sub>2A</sub>R and the A<sub>2A</sub>R-A-SMALPs were purified, concentrated and dialysed, biophysical experiments were performed to confirm the pharmacology of the receptor within the SMALP and to test novel properties of the SMA-solubilised receptors. The initial tests were radioligand binding experiments, to confirm whether the solubilised and purified A<sub>2A</sub>R had the same binding affinity for the inverse agonist, ZM241385, when compared with the membranes expressing the dghA<sub>2A</sub>R (**figure 39A**). As there was no significant difference between the binding affinity of membranes and SMALPs, this ensured that the solubilisation and purification process did not impact receptor pharmacology (**figures 39A and 39C**). The same was done for the modified A<sub>2A</sub>R-A construct, which also showed no significant difference in the binding affinities of the purified SMALPs and membranes (**figures 39B and 39C**). Modifying the A<sub>2A</sub>R-A construct with extra tags did not impede the receptors pharmacology, as shown by the pK<sub>i</sub> being statistically identical to the dghA<sub>2A</sub>R (**figure 39C**). The solubilisation and purification protocols were similar to the protocols used in Jamshad *et al.* (2015), where the binding affinity of ZM241385 was pK<sub>i</sub> 7.95 (± 0.45) for the A<sub>2A</sub>R expressed in *Pichia pastoris* membranes and 7.79 (± 0.14) for A<sub>2A</sub>R-SMALPs (Jamshad *et al.*, 2015). The pK<sub>i</sub> affinity for ZM241385 for the dghA<sub>2A</sub>R and

the A<sub>2A</sub>R-A in membranes and SMALPs was slightly higher in **figure 39C**. The slight difference in affinity may be due to alterations in the preparation of the yeast cells and membranes. Comparing the data in **figure 39** with the data from Jamshad *et al.*, (2015) demonstrated that the solubilisation and purification method was reproducible, yielding functional receptors.

The emerging DIBMA copolymer is of huge interest in SMALP research as it is modified to have less impact on the phospholipid core and is tolerable to divalent cations (Oluwole *et al.*, 2017; Danielczak *et al.*, 2019). The tolerance to Mg<sup>2+</sup> is essential for experiments involving enzymes requiring ATP and magnesium, where SMALPs would precipitate to exposure to high concentrations of divalent cations (Pollock *et al.*, 2018). The BirA500 enzyme buffers include high concentrations of magnesium as shown in **chapter 6**, where the A<sub>2A</sub>R-A-SMALP was unsuccessfully biotinylation, unless the magnesium concentration was lowered under 4 mM. Replacing SMALPs with DIBMALPs could possibly overcome this limitation, however there is limited evidence to support the applicability of DIBMA in the solubilisation of conformationally stable receptors (Oluwole *et al.*, 2017; Oluwole *et al.*, 2017; Danielczak *et al.*, 2019). To investigate the suitability of DIBMALPs in solubilising pharmacologically active receptors, the DIBMA and SMA copolymers were compared to observe the stability of the dghA<sub>2A</sub>R over time (**figure 41**). The binding affinity of the ligand remained the same between the SMALP and DIBMALP conditions (**figure 41D**), which showed that DIBMA solubilisation does not alter the pharmacology of the receptor, with respect to the pK<sub>i</sub> observed in SMALPs and the yeast membranes (**figures 39C and figure 41D**). The receptor in DIBMALPs is not as stable as SMALPs over time, where over 50% of the dghA<sub>2A</sub>R-DIBMALPs lost their binding capability after 6 days of being stored at 4°C (**figures 41A-C**). The loss of B<sub>max</sub> for the dghA<sub>2A</sub>R-DIBMALPs could be due to the lower effect the DIBMA copolymer has on lipid packing than the SMA copolymer (Gulamhussein *et al.*, 2020; Oluwole *et al.*, 2017). This would render the DIBMALP less stable than SMALPs, resulting in a reduction in binding over time. This could be problematic for the phage display biopanning rounds, which take over a week to develop target-specific phage (**Chapter 5**), requiring the DIBMALPs to be stored in the fridge for over a week. The reduction of folded receptors in DIBMALPs would result in a reduction of conformational epitopes for nanobody discovery and can potentially raise nanobodies

targeting unfolded receptors. The DIBMA copolymer should therefore be modified to sustain conformational stability over time.

Fluorescence correlation spectroscopy has been implemented to discover the pharmacodynamics of a single ligand, binding to a single SMA-solubilised receptor. This technique was successful in analysing binding dynamics using a small concentration (5  $\mu\text{g/ml}$ ) of sample, which was analysed in real-time. The total and non-specific binding of the fluorescent ligand, CA200645, was observed for both the dghA<sub>2A</sub>R-SMALP and the A<sub>2A</sub>R-A-SMALP (**figures 42A and 43B**), demonstrating that the technique can be an alternative to radioligand binding techniques. The diffusion coefficient of the ligand-bound dghA<sub>2A</sub>R-SMALP was slower by approximately 2  $\mu\text{m}^2/\text{s}$  than the A<sub>2A</sub>R-A-SMALP for the total and non-specific binding conditions (**figure 43A**). The diffusion of CA200645 was slightly faster when ZM241385 was present. To improve the results, the diffusion coefficient of the free ligand without the presence of SMALPs could demonstrate the non-specific interaction between the ligand and the SMALP.

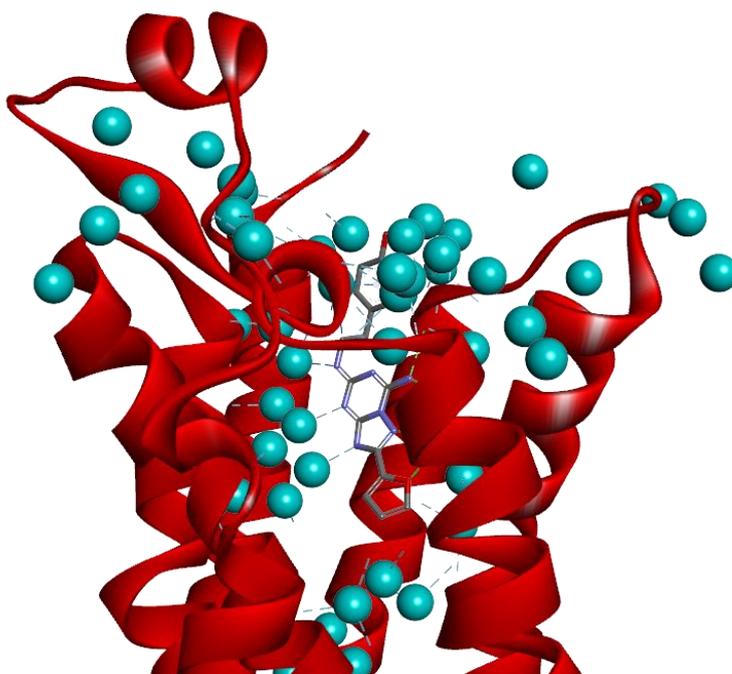
When observing the concentration of unbound CA200645, the levels of fluorescence were the same for the total binding and non-specific binding conditions (**figure 42B**). The potential reason for this could be the lipophilicity of CA200645 and its interaction with the lipid bilayer. The bulk ligand concentration a receptor is exposed to can differ from the actual concentration, as ligands can interact non-specifically with lipids. With regards to the fluorescent ligand used in the FCS experiments, CA200645 consists of the fluorophore, BODIPY, which is lipophilic (Stoddart *et al.*, 2015; Stoddart *et al.*, 2015), explaining why the total and non-specific binding conditions showed a similar concentration of CA200645 (**figure 42B**).

XRF was implemented to study the water networks within an SMA-solubilised GPCR, as water molecules are essential for the stability of different conformations of the receptor as well as receptor function (Venkatakrisnan *et al.*, 2019). Structures have previously shown the influence of water networks on GPCR structures, however as the proteins were processed at cryogenic temperatures, the physiological characteristics of the water molecules were not analysed (Venkatakrisnan *et al.*, 2019; Rucktooa *et al.*, 2018). As XRF utilised a short exposure x-ray beam to modify water-exposed regions

of the GPCR without any freezing involved, the water-protein interactions were more physiological than crystallography-based methods. Solubilising receptors into SMALPs provided a native structure of the protein to detect water-interacting residues. The structure in **figure 44**, illustrates a majority of the TM bundle, with exception to TM1, being hydroxyl modified, indicating its exposure to water molecules. The positioning of the crystal water molecules are in close proximity to the water-exposed residues of the receptor **figure 44** (Rucktooa *et al.*, 2018) and this was further shown by the potential molecular contacts being made by the crystal water-molecules and the water-exposed regions of the receptor (**figure 44B**). There was a large cluster of water molecules at the extracellular surface of the A<sub>2A</sub>R crystal structure which were not detected in the dghA<sub>2A</sub>R-SMALP XRF structure, where the residues were either unmodified or undetected by mass spectrometry (**figure 44**). The ECL2 of the crystal structure makes polar contacts with a water network, although this was not detected by XRF (**figure 44A**) (Rucktooa *et al.*, 2018). The similarities between the crystal structure and the XRF-derived structure demonstrate that XRF is a useful tool to detect water-exposed residues of membrane receptors solubilised in SMALPs. The possible hydrogen bonding in **figure 44B** demonstrates that the XRF technique is useful to study residues of receptors which interact with water under physiological conditions

Literature has mentioned differences in the conserved water-mediated interactions between active and inactive conformations of a GPCR (Venkatakrishnan *et al.*, 2019; Sabbadin *et al.*, 2014). The ligand binding pocket of the A<sub>2A</sub>R is within the upper portion of the TM bundle, where water molecules are highly mobile, with some stable water-mediated interactions (Venkatakrishnan *et al.*, 2019). Analysis of the A<sub>2A</sub>R-StaR-BRIL-ZM241385 structure (PDB ID: 5OLG) illustrated ZM241385 binding the upper portion of the TM bundle and can make electrostatic contacts with the water molecules (**figure 45**) (Rucktooa *et al.*, 2017). Water molecules may play a role in facilitating the binding of the inverse agonist, ZM241384, to the receptor (Carpenter and Lebon, 2017; Rucktooa *et al.*, 2017). Water molecules for water mediated interactions have been found in the binding site for caffeine and NECA, which are similar in structure to ZM241385 (Carpenter and Lebon, 2017). Molecular dynamic

simulations have shown a difference in water-mediated interactions in the ligand binding pocket of different GPCRs (Venkatakrishnan *et al.*, 2019).



**Figure 45** | Crystal structure of the A<sub>2A</sub>R (PDB ID: 5OLG), complexed with ZM241385. The ligand binding site of the receptor resides in the upper portion of the TM bundle. The chemical structure of ZM241385 is shown as a stick diagram. Electrostatic contacts between the ligand, water molecules and the receptor are shown as dashed lines. The receptor is shown in red and the water molecules are blue spheres. Structure analysed using BIOVIA Discovery Studio. Diagram taken and modified from (Rucktooa *et al.*, 2017).

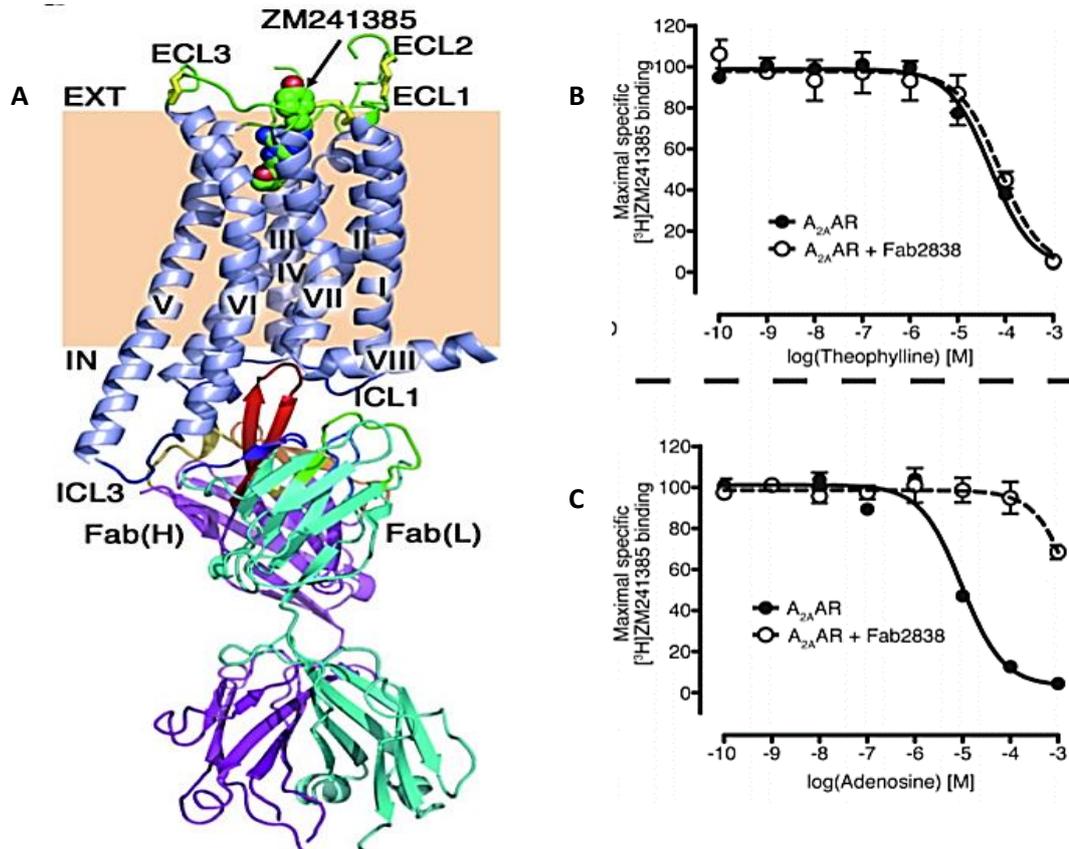
MD simulations have illustrated that the conserved water-mediated interactions between TM1 and TM2 are essential for G protein docking (Venkatakrishnan *et al.*, 2018). This was also highlighted in the A<sub>2A</sub>R structure (**figure 44**), suggesting that the intracellular surface of the receptor may interact with water molecules in order for a G protein to bind. Further XRF experiments with the A<sub>2A</sub>R-SMALP with an agonist or an antagonist bound to the receptor would provide novel information on the changes of the water-mediated interactions within the receptor upon ligand binding and when solubilised in a SMALP.

## Chapter 5: Using phage display against A<sub>2A</sub>R-SMALPs to generate A<sub>2A</sub>R-specific nanobodies

### 5.1: Introduction

Antibodies have proved to become as essential tool to study GPCR structure and they are useful to localise GPCRs. They are also advantageous to study GPCRs as they can have a pharmacological effect on the receptors. They can be used to lock a GPCR in one conformational state for structural studies.

Crystal structures of the beta-2 adrenergic receptor and the A<sub>2A</sub>R have been made using antibodies to stabilise crystal contacts. In 2012, a crystal structure of the A<sub>2A</sub>R with the monoclonal antibody fragment, Fab2838, showed the variable region of the antibody interacting with the intracellular side of the receptor (**figure 46**) (Hino *et al.* 2012). This antibody locked the receptor in an inactive state by forming multiple hydrogen bonds with the receptor, reinforcing the ionic lock. The antagonists' affinity to bind the receptor was not affected by Fab2838, whereas the agonist, adenosine lost binding affinity (**figure 46**) (Hino *et al.* 2012). The antibody bound allosterically to the receptor, without making contact with the orthosteric ligand binding site (Hino *et al.*, 2012). Fab2838 was therefore termed an allosteric inverse agonist as it pushed the A<sub>2A</sub>R into an inactive conformation without binding to the orthosteric site (Hino *et al.* 2012). The study demonstrated the effectiveness of antibodies to generate GPCR structures as well as to alter the pharmacological properties of the receptors. Antibodies therefore have a therapeutic value in treating receptor-related diseases. The ability of an antibody to alter GPCR pharmacology proves the use of antibodies as a potential therapeutic tool.

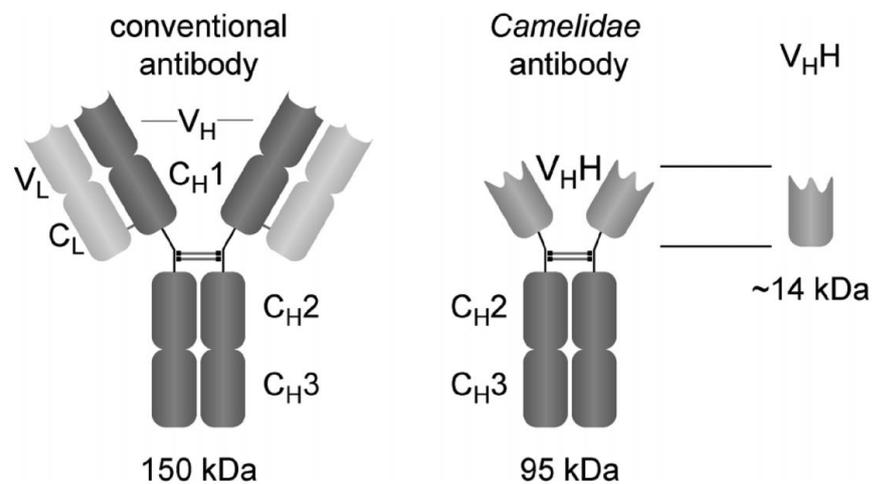


**Figure 46** | **A**) Crystal structure showing the A<sub>2A</sub>R (blue TMD) complexed with the Fab portion of an intracellular monoclonal antibody (Fab2838). **B**) Radioligand binding assay demonstrating the ligand-binding profile of the A<sub>2A</sub>R antagonist, Theophylline, in wild type A<sub>2A</sub>R and A<sub>2A</sub>R-Fab2838. The binding profiles for both A<sub>2A</sub>R and A<sub>2A</sub>R-Fab2838 were similar. **C**) Radioligand binding assay demonstrating the ligand-binding profile of the A<sub>2A</sub>R agonist, Adenosine, in wild type A<sub>2A</sub>R and A<sub>2A</sub>R-Fab2838. Fab2838 significantly reduced the affinity of adenosine to bind A<sub>2A</sub>R due to stabilising the inactive state of the receptor (Hino *et al.* 2012).

Besides the adverse physiological effects associated with small molecule drugs, targeting receptors with small molecule drugs is difficult to achieve, as they have limited target specificity and have an increased chance of non-specific interactions (Muratspahic *et al.*, 2019). The non-specific interactions could be due to the high lipophilicity of a ligand or their small molecular weight (Hutchings *et al.* 2017). Discovering drugs with high specificity to a receptor is important, where nanobodies can create more contacts with a receptor than a small molecule ligand. In the case of the CGRPR, the small molecule drugs, olcegepant and telcagepant antagonise the CGRPR and the amylin 1 (AMY1) receptor (Walker *et al.*, 2018). Phage display can potentially be used to discover nanobodies which specifically interact with the CGRPR and not the AMY1 receptor.

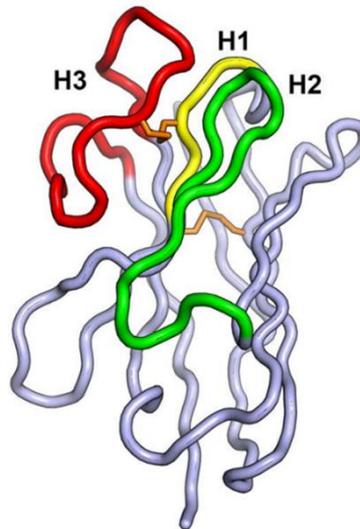
Monoclonal antibody therapeutics have emerged as an alternative to small-molecule drugs as they overcome the disadvantages seen with conventional drugs. They can be developed to become highly specific to one receptor as they target unique epitopes of a receptor. Monoclonal antibodies are stable in serum, therefore they have a long lifespan within the body (Jo and Jung, 2016), which reduces the dosage frequency required to treat a patient. Antibodies can prevent a natural GPCR ligand from binding to a receptor and can stabilise an active or inactive state of a receptor and act as an allosteric or orthosteric modulator, which can affect ligand binding as seen in figure 9 (Hutchings *et al.* 2017, Hino *et al.* 2012). They can also stabilise a ligand-bound state of a GPCR as well as preventing receptor degradation (Hutchings *et al.* 2017). GPCR oligomerization can potentially be stabilised by monoclonal antibodies (Hutchings *et al.* 2017). The natural ligands for family B GPCRs are peptides, therefore creating small molecule drugs against family B GPCRs is less feasible when compared with antibody-based therapeutics. Small molecule antagonists have been created against the CGRPR, however they are currently not available for clinical use (Tringali and Navarra, 2019). In response, antibody-based methods were implemented to create an anti-CGRPR therapeutic antibody (Jain *et al.*, 2018). The CGRPR antagonist, Erenumab is the first antibody to be clinically approved to treat migraine (Raidler, 2019; Jain *et al.*, 2018). Antibodies therefore show great promise to specifically target a GPCR in order to tackle diseases (Jo and Jung, 2016).

Phage display has been incorporated in the project in an attempt to discover nanobodies to alter GPCRs. Nanobodies are derived from camelid heavy-chain antibodies (HCAb), composed of the antigen-recognition variable domain (VHH), which are attached to heavy chains (CH) via a hinge region (Muyldermans, 2013). There are two types of HCABs, (IgG2 and IgG3) in the sera of *Camelidae*, where both are similar in composition, with exception to the hinge region being larger in IgG2. The difference between HCABs and conventional IgGs is the lack of the L chain polypeptide and CH1 in the antigen-recognition region of the HCABs (Muyldermans, 2013). Instead of having four domains making up one section of the Fab region of an IgG, the HCABs have only one single variable domain, termed VHH (**figure 47**) (Muyldermans, 2013; Resemann *et al.*, 2010). Another difference is the smaller structure of the HCABs when compared with IgG, which may be better adapted to target crevices and hidden targets in which conventional IgG may not be able to access (Muyldermans, 2013). The VHH of a HCABs functions similarly to the Fab region of an IgG in antigen-recognition and can be isolated as nanobodies for therapeutic purposes (Muyldermans, 2013).



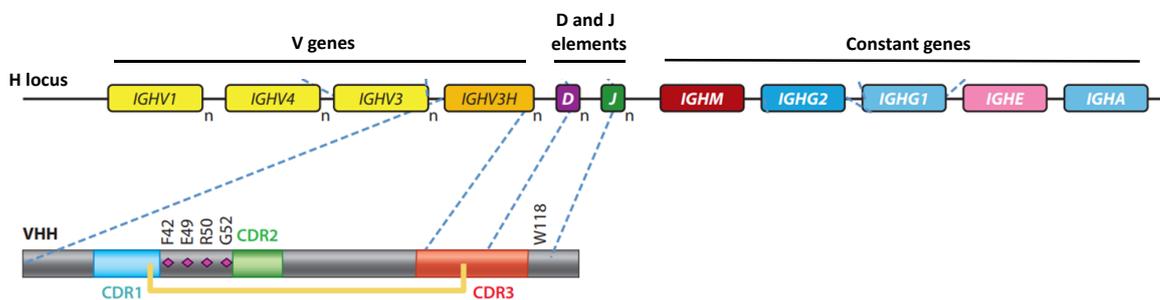
**Figure 47** | Schematic representation comparing the sizes of a conventional IgG antibody, a *Camelidae* HCABs heavy chain antibody and a VHH nanobody. Illustration obtained from (Resemann *et al.*, 2010).

A VHH nanobody is composed of a framework of nine  $\beta$ -strands, configured into a four-stranded  $\beta$ -sheet and a five-stranded  $\beta$ -sheet (Muyldermans, 2013). The sheets are joined by loops (H1-H3) and two conserved cysteines, which form a disulphide bond (**figure 48**) (Koch *et al.*, 2017; Muyldermans, 2013). The variable sequences are compartmentalised into three hypervariable (HV) regions, which reside in the loops (H1-H3) that are connected to the  $\beta$ -strand framework (Muyldermans, 2013). The HV region forms a continuous surface at the N-terminal domain of the VHH (Muyldermans, 2013). Since the HV is complementary to the target epitopes, the HV loops are termed the complementarity-determining region (CDR). Loops H1 and H2 are restricted in certain locations of the VHH, where H3 can vary in size (**figure 48**) (Koch *et al.*, 2017; Muyldermans, 2013).



**Figure 48** | Crystal structure of the VHH-5 nanobody (PDB ID: 5U64). The  $\beta$ -strand framework is represented by the grey strands. The CDR loops are shown as yellow (H1), green (H2) and red (H3) structures. The conserved disulphide bonds is coloured in orange. Structure was taken from (Koch *et al.*, 2017).

The H locus of the *Camelidae* genome encodes the genes which make up the antibodies (**figure 49**) (Muhyldermands, 2013). The VHH gene can be digested and cloned into a phagemid vector, with the VHH gene being fused to a viral coat protein gene (Wesolowski *et al.*, 2009). The vector is then transfected into *E.coli*, alongside a helper M13 phage to assemble the new phages expressing the VHH gene on its surface (Wesolowski *et al.*, 2009). The naïve VHH-conjugated phage is then ready for phage panning to select target-specific phage for nanobody generation.



**Figure 49** | Diagram of the H locus and the composition of the VHH gene. Within the H locus of the *Camelidae* genome, there are V genes, D and J gene elements and the constant genes. The ‘n’ represents variants of a gene. In B cells, the IGHV3H is reconfigured to one D gene element and one J gene element to form the VHH. The grey segments of VHH encode the framework  $\beta$ -strands. The blue, green and red segments encode the CDR, containing the hypervariable sequences which complement the target epitopes. F42, E49, R50, G52 and W118 are hallmark amino acids, encoded by the IGHV and IGHVH genes. Diagram taken from (Muhyldermands, 2013). The V genes and D and J elements can be rearranged to alter the VHH, through VDJ recombination (Roth, 2014; Achour *et al.*, 2008).

There are many antibody production methods available to create antibodies. Some techniques require a lengthy time period to generate target-specific antibodies. The method utilised in this project was phage display, where a phagemid vector containing the VHH genes and the helper phage were transfected into *E.coli* to generate phage expressing the VHH (Wesolowski *et al.*, 2009; Yan *et al.*, 2014; Hoogenboom, 2005). Prior to VHH-phage conjugation, the VHH genes were extracted from non-immunized Llamas, where the genes were amplified by PCR to generate the naïve VHH library (Sabir *et al.*, 2014) The filamentous M13 phage used in this project interacts with the F pilus of certain strains of *E.coli*, via the phage coat protein (Ledsgaard *et al.*, 2018). The phage does not cause cell lysis upon infection, instead

generating and releasing new phages (Ledsgaard *et al.*, 2018). The single-stranded DNA of 6407 bp makes up the genome of the phage, with nine genes encoding coat proteins, phage assembly proteins and replication proteins (Ledsgaard *et al.*, 2018). Once the phage is assembled, the filamentous structure of the phage is 900 nm in length and 6.5 nm in width (Ledsgaard *et al.*, 2018). In phage display, phagemids are constructed with the VHH genes fused to gene III of the phagemid, which encodes the G3P attachment protein, which is associated with phage coat protein generation (Ledsgaard *et al.*, 2018). The phagemid is then transfected into bacteria. The phagemid-positive bacteria are infected with helper phage, which transcribe and translate the M13 genome as well as VHH-gene III fusion construct of the phagemid (Ledsgaard *et al.*, 2018). As the phagemid does not contain the full M13 genome, the helper phage are required for virus assembly and replication (Ledsgaard *et al.*, 2018). Once the VHH-conjugated bacteriophage is extracted from the *E.coli*, a phage library is formed, with each clone having different VHH amino acid combinations.

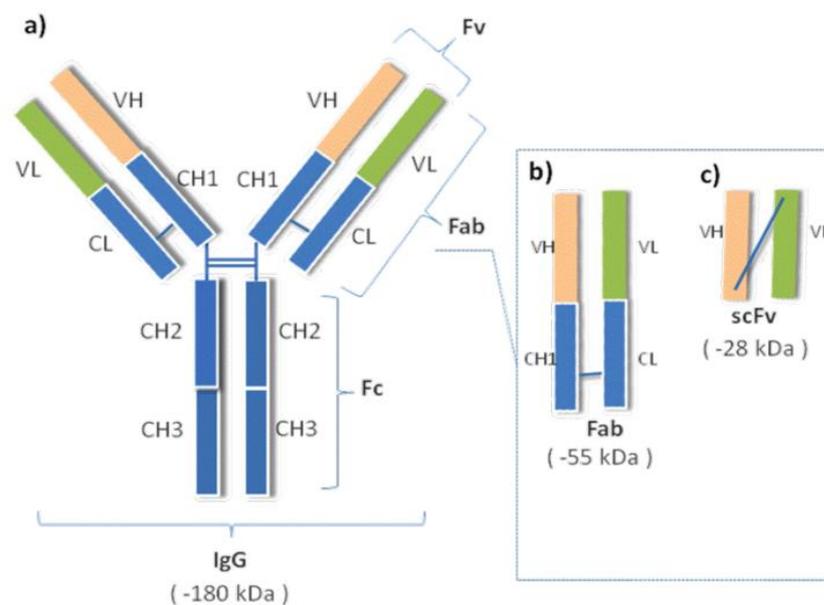
The benefits of phage display are: the rapid time to create antibodies, the use of bacterial cells to grow and clone the phage and the starting phage library containing  $\sim 10^{12}$  naïve phages to target different epitopes (Hoogenboom, 2005). This makes phage display a versatile and rapid tool for nanobody production. The naïve phage used in this project express the camelid VHH (V-like domains in camelids) fragments, which are significantly smaller than conventional IgG-based therapeutics. This allows the llama VHH expressed in phage and later, nanobodies, to bind to crevices in a membrane receptor (Holliger and Hudson, 2005). Research has shown improved solubility and enhanced penetration of camelid VHH compared to mouse VH domains of antibodies, therefore camelids would be useful to target ligands-binding domains and crevices of the extracellular surface of a GPCR (Holliger and Hudson, 2005).

Nanobodies are of current interest to lock GPCRs into particular conformations (Manglik *et al.*, 2017). Crystallisation attempts of GPCRs with nanobodies has been successful, as observed with the  $\beta_2$ AR and the high-affinity, agonistic Nb80 nanobody (Rasmussen *et al.*, 2011). Nanobodies have also been fused with GFP to probe the  $\beta_2$ AR signalling process (Irannejad *et al.*, 2013). Therapeutic nanobodies have also been developed against the CXCR4 GPCR to prevent viral activity (Manglik *et al.*, 2017).

Nanobodies have gained popularity amongst GPCR research due to their compact, spherical and smaller shape, allowing them to bind to crevices of a receptor which may be impossible for standard antibodies (Manglik *et al.*, 2017). This will be ideal for nanobody generation against family B GPCRs, such as the CGRPR, where gepant antagonists bind the extracellular peptide-binding cleft of the receptor and not within the chalice (Ter Haar *et al.*, 2010). The high variability and length of the H3 loop, or CDR3, can be advantageous for VHH nanobodies to target difficult epitopes (Manglik *et al.*, 2017; Muhyldermands, 2013). Multiple monomeric nanobodies can be joined to become highly specific to a target or can conformationally lock or constrain two proteins, such as CLR and RAMP1 (Muhyldermands, 2013). As the SMA copolymer was used to solubilise GPCRs in SMALPs, the native conformation of the receptor will enable the discovery of conformation-specific VHH-phage (Manglik *et al.*, 2017).

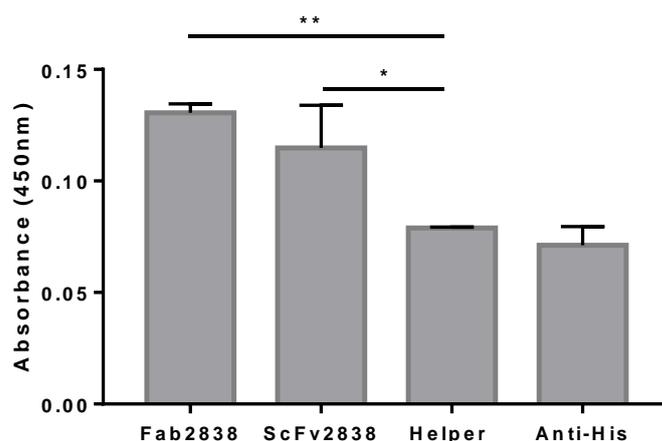
## 5.2: Results

### 5.2.1: Assessing the A<sub>2A</sub>R specific binding of phage, conjugated with the positive control Fab2838 and ScFv2838.



**Figure 50| A)** An illustration displaying the components which make up an IgG. **B)** The variable Fab-fragment of an IgG. **C)** A single-chain fragment variable. Diagram obtained from (Jara-Acevedo *et al.*, 2016).

Initial tests were conducted to observe whether the M13 bacteriophage, conjugated with either Fab2838 or ScFv2838, bound specifically to the dghA<sub>2A</sub>R. Fab2838 is a Fab-fragment from a mouse monoclonal antibody, which was specific to the A<sub>2A</sub>R (Hino *et al.*, 2012). The size of a Fab-fragment is around 55 kDa (Jara-Acevedo *et al.*, 2016). Fab2838 has previously been used in structural studies to obtain the crystal structure of the A<sub>2A</sub>R with the ICL3 intact (Hino *et al.*, 2012) and also displaying an allosteric role in antagonising the GPCR. ScFv2838 is a single-chain fragment variable (ScFv), which is a short peptide linker made from fusing the V<sub>H</sub> and V<sub>L</sub> antibody genes together (Ahmad *et al.*, 2012). ScFvs are half the size of Fab-fragments as shown in **figure 50**. The ScFv2838 was generated by UCB and also bound specifically to the dghA<sub>2A</sub>R. Both the Fab2838 and ScFv2838 fragments were conjugated to M13 bacteriophage. The phage were used to see if they bound to the dghA<sub>2A</sub>R-SMALP, which was immobilised onto a NUNC maxisorp plate, as well to see if they were detectable through indirect ELISA.

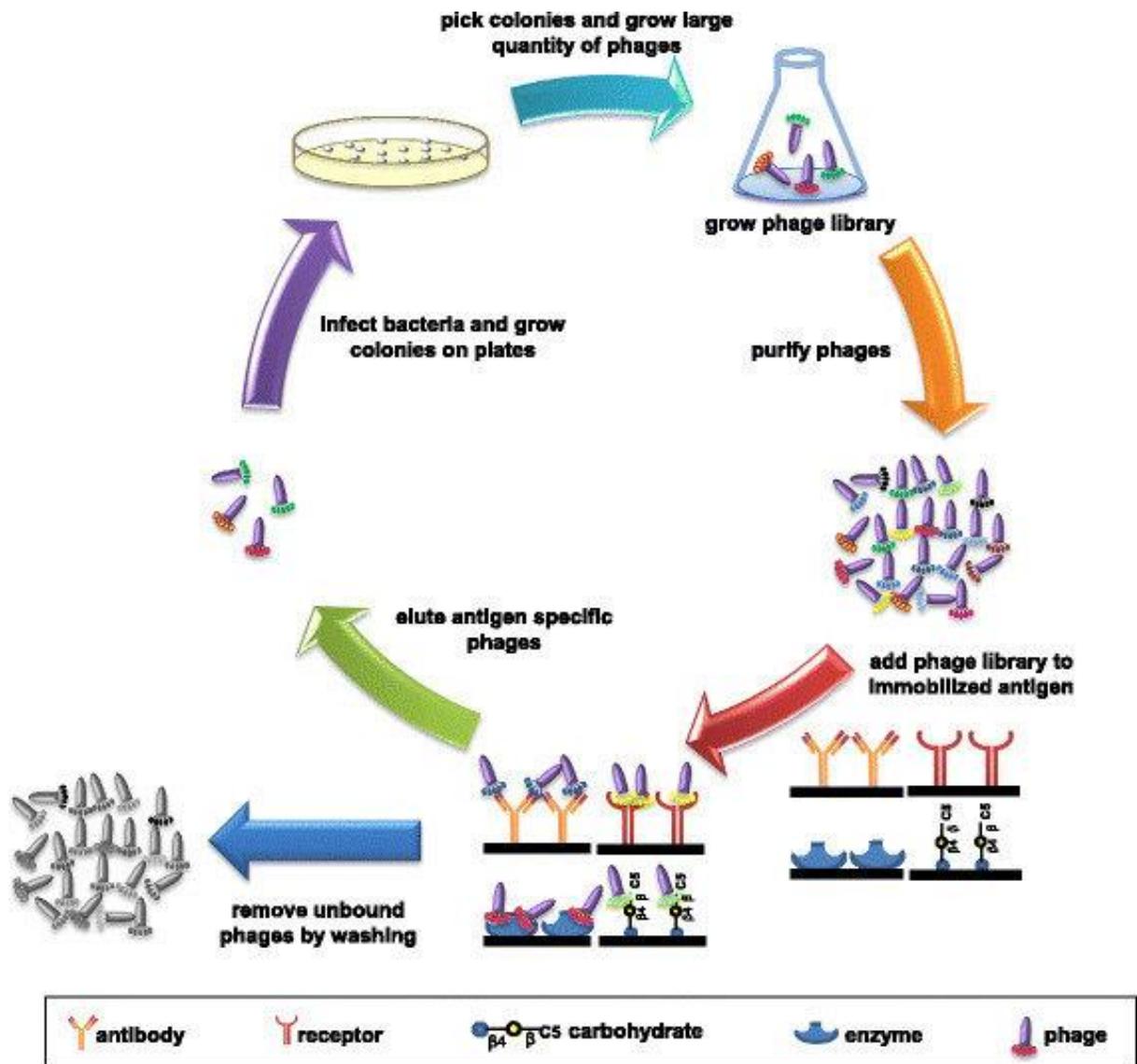


**Figure 51| M13 bacteriophage conjugated with either Fab2838 or ScFv2838 bound the SMA solubilised A<sub>2A</sub>R.** The A<sub>2A</sub>R-SMALP, at a concentration of 5 µg/ml, was captured onto a NUNC maxisorp ELISA plate. 10<sup>11</sup> phage/ml of M13 bacteriophage were incubated with the immobilised A<sub>2A</sub>R-SMALP. The phage were conjugated with either Fab2838 or ScFv2838, which were the variable fragments of an antibody. The helper M13 bacteriophage were used as a negative control. The HRP-conjugated goat anti-M13 antibody was used to bind to the M13 bacteriophage. A mouse anti-his IgG was used to bind the histidine tag of the A<sub>2A</sub>R. The secondary antibody, HRP-conjugated goat anti-mouse IgG, bound to the anti-his antibody. OPD was used as the substrate for the HRP-conjugated antibodies and the absorbance was detected at a 450 nm wavelength. The graph represents three independent experiments. The one way ANOVA test, followed by Dunnett's multiple comparisons (Graphpad Prism 7) was used to compare the means of each bar (\*, \*\*, p ≤ 0.05; p ≤ 0.01). The error bars represent the standard error of the mean.

The M13 bacteriophage conjugated with either Fab2838 or ScFv2838 showed a higher signal than the negative control helper phage condition, when binding the dghA<sub>2A</sub>R-SMALP (**figure 51**). There was no significant difference between the absorbance of the M13 bacteriophage conjugated with Fab2838 or ScFv2838 (**figure 51**). The anti-His detection was significantly lower than the M13 phage, conjugated with the antibody fragments, and was similar to the helper phage negative control. (**figure 51**).

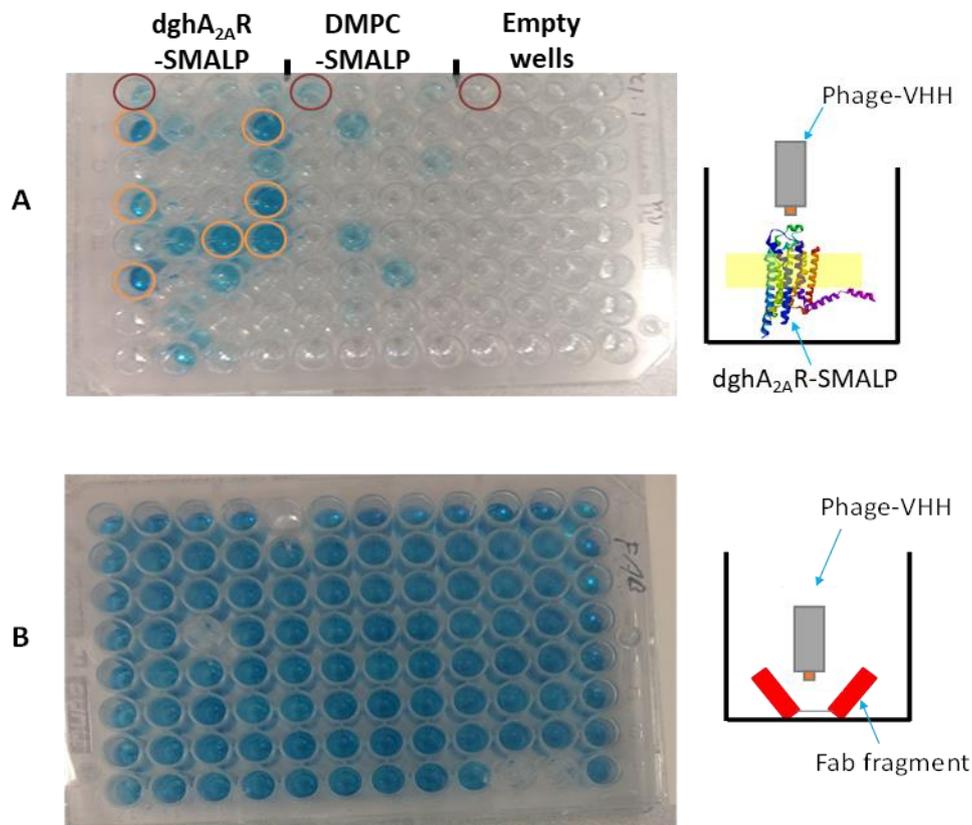
### **5.2.2: Discovering A<sub>2A</sub>R-specific, VHH-conjugated phage using phage display**

The VHH nanobodies were expressed on the M13 phage. Phage display was used to generate phage-VHH, which were specific to the dghA<sub>2A</sub>R, using the procedure in **figure 52**.



**Figure 52|** Illustration of the phage display panning strategy. Initially, the target antigens, which could be antibodies, receptors, enzymes and sugars are immobilised onto an ELISA plate. Naïve phage libraries are added to the immobilised targets for phage binding. The ELISA plate is washed to remove unbound phages. The target-specific phage are eluted and are incorporated into bacteria to expand the phage. The target-specific phage are isolated from the bacteria. The panning process is repeated two more times to ensure the phage are highly specific to the target. Diagram taken from (Wu *et al.*, 2016).

After the three biopanning rounds, the bound phage was eluted from the target SMALPs and expanded in bacteria. The phage were tested using phage ELISA to see which VHH-phage bound specifically to the A<sub>2A</sub>R (**figure 53**).



**Figure 53| Phage ELISA plates displaying phage binding to their target protein.** Naïve llama VHH phage library was incubated with target protein: dghA<sub>2A</sub>R-SMALP or Fab, which were immobilised onto a NUNC maxisorp ELISA plate. Non-binding phage were washed away during biopanning. After three biopanning rounds, phage were expanded in TG1 bacteria and single colonies containing monoclonal phage were tested in phage ELISA. HRP-linked anti-phage antibody was used to bind the phage for ELISA. **A)** Three conditions were used to identify dghA<sub>2A</sub>R-specific phage: dghA<sub>2A</sub>R-SMALP, DMPC-SMALP and empty wells washed with PBS. The A<sub>2A</sub>R-SMALP and DMPC-SMALP were immobilised onto NUNC maxisorp plates. Monoclonal phage was applied to each well of one condition, which was symmetrical for all other conditions. Polyclonal phage was also applied to the ELISA plates (red circles). Orange circles represent phage binding to the A<sub>2A</sub>R in the A<sub>2A</sub>R-SMALP condition and not the other conditions. **B)** A positive control with Fab protein immobilised onto NUNC maxisorp ELISA plates. Fab-specific phage was added to each well and phage ELISA confirmed their binding.

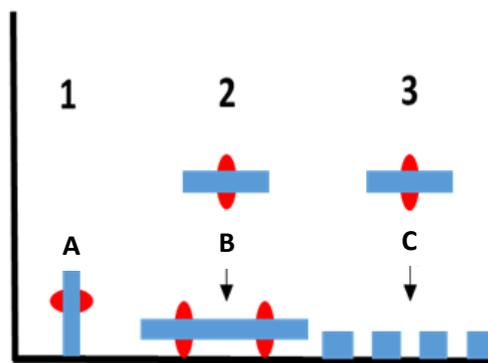
The dghA<sub>2A</sub>R-SMALP was immobilised onto NUNC maxisorp plates and served as the target antigen for the phage expressing llama VHH (which underwent biopanning rounds to become dghA<sub>2A</sub>R-specific). DMPC-SMALPs were also immobilised onto the ELISA plates to act as a negative control to detect SMALP/SMA binding phage. The final condition was wells washed with PBS to detect plastic binding of the phage and to detect background signal (**figure 53A**). The monoclonal phages were added to every well in each condition. The position of a monoclonal phage in the wells was kept the same between all three conditions. Polyclonal phage, derived from a TG1-phage pellet was applied to the first wells of each condition (red circle) and phage ELISA confirmed binding of the polyclonal phage in the A<sub>2A</sub>R-SMALP and DMPC-SMALP conditions. Many wells were bound by phage ELISA in both the A<sub>2A</sub>R-SMALP and the DMPC-SMALP conditions and no phage were detected in the PBS washed wells. However seven wells, (22.5% of the total monoclonal phage (highlighted in orange)), bound to the target but not DMPC-SMALPs and were further investigated (**figure 53A**).

For the positive control, Fab fragments were immobilised onto an NUNC maxisorp ELISA plate. Monoclonal Fab-specific phage was applied to the immobilised Fab. Monoclonal phage bound to Fab in all but one of the 48 wells (**figure 53B**).

### 5.3: Discussion

Phage display is a technique used particularly by the pharmaceutical industry to rapidly produce antibodies or nanobodies for therapeutic use against target antigens. In this case, GPCR-SMALPs were the antigens for the naïve phage library. The addition of SMALPs, which provide a native environment for the GPCRs, enables nanobody discovery against receptors in their native conformation (Wheatley *et al.*, 2016). The experiments within this chapter assessed the suitability of SMALPs for phage binding and detection. **Figure 51** demonstrated that existing anti-A<sub>2A</sub>R antibody fragments can be conjugated to the M13 phage and bind specifically to the A<sub>2A</sub>R. It was also essential for the phage to be detected during ELISA, where the anti-M13 antibodies conjugated with HRP successfully bound the dghA<sub>2A</sub>R-specific phage. The signal achieved by the dghA<sub>2A</sub>R-specific phage was greater than the negative control helper phage, suggesting that the signal obtained by the Fab2838 and ScFv2838 was greater than background levels. The signals of the anti-his antibodies however was similar to the helper phage signal, suggesting

that the absorbance achieved by the anti-his antibodies was similar to background binding. Therefore the anti-his tag antibody did not bind as efficiently as the Fab2838 and ScFv2838 conjugated bacteriophage. According to Dr A. Scott-Tucker, (UCB, Slough, UK), the expected absorbance should be around 4.5, whereas the highest signal, achieved by Fab2838 was 0.13, even though the standard concentration of protein was used in the assay. The low signal may have been due to the following issues of direct capture of SMALPs onto a NUNC maxisorp plate, as shown in **figure 54** The SMALPs could potentially bind the ELISA wells in an innapropriate orientation, prevent access for the phage to bind the receptor. The size of the SMALP could prevent unbound SMALPs from making contact with the surface of the ELISA wells. Excess SMA present in the buffer could also pre-coat the wells, limiting the properties of the NUNC maxisorp plate from immobilising the SMALP.



**Figure 54** | Potential issues of immobilising SMALPs onto a NUNC maxisorp plate for phage display. A) The SMALP could be captured in its wrong orientation, limiting the accessibility of the phage to bind the appropriate targets of the receptor. B) Steric hindrance of the immobilised SMALP may prevent additional SMALPs from being captured. C) The SMA copolymer within the buffer may pre-coat the ELISA wells, interfering with the immobilisation of the SMALPs. As NUNC maxisorp plates have a high affinity towards hydrophilic molecules (ThermoFisher Scientific [online] last accessed: 06.08.2020), the SMA copolymer could possibly be captured onto the plate through its maleic acid component.

Although a low signal was achieved by the positive control assay (**figure 51**), the dghA<sub>2A</sub>R-SMALP was detectable by the dghA<sub>2A</sub>R-specific phage. However, the differences between Fab2838 and the controls were not significant enough to demonstrate that there was enough dghA<sub>2A</sub>R-SMALPs captured onto the plate for phage display. Also, the anti-his antibody was not able to detect the his-tagged dghA<sub>2A</sub>R-SMALP, which may be due to electrostatic interactions formed by the his-tag and the maleic acid of the SMALP. Nevertheless, an attempt was made to discover dghA<sub>2A</sub>R-specific VHH-conjugated phage. From the phage biopanning rounds and the phage ELISA, seven monoclonal VHH expressing phage were detected, which bind to the dghA<sub>2A</sub>R (**figure 53A**). Two negative controls were used in phage ELISA: immobilised DMPC-SMALPs and empty wells washed with PBS. The negative controls were used to identify phage binding to the SMA and lipid as well as plastic-binding phage. No monoclonal phages bound to the plastic as shown in the PBS negative control (**figure 53A**), however many phages bound to the DMPC-SMALP and so were eliminated from further study. The seven monoclonal phages binding to the dghA<sub>2A</sub>R did not bind to DMPC-SMALPs nor the PBS, hence the phages were specific to the dghA<sub>2A</sub>R only. Phages specific to Fab-fragments were also produced in parallel to phages against the dghA<sub>2A</sub>R (**figure 53B**). The positive control Fab-fragments were immobilised on ELISA plates and were used to test if the phage display was working according to the specifications provided by UCB. 96% of the phage were specific to the Fab-fragment, which was greater than the amount of phage specific to the dghA<sub>2A</sub>R, which was 22% of the total naïve phage used. According to UCB, a highly successful phage display yields an 80% enrichment of phage specific to the target antigen. The lower percentage of dghA<sub>2A</sub>R-specific phage may have been due to the immobilisation issues illustrated in **figure 54** as well as the wash cycles in the biopanning rounds, which may have removed the SMALPs from the ELISA plate. Avi-tagged GPCR constructs have been designed to eliminate the immobilisation issues associated with the SMALPs, which will be described in **chapter 6**. The phage which showed specificity for the dghA<sub>2A</sub>R was reformatted into a nanobody, without the phage intact, within UCB. The nanobodies did bind the A<sub>2A</sub>R, however they also displayed non-specific binding to phospholipids. This could likely be due to the negative control being DMPC-SMALPs and not SMALPs from a wild type yeast membrane.

## Chapter 6: The biotinylation and immobilisation of the avi-tagged A<sub>2A</sub>R-A.

### 6.1: Introduction

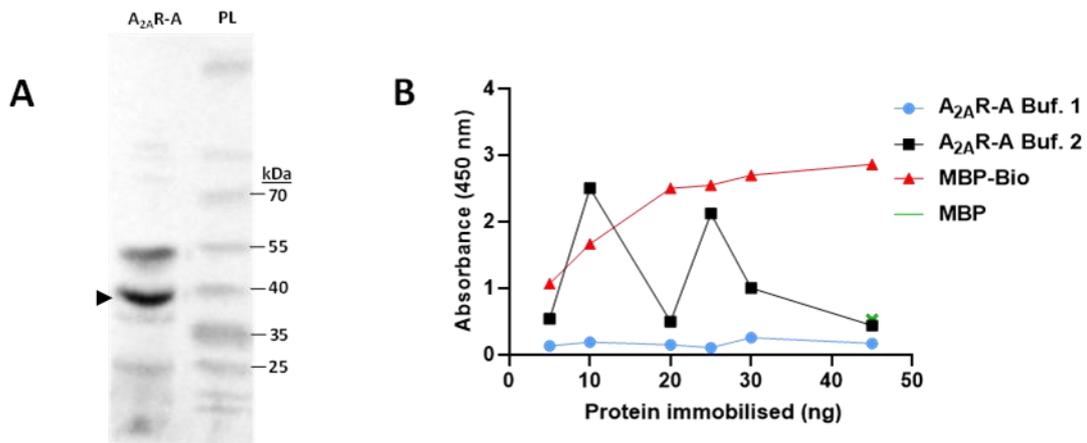
Due to the low enrichment of dghA<sub>2A</sub>R-specific VHH-phage in chapter 6, an avi-tagged A<sub>2A</sub>R construct was designed to enhance the immobilisation of the receptor onto ELISA plates for phage display. Site-specific protein biotinylation is often preferred for the immobilisation of proteins as fusing biotin to a protein is less intrusive to the proteins structure than methods involving antibody or GFP capture (Graslund *et al.*, 2017). A recent phage display-focussed study demonstrated the benefits of biotinylating a protein-nanodisc via an avi-tag, with a 90% efficiency in streptavidin pull-down for immobilisation (Dominik *et al.*, 2016). The avi-tag consists of the amino acid residues: GLNDIFEAKIEWHE and is a highly specific substrate for the enzyme, biotin ligase (BirA). BirA is an enzyme found in *Escherichia coli*, which can modify specific peptides of a protein sequence by adding biotin (Fairhead and Howarth, 2015). Biotinylation of an avi-tag is irreversible, where proteolytic conditions are the only way to break the biotin-avitag complex (Fairhead and Howarth, 2015). The SMA solubilised A<sub>2A</sub>R-A was biotinylated for immobilisation on a streptavidin-coated ELISA plate. Streptavidin is 60 kDa protein, originating from *Streptomyces avidinii*. It is known for its high affinity for binding biotin, with a K<sub>D</sub> of 4x10<sup>-14</sup> M (Fairhead and Howarth, 2015). This interaction is rapid and specific and the high affinity of the biotin-streptavidin interaction makes it ideal for SMALP immobilisation (Fairhead and Howarth, 2015, Holmberg *et al.*, 2005).

A flexible 15 amino acid poly-GS linker was incorporated into the C terminus of the A<sub>2A</sub>R-A, positioned before the avi-tag sequence. This was to prevent the steric hindrance between the tag and the A<sub>2A</sub>R-A-SMALPs, which could reduce the amount of immobilised SMALPs. Experiments have shown that different linker sizes with a biotinylated avi-tag at the end of the C-terminus can affect immobilisation, where the longer the linker sequence, the better the immobilisation (Ikonomova *et al.*, 2018). This could be due to steric hindrance of the biotin being too close to the protein, preventing the access of biotin to bind streptavidin (Ikonomova *et al.*, 2018).

## 6.2: Results

### 6.2.1: Assessing whether the biotinylated A<sub>2A</sub>R-A-SMALP can interact with streptavidin.

An indirect ELISA assay was used to test whether the directly immobilised, biotinylated A<sub>2A</sub>R-A-SMALP can interact with free streptavidin- alkaline phosphatase (SV-AP). pNPP was used as the substrate for the alkaline phosphatase, producing a yellow colour upon catalysis. Avi-tagged myelin basic protein (MBP) was bought from Avidity and was used as a positive control (MBP-biotin) and a negative control (MBP). The assay investigated whether the biotinylated A<sub>2A</sub>R-A-SMALP can interact with streptavidin, as well as the extent of biotin-streptavidin interaction when compared with the controls.



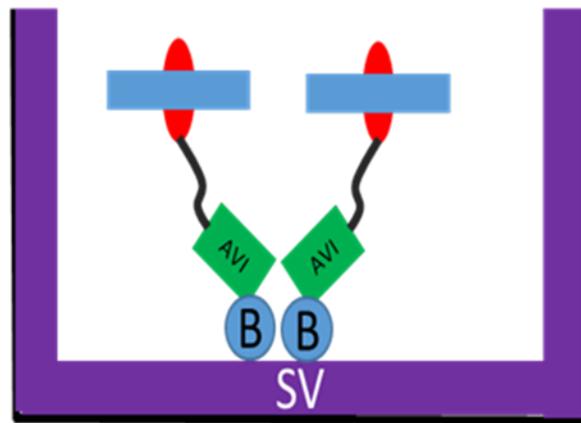
**Figure 55| The A<sub>2A</sub>R-A-SMALP was biotinylated through its avi-tag and can interact with streptavidin.** **A)** A western blot illustrating the presence of the avi-tag on the A<sub>2A</sub>R-A. The X33 yeast membranes expressing A<sub>2A</sub>R-A were used for the western blot. The primary antibody was the mouse anti-avi-tag IgG. The secondary antibody was the HRP-conjugated goat anti-mouse IgG. EZ-ECL was used as the substrate for chemoluminescence. The arrowhead indicates the band representing the A<sub>2A</sub>R-A. **B)** Graph illustrating the amount of SV-AP complexing with the biotin of the A<sub>2A</sub>R-A-SMALPs. The biotinylation reaction was conducted in either the avidity buffer (buffer 1), or the manually constructed buffer (buffer 2). The avi-tagged MBP-biotin was used as a positive control. 45 ng of MBP was used as the negative control, as it was not biotinylated. pNPP was used as the substrate for SV-AP. After 20 minutes of pNPP incubation, the ELISA plate was read at a 450 nm wavelength. The graph is based on one experiment.

The avi-tag was confirmed to be present on the A<sub>2A</sub>R-A, when expressed in *Pichia pastoris* membranes. Multiple bands were detected in the western blot (**figure 55A**), however the densest band was around 37 kDa. Another dense band was detected at 55 kDa two faint bands above the 70 kDa marker. There were also faint bands below the 37 kDa point, which may be degradative products (**figure 55A**). After solubilisation of the A<sub>2A</sub>R-A-expressing membrane with SMA2000, nickel resin was used to purify the A<sub>2A</sub>R-A-SMALPs. The Bir500 biotin ligase was used to biotinylated the avi-tag of the A<sub>2A</sub>R-A. A pNPP indirect ELISA assay was used to assess the formation of biotin-streptavidin complexes, where the streptavidin was conjugated with alkaline phosphatase (SV-AP) (**figure 55B**). The absorbance signal represented the amount of biotin-SV-AP complexes formed. The MBP-biotin positive control displayed a hyperbolic curve, where increasing the amount of MBP-biotin onto the ELISA plate increased the amount of biotin-streptavidin complexes being formed (**figure 55B**). Saturation was reached after 30 ng of the positive control was applied to the wells.

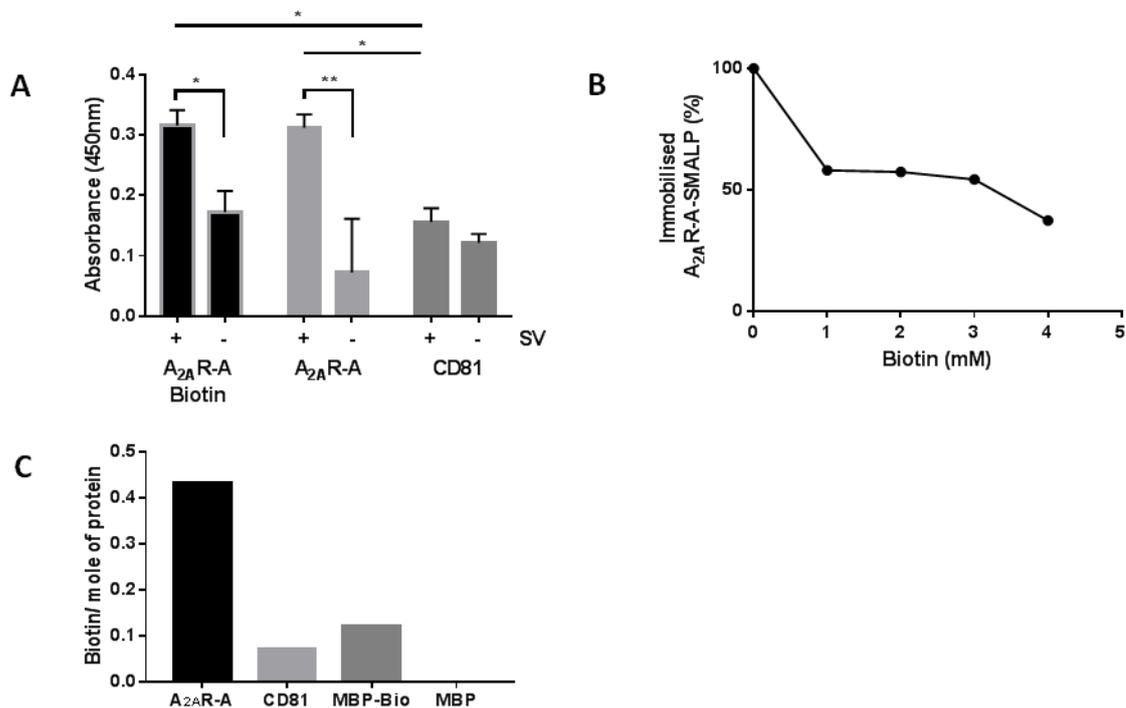
The biotinylated A<sub>2A</sub>R-A-SMALP in the avidity-supplied buffer (buffer 1) showed a flat signal, below the MBP negative control (**figure 55B**). The increasing trend with MBP-biotin was not observed with the biotinylated A<sub>2A</sub>R-A-SMALP, instead a broadly biphasic trend was observed (**figure 55B**). The highest signal was observed when 10 ng of SMALP was captured onto the NUNC maxisorp plate, where the absorbance signal was slightly higher than 10 ng of MBP-biotin (**figure 55B**). Fluctuations of the absorbance were observed between 10 ng and 25 ng of SMALP capture. Further increasing beyond 25 ng of SMALP capture reduced the amount of biotin-SV-AP complexes formed (**figure 55B**). The lowest signal observed for the A<sub>2A</sub>R-A-SMALP was when 5 ng and 45 ng was captured onto the ELISA plate, where the signal was similar to the negative MBP control, indicating that no biotin-SV-AP complexes were formed.

### 6.2.2: The immobilisation of the biotinylated A<sub>2A</sub>R-A-SMALP

Once the biotin-streptavidin interaction was confirmed (**figure 55B**), the A<sub>2A</sub>R-A-SMALP was immobilised onto streptavidin-coated NUNC maxisorp plates. The illustration in **figure 56** explains the theory of the immobilisation of the A<sub>2A</sub>R-A-SMALP.



**Figure 56** | Wells of the NUNC maxisorp ELISA plate were pre-coated with streptavidin (SV). The purified, biotinylated A<sub>2A</sub>R-A-SMALP can be immobilised onto the streptavidin-coated plate, where the biotin forms a complex with streptavidin. The biotin was conjugated to the cytoplasmic, C terminal avi-tag. A flexible GS linker upstream the avi-tag provides flexibility to the SMALP to reduce steric hindrance. The immobilised A<sub>2A</sub>R-A-SMALP can be exposed to antibodies or phage.



**Figure 57| Streptavidin-coated ELISA wells enhanced the immobilisation of the A<sub>2A</sub>R-A-SMALP.**

**A)** Bar chart illustrating the immobilisation of purified protein-SMALPs onto a NUNC maxisorp ELISA plate. The wells were either coated with SV (+) or were covered with PBS (-). The biotinylated A<sub>2A</sub>R-A-SMALP and the A<sub>2A</sub>R-A-SMALP were captured onto the wells. CD81-SMALP was used as a negative control due to the non-existence of the avi-tag. 4 µg/ml of the SMALPs were pipetted onto the wells. The graph represents three independent experiments, with the error bars representing the standard error of the mean. The two-way ANOVA statistical test, follow by Sidak's multiple comparisons test (Graphpad Prism 7) was used to compare the mean of each bar (\* p ≤ 0.05; \*\* p ≤ 0.01). **B)** Graph illustrating the displacement of the biotinylated A<sub>2A</sub>R-A-SMALP when increasing concentrations of free biotin was added to the streptavidin-coated ELISA wells. Prior to SMALP immobilisation, free biotin at different concentrations was applied to the streptavidin-coated wells and was incubated for 1 hour at room temperature. The supernatant was washed from the wells and 2 µg/ml of the A<sub>2A</sub>R-A-SMALP was added to each well. For panels A and B, the mouse anti-his IgG was used to bind the histidine tags of the proteins and the HRP-conjugated goat anti-mouse antibody targeted the mouse IgG. OPD was used as the substrate for the HRP, which was detected at a 450 nm wavelength. Graph is representative of one experiment. **C)** The A<sub>2A</sub>R-A-SMALP showed high levels of interaction than the CD81-SMALP negative control and the MBP-bio positive control. The A<sub>2A</sub>R-A-SMALP and the CD81-SMALP were mixed in a 1 ml cuvette tube containing HABA/avidin. The MBP-biotin was the positive control and the MBP was the negative control, also mixed in a cuvette tube containing HABA/avidin. Absorbance was read at a 500 nm wavelength. The HABA calculator website from Thermofisher Scientific was used to convert the absorbance to biotin/ mole of protein. Bar chart is from one experiment.

The A<sub>2A</sub>R-A-SMALP, biotinylated and non-biotinylated, showed a significantly higher signal than the CD81-SMALP control, when immobilised onto SV-coated wells (**figure 57A**). The immobilisation of the biotinylated A<sub>2A</sub>R-A-SMALP was more than double the immobilisation of the CD81-SMALP on the streptavidin-coated wells (**figure 57A**). There was however no significant difference between the CD81-SMALP negative control and the biotinylated A<sub>2A</sub>R-A-SMALP, when observing the wells that were not coated with streptavidin. Unexpectedly, The A<sub>2A</sub>R-A-SMALP demonstrated a higher level of immobilisation on streptavidin-coated wells than the wells that were left blank (**figure 57A**). Furthermore, a higher signal was observed for the A<sub>2A</sub>R-A-SMALP than the CD81-SMALP in the streptavidin conditions. This indicated that the biotinylation of the A<sub>2A</sub>R-A-SMALP, using BirA500 biotin ligase (from Avidity), did not make a significant difference to SMALP immobilisation when comparing the absorbance with the non-biotinylated A<sub>2A</sub>R-A-SMALP. After pre-incubating the streptavidin-coated wells with increasing concentrations of free D-biotin, the amount of immobilised biotinylated A<sub>2A</sub>R-A-SMALP declined (**figure 57B**). Almost half of the total SMALPs were not captured by streptavidin when 1 mM of Biotin was pre-incubated in the ELISA plate. However increasing the concentration of the biotin to 3 mM did not cause further reduction in SMALP immobilisation (**figure 57B**). Applying 4 mM free biotin to the streptavidin wells further reduced the immobilisation of SMALPs (**figure 57B**). The amount of biotin molecules attached to the A<sub>2A</sub>R-A was analysed using the HABA/avidin assay (**figure 57C**). The A<sub>2A</sub>R-A-SMALP displayed a higher level of biotinylation than the MBP-Bio positive control. The CD81-SMALP also showed a signal indicating biotinylation however the values were lower than the positive control and therefore may be a background result (**figure 57C**).

### 6.3: Discussion

Due to the low enrichment of A<sub>2A</sub>R-specific phage in chapter 6, an attempt to rectify the immobilisation issue was conducted by modifying the A<sub>2A</sub>R construct. A flexible linker sequence and an avi-tag was fused to the C terminus of the A<sub>2A</sub>R construct to enhance immobilisation of the receptor-SMALP, as previous research had used nanodiscs with proteins containing a flexible linker and an avi-tag with success in immobilisation (Ikonomova *et al.*, 2018). The pPICZ $\alpha$  A-A<sub>2A</sub>R-A construct was successfully integrated into the genome of X33 *Pichia pastoris*, as shown in **chapter 3**, and the avi-tag was also present in the C terminus of the receptor (**figure 55A**). After solubilisation of the A<sub>2A</sub>R-A using SMA2000 and purification using affinity chromatography, the A<sub>2A</sub>R-A-SMALP was biotinylated using the Bir500 Avidity biotinylation kit. Initially the original buffers, supplied by Avidity, were used to biotinylate the A<sub>2A</sub>R-A-SMALP, however only background signal was observed (**figure 55B**). The final reaction mixture from the buffers provided by Avidity lead to the final concentrations: 10 mM ATP, 10 mM Mg(OAc)<sub>2</sub> and 50  $\mu$ M d-Biotin, using the protocol from Avidity (Avidity [online] last accessed 9.12.2019). As the final reaction contained 10 mM Mg<sup>2+</sup>, SMALP precipitation was likely to occur as the maximum Mg<sup>2+</sup> concentration SMA could tolerate is 4 mM (Oluwole *et al.*, 2017). The Avidity buffers were replaced with ingredients (**Chapter 2, table 16**) to counteract the negative effects of Mg<sup>2+</sup> on the SMALPs, where the reaction mixture contained: 2.5 mM MgCl, 10 mM ATP and 0.3 mM D-Biotin. The reaction condition was optimised based on the procedure used in Li and Sousa, (2012). There was some success in the interaction of biotin and streptavidin, when biotin was conjugated to the receptor, where the amount of biotin-streptavidin complexes were initially similar to the positive control (MBP-biotin) when low amounts of receptor were used to coat the plate (**figure 55B**). However, adding more than 10 ng of biotinylated A<sub>2A</sub>R-A-SMALP to an ELISA well reduced the amount of biotin-streptavidin complexes formed, when comparing the signals with the positive control (**figure 55B**). As the SMALPs were directly captured onto the NUNC maxisorp ELISA plate, steric hindrance and the washing steps may be responsible for the deficit in biotin-streptavidin complexes being formed, which may explain the low enrichment of A<sub>2A</sub>R-specific phage in **chapter 5**.

Once the biotin-streptavidin complexes have been confirmed, the immobilisation assay was conducted to assess whether the biotinylated A<sub>2A</sub>R-A-SMALP can be captured onto the NUNC maxisorp ELISA plate, pre-coated with streptavidin. The data has demonstrated that adding streptavidin to the ELISA plate improved the capture of the A<sub>2A</sub>R-A-SMALP (**figure 57A**). Interestingly, the biotinylation of the avi-tag using Bir500 did not make a significant difference to the immobilisation of the A<sub>2A</sub>R-A-SMALP onto streptavidin wells (**figure 57A**). However, the incorporation of the avi-tag may have made a difference to A<sub>2A</sub>R-A-SMALP immobilisation, which was significantly higher than the non-avi-tagged CD81-SMALPs. This may suggest that a biotin ligase enzyme could possibly be present in the *Pichia pastoris* that may cause in vivo biotinylation of the A<sub>2A</sub>R-A. Alternatively, the A<sub>2A</sub>R-A-SMALP may form non-specific interactions with the streptavidin-coated wells, as indicated by the high level of immobilisation of the non-BirA500 biotinylated A<sub>2A</sub>R-A-SMALP to the streptavidin wells (**figure 57A**). The HABA/avidin biotin quantitation assay was incorporated into this study to observe the extent of biotinylation of the A<sub>2A</sub>R-A-SMALPs, prior to BirA500 treatment, to determine whether the A<sub>2A</sub>R-A was endogenously biotinylated in *Pichia pastoris* (**figure 57C**). The A<sub>2A</sub>R-A-SMALP showed higher levels of biotinylation than the MBP-biotin positive control, which may indicate the GPCR being biotinylated when expressed in *Pichia pastoris*. However as the experiment was conducted once, further investigation is required to obtain solid evidence of protein biotinylation in the yeast strain. The biotin titration data in **figure 57B** has demonstrated that the A<sub>2A</sub>R-A-SMALP was biotinylated by the Bir500, since increasing the concentration of free biotin prevented the immobilisation of the biotinylated GPCR-SMALP. This is due to the free biotin occupying the streptavidin and as this reaction is almost non-reversible, this will most likely prevent the immobilisation of the biotinylated A<sub>2A</sub>R-A-SMALP. An overall improvement of the immobilisation of the A<sub>2A</sub>R-SMALP was observed with the avi-tagged construct, which should improve the enrichment of A<sub>2A</sub>R-specific VHH-phage during phage display.

## **Chapter 7: The expression of the CGRP receptors in HEK293T mammalian cells**

### **7.1: Introduction**

The CGRPR-avi-tag (CGRPR-A) was developed to immobilise the receptor using biotin and streptavidin, in order to use the receptor for phage display. In this chapter, the CGRPR-A, consisting of the CLR-avi-tag (CLR-A) and RAMP1 were expressed in mammalian cells. The receptor was solubilised using SMA2000 to form CGRPR-A-SMALPs, with the surrounding mammalian lipid bilayer intact, maintaining the receptor in its native conformation. A functional assay and western blotting was used to detect the presence of the tags of the receptor and to compare its function to the T8-HA-CLR, used in previous studies (Barwell *et al.*, 2012; Simms *et al.*, 2018).

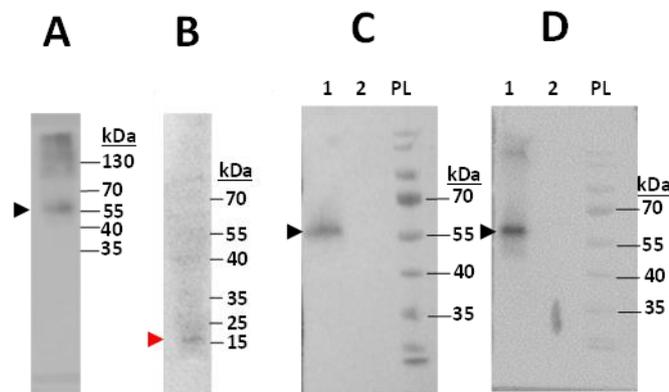
Human embryonic kidney (HEK) cells have proven invaluable in the expression of recombinant proteins, as its internal machinery allows the cell to process the post-translational modifications of human proteins appropriately (Thomas and Smart, 2004). Compared to alternative mammalian cells, the HEK293 cells are often the first choice for protein expression (Thomas and Smart, 2014). The cell line can be transiently transfected with a plasmid constructs, which is beneficial for short-scale growth and analysis of multiple mutant proteins (Thomas and Smart, 2004). Transient transfection protocols are often quick to express the protein of interest within HEK293 cells. The HEK293 cell lines can also be stably transfected for large scale growth of cells and high yield production of recombinant proteins (Liu *et al.*, 2014). However creating stable cell lines can be time consuming due to the selection of clones with the integrated gene of interest (Liu *et al.*, 2014).

HEK293T cells were used to transiently express CLR or CLR-avi-tag (CLR-A) with RAMP1. The cell line was chosen due to previous work with the CGRPR, which was successfully expressed in previous literature (Simms *et al.*, 2018; Zhu *et al.*, 2020). There was also no reported endogenous expression of the CGRPR within HEK293T cells, which was ideal for expression of recombinant CGRPR (McLatchie *et al.*, 1998). Also established in HEK293T cells was that the CLR cannot be expressed without RAMP1, which is advantageous for purifying the CGRPR, knowing that the CLR required RAMP1 for its expression (McLatchie *et al.*, 1998).

## 7.2. Results

### 7.2.1: Expression and solubilisation of CGRPR-A from HEK293T cell membranes

HEK293T cells were transiently transfected with pCDNA3.1+- T8-HA-CLR-A and pCDNA3.1+- FLAG- RAMP1, using Lipofectamine 3000. After two days of growth, the cells were homogenized and solubilised using SMA2000. The expression of the CGRPR-A was observed by western blotting between the membranes and after solubilisation.

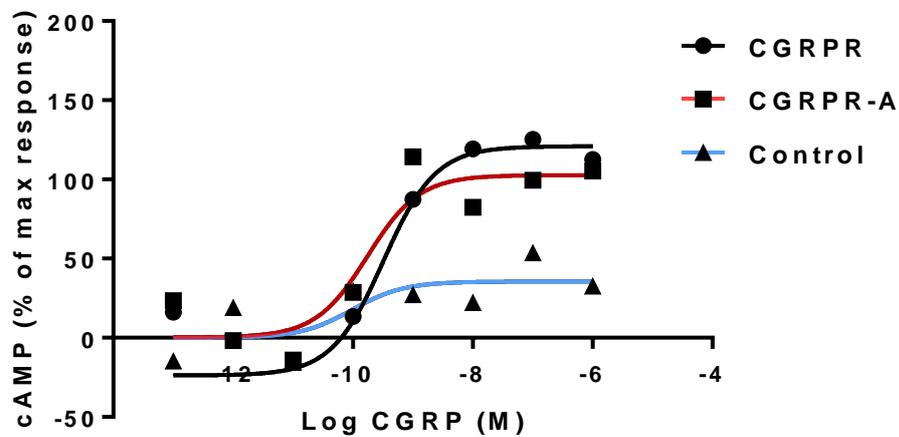


**Figure 58| Expression of the CGRPR-A in transiently transfected HEK293T cell membranes and in SMALPs.** Panels **A** and **B** are western blots representing the expression of **A)** T8-HA-CLR-A and **B)** FLAG-RAMP1 in HEK293T cell membranes. **A)** The primary mouse anti-HA antibody was used to detect the CLR-A and the HRP-conjugated goat anti-mouse antibody targeted the primary antibody. **B)** The primary rabbit anti-FLAG antibody was used to detect RAMP1 and the HRP-conjugated goat anti-rabbit antibody targeted the primary antibody. Panels **C** and **D** are western blots representing the presence of the CLR-A after solubilising the homogenized HEK293T cell membranes using 2.5% (w/v) SMA2000. **C)** The primary mouse anti-HA antibody was used to detect the CLR-A and the HRP-conjugated goat anti-mouse antibody targeted the primary antibody. 1 represents the CGRPR-A-SMALP. 2 represents the non-transfected HEK293T membrane-SMALP. PL represents the protein ladder. **D)** The primary mouse anti-avi-tag antibody was used to detect the CLR-A and the HRP-conjugated goat anti-mouse antibody targeted the primary antibody. 1 represents the CGRPR-A-SMALP. 2 represents the non-transfected HEK293T membrane-SMALP. PL represents the protein ladder. EZ-ECL was used as the substrate for HRP. The western blots represent three independent experiments. For all gels, the black arrowheads indicate the bands which represent the CLR-A and the red arrowhead indicates the band representing the FLAG-RAMP1.

The T8-HA-CLR-A and FLAG-RAMP1 constructs were transiently co-transfected into HEK293T cells using lipofectamine 3000. The membrane was extracted from the cells and was resuspended to 80 mg/ml, using buffer A. The expression of the CGRPR-A was confirmed in **figure 58**, where a band was detected at around 55 kDa (**figure 58A**), representing the CLR-A. RAMP1 was also expressed alongside CLR-A, where a band was detected at around 15 kDa (**figure 58B**). After confirmation of CGRPR-A expression, the membranes were solubilised using SMA2000 to form SMALPs. CLR-A was confirmed to be in the SMALPs by western blotting via the HA tag on CLR-A, where a band was detected at 55kDa (**figure 58C**), similarly to **figure 58A**. The C terminal avi-tag of the CLR-A was also confirmed to be intact with the SMA solubilised receptor, where a band was detected at 55 kDa, using an anti-avi-tag antibody (**figure 58D**). No bands were detected in the membrane-SMALP control (**figures 58C and 58D**).

### 7.2.2: cAMP production upon CGRP binding to CGRPR and CGRPR-A

To test the functionality of the CGRPR-A, the LANCE cAMP accumulation ELISA assay was used to detect cAMP release upon ligand incubation. This was to confirm whether the modified CGRPR-A construct was functional upon ligand binding. The functionality of CGRPR-A was compared with the CGRPR construct, consisting of T8-HA-CLR and FLAG-RAMP1, and non-transfected HEK293T cells to test the extent of CGRPR-A functionality



**Figure 59** | Dose-response curves of CGRPR and CGRPR-A showed an increased cAMP stimulation upon increasing the  $\alpha$ CGRP concentration. The sigmoidal dose-response curves were compared between the CGRPR and the CGRPR-A, which were expressed in HEK293T cells. The non-transfected HEK293T cells were used as a control. The curves represent cAMP stimulation after activation of the receptors using  $\alpha$ CGRP. Graph is representative of one experiment.

The cAMP accumulation ELISA assay was used to detect the functional response of the CGRPR and the CGRPR-A (**figure 59**). The CGRPR and the CGRPR-A showed a high cAMP stimulation after receptor activation by the  $\alpha$ CGRP. At a  $10^{-6}$  M concentration of the peptide, the  $E_{max}$  of the CGRPR and the CGRPR-A were both similar to that of 10  $\mu$ M forskolin. However CGRPR-A displayed an approximate 20% decrease in its  $E_{max}$  when compared with the CGRPR. The  $pEC_{50}$  values for the

receptors were similar, being: 9.5 for the CGRPR and 9.8 for the CGRPR-A. There was some cAMP stimulation detected for the control cells, which increased as the concentration of  $\alpha$ CGRP increased. However the percentage of the maximum forskolin response was 75-95% lower than the CGRPR and CGRPR-A.

### 7.3: Discussion of chapter

The CGRPR-A was modified with an avi-tag for downstream applications of the construct, such as biotinylation for streptavidin immobilisation and phage display, as shown in **chapter 2**. Once the construct was manufactured by Genscript, the gene was expressed transiently in HEK293T cells. The expression of the construct in the cell membrane was confirmed by western blotting (**figure 58A and 58B**). As the SDS present in the sampling buffer and the gel break the tertiary structures of proteins, the presence of the CLR-A band would be present as a monomer and not as part of the CGRPR-A (Jensen, 2002). The CLR-A was approximately 60 kDa in size as shown by the band in the blot (**figure 58A**), which is similar to the tagged CLR detected in Buhlmann *et al.*, (2000) and Hilairet *et al.*, (2001). Glycosylation events, such as N-linked glycosylation, can occur during the maturation of the CLR through the endoplasmic reticulum and the golgi apparatus (Hilairet *et al.*, 2001). The CLR-A in this project was likely to be in its glycosylated form, as the Mw of the CLR-A (~60 kDa) (**figure 58**) was close to the 66 kDa myc-tagged, N-linked glycosylated CLR in Hilairet *et al.*, (2001). Deglycosylation of the myc-tagged, glycosylated CLR reduced the size of the receptor to 50 kDa (Hilairet *et al.*, 2001). Research has shown that there are three putative, extracellular glycosylation sites in the CLR, with Asn123 being essential for ligand binding and signalling (Kamitani and Sakata. 2001). RAMP 1 allows the transit of the CLR through the Golgi apparatus, allowing the glycosylation of the receptor (Hilairet *et al.*, 2001). The RAMP1, which was co-expressed with the CLR-A, was detected at approximately 15-20 kDa (**figure 58B**). A band was detected just above the 15 kDa marker, where literature has confirmed the size of a monomeric RAMP1 to be around 15 kDa (McLatchie *et al.*, 1998). According to Buhlmann *et al.* (2000), RAMP1 adds 10 kDa to the size of the CLR.

After solubilisation of the CGRPR-A from the HEK193T cell membranes using the SMA copolymer, western blotting was used to confirm the presence of the CGRPR-A in the SMALP as well as the C terminal avi-tag, which was important for downstream experiments. The CLR-A was confirmed to be present in the SMALP in **figure 58C**, where the band around 55 kDa was detected. This was similar to the size of CLR in literature (Buhlmann *et al.*, 2000; McLatchie *et al.*, 1998) and the CLR-A expressed in the cell membrane (**figures 58A and 58C**). The anti-avi-tag western blot confirmed the C terminal avi-tag to be intact to the CLR, as shown by the band around the 55 kDa in **figure 58D**. The experiments for **figure 58** therefore demonstrated the successful expression and solubilisation of the CGRPR-A and the presence of the avi-tag allows biotinylation of the receptor for phage display.

The functionality of the CGRPR and the CGRPR-A was also assessed in HEK293T cells, using a cAMP assay to observe cAMP stimulation upon agonist binding (**figure 59**). The concentration-response curves for CGRPR and CGRPR-A were similar, with the pEC<sub>50</sub> being close in value for the constructs. This showed that the modification of the CLR-A did not impede the receptors ability to bind  $\alpha$ CGRP and its activity. The pEC<sub>50</sub> of the CGRPR is often reported to be around 9.5 (Hay *et al.*, 2018; Salvatore *et al.*, 2009), which is similar to the pEC<sub>50</sub> values obtained for the CGRPR and CGRPR-A used in this project (**figure 59**). Therefore the dose-response curves and the binding affinity of the agonist to the CGRPR constructs were concordant with previous literature. The non-transfected controls showed a slight increase of cAMP stimulation upon increasing the concentration of the peptide agonist, which may indicate endogenous expression of the receptor. This may contradict the literature which state that the HEK293T cells do not express CGRP or the calcitonin subfamily of receptors (McLatchie *et al.*, 1998). As the data was based on one experiment, further repeats are needed to see if the observations are consistent.

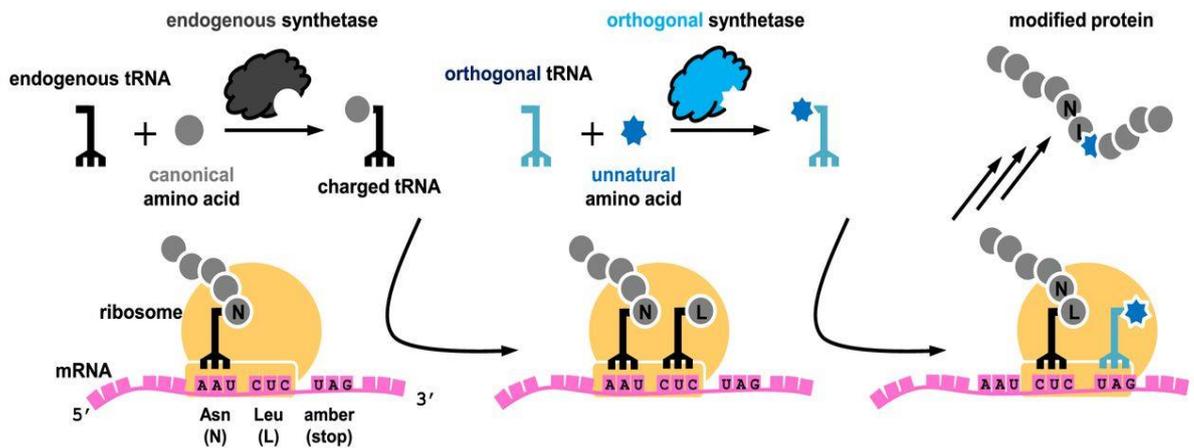
As the expression and solubilisation of the CGRPR-A has been confirmed by western blotting, and its activity being similar to the CGRPR used in previous studies, the modified receptor construct can be used for further experiments, such as receptor immobilisation for phage display.

## Chapter 8: Analysing ligand-receptor binding interactions using photoaffinity cross-linking

### 8.1: Introduction

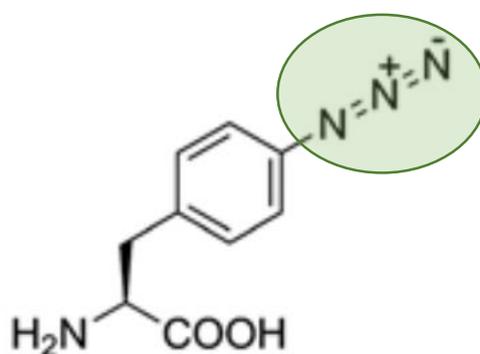
Photoaffinity cross-linking can be used to identify amino acid residues of a receptor which interact with its ligand (Grunbeck and Sakmar, 2013). The technique was implemented in this study to assess the binding topology of ECLs 1 and 3 of the CGRPR, when bound to the fluorescent CGRP. The technique is also useful to lock the receptor in an active conformation, which could then be SMA solubilised for phage display

The methodology involves the introduction of a cross-linking unnatural amino acid, such as p-azido-phenylalanine (AzF), into the protein sequence of the receptor (Wang *et al.*, 2009; Grunbeck and Sakmar, 2013; Simms *et al.*, 2018). This is achieved by creating a tRNA and its complementary aminoacyl synthetase to recognise a termination codon and incorporating an unnatural amino acid upon translation (**figure 60**) (Wang *et al.*, 2009; Nodling *et al.*, 2019). The least frequently used termination codon in mammalian cells is the amber stop codon, TAG/ UAG, which is frequently used in unnatural amino acid incorporation studies (Nodling *et al.*, 2019; Schmied *et al.*, 2014).



**Figure 60** For the translation of an endogenous protein, the tRNA recognises specific codons of the mRNA sequence in order to incorporate the correct amino acid into the protein sequence. An endogenous aminoacyl synthetase recognises its complementary tRNA in order to charge it with the correct amino acid. In an endogenous system, an amber stop codon present in the mRNA sequence will terminate translation. For unnatural amino acid incorporation, orthogonal tRNA and synthetases can be designed to recognise a stop codon and apply an unnatural amino acid to the translated sequence instead of termination. After translation has been completed, the resulting protein sequence will be modified with the unnatural amino acid. Diagram taken from (Nodling *et al.*, 2019).

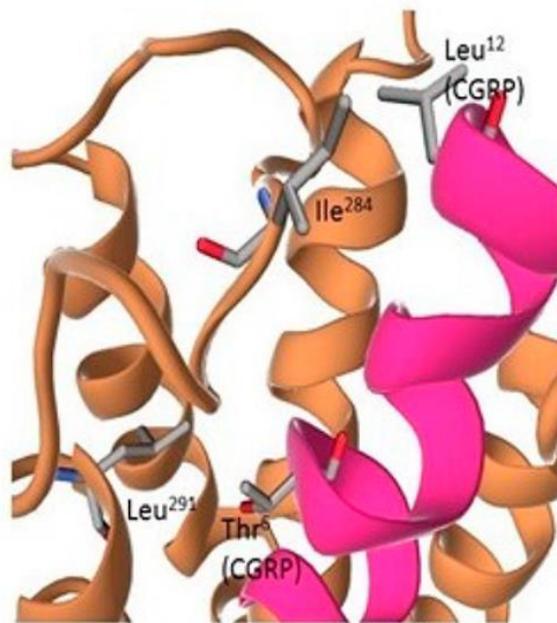
AzF can be crosslinked in-vivo and in-vitro (Wang *et al.*, 2009) In cross-linking studies, it is used as a nitrene-generating cross-linking amino acid, incorporated into the receptor structure (**figure 61**) (Shao *et al.*, 2015; Seidel and Coin, 2018). The photo-activatable region of AzF in physiological conditions is chemically inert, however upon activation by UV exposure, AzF generates the highly reactive nitrene (Seidel and Coin, 2018; Shao *et al.*, 2015). The nitrene can form non-specific, covalent interactions with surrounding molecules in close proximity (Seidel and Coin, 2018). In order for covalent bonds to form, the surrounding molecules must be within a 3 Å radius away from the photo-activated AzF (Seidel and Coin, 2018).



**Figure 61** | Chemical structure of p-Azido-L-phenylalanine (AzF). The moiety highlighted in a green circle is the photo-activatable nitrene-generating group. Diagram taken and modified from (Seidel and Coin, 2018).

In this project, the extracellular loops 1 and 3 of the CLR were modified with the unnatural amino acid AzF. The CLR mutants were co-expressed with RAMP1 in HEK293T cells. The fluorescent ligand, [Lys(5(6)-carboxyfluorescein)15]- $\alpha$ CGRP (F-CGRP), was incubated with the cells expressing the CGRPR mutants. After the ligand binding incubation period, ultraviolet light at 254 nm wavelength was used to crosslink the AzF residue of the CGRPR to the bound ligand. Previous research by Simms *et al.* (2018) identified the residues of the ECL2, the ECD and the TMs of the CGRPR which make contact with the peptide. The experiments have shown that I284, S286, T288, L290, L291 and Y292 of the ECL2 make contact with the peptide, with I284 and L291 forming the most significant contacts (Simms

*et al.*, 2018). Tyrosine scanning mutagenesis confirmed that the residues substituted by tyrosine did not drastically change the pharmacological properties of the receptor (Simms *et al.*, 2018). As the AzF substituted the same sites used in tyrosine scanning, this suggested that the incorporation of AzF would not alter the properties of the receptor (Simms *et al.*, 2018). A molecular model was produced, prior to the availability of the cryo-EM CGRPR structure by Liang *et al.*, (2018), to visualise the molecular contacts made by the cross-linking residues and the CGRP (**figure 62**) (Simms *et al.*, 2018). As the CGRPR:CGRP structure is available in the Protein DataBase, the cross-linked residues in ECL1 and ECL3 can now be compared with the cryo-EM structure, generated by Liang *et al.*, (2018).

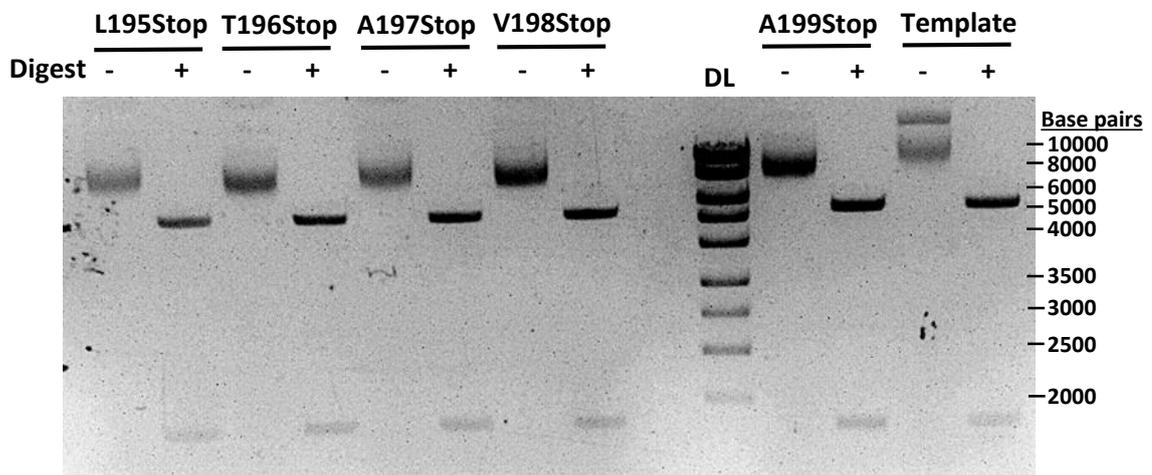


**Figure 62** | Molecular model illustrating the contacts made by residues of the CGRP and the CGRPR. Ribbons in orange highlight the CLR and the ribbon in pink illustrates the CGRP. I284 of the CLR ECL2 is in close proximity to L12 of CGRP. L291 makes contact with T6 of CGRP. Diagram taken from (Simms *et al.*, 2018).

## 8.2: Results

### 8.2.1: Generation of the human CLR mutants with the amber stop codon

The T8-HA-CLR construct was used as a template to produce the amber stop codon mutants. Primers containing the stop codon were designed against the ECL1 and ECL3 of the CLR construct (**Appendix**). Quick-change PCR was used to produce the mutant codons, which were transformed into XL10 Gold *E.coli*, which were subsequently minipreped. A diagnostic digest was performed to observe whether the mutant constructs were incorporated into the plasmid and sequencing was done to identify whether the stop codon was present in the mutant DNA sequence.



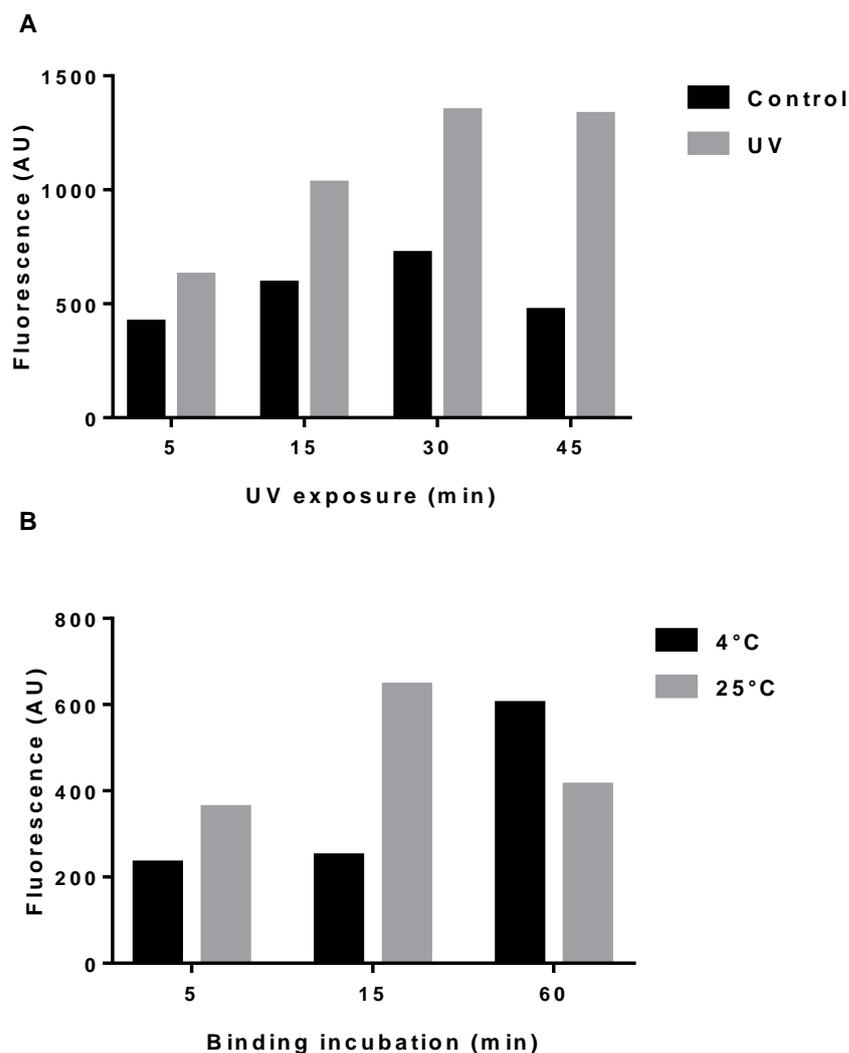
**Figure 63** | The pcDNA3.1+ vector successfully incorporated the TAG-mutant hCLR construct. A diagnostic digest of hCLR mutants: L195Stop, T196Stop, A197Stop, V198Stop, A199Stop. The template was the T8-HA-hCLR construct used to generate the amber stop codon mutants. The restriction enzymes: EcoRI and BamHI were used for the restriction digest. 100 ng of digested (+) and non-digested (-) DNA were loaded into each well of an agarose gel. Red Safe was used to bind to the DNA for UV detection. DL represents the DNA ladder (Gene ruler). The gel represents a snapshot of selected mutants, where the same band pattern was observed for the rest of the ECL1 and ECL3 mutants.

The diagnostic digest demonstrated that the mutant DNA were all the same size when non-digested (**figure 63**). The size of the un-cut plasmid construct was positioned at the 7000 base pairs (bp). The template DNA however was slightly larger than the mutant DNA, which was positioned at 9000 bp, with an additional band on top of the 9000 bp band (**figure 63**). When the template and mutant plasmid constructs were digested using EcoRI and BamHI, the bands were all equal in size. The larger band,

representing the plasmid, was around 5000 bp, whilst the smaller band, representing the mutant and template hCLR, was around 1500 bp (**figure 63**). The total bp of the plasmid and the mutant constructs was approximately 7000 bp, which was equal to the size of the band for the un-cut mutant DNA (**figure 63**).

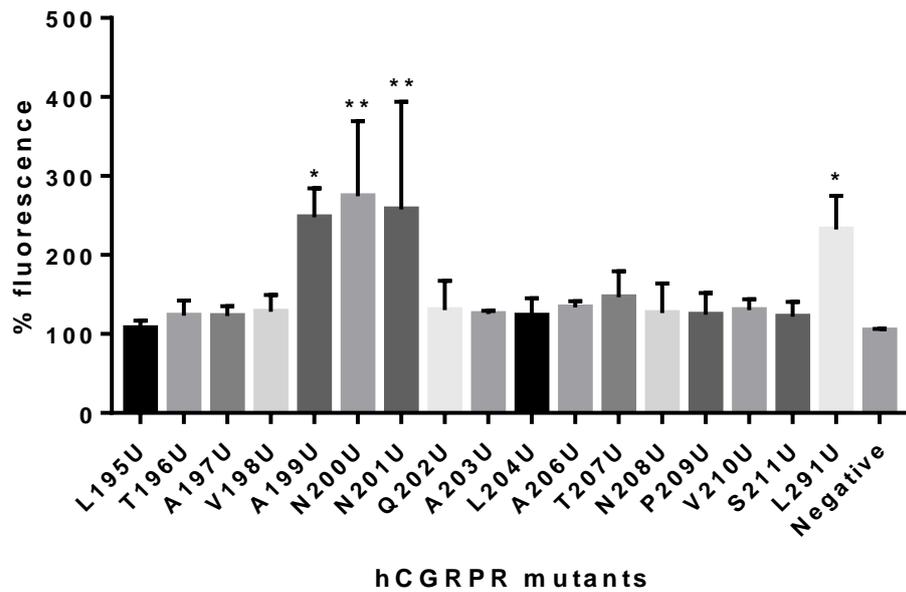
### **8.2.2: Using photoaffinity cross-linking to decipher the ligand binding sites of ECLs 1 and 3 of the CGRPR.**

The orthogonal tRNA (code 136) and aminoacyl synthetase (code 192) were kindly provided by Prof. Tom Sakmar (The Rockefeller University, NYC). The tRNA was designed to recognise the amber stop codon of the mutant hCLR constructs and the aminoacyl synthetase charged the tRNA with AzF. Once the mutant constructs were co-expressed with RAMP1 in HEK293T cells, with incubation with AzF, photoaffinity cross-linking was performed to observe binding of F-CGRP. The initial experiments were conducted to optimise the UV exposure and binding conditions to achieve optimum crosslinking of the fluorescent ligand to the receptor. The L291U CGRPR mutant was selected for initial experiments and was chosen as a positive control due to giving the highest cross-linking signal out of ECL2 in previous photoaffinity cross-linking experiments (Simms *et al.*, 2018)



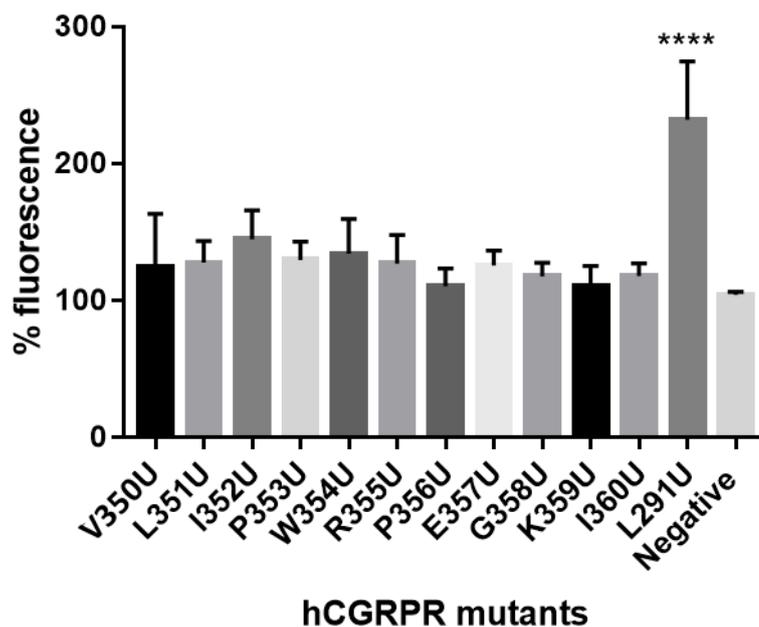
**Figure 64| 30 minutes UV exposure achieved the optimum signal and a change in temperature affects ligand-receptor interactions over time.** **A)** Increasing the UV exposure time enhances the cross-linking reaction between AzF and F-CGRP. hCLR L291U/RAMP1 was expressed in HEK193T. F-CGRP bound the mutant CGRPR for 15 minutes at 37°C, 5% CO<sub>2</sub>. The cells were exposed to UV radiation for 5-45 minutes. Samples for the control conditions were covered in foil to prevent their UV exposure. **B)** A 15 minute binding incubation time of F-CGRP to hCLR L291U/RAMP1 at room temperature produced the highest cross-linking signal. The hCLR mutant L291U/RAMP1 was expressed in HEK193T. F-CGRP bound the mutant CGRPR for 5-60 minutes at either 4°C or 25°C. The cells were exposed to UV radiation for 30 minutes. Samples for the control conditions were covered in foil to prevent their UV exposure and their values were subtracted from the samples exposed to UV light. ‘U’ represents the AzF substitution. The graphs are from one experiment.

The UV exposure for cross-linking varies amongst different literature, which is probably due to differences in the equipment used to emit UV light and the mutant receptor being studied (Simms *et al.*, 2018; Naganathan *et al.*, 2013). Therefore the initial experiment involved exposing the CGRPR mutant to UV light for different time periods (**figure 64A**). The L291U CGRPR mutant was used for optimization due to forming the most cross-linked ligand-receptor complexes out of the ECL2 mutants used in Simms *et al.* (2018). Increasing the UV exposure time from 5 minutes to 30 minutes increased the fluorescence absorbance of the samples exposed to UV light (**figure 64A**). However increasing the UV exposure from 30 minutes to 45 minutes did not lead to further increase in fluorescence. The signal emitted by the UV-exposed samples at 30 and 45 minute exposure time was approximately double the signal achieved by the control conditions which were not exposed to UV light (**figure 64A**). Once the optimum UV exposure time was established, the binding incubation time and temperature were assessed (**figure 64B**). Ligand incubation at room temperature (25°C) achieved the highest fluorescence at 15 minutes, which was quicker than ligand incubation on ice (4°C). The ligand incubation condition was changed to 15 minutes in a cell incubator set at 37°C, 5% CO<sub>2</sub>, as the same conditions were used in a photoaffinity cross-linking study and for in vivo ligand incubation with CGRP (Woolley *et al.*, 2017; Naganathan *et al.*, 2013). The highest fluorescence absorbance was achieved with these conditions.



**Figure 65| The CGRPR with the ECL1 A199U, N200U and N201U mutations cross-linked with F-CGRP.** The hCGRPR mutants contained the AzF unnatural amino acid and were expressed in HEK293T cells. F-CGRP was used to bind the mutant receptors for 15 minutes at 37°C, 5% CO<sub>2</sub>. The cells were exposed to UV radiation for 30 minutes. Cells which were not transfected were used as a negative control. Values represent the percentage increase of the control samples not exposed to UV light. The graph represents six independent experiments and the error bars represent the standard error of the mean. The one-way ANOVA statistical test, followed by Dunnett's multiple comparisons test (Graphpad prism 7) was used to compare the fluorescence of the mutants with the negative control (\*  $p \leq 0.05$ ; \*\*  $p \leq 0.01$ ). 'U' represents the AzF substitution.

Once the binding and exposure conditions were established with the positive control, L291U CGRPR, photoaffinity cross-linking was used on the ECL1 CGRPR mutants, containing the AzF substitution (**figure 65**). The mutants were either exposed to UV light or were covered in aluminium foil during the exposure. The percentage increase of the UV exposed samples were calculated using the fluorescence absorbance values of the samples covered in foil. The negative control was the non-transfected HEK293T cells, which were either exposed to UV radiation or were not exposed. The purpose of the negative control was to detect the background signal of the assay in order to compare the value with the mutants. Mutants N200U and N201U displayed a significant increase in fluorescence when compared with the negative control (**figure 65**). Mutation A199U also displayed a significant increase in fluorescence, with a similar level to the positive control (**figure 65**). The positive control was also significantly higher in fluorescence when compared with the negative control. The rest of the mutants displayed a percentage fluorescence that was similar to the background noise, showing no cross-linking with the ligand (**figure 65**).



**Figure 66| No ECL3 CGRPR mutants cross-linked with F-CGRP.** The hCGRPR mutants contained the AzF unnatural amino acid and were expressed in HEK293T cells. F-CGRP was used to bind the mutant receptors for 15 minutes at 37°C, 5% CO<sub>2</sub>. The cells were exposed to UV radiation for 30 minutes. Cells which were not transfected were used as a negative control. Values represent the percentage increase of the control samples not exposed to UV light. The graph represents three independent experiments and the error bars represent the standard error of the mean. The one-way ANOVA statistical test, followed by Dunnett’s multiple comparisons test (Graphpad prism 7) was used to compare the fluorescence of the mutants with the negative control (\*\*\*\*  $p \leq 0.0001$ ). ‘U’ represents the AzF substitution.

The ECL3 CGRPR mutants containing the AzF substitutions were also analysed to observe their cross-linking to the fluorescent CGRP (**figure 66**). The negative control was used to obtain the background signal of the assays and the positive control, L291U CGRPR, was used to compare the levels of cross-linking between the mutants. All the ECL3 mutants demonstrated a percentage fluorescence that was equal to the background noise, as there was no significant change in fluorescence when compared with the negative control (**figure 66**).

### **8.3: Discussion of chapter**

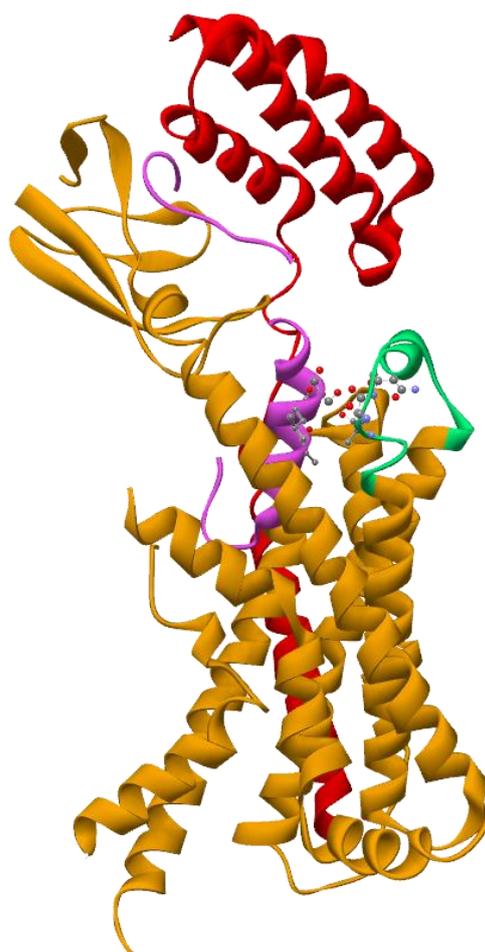
Incorporation of an unnatural amino acid into a protein's primary structure has proven useful for many biological applications, such as identifying binding site topology and protein-protein interactions (Gagnon *et al.*, 2019; Simms *et al.*, 2018; Naganathan *et al.*, 2013; Wu and Wu., 2015; Wang *et al.*, 2009). To incorporate a non-canonical amino acid into a protein, a stop codon must replace the codon of interest using PCR. As the whole of ECL1 and ECL3 of the hCGRPR was being investigated, 28 CLR mutations were made to have the amber stop codon to replace the residues of the extracellular loops. Once the polymerase chain reaction was complete and the transformation of the *E.coli* was successful, an agarose gel with ethidium bromide was used to confirm the presence of the mutant CLR construct in the pcDNA3.1+ plasmid (**figure 63**). According to Uniprot and sequencing data, the size of pcDNA3.1+ is 5248 bp and the size of the T8-HA-CLR construct was 1488 bp, which is concordant with the size of the digested bands in **figure 63**. Integration of the CLR mutant construct into the plasmid gives a total size of 6736 bp. Sequencing of the mutant constructs have confirmed the presence of the mutant within the construct, allowing the progression of AzF incorporation.

For understanding the topology of the binding sites of receptors, cross-linkable unnatural amino acids can be incorporated into a receptor. This has been conducted previously for the CGRPR by Simms *et al.* (2018), where certain residues of the ECD, ECL2 and the TMs have been identified to be in close proximity to the residues of  $\alpha$ CGRP. The ECL2 mutation at residue L291 showed the highest cross-linking signal with the fluorescent CGRP (Simms *et al.*, 2018). The L291 CLR mutant was used as a positive control to assess ECL1 and ECL3 photoaffinity cross-linking. Preliminary experiments were conducted to test the pharmacodynamics aspects of the assay, using the positive control (**figure 64**). Increasing binding incubation and UV exposure time lengths were used to identify the optimum conditions to achieve the highest signal, where a 15 minute ligand binding incubation at 37°C, 5% CO<sub>2</sub> and UV exposure for 30 minutes was sufficient for subsequent assays (**figure 64**). One valuable point from the binding incubation experiment (**figure 64B**) was that the ligand-receptor interaction can be slowed down at 4°C. This will be useful for assessing the ligand binding dynamics for different residues of the receptor binding site. Analysing the residue the ligand makes contact with first can be observed using the cross-linking assay, thus mapping the pathway the ligand takes to reach its final binding site (Bower *et al.*, 2018). Analysing the binding site topology can also be assessed, where residues interacting with CGRP may differ from adrenomedullin and the speed of residue contacts may also differ.

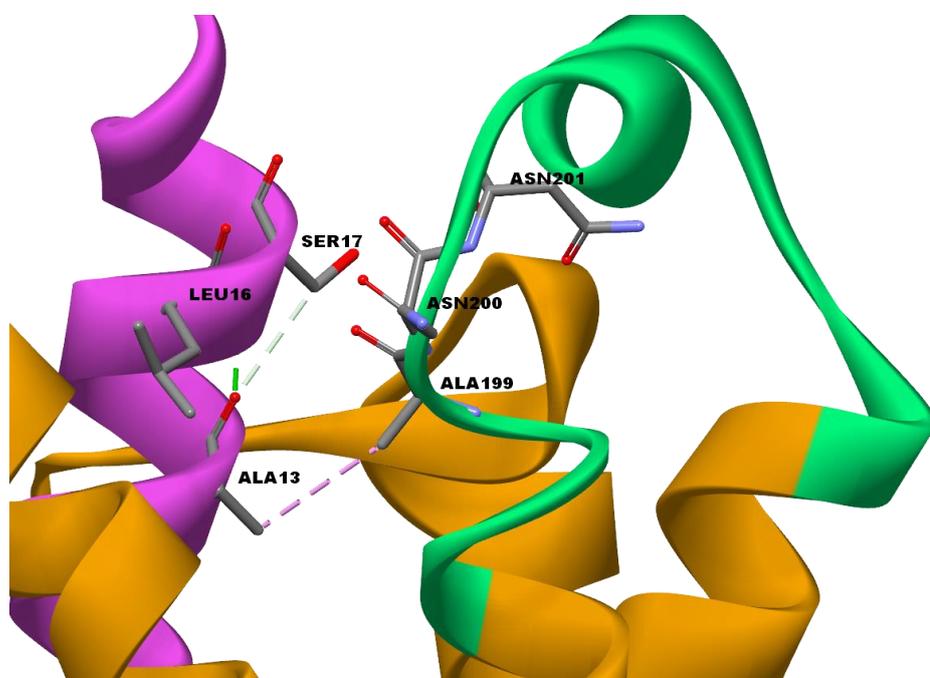
Once the ideal reaction conditions were established, the ECL1 and ECL3 CGRPR mutants were used for the cross-linking assay. A clear interaction was observed for three residues of the ECL1, where mutants A199U, N200U and N201U demonstrated a cross-linking signal that was similar to the positive control (**figure 65**). The rest of the mutant receptors may not have formed cross-links with the ligand or were not well expressed on the cell surface, where western blotting is required to confirm expression of all the AzF mutants. The recent cryo-EM structure of the CLR:RAMP1, complexed with  $\alpha$ CGRP (PDB ID: 6E3Y) was used to establish the positioning of the ligand and the ECL1 residues, A199-N201 (**figure 67**). The three residues of the ECL1 were in close proximity to the A13, L16 and S17 residues of the peptide, which may allow intermolecular contacts to be made between the peptide and the ECL1 residues, according to the cryo-EM structure (**figure 67B**). There were no intermolecular hydrogen

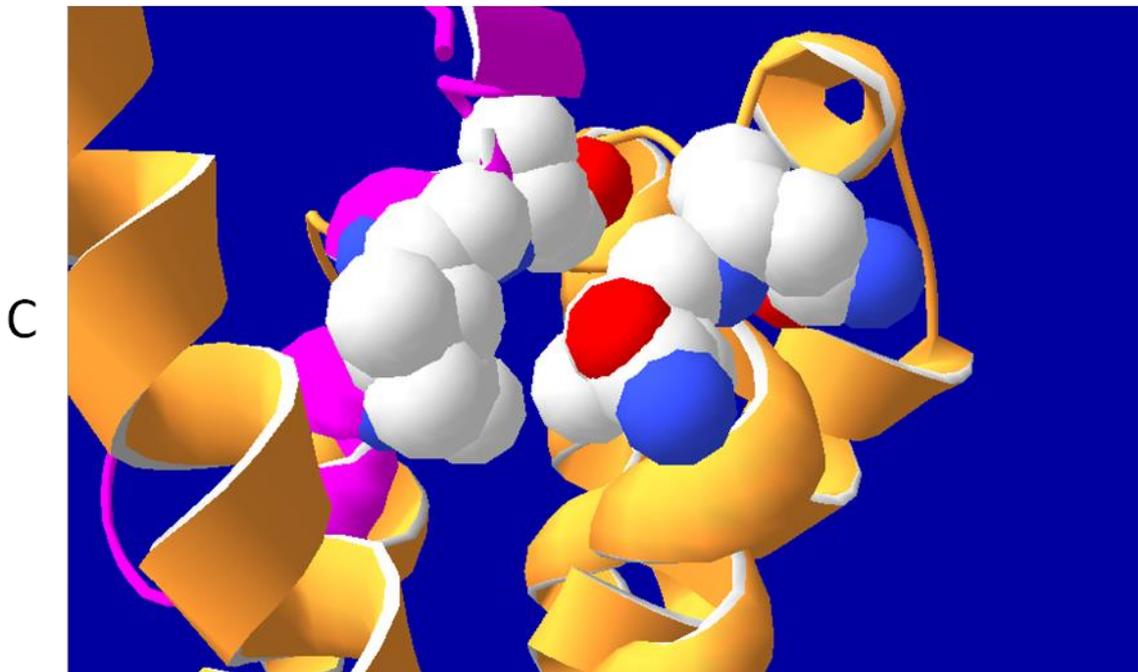
bonds formed between the residues of the peptide and the ECL1 of the receptor, according to the cryo-EM structure and the bond-interaction feature of BIOVIA Discovery Studio (**figure 67B**). Only one side-chain interaction was found between residue A13 of the peptide and A199 of ECL1, where an alkyl hydrophobic interaction can potentially form between the methyl groups of the two alanine residues (**figure 67B**). The bond length between the two residues is 4.17 Å, which was measured using Discovery Studio and Swiss PDB viewer (**figure 67B**). Residues N200 and N201 do not make any contacts with the peptide through their side chains, where the side chain of N201 was facing away from the peptide (**figure 67A-B**). The van der Waals space filling tool from Swiss PDB viewer identified possible interactions through van der Waals forces between N200 and N201 of the ECL1 and L16 and S17 of the peptide, via the backbone of the residues (**figure 67C**). When observing the side view of the CGRPR structure, the C terminus of the CGRP is leaning towards the ECL1 (**figure 67A**), with the residues, A199-N201, being closest to the peptide (**figure 67A-B**). As the cryo-EM structure is only a snapshot of one configuration of the frozen CGRPR, it is probable that the torsion angles of the residues are dynamic, where potential contacts may be missing from the structure (Liang *et al.*, 2018). Nevertheless, the orientation of the peptide and ECL1 and their potential molecular contacts agree with the cross-linking data, showing that the three residues make contact with the peptide (**figure 65 and figure 67**). As AzF residues need to be less than 3 Å away from the closest molecule to make contact with the peptide, the distance between the A199U, N200U and N201U substitutions are around 3 Å away from the peptide. Swiss PDB viewer showed that residues N200 and N201 of the ECL1 are around 4-5 Å away from residues L16 and S17 of the peptide. As the residues were mutated to AzF, this may have an influence on the distance between the reactive moiety of its side chain and residues of the peptide.

A



B





**Figure 67| Ala99, Asn200 and Asn201 of the CLR are in close proximity to Leu16 and Ser17 of  $\alpha$ CGRP.** The CLR:RAMP1 cryo-EM structure (PDB ID: 6E3Y by Liang *et al.*, (2018)) was analysed using BIOVIA Discovery Studio. Images were taken of the: **A)** A side-view of the CGRPR structure. **B)** Close-up interactions between residues Ala199, Asn200, Asn201, Ala13, Leu16 and Ser17. The orange secondary structure represents the CLR and the red secondary structure represents the RAMP1. The structure highlighted in green represents the ECL1 and the purple structure is the CGRP. The grey sticks represent carbon chains. The red sticks and spheres mark the OH and CO groups. The blue sticks and spheres represent NH groups. The green dashed lines are hydrogen bonds and the purple dashed line is a hydrophobic interaction **C)** Van der Waals contacts between Ala199, Asn200, Asn201 of the ECL1 and Ala13, Leu16 and Ser17 of the peptide. The pink ribbon is the CGRP and the orange ribbon is the CLR. The red, white and blue spheres mark the van der Waals forces from the residues of interest. The van der Waals space filling tool was used to observe the forces and were rendered solid 3D. Image created using Swiss PDP viewer.

Evidence from literature has stated that there are limited contacts between ECL1 residues and the peptide (Liang *et al.*, 2018) such as the contact between L16 and S17 of the peptide and A199, N200, Q202 and V205 of the ECL1 (Liang *et al.*, 2018). As residues A199 and N200 cross-linked with the peptide and is concordant with the cryo-EM structure and supporting literature (Liang *et al.*, 2018), the photoaffinity cross-linking assay proved to be an accurate and effective technique to identify the ligand binding site topology. Q202 was stated to form a weak hydrogen bond with the S17 oxygen backbone, however the

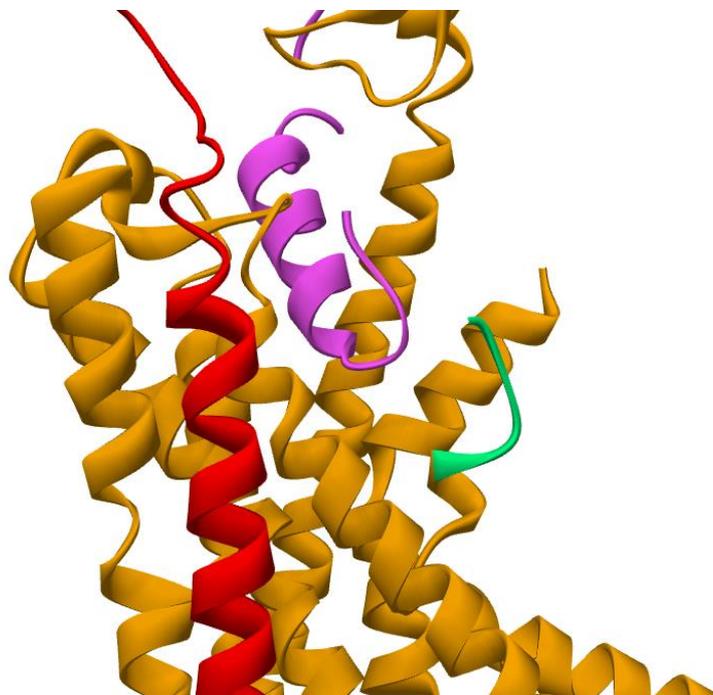
interaction between A199 and N200 of the ECL1 and the backbone of L16 and S17 of the peptide was not specified (Liang *et al.*, 2018). Therefore it is probable that A199U, N200 and N201 may form non-covalent interactions with the peptide. Van der Waals forces and a hydrophobic interaction are possible interactions that are formed between the ECL1 residues and the peptide, according to the bond-identifier tool of Discovery Studio and the van der Waals space filling tool from Swiss PDB viewer (**figure 67C**).

The cross-linking data did not show any interaction between residues Q202 and V205 (**figure 65**) and the peptide although Liang *et al.* (2018) has stated that there was. This absence of cross-linking with the Q202U and V205U mutants may be due to the AzF nitrine group not being in close proximity to the peptide to form a cross-link, or there may have been an issue with the expression of the two mutants. A tyrosine scanning mutagenesis experiment needs to be done, similarly to Simms *et al.*, (2018), to assess whether the incorporation of AzF impacts ligand binding. The reason tyrosine can be used for scanning mutagenesis is because the orthogonal aminoacyl tRNA synthetase also accepts tyrosine, as well as AzF (Simms *et al.*, 2018). As tyrosine shows the highest structural similarity to AzF out of the amino acids, the tyrosine-mutated receptors should be able to bind the CGRP (Simms *et al.*, 2018). Alanine scanning can also be used to replace the AzF residues with alanine to see if ligand binding is abolished when substituting AzF with alanine.

From existing literature, alanine mutations of N200 did not affect the potency of CGRP, therefore the amino acid side chain has limited importance in the binding of the peptide (Barwell *et al.*, 2011). An A199 mutation to leucine however did reduce the potency of CGRP, as indicated by the reduction of cAMP (Barwell *et al.*, 2012). There have been speculations where residue A199 can form a hydrophobic cluster with residue L195 as both residues have a hydrophobic side chain (Barwell *et al.*, 2012). This cluster was suggested to be of importance in CGRP binding. There was uncertainty on whether residues L195 and A199 of ECL1 made contact with the peptide, prior to the discovery of the cryo-EM structure by Liang *et al.* (2018). However the photoaffinity cross-linking data demonstrated that only one residue of the hydrophobic cluster, A199, interacted with the ligand (**figure 65 and 67B**). This was also supported by Liang *et al.* (2018), where residue A199, of the hydrophobic cluster, formed an interaction with the peptide. Based on mutagenesis studies, residues A203 and A206 of ECL1 were shown to impact

CGRP potency (Barwell *et al.*, 2012). It was speculated that A203 and A206 make contact with the ligand due to being positioned in the middle of ECL1, however based on the cryo-EM structure (**figure 67**), the residues would be positioned further away from the peptide when compared with residues A199-N201. Liang *et al.* (2018) however did not specify that residues A203 and A206 make contact with CGRP.

The photoaffinity cross-linking assay did not indicate any cross-linking occurring between the ECL3 CLR mutants and the fluorescent CGRP (**figure 66**). This could be due to a lack of expression of the mutants on the cell surface, or the ECL3 may not be in close proximity to the ligand. Western blotting is required to confirm the expression of the ECL3 mutant CGRPR. When observing the cryo-EM structure of the CGRPR in complex with  $\alpha$ CGRP, the ECL3 residues seem to be positioned away from the peptide ligand (**figure 68**). Unfortunately, the full ECL3 structure was not resolved in the cryo-EM structure (**figure 68**) (Liang *et al.*, 2018). Therefore there are uncertainties on whether ECL3 can interact with the peptide.



**Figure 68| ECL3 does not form molecular contacts with  $\alpha$ CGRP.** The CLR:RAMP1 cryo-EM structure (PDB ID: 6E3Y by Liang *et al.*, (2018)) was analysed using Discovery Studio. A side profile image was taken of the receptor. The orange secondary structure represents the CLR and the red secondary structure represents the RAMP1. The structure highlighted in green represents the ECL3 and the purple structure is the CGRP.

Existing literature also showed uncertainty on whether ECL3 of the CLR makes contact with  $\alpha$ CGRP. Mutagenesis experiments involving the ECL3 suggested a few residues having an impact on CGRP potency, with I360 being the main residue of ECL3 (Barwell *et al.*, 2012). Experiments involving the calcitonin receptor showed that L368, (V364 being the equivalent residue of CLR), was a photoaffinity contact for salmon calcitonin (Barwell *et al.*, 2012). It was hypothesized that the top of the ECL3 of the CLR could be in close proximity to the CGRP (Barwell *et al.*, 2012). However the contact distance between the top of ECL3 and the CGRP may be over 10 Å (Barwell *et al.*, 2012), which may explain why no cross-linking was detected between the ECL3 residues of CGRPR and the peptide, since AzF must be below 5 Å away (Simms *et al.*, 2018) from a molecule in order to cross-link (**figure 67 and 68**). As no binding interactions were established between ECL3 of CGRPR and the  $\alpha$ CGRP; it was suggested that the ECL3 residues may provide a structural support for the CLR to associate with RAMP1 (Barwell *et al.*, 2012). The full structure of ECL3 was not resolved in the cryo-EM structure (PDB ID: 6E3Y) due to the high mobility of the loop (**figure 68**) (Liang *et al.*, 2018). MD simulation programme PLOP50 was used to generate a full hypothetical ECL3 of the cryo-EM structure, which hypothesised that CGRP makes contact with the region of the ECL3 that was missing (**figure 68**) (Liang *et al.*, 2018). No persistent interactions from the simulation were observed as the loop still demonstrated fluidity. As UV cross-linking forms strong covalent bonds between the unnatural amino acid, AzF, and the fluorescent peptide, a signal would probably be detected if ECL3 did interact with AzF. The high mobility and heterogeneity of the ECL3 structure would not have been an issue as the F-CGRP would not dissociate from the loop upon cross-linking. As no interaction was observed between the ECL3 mutants and F-CGRP (**figure 67**), it is unlikely that the ECL3 residues contribute to the CGRP binding site.

The photoaffinity cross-linking assay has proven useful to decipher the residues of the extracellular loops of the CGRPR which make up part of the CGRP binding site. As experiments have shown CGRP binding to residues of the ECD, ECL2 and some of the TMs (Simms *et al.*, 2018), the contribution of identifying the ECL1 and ECL3 binding site residues will further map the topology of the ligand binding site. The assay was also successful in determining the residues which interact with the peptide ligand, as comparisons with the recent cryo-EM structure showed high similarity in the residues of the ECL1 which make up the binding site. Another important feature of photoaffinity cross-linking is that covalent bonds can be made between AzF and CGRP, preventing dissociation of the peptide ligand. This will be useful in permanently shifting the conformation of the CGRPR to an active state, which will be useful for SMA solubilisation and phage display to generate nanobodies against a receptor in its active conformation.

## Chapter 9: General discussion and future experimental considerations

Detergents have historically been used to solubilise membrane proteins for structural studies, which have provided target sites for drug discovery. Detergents can alter the the native lipid bilayer surrounding the protein, which could affect the conformation of the protein or remove the allosteric effects of the lipids. In response, SMALPs have been used to retain the native conformation of a membrane protein, with the lipid bilayer intact. The A<sub>2A</sub>R has previously been solubilised into SMALPs and showed binding capability after solubilisation and purification (Jamshad *et al.*, 2015). The aim of this project was to further characterise the properties of the SMA-solubilised A<sub>2A</sub>R and the CGRPR using various techniques. SMALPs were also tested in the nanobody discovery technique, phage display, to assess the suitability of receptor-SMALPs, as a conformational epitope for therapeutic nanobodies.

The SMA polymer has been used to solubilise the A<sub>2A</sub>R receptor by Jamshad *et al.* (2015). The deglycosylated human A<sub>2A</sub>R construct, designed by Fraser *et al.* (2006), was expressed in yeast and mammalian cells to characterise the pharmacology of the receptor when solubilised into SMALPs. The A<sub>2A</sub>R was successfully solubilised in SMALPs and was characterised using different ligands to confirm the conformational stability and binding characteristics of the receptor (Jamshad *et al.*, 2015). The A<sub>2A</sub>R-SMALP was successfully purified and concentrated into milligram amounts, using the *Pichia pastoris* cell line for high protein expression (Jamshad *et al.*, 2015). Furthermore, the elution of the receptor and its purity was characterised using silver staining (Jamshad *et al.*, 2015). The same deglycosylated A<sub>2A</sub>R construct using in Jamshad *et al.* (2015) and an avi-tag modified A<sub>2A</sub>R-A construct were were used for solubilisation and purification. Chapter 3 has demonstrated that the dghA<sub>2A</sub>R and the A<sub>2A</sub>R-A SMALPs were successfully purified, showing high purity with a few additional bands which are yet to be identified using mass spectrometry. As coomassie staining has been routinely used to assess the purity of the receptors, the data cannot be accurately compared to the silver stained gel using in Jamshad *et al.* (2015), which can be implemented in future studies. To remove the remaining contaminating proteins from the purified samples, size exclusion chromatography can also be used as an additional purification step, eluting the dghA<sub>2A</sub>R or the A<sub>2A</sub>R-A SMALPs based on their size.

The CGRPR-A receptor was expressed in HEK293T cells in order to SMA solubilise the receptor with a mammalian lipid bilayer intact. There is less evidence in the solubilisation of the CLR:RAMP1 in literature, however **chapter 7** has demonstrated that the receptor was successfully solubilised in SMALPs. A cAMP assay was also implemented to compare the pharmacology and functionality of the CGRPR-A, comparing the data with an existing CGRPR construct, used previously in Salvatore *et al.* (2010). The CGRPR-A modifications did not affect the pharmacology of the receptor and that the EC50 was similar to existing research. The next step for the CGRPR-A-SMALP would have been its purification using the flag-tag of RAMP1, in order to isolate CLR and RAMP1 as a complex. However previous purification attempts did not yield enough protein for downstream experiments, due to the construct being transiently transfected into HEK293T cells. The CGRPR-A construct should instead be incorporated into a zeocin mammalian plasmid vector, with an inducible gene to create a stable cell line that can be grown at a high yield with a high expression of the receptor. A radioligand binding assay must also be used to compare the binding capabilities of the receptor in the membrane, after solubilisation and purification. This will give a better idea of whether it is properly folded after SMA solubilisation.

Once the dghA<sub>2A</sub>R and the A<sub>2A</sub>R-A SMALPs have been purified, a competitive radioligand binding assay was used to determine the pharmacology of the receptors. ZM241385 showed a similar affinity towards the dghA<sub>2A</sub>R and the A<sub>2A</sub>R-A expressed in membranes and after SMA solubilisation. This was concordant with the binding data obtained by Jamshad *et al.* (2015), therefore the methodology of SMA solubilisation and purification was reproducible.

In addition to radioligand binding assay, further techniques have been implemented to study biophysics of the SMA-solubilised A<sub>2A</sub>R. FCS is a recently developed technique to understand the binding kinetics of a receptor in real time as well as to identify the non-specificity of ligand binding to a lipid environment. FCS has previously been used to study GPCR binding kinetics in living kinetics and has more recently been used to study receptors in membrane microdomains (Briddon *et al.*, 2018). In this project, the A<sub>2A</sub>R-SMALP was the first GPCR to be studied under FCS, where a clear displacement of the ligand was identified. An important result of FCS was that the fluorescent ligand bound non-

specifically to the SMALP, which was most likely due to the lipophilic BODIPY that was conjugated to the ligand. This finding indicated that in an in vivo environment, the actual concentration of a ligand to which a receptor is exposed to may differ from the bulk concentration of ligand administered to a live cell sample, as the ligand may non-specifically bind lipids. However this depends on whether lipophilic moieties are present on the ligand. Using SMALPs in FCS can be used to screen the lipophilicity of different drugs, which is important for drug development. The technique can also be used to conduct single molecule studies on the purified A<sub>2A</sub>R to observe the time length of ligand binding as well as the oligomerization capabilities of the receptor. The FCS data in **chapter 4** represents the initial studies conducted by myself, with assistance from the hosts at University of Nottingham. Further experiments have been conducted in collaboration with University of Nottingham and University of Birmingham, where a paper has been published (Grime *et al.*, 2020).

X-ray radiolytic footprinting is becoming a valuable asset in understanding the structural pharmacology of receptors. The technique exposes residues of membrane proteins to irradiated water molecules, highlighting the regions of the receptor which can interact with water (Gupta *et al.*, 2016). Changes in the water networks have been observed when rhodopsin transitions from an inactive to an active state, demonstrating the importance of water molecules in receptor conformation (Gupta *et al.*, 2016). Water molecules have shown importance in directing the sodium ion to its binding site within the A<sub>2A</sub>R, where the ion acts as a negative allosteric modulator (Carpenter and Lebon, 2017). Although SMALPs have been suggested to render a receptor in its native conformation, the effect SMA exerts on the receptor and the surrounding lipids is unknown. XRF has been used in this project to assess whether water-exposed residues are present on the receptor in a SMALP context. The upper portion of the TMs and the intracellular surface were identified as interaction sites for water molecules, demonstrating that using SMALPs in XRF is possible. Further experiments are required to test whether the water-exposed residues change upon agonist and antagonist binding, which can then be compared with the rhodopsin data to analyse changes in similar motifs of family A receptors. Using the ligands will also inform which conformational state the A<sub>2A</sub>R is in when SMA solubilised.

After the SMA-solubilised dghA<sub>2A</sub>R demonstrated binding capability with ZM241385, phage display was used to generate dghA<sub>2A</sub>R-specific phage as shown in **chapter 5**. Although seven dghA<sub>2A</sub>R-specific phages were detected using phage ELISA, the enrichment of specific phage was low when comparing the total monoclonal phage bound to the dghA<sub>2A</sub>R-SMALP and the total monoclonal phage bound to the Fab positive control. The most probable issue was that the SMALP was washed away from the ELISA plate during the bio-panning wash cycles, therefore an attempt was made to enhance immobilisation. In **chapter 6**, the A<sub>2A</sub>R-A was biotinylated, in order to immobilise the receptor onto a streptavidin-coated ELISA plate. The A<sub>2A</sub>R-A displayed better immobilisation onto streptavidin-coated wells, however the absorbance levels of the ELISA assays are not as high as the required levels by UCB. As the biotinylation conditions were sub-optimal, tests need to be conducted with different buffer compositions, in particular the ATP, the magnesium and the D-biotin content. A recombinant biotin ligase gene should also be expressed in a bacterial vector for an economical supply of the enzyme, instead of purchasing a bulk of the enzyme from Avidity. Therefore further optimization experiments can be conducted without the worry of limitation of the enzyme. Alternatively, the *Pichia pastoris* strain can be modified to express biotin ligase, prevented the need to use the enzyme after SMA solubilisation. Another important consideration is to test different polymers for solubilisation which are not limited by their intolerance to divalent cations. Diisobutylene maleic-acid (DIBMA) is one of the latest polymers designed after the solubilisation properties of SMA have been discovered. The polymer still renders the solubilised protein with a surrounding lipid bilayer intact, however the DIBMA copolymer is less disruptive to the lipid acyl-chain order and is less prone to precipitation by divalent cations (Oluwole *et al.*, 2017). Therefore the magnesium concentration required for efficient biotin ligase activity will not affect the solubility of the copolymer. The experiments conducted in **chapter 4** have clearly shown that the DIBMA copolymer can solubilise the receptor and retain binding affinity of the ligand, however the copolymer was less stable over time when compared with the SMA2000 copolymer. The DIBMA copolymer is still in its early stages of discovery and further polymer modifications may be required to improve the stability of the DIBMALP.

The photoaffinity cross-linking assay was adopted to study the binding site topology of the CGRPR extracellular loops. The assay can also be used to covalently bind the CGRP ligand to the receptor, stabilising the conformation of the receptor into an active state. A recently published research article demonstrated that the CGRP ligand can interact with residues of the ECL2, the TMs and the ECD (Simms *et al.*, 2018). Residues I284U and L291U of the ECL2 formed the most prominent contacts with the CGRP, which was concordant with the data in the latest cryo-EM structure (Simms *et al.*, 2018; Liang *et al.*, 2018). This established the validity of the cross-linking assay to make evident the residues of the CGRPR which form the CGRP binding site. The technique was used in **chapter 8** to identify the residues of ECL1 and ECL3 of the CGRPR. ECL1 formed limited contacts with CGRP, where residues A199 and N200 were in close proximity to the peptide, as made evident by the cryo-EM structure (Liang *et al.*, 2018) and the data in **chapter 8**. N201U also cross-linked with CGRP and was in close proximity to the peptide according to the cryo-EM structure (Liang *et al.*, 2018). ECL3 has been predicted to not form close contacts with the peptide. In this study, no cross-linking was observed between the ECL3 AzF mutant residues and the fluorescent CGRP, which may be due to an issue with the expression of the mutants. Although chapter 8 provided valuable information on the residues of the extracellular loops which make contact with the peptide, the recent cryo-EM structure has established some of the residues of the CGRPR which form molecular contacts with the ligand (Liang *et al.*, 2018). As the cryo-EM structure is based on a frozen sample, the interactions in the structure may not fully reflect the interactions which occur within the cell membrane at 37°C, therefore ligand binding experiments are essential in identifying interactions which occur *in vivo*. Further photoaffinity cross-linking experiments are required to characterise molecular interactions of the CGRPR. The CGRPR residues which interact with the G protein and the RCP are not well established and require further analysis to identify the intracellular binding topology. A G protein or the RCP can be fluorescently tagged and intracellular AzF mutants of the CGRPR can be made to see if the G protein and the RCP can cross-link with the receptor. Since the CGRPR:CGRP interaction has been heavily studied, the CGRPR:adrenomedullin can also be observed using photoaffinity cross-linking, as there are limited resources investigating this interaction. This is mainly due to adrenomedullin having a lower affinity towards the CLR:RAMP1 complex and therefore dissociating faster than the CGRP (Miret *et al.*, 2002). Photoaffinity cross-linking can observe

this by covalently binding a fluorescent adrenomedullin to the AzF-CLR:RAMP1 mutants. The data can be compared to **chapter 8** and Simms *et al.* (2018). Different RAMPs can also be co-expressed with the mutant CLR constructs to observe differences in photo cross-linking with adrenomedullin and CGRP. In relation to the project objective, since covalent bonds are formed between the peptide and the receptor upon UV exposure, the receptor can be stabilised in different conformations. The conformationally stabilised receptor can then be SMA solubilised and phage display can be used to generate nanobodies against an inactive and active conformation of the CGRPR, with the native lipid bilayer intact. This will generate nanobodies which are biased to certain conformations of the receptor and potentially be biased towards a one signalling pathway of the GPCR, whilst inhibiting another pathway. This demonstrates the usefulness of combining SMALPs with photoaffinity cross-linking and phage display to generate conformation-selective therapeutic nanobodies.

In conclusion, the use of detergents to solubilise membrane proteins can alter the conformation of the protein as the lipid bilayer can be removed (Wheatley *et al.*, 2016). SMA copolymers have been designed to overcome the limitations of detergents by solubilising receptors with the lipid bilayer intact. SMA has been incorporated in this study to solubilise the GPCRs: A<sub>2A</sub>R and CGRPR. The properties of the A<sub>2A</sub>R-SMALP was further analysed using techniques such as XRF and FCS. Phage display was also incorporated into this project to assess whether the SMALPs were suitable for nanobody generation and drug discovery. Currently, there are issues in SMALP immobilisation where an attempt was made to use the biotin/streptavidin technique to reduce removal of the SMALPs during phage display panning. Generating nanobodies against family B GPCRs would be ideal as to target the v-shaped chalice of the receptors, which contains the ligand binding site (Culhane *et al.*, 2015; Karageorgos *et al.*, 2018). Production of small molecule drugs to mimic peptide ligands would not be as feasible as generating a nanobody which covers all the binding site residues an endogenous ligand would occupy. The photoaffinity cross-linking assay demonstrated that the binding topology and the CGRPR can be analysed, and this technique could be useful in locking receptors in one conformation for downstream applications.

## References

- Abu-Thabit. (2017) Thermochemistry of Acrylamide Polymerization: An Illustration of Auto acceleration and Gel Effect. *World Journal of Chemical Education*. **5**(3): 94-101
- Achour, I., Cavelier, P., Tichit, M., Bouchier, C., Lafaye, P., Rougeon, F. (2008) Tetrameric and Homodimeric Camelid IgGs Originate from the Same IgH Locus. *The Journal of Immunology*. **181**(3): 2001–2009.
- Adam, A. C., González-Blasco, G., Rubio-Teixeira, M. & Polaina, J. (1999). Transformation of *Escherichia coli* with DNA from *Saccharomyces cerevisiae* Cell Lysates. *Appl Environ Microbiol*, **65**, 5303.
- Ahmad, Z.A., Yeap, S.K., Ali, A.M., Ho, W. Y., Alitheen, N. B. M., Hamid, M. (2012) ScFv antibody: Principles and clinical application. *Clinical and Developmental Immunology*, **2012**: 980250
- Akhtar, A. (2019) The role of Anti-calcitonin Gene-related peptide in migraine and its implication in developing countries: a reasonable option to consider despite higher cost. *cureus*. **11**(6): e4796
- Asif, A., Mohsin, H., Tanvir, R. & Rehman, Y. (2017). Revisiting the Mechanisms Involved in Calcium Chloride Induced Bacterial Transformation. *Frontiers in Microbiology*, **8**, 2169.
- Arbabi-Ghahroudi, M. (2017) Camelid single-domain antibodies: Historical perspective and future outlook. *Frontiers in Immunology*. **8**: 1–8.
- Armstrong, J. F., Faccenda, E., Harding, S. D., Pawson, A. J., Southan, C., Sharman, J. L., Campo, B., Cavanagh, D. R., Alexander, S. P. H., Davenport, A. P., Spedding, M., Davies, J. A; NC-IUPHAR (2019) The IUPHAR/BPS guide to pharmacology in 2020: extending immunopharmacology content and introducing the IUPHAR/MMV guide to malaria pharmacology. *Nucl. Acids Res*. pii: gkz951. doi: 10.1093/nar/gkz951 [Epub ahead of print].
- Avidity. Reaction conditions for BirA biotin ligase. [online] Last accessed: 9.12.2019. <<https://www.avidity.com/showpdf.asp?N=B9B6C28E-1976-4CFE-89DE-E37D1D8DB2CA>>
- Barwell, J., Woolley, M. J., Wheatley, M., Conner, A. C., Poyner, D. R. (2012) The role of the extracellular loops of the CGRP receptor, a family B GPCR. *Biochemical Society Transactions*. **40**(2): 433–437.
- Barwell, J., Conner, A., Poyner, D. R. (2011) Extracellular loops 1 and 3 and their associated transmembrane regions of the calcitonin receptor-like receptor are needed for CGRP receptor function. *Biochim Biophys Acta*. **1813**: 1906-1916
- Bawa, Z., Routledge, S. J., Jamshad, M., Clare, M., Sarkar, D., Dickerson, I., Ganzlin, M., Poyner, D. R., Bill, R. M. (2014) Functional recombinant protein is present in the pre-induction phases of *Pichia pastoris* cultures when grown in bioreactors, but not shake-flasks. *Microbial Cell Factories*. **13**: 127
- Behera, B. C., Yadav, H., Singh, S. K., Sethi, B. K., Mishra, R. R., Kumari, S. & Thatoi, H. (2017). Alkaline phosphatase activity of a phosphate solubilizing *Alcaligenes faecalis*, isolated from Mangrove soil. *Biotechnology Research and Innovation*, **1**, 101-111.
- Zhao, P., Liang, Y. L., Belousoff, M. J., Deganutti, G., Fletcher, M. M., Willard, F. S., Bell, M. G., Christe, M. E., Sloop, K. W., Inoue, A., Truong, T. T., Clydesdale, L., Funnell, S. G. B., Christopoulos, A., Wang, M. W., Miller, L. J., Reynolds, C. A., Danev, R., Sexton, P. M., Wootten, D. (2020) Activation of the GLP-1 receptor by a non-peptidic agonist. *Nature*. **577**: 432-436

- Block, H., Maertens, B., Spriestersbach, A., Brinker, N., Kubicek, J., Fabis, R., Labahn, J. & Schäfer, F. (2009). Chapter 27 Immobilized-Metal Affinity Chromatography (IMAC): A Review. *In: BURGESS, R. R. & DEUTSCHER, M. P. (eds.) Methods Enzymol.* Academic Press.
- Boerrigter, G., Lark, M. W., Whalen, E. J., Soergel, D. G., Violin, J. D. & Burnett, J. C., Jr. (2011). Cardiorenal actions of TRV120027, a novel ss-arrestin-biased ligand at the angiotensin II type I receptor, in healthy and heart failure canines: a novel therapeutic strategy for acute heart failure. *Circ Heart Fail*, **4**, 770-8.
- Boerrigter, G., Soergel, D. G., Violin, J. D., Lark, M. W. & Burnett, J. C., Jr. (2012). TRV120027, a novel beta-arrestin biased ligand at the angiotensin II type I receptor, unloads the heart and maintains renal function when added to furosemide in experimental heart failure. *Circ Heart Fail*, **5**, 627-34.
- Booe, J. M., Walker, C. S., Barwell, J., Kuteyi, G., Simms, J., Jamaluddin, M. A., Warner, M. L., Bill, R. M., Harris, P. W., Brimble, M. A., Poyner, D. R., Hay, D. L. & Pioszak, A. A. (2015). Structural Basis for Receptor Activity-Modifying Protein-Dependent Selective Peptide Recognition by a G Protein-Coupled Receptor. *Mol Cell*, **58**, 1040-52.
- Bornhorst, J. A. & Falke, J. J. 2000. [16] Purification of proteins using polyhistidine affinity tags. *Methods Enzymol.* Academic Press.
- Bortolato, A., Dore, A. S., Hollenstein, K., Tehan, B. G., Mason, J. S. & Marshall, F. H. (2014). Structure of Class B GPCRs: new horizons for drug discovery. *Br J Pharmacol*, **171**, 3132-45.
- Bower, R. L., Yule, L., Rees, T. A., Deganutti, G., Hendrikse, E. R., Harris, P. W. R., Kowalczyk, R., Ridgway, Z., Wong, A. G., Swierkula, K., Raleigh, D. P., Pioszak, A. A., Brimble, M. A., Reynolds, C. A., Walker, C. S., Hay, D. L. (2018) Molecular Signature for Receptor Engagement in the Metabolic Peptide Hormone Amylin. *ACS Pharmacology & Translational Science*. **1**(1): 32–49.
- Briddon, S. J., Kilpatrick, L. E. & Hill, S. J. (2018). Studying GPCR Pharmacology in Membrane Microdomains: Fluorescence Correlation Spectroscopy Comes of Age. *Trends Pharmacol Sci*, **39**, 158-174.
- Bühlmann, N., Leuthäuser, K., Muff, R., Fischer, J. A. & Born, W. (1999). A Receptor Activity Modifying Protein (RAMP)2-Dependent Adrenomedullin Receptor Is a Calcitonin Gene-Related Peptide Receptor when Coexpressed with Human RAMP1\*. *Endocrinology*, **140**, 2883-2890.
- Bühlmann, N., Aldecoa, A., Leuthäuser, K., Gujer, R., Muff, R., Fischer, J. A., Born, W. (2000) Glycosylation of the calcitonin receptor-like receptor at Asn60 or Asn112 is important for cell surface expression. *FEBS Letters*. **486**(3): 320–324.
- Burlingham, B. T. & Widlanski, T. S. (2003). An Intuitive Look at the Relationship of  $K_i$  and  $IC_{50}$ : A More General Use for the Dixon Plot. *Journal of Chemical Education*, **80**, 214.
- Byrne, B. (2015). *Pichia pastoris* as an expression host for membrane protein structural biology. *Curr Opin Struct Biol*, **32**, 9-17.
- Cao, E., Liao, M., Cheng, Y. & Julius, D. (2013). TRPV1 structures in distinct conformations reveal activation mechanisms. *Nature*, **504**, 113-8.
- Cao, G.-J., Jiang, X., Zhang, H., Croley, T. R. & Yin, J.-J. (2017). Mimicking horseradish peroxidase and oxidase using ruthenium nanomaterials. *RSC Advances*, **7**, 52210-52217.
- Carmen, S. & Jermutus, L. (2002). Concepts in antibody phage display. *Briefings in Functional Genomics*, **1**, 189-203.

- Carpenter, B., Lebon, G. (2017) Human adenosine A2A receptor: Molecular mechanism of ligand binding and activation. *Frontiers in Pharmacology*, **8**: 1–15.
- Carpenter, E., Beis, K., Cameron, A. & Iwata, S. (2008). Overcoming the challenges of membrane protein crystallography. *Curr Opin Struct Biol*, **18**, 581-586.
- Castro-Mejia, J. L., Muhammed, M. K., Kot, W., Neve, H., Franz, C. M., Hansen, L. H., Vogensen, F. K. & Nielsen, D. S. (2015). Optimizing protocols for extraction of bacteriophages prior to metagenomic analyses of phage communities in the human gut. *Microbiome*, **3**, 64.
- Chandel, T. I., Zaman, M., Khan, M. V., Ali, M., Rabbani, G., Ishtikhar, M. & Khan, R. H. (2018). A mechanistic insight into protein-ligand interaction, folding, misfolding, aggregation and inhibition of protein aggregates: An overview. *Int J Biol Macromol*, **106**, 1115-1129.
- Chang, C. H., Hsiung, H. A., Hong, K. L. & Huang, C. T. (2018). Enhancing the efficiency of the *Pichia pastoris* AOX1 promoter via the synthetic positive feedback circuit of transcription factor Mxr1. *BMC Biotechnol*, **18**, 81.
- Charles, A., Pozo-Rosich, P. (2019) Targeting calcitonin gene-related peptide: a new era in migraine therapy. *Lancet*. **394**(10210): 1765-1774
- Cheng, Y. (2018). Membrane protein structural biology in the era of single particle cryo-EM. *Curr Opin Struct Biol*, **52**, 58-63.
- Cheng, Y., Grigorieff, N., Penczek, P. A. & Walz, T. (2015). A primer to single-particle cryo-electron microscopy. *Cell*, **161**, 438-449.
- Chevalier, F. (2010). Standard Dyes for Total Protein Staining in Gel-Based Proteomic Analysis. *Materials (Basel)*, **3**, 4784-4792.
- Clark, K. L., Zeng, Z., Langford, A. L., Bowen, S. M. & Todd, S. C. (2001). PGRL is a major CD81-associated protein on lymphocytes and distinguishes a new family of cell surface proteins. *J Immunol*, **167**, 5115-21.
- Congreve, M., Murray, C. W. & Blundell, T. L. (2005). Keynote review: Structural biology and drug discovery. *Drug Discovery Today*, **10**, 895-907.
- Currie, G. M. (2018). Pharmacology, Part 1: Introduction to Pharmacology and Pharmacodynamics. *J Nucl Med Technol*, **46**, 81-86.
- Cournia, Z., Allen, T. W., Andricioaei, I., Antonny, B., Baum, D., Brannigan, G., Buchete, N. V., Deckman, J. T., Delemotte, L., Del Val, C., Friedman, R., Gkeka, P., Hege, H. C., Henin, J., Kasimova, M. A., Kolocouris, A., Klein, M. L., Khalid, S., Lemieux, M. J., Lindow, N., Roy, M., Selent, J., Tarek, M., Tofoleanu, F., Vanni, S., Urban, S., Wales, D. J., Smith, J. C. & Bondar, A. N. (2015). Membrane Protein Structure, Function, and Dynamics: a Perspective from Experiments and Theory. *J Membr Biol*, **248**, 611-40.
- Culhane, K. J., Liu, Y., Cai, Y. & Yan, E. C. (2015). Transmembrane signal transduction by peptide hormones via family B G protein-coupled receptors. *Front Pharmacol*, **6**, 264.
- Danielczak, B., Meister, A., Keller, S. (2019) Influence of Mg<sup>2+</sup> and Ca<sup>2+</sup> on nanodisc formation by diisobutylene/maleic acid (DIBMA) copolymer. *Chemistry and Physics of Lipids*. **221**: 30–38.

- Dawaliby, R., Trubbia, C., Delporte, C., Masureel, M., Van Antwerpen, P., Kobilka, B. K. & Govaerts, C. (2016). Allosteric regulation of G protein-coupled receptor activity by phospholipids. *Nat Chem Biol*, **12**, 35-9.
- De Graaf, C., Song, G., Cao, C., Zhao, Q., Wang, M. W., Wu, B. & Stevens, R. C. (2017). Extending the Structural View of Class B GPCRs. *Trends Biochem Sci*, **42**, 946-960.
- Digby, G. J., Shirey, J. K. & Conn, P. J. (2010). Allosteric activators of muscarinic receptors as novel approaches for treatment of CNS disorders. *Molecular BioSystems*, **6**, 1345-1354.
- Dobson, L., Remenyi, I. & Tusnady, G. E. (2015). The human transmembrane proteome. *Biol Direct*, **10**, 31.
- Dominik, P.K., Borowska, M.T., Dalmas, O., Kim, S. S., Perozo, E., Keenan, R. J., Kossiakoff, A. A. (2017) *Proteins Using Antibody Phage Display With Nanodiscs*. **24**(2): 300–309.
- Dong, M., Pinon, D. I., Cox, R. F. & Miller, L. J. (2004). Importance of the amino terminus in secretin family G protein-coupled receptors. Intrinsic photoaffinity labeling establishes initial docking constraints for the calcitonin receptor. *J Biol Chem*, **279**, 1167-75.
- Dorr, J. M., Scheidelaar, S., Koorengevel, M. C., Dominguez, J. J., Schafer, M., Van Walree, C. A. & Killian, J. A. (2016). The styrene-maleic acid copolymer: a versatile tool in membrane research. *Eur Biophys J*, **45**, 3-21.
- Fairhead, M., Howarth, M. (2014) Site-specific biotinylation of purified proteins using BirA. *Site-specific protein labelling* **1266**: 171-184
- Fraser, N. J. (2006). Expression and functional purification of a glycosylation deficient version of the human adenosine 2a receptor for structural studies. *Protein Expr Purif*, **49**, 129-37.
- Fredriksson, R. & Schioth, H. B. (2005). The repertoire of G-protein-coupled receptors in fully sequenced genomes. *Mol Pharmacol*, **67**, 1414-25.
- Fu, Y., Zhao, J. & Chen, Z. (2018). Insights into the Molecular Mechanisms of Protein-Ligand Interactions by Molecular Docking and Molecular Dynamics Simulation: A Case of Oligopeptide Binding Protein. *Comput Math Methods Med*, **2018**, 3502514.
- Gagnon, L., Cao, Y., Cho, A., Sedki, D., Huber, T., Sakmar, T. P., Laporte, S. A. (2019) Genetic code expansion and photocross-linking identify different  $\beta$ -arrestin binding modes to the angiotensin II type 1 receptor. *Journal of Biological Chemistry*. **294**(46): 17409-17420
- Gao, Y., Cao, E., Julius, D. & Cheng, Y. (2016). TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. *Nature*, **534**, 347-51.
- Gao, Y., Hu, H., Ramachandran, S., Erickson, J. W., Cerione, R. A. & Skiniotis, G. (2019). Structures of the Rhodopsin-Transducin Complex: Insights into G-Protein Activation. *Mol Cell*, **75**, 781-790.e3.
- Gao, Z.-G. & Ijzerman, A. P. (2000). Allosteric modulation of A2A adenosine receptors by amiloride analogues and sodium ions. *Biochemical Pharmacology*, **60**, 669-676.
- Gonzalez, E., Shepherd, L. M., Saunders, L. & Frey, M. W. (2016). Surface Functional Poly(lactic Acid) Electrospun Nanofibers for Biosensor Applications. *Materials (Basel)*, **9**.

- Goutelle, S., Maurin, M., Rougier, F., Barbaut, X., Bourguignon, L., Ducher, M. & Maire, P. (2008). The Hill equation: a review of its capabilities in pharmacological modelling. *Fundam Clin Pharmacol*, **22**, 633-48.
- Graslund, S., Savitsky, P., Muller-Knapp, S. (2017) In vivo biotinylation of antigens in *E.coli*. *Methods Mol Biol*. **1586**: 337-344
- Graphpad. Statistics guide. Last accessed: 19.03.2020. [online]  
<[https://www.graphpad.com/guides/prism/7/statistics/stat\\_holms\\_multiple\\_comparison\\_test.htm](https://www.graphpad.com/guides/prism/7/statistics/stat_holms_multiple_comparison_test.htm)>
- Green, R. & Rogers, E. J. (2013). Transformation of chemically competent *E. coli*. *Methods Enzymol*, **529**, 329-36.
- Grime, R., Goulding, J., Uddin, R., Stoddart, L., Hill, S., Poyner, D. R., Briddon, S., Wheatley, M. (2020) Single molecule binding of a ligand to a G-protein-coupled-receptor in real time using fluorescence correlation spectroscopy, rendered possible by nano-encapsulation in styrene maleic acid lipid particles. *Nanoscale*.
- Grunbeck, A., Sakmar, T. P. (2013) Probing G protein-coupled receptor - Ligand interactions with targeted photoactivatable cross-linkers. *Biochemistry*. **52**(48): 8625–8632.
- Guex, N. (1999) Protein modelling for all. *Trends Biochem. Sci.* **24**, 364–367
- Gulamhussein, A. A., Meah, D., Soja, D. D., Fenner, S., Saidani, Z., Akram, A., Lallie, S., Mathews, A., Painter, C., Liddar, M. K., Mohammed, Z., Chiu, L. K., Sumar, S. S., Healy, H., Hussain, N., Patel, J. H., Hall, S. C. L., Dafforn, T. R. & Rothnie, A. J. (2019). Examining the stability of membrane proteins within SMALPs. *European Polymer Journal*, **112**, 120-125.
- Gulamhussein, A. A., Uddin, R., Tighe, B. J., Poyner, D. R., Rothnie, A. J. (2020) A comparison of SMA (styrene maleic acid) and DIBMA (di-isobutylene maleic acid) for membrane protein purification. *Biomembranes*. <https://doi.org/10.1016/j.bbmem.2020.183281>
- Gupta, S., Feng, J., Chan, L. J., Petzold, C. J. & Ralston, C. Y. (2016). Synchrotron X-ray footprinting as a method to visualize water in proteins. *J Synchrotron Radiat*, **23**, 1056-69.
- Gupta, S., Sutter, M., Remesh, S. G., Dominguez-Martin, M. A., Bao, H., Feng, X. A., Chan, L. G., Petzold, C. J., Kerfeld, C. A. & Ralston, C. Y. (2019). X-ray radiolytic labeling reveals the molecular basis of orange carotenoid protein photoprotection and its interactions with fluorescence recovery protein. *J Biol Chem*, **294**, 8848-8860.
- Gurevich, V. V. & Gurevich, E. V. (2017). Molecular Mechanisms of GPCR Signaling: A Structural Perspective. *Int J Mol Sci*, **18**.
- Hanlon, C. D. & Andrew, D. J. (2015). Outside-in signaling – a brief review of GPCR signaling with a focus on the <em>Drosophila</em> GPCR family. *Journal of Cell Science*, **128**, 3533.
- Hardy, D., Bill, R. M., Rothnie, A. J. & Jawhari, A. (2019). Stabilization of Human Multidrug Resistance Protein 4 (MRP4/ABCC4) Using Novel Solubilization Agents. *SLAS Discov*, **24**, 1009-1017.
- Hay, D. L. & Pioszak, A. A. (2016). Receptor Activity-Modifying Proteins (RAMPs): New Insights and Roles. *Annu Rev Pharmacol Toxicol*, **56**, 469-87.

- Hay, D.L., Garelja, M.L., Poyner, D.R., Walker, C. S. (2018) Update on the pharmacology of calcitonin/CGRP family of peptides: IUPHAR Review 25. *British Journal of Pharmacology*. **175**(1): 3–17.
- Hilairret, S., Foord, S.M., Marshall, F.H., Bouvier, M. (2001) Protein-Protein Interaction and Not Glycosylation Determines the Binding Selectivity of Heterodimers between the Calcitonin Receptor-like Receptor and the Receptor Activity-modifying Proteins. *Journal of Biological Chemistry*. **276**(31): 29575–29581.
- Hino, T., Arakawa, T., Iwanari, H., Yurugi-Kobayashi, T., Ikeda-Suno, C., Nakada-Nakura, Y., Kusano-Arai, O., Weyand, S., Shimamura, T., Nomura, N., Cameron, A. D., Kobayashi, T. Hamakubo, T., Iwata, S. & Murata, T. (2012). G-protein-coupled receptor inactivation by an allosteric inverse- agonist antibody. *Nature*, **482**, 237-240.
- Hollenstein, K., De Graaf, C., Bortolato, A., Wang, M. W., Marshall, F. H. & Stevens, R. C. (2014). Insights into the structure of class B GPCRs. *Trends Pharmacol Sci*, **35**, 12-22.
- Holliger, P., Hudson, P. J. (2005) Engineered antibody fragments and the rise of single domains. *Nature Biotechnology*. **23**(9): 1126–1136.
- Holmberg, A., Blomstergren, A., Nord, O., Lukacs, M., Lundeberg, J., Uhlen, M. (2005) The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures. *Electrophoresis*. **26**(3): 501-510
- Hoogenboom, H. R. (2005). Selecting and screening recombinant antibody libraries. *Nature Biotechnology*, **23**, 1105-1116.
- Hu, G. M., Mai, T. L. & Chen, C. M. (2017). Visualizing the GPCR Network: Classification and Evolution. *Sci Rep*, **7**, 15495.
- Hutchings, C. J., Koglin, M., Olson, W. C., Marshall, F. H. (2017) Opportunities for therapeutic antibodies directed at G-protein-coupled receptors. *Nature Reviews Drug Discovery*. **16**(11): 787–810.
- Ikonomova, S. P., Le, M. T., Kalla, N., Karlsson, A. J. (2018) Effect of linkers on immobilisation of ScFvs with biotin-streptavidin interaction. *Biotechnology and applied biochemistry*. **65**(4): 580-585
- Ingham, K. C. (1984) Protein precipitation with polyethylene glycol. *Methods Enzymol*. **104**: 351-356
- Inoue, A., Raimondi, F., Kadji, F. M. N., Singh, G., Kishi, T., Uwamizu, A., Ono, Y., Shinjo, Y., Ishida, S., Arang, N., Kawakami, K., Gutkind, J. S., Aoki, J. & Russell, R. B. (2019). Illuminating G-Protein- Coupling Selectivity of GPCRs. *Cell*, **177**, 1933-1947 e25.
- Irannejad R, Tomshine J. C., Tomshine J. R., Chevalier M., Mahoney JP., Steyaert J, Rasmussen S. G., Sunahara R. K., El-Samad H., Huang B., von Zastrow M. (2013). Conformational biosensors reveal GPCR signalling from endosomes. *Nature*. **495**(7442): 534-538
- J Gingell, J., Simms, J., Barwell, J., Poyner, D. R., Watkins, H. A., Pioszak, A. A., Sexton, P. M. & Hay, D. L. (2016). An allosteric role for receptor activity-modifying proteins in defining GPCR pharmacology. *Cell Discovery*, **2**, 16012.
- Jara-Acevedo, R., Diez, P., Gonzalez-Gonzalez, M., Degano, R. M., Ibarrola, N., Gongora, R., Orfao, A., Fuentes, M.(2016) Methods for Selecting Phage Display Antibody Libraries. *Current Pharmaceutical Design*. **22**(43): 6490–6499.

- Jaakola, V.-P., Griffith, M. T., Hanson, M. A., Cherezov, V., Chien, E. Y. T., Lane, J. R., Ijzerman, A. P. & Stevens, R. C. (2008). The 2.6 Angstrom Crystal Structure of a Human A2A Adenosine Receptor Bound to an Antagonist. *Science*, **322**, 1211.
- Jacobson, K. A. & Gao, Z.-G. (2006). Adenosine receptors as therapeutic targets. *Nature Reviews Drug Discovery*, **5**, 247-264.
- Jain, S., Yuan, H., Spare, N. & Silberstein, S. D. (2018). Erenumab in the treatment of migraine. *Pain Management*, **8**, 415-426.
- Jakubík, J., Randáková, A., El-Fakahany, E. E. & Doležal, V. (2019). Analysis of equilibrium binding of an orthosteric tracer and two allosteric modulators. *PLoS One*, **14**, e0214255.
- Jamshad, M., Charlton, J., Lin, Y.-P., Routledge, Sarah j., Bawa, Z., Knowles, Timothy j., Overduin, M., Dekker, N., Dafforn, Tim r., Bill, Roslyn m., Poyner, David r. & Wheatley, M. (2015a). G-protein coupled receptor solubilization and purification for biophysical analysis and functional studies, in the total absence of detergent. *Bioscience Reports*, **35**.
- Jamshad, M., Grimard, V., Idini, I., Knowles, T. J., Dowle, M. R., Schofield, N., Sridhar, P., Lin, Y. P., Finka, R., Wheatley, M., Thomas, O. R., Palmer, R. E., Overduin, M., Govaerts, C., Ruyschaert, J. M., Edler, K. J. & Dafforn, T. R. (2015b). Structural analysis of a nanoparticle containing a lipid bilayer used for detergent-free extraction of membrane proteins. *Nano Res*, **8**, 774-789.
- Jamshad, M., Rajesh, S., Stamataki, Z., Mckeating, J. A., Dafforn, T., Overduin, M. & Bill, R. M. (2008). Structural characterization of recombinant human CD81 produced in *Pichia pastoris*. *Protein Expr Purif*, **57**, 206-216.
- Jensen, E. C. (2012). The Basics of Western Blotting. *The Anatomical Record*, **295**, 369-371.
- Jo, M. & Jung, S. T. (2016). Engineering therapeutic antibodies targeting G-protein-coupled receptors. *Experimental & Molecular Medicine*, **48**, e207-e207.
- Kamitani, S., Sakata, T. (2001) Glycosylation of human CRLR at Asn123 is required for ligand binding and signaling. *Biochimica et Biophysica Acta - Molecular Cell Research*. **1539**(1-2): 131-139.
- Karageorgos, V., Venihaki, M., Sakellaris, S., Pardalos, M., Kontakis, G., Matsoukas, M. T., Gravanis, A., Margioris, A. & Liapakis, G. (2018). Current understanding of the structure and function of family B GPCRs to design novel drugs. *Hormones (Athens)*, **17**, 45-59.
- Kee, Z., Kodji, X. & Brain, S. D. (2018). The Role of Calcitonin Gene Related Peptide (CGRP) in Neurogenic Vasodilation and Its Cardioprotective Effects. *Front Physiol*, **9**, 1249.
- Kenakin, T. (2004). Principles: receptor theory in pharmacology. *Trends Pharmacol Sci*, **25**, 186-92
- Kenakin, T. (2016). The mass action equation in pharmacology. *British Journal of Clinical Pharmacology*, **81**, 41-51.
- Kilpatrick, L. E., Hill, S. J. (2016) The use of fluorescence correlation spectroscopy to characterize the molecular mobility of fluorescently labelled G protein-coupled receptors. *Biochemical Society Transactions*. **44** (2): 624-629.
- Kim, T. K. (2014) Analysis of variance (ANOVA) comparing means of more than two groups. *Restorative Dentistry and endodontics*. **39**(1): 74-77

- Knobloch, J., Suhendro, D. K., Zieleniecki, J. L., Shapter, J. G., Koper, I. (2015) Membrane-drug interactions studied using model membrane systems. *Saudi Journal of Biological Sciences*. **22**(6): 714–718.
- Knowles, T.J. Finka, R. Smith, C. Lin, Y.P. Dafforn, T. Overduin, M. (2009) Membrane proteins solubilised intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *J Am Chem Soc*. **131**(22): 7484-5
- Koch, K., Kalusche, S., Torres, J. L., Stanfield, R. L., Danquah, W., Khazanehdari, K., von Briesen, H., Geertsma, E. R., Wilson, I. A., Wernery, U., Koch-Nolte, F., Ward, A. B., Dietrich, U. (2017) Selection of nanobodies with broad neutralizing potential against primary HIV-1 strains using soluble subtype C gp140 envelope trimers. *Scientific Reports*. **7**(1): 1–15.
- Korotych, O., Mondal, J., Gattas-Asfura, K. M., Hendricks, J., Bruce, B. D. (2019) Evaluation of commercially available styrene-co-maleic acid polymers for the extraction of membrane proteins from spinach chloroplast thylakoids. *European Polymer Journal*. **114**: 485-500
- Kupcsulik, B., Sevela, B. (2004) Effect of methanol concentration on the recombinant pichia pastoris mut<sup>s</sup> fermentation. *Periodica Polytechnica Chemical Engineering*. **48**(2): 73-87.
- Kusano, S., Kukimoto-Niino, M., Hino, N., Ohsawa, N., Okuda, K., Sakamoto, K., Shirouzu, M., Shindo, T. & Yokoyama, S. (2012). Structural basis for extracellular interactions between calcitonin receptor- like receptor and receptor activity-modifying protein 2 for adrenomedullin-specific binding. *Protein Sci*, **21**, 199-210.
- Lagerstrom, M.C. Schioth, H.B. (2008) Structural diversity of G protein-coupled receptors and significance for drug discovery. *Nat Rev Drug Discov*. **7**(4):339-357
- Lassen, L. H., Jacobsen, V. B., Haderslev, P. A., Sperling, B., Iversen, H. K., Olesen, J. & Tfelt-Hansen, P. (2008). Involvement of calcitonin gene-related peptide in migraine: regional cerebral blood flow and blood flow velocity in migraine patients. *The Journal of Headache and Pain*, **9**, 151-157.
- Lebon, G. Warne, T. Edwards, P.C. Bennett, K. Langmead, C.J. Leslie, A.G.W. Tate, C.G. (2011) Agonist-bound adenosine A<sub>2A</sub> receptor structures reveal common features of GPCR activation. *Nature*. **474**: 521-525
- Lebon, G., Edwards, P. C., Leslie, A. G. & Tate, C. G. (2015). Molecular Determinants of CGS21680 Binding to the Human Adenosine A<sub>2A</sub> Receptor. *Mol Pharmacol*, **87**, 907-15.
- Ledsgaard, L., Kilstrup, M., Karatt-Vellatt, A., Mccafferty, J. & Laustsen, A. H. (2018). Basics of Antibody Phage Display Technology. *Toxins (Basel)*, **10**.
- Lee, S. C., Knowles, T. J., Postis, V. L. G., Jamshad, M., Parslow, R. A., Lin, Y., Goldman, A., Sridhar, P., Overduin, M., Muench, S. P. and Dafforn, T. R. (2016) A method for detergent-free isolation of membrane proteins in their local lipid environment. *Nature protocols*. **11**(7): 1149–1162.
- Lee, P. Y., Costumbrado, J., Hsu, C. Y. & Kim, Y. H. (2012). Agarose gel electrophoresis for the separation of DNA fragments. *J Vis Exp*.
- Lee, S. C., Knowles, T. J., Postis, V. L., Jamshad, M., Parslow, R. A., Lin, Y. P., Goldman, A., Sridhar, P., Overduin, M., Muench, S. P. & Dafforn, T. R. (2016). A method for detergent-free isolation of membrane proteins in their local lipid environment. *Nat Protoc*, **11**, 1149-62.
- Leone, R. D., Lo, Y. and Powell, J. D. (2015) ‘A<sub>2A</sub>R antagonists : Next generation checkpoint blockade for cancer immunotherapy’, *CSBJ. Elsevier B.V*. **13**:265–272.

- Li, F. & Mullins, J. I. 2002. Site-Directed Mutagenesis Facilitated by DpnI Selection on Hemimethylated DNA. *In: BRAMAN, J. (ed.) In Vitro Mutagenesis Protocols*. Totowa, NJ: Humana Press.
- Kumar, A., Basak, S., Rao, S., Gicheru, Y. W., Mayer, M., Sansom, M. S., Chakrapani, S. (2020) Mechanisms of activation and desensitization of full-length glycine receptor in lipid nanodiscs. *Nat Commun.* **11**(3752)
- Li, Y. & Sousa, R. (2012). Expression and purification of E. coli BirA biotin ligase for in vitro biotinylation. *Protein Expr Purif*, **82**, 162-7.
- Liang, Y. L., Khoshouei, M., Deganutti, G., Glukhova, A., Koole, C., Peat, T. S., Radjainia, M., Pnitzko, J. M., Baumeister, W., Miller, L. J., Hay, D. L., Christopoulos, A., Reynolds, C. A., Wootten, D. & Sexton, P. M. (2018). Cryo-EM structure of the active, Gs-protein complexed, human CGRP receptor. *Nature*, **561**, 492-497.
- Lin-Cereghino, J., Wong, W. W., Xiong, S., Giang, W., Luong, L. T., Vu, J., Johnson, S. D. & Lin-Cereghino, G. P. (2005). Condensed protocol for competent cell preparation and transformation of the methylotrophic yeast *Pichia pastoris*. *BioTechniques*, **38**, 44-48.
- Liu, W., Chun, E., Thompson, A. A., Chubukov, P., Xu, F., Katritch, V., Han, G. W., Roth, C. B., Heitman, L. H., Ap, I. J., Cherezov, V. & Stevens, R. C. (2012). Structural basis for allosteric regulation of GPCRs by sodium ions. *Science*, **337**, 232-6.
- Liu, X., Ping, H., Zhang, C. (2014) Rapid establishment of a HEK 293 cell line expressing FVIII-BDD using AAV site-specific integration plasmids. *BMC Research Notes*. **7**(1): 1–6.
- Loll, P. J. (2003). Membrane protein structural biology: the high throughput challenge. *Journal of Structural Biology*, **142**, 144-153.
- Luttrell, L. M. & Lefkowitz, R. J. (2002). The role of  $\beta$ -arrestins in the termination and transduction of G-protein-coupled receptor signals. *Journal of Cell Science*, **115**, 455.
- Ma, L. & Pei, G. (2007).  $\beta$ -arrestin signaling and regulation of transcription. *Journal of Cell Science*, **120**, 213.
- Manglik, A., Kobilka, B. K. & Steyaert, J. (2017). Nanobodies to Study G Protein–Coupled Receptor Structure and Function. *Annu Rev Pharmacol Toxicol*, **57**, 19-37.
- Marinus, M. & Løbner-Olesen, A. (2014). DNA Methylation. *EcoSal Plus*.
- Massink, A., Gutierrez-De-Teran, H., Lenselink, E. B., Ortiz Zacarias, N. V., Xia, L., Heitman, L. H., Katritch, V., Stevens, R. C. & Ap, I. J. (2015). Sodium ion binding pocket mutations and adenosine A2A receptor function. *Mol Pharmacol*, **87**, 305-13.
- Mcgraw, C., Yang, L., Levental, I., Lyman, E. & Robinson, A. S. (2019). Membrane cholesterol depletion reduces downstream signaling activity of the adenosine A2A receptor. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, **1861**, 760-767.
- Mclatchie, L. M., Fraser, N. J., Main, M. J., Wise, A., Brown, J., Thompson, N., Solari, R., Lee, M. G. & Foord, S. M. (1998). RAMPs regulate the transport and ligand specificity of the calcitonin-receptor-like receptor. *Nature*, **393**, 333-9.

- Miller, P. S., Barwell, J., Poyner, D. R., Wigglesworth, M. J., Garland, S. L., Donnelly, D. (2010) Non-peptidic antagonists of the CGRP receptor, BIBN4096BS and MK-0974, interact with the calcitonin receptor-like receptor via methionine-42 and RAMP1 via tryptophan-74. *Biochem Biophys Res Commun.* **391**(1): 437-442
- Milligan, G. & Kostenis, E. (2006). Heterotrimeric G-proteins: a short history. *Br J Pharmacol*, **147 Suppl 1**, S46-55.
- Moraes, I., Evans, G., Sanchez-Weatherby, J., Newstead, S. & Stewart, P. D. (2014). Membrane protein structure determination - the next generation. *Biochim Biophys Acta*, **1838**, 78-87.
- Mishra, P., Singh, U., Pandey, C. M., Mishra, P., Pandey, G. (2019) Application of student's t-test, analysis of variance, and covariance. *Annals of cardiac anaesthesia*, **22**(4): 407-411
- Morrison, K. A., Akram, A., Mathews, A., Khan, Z. A., Patel, J. H., Zhou, C., Hardy, D. J., Moore-Kelly, C., Patel, R., Odiba, V., Knowles, T. J., Javed, M. U., Chmel, N. P., Dafforn, T. R. & Rothnie, A. J. (2016). Membrane protein extraction and purification using styrene-maleic acid (SMA) copolymer: effect of variations in polymer structure. *Biochem J*, **473**, 4349-4360.
- Motulsky, H., Neubig, R. (2002) Analyzing radioligand binding data. *Curr Protoc Neurosci*.
- Muratspahic, E., Freissmuth, M., Gruber, C. W. (2019) Nature-derived peptides: a growing niche for GPCR ligand discovery. *Trends in Pharmacological sciences*. **40**(5): 309-326
- Muyldermans, S. (2013). Nanobodies: Natural Single-Domain Antibodies. *Annual Review of Biochemistry*, **82**, 775-797.
- Naganathan, S., Grunbeck, A., Tian, H., Huber, T., Sakmar, T. P. (2013) Genetically-encoded molecular probes to study G protein-coupled receptors. *Journal of Visualized Experiments*, (79): 1–9.
- Neale, C., Herce, H. D., Pomes, R. & Garcia, A. E. (2015). Can Specific Protein-Lipid Interactions Stabilize an Active State of the Beta 2 Adrenergic Receptor? *Biophys J*, **109**, 1652-62.
- Ngo, T., Ilatovskiy, A. V., Stewart, A. G., Coleman, J. L., Mcrobb, F. M., Riek, R. P., Graham, R. M., Abagyan, R., Kufareva, I. & Smith, N. J. (2017). Orphan receptor ligand discovery by pickpocketing pharmacological neighbors. *Nat Chem Biol*, **13**, 235-242.
- Nödling, A. R., Spear, L. A., Williams, T. L., Luk, L. Y. P., Tsai, Y. H. (2019) Using genetically incorporated unnatural amino acids to control protein functions in mammalian cells. *Essays in Biochemistry*. **63**(2): 237–266.
- Nuber, S., Zabel, U., Lorenz, K., Nuber, A., Milligan, G., Tobin, A. B., Lohse, M. J. & Hoffmann, C. (2016).  $\beta$ -Arrestin biosensors reveal a rapid, receptor-dependent activation/deactivation cycle. *Nature*, **531**, 661-664.
- Okada, T., Sugihara, M., Bondar, A. N., Elstner, M., Entel, P., Buss, V. (2004) The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *Journal of Molecular Biology*, **342** (2): 571–583.
- Oluwole, A. O., Danielczak, B., Meister, A., Babalola, J. O., Vargas, C., Keller, S. (2017) Solubilization of Membrane Proteins into Functional Lipid-Bilayer Nanodiscs Using a Diisobutylene/Maleic Acid Copolymer. *Angewandte Chemie - International Edition*. **56**(7): 1919–1924.

- Oluwole, A. O., Klingler, J., Danielczak, B., Babalola, J. O., Vargas, C., Pabst, G., Keller, S. (2017) Formation of Lipid-Bilayer Nanodiscs by Diisobutylene/Maleic Acid (DIBMA) Copolymer. *Langmuir*, **33**(50): 14378–14388.
- Orban, T., Gupta, S., Palczewski, K. & Chance, M. R. (2010). Visualizing water molecules in transmembrane proteins using radiolytic labeling methods. *Biochemistry*, **49**, 827-34.
- Overington, J. P. Al-lazikani, B. Hopkins, A. L. (2006) How many drug targets are there? *Nature review*. **5**, 993-996
- Perkin Elmer. LANCE cAMP assays. [online] Last accessed: 23.12.2019 <<https://www.perkinelmer.com/lab-products-and-services/application-support-knowledgebase/lance/lance-camp-asay.html>>
- Pollock, N. L., Lee, S. C., Patel, J. H., Gulamhussein, A. A., Rothnie, A. J.(2018) Structure and function of membrane proteins encapsulated in a polymer-bound lipid bilayer. *Biochimica et Biophysica Acta – Biomembranes*. **1860**(4): 809–817.
- Rasmussen S. G, Choi H. J., Fung J. J., Pardon E., Casarosa P., Chae P. S., Devree B. T., Rosenbaum D. M., Thian F. S., Kobilka T. S., Schnapp A., Konetcki I., Sunahara R. K., Gellman S. H., Pautsch A., Steyaert J., Weis W. I., Kobilka B. K. (2011) Structure of a nanobody-stabilized active state of the  $\beta$ 2 adrenoceptor. *Nature*. **469**(7329): 175–180.
- Rahimzadeh, M., Sadeghizadeh, M., Najafi, F., Arab, S. & Mobasheri, H. (2016). Impact of heat shock step on bacterial transformation efficiency. *Molecular biology research communications*, **5**, 257-261.
- Resemann, A., Wunderlich, D., Rothbauer, U., Warcheid, B., Leonhardt, H., Fuchser, J., Kuhlmann, K., Suckau, D. (2010) Top-down de novo protein sequencing of a 13.6 kDa camelid single heavy chain antibody by matrix-assisted laser desorption ionization-time-of-flight/ time-of-flight mass spectrometry. *Analytical Chemistry*. **82**(8): 3283–3292.
- Roth, B. L. and Kroeze, W. K. (2015) Integrated Approaches for Genome-wide Interrogation of the Druggable Non-olfactory G Protein-coupled Receptor. *Journal of Biological Chemistry*. **290**(32): 19471–19477.
- Roth, D. B. (2014) V(D)J Recombination: Mechanism, Errors and Fidelity. *Microbiol Spectr*. **2**(6): doi: 10.1128/microbiolspec.MDNA3-0041-2014
- Routledge, S. J., Mikaliunaite, L., Patel, A., Clare, M., Cartwright, S. P., Bawa, Z., Wilks, M. D., Low, F. Hardy, D., Rothnie, A. J. & Bill, R. M. (2016). The synthesis of recombinant membrane proteins in yeast for structural studies. *Methods*, **95**, 26-37.
- Rovati, G.E. Capra, V. Neubig, R.R. (2007) The highly conserved DRY motif of class A G protein-coupled receptors: beyond the ground state. *Mol Pharmacol*. **71**(4): 959-64
- Rucktooa, P., Cheng, R.K.Y., Segala, E., Geng, T., Errey, J. C., Brown, G. A., Cooke, R. M., Marshall, F. H., Dore, A. S. (2018) Towards high throughput GPCR crystallography: In Meso soaking of Adenosine A2A Receptor crystals. *Scientific Reports*. **8**(1): 2–8.
- Ruiz, M.d.L. Lim, Y-H. Zheng, J. (2013) Adenosine A<sub>2A</sub> Receptor as a Drug Discovery Target. *J. Med. Chem*. **57**(9): 3623-3650
- Russell, F. A., King, R., Smillie, S. J., Kodji, X. & Brain, S. D. (2014). Calcitonin Gene-Related Peptide: Physiology and Pathophysiology. *Physiological Reviews*, **94**, 1099-1142.

- Sabir, J. S. M., Atef, A., El-Domyati, F. M., Edris, S., Hajrah, N., Alzohairy, A. M., Bahieldin, A. (2014) Construction of naïve camelids VHH repertoire in phage display-based library. *Immunology* **337**(4): 244-249
- Sabbadin, D., Ciancetta, A., Moro, S. (2014) Bridging molecular docking to membrane molecular dynamics to investigate GPCR-ligand recognition: The human A2A adenosine receptor as a key study. *J Chem Inf Model.* **54**(1): 169-183
- Sato, J., Makita, N. & Iiri, T. (2016). Inverse agonism: the classic concept of GPCRs revisited [Review]. *Endocrine Journal*, **63**, 507-514.
- Salvatore, C. A., Moore, E. L., Calamari, A., Cook, J. J., Michener, M. S., O'Malley, S., Miller, P. J., Williams, D. L. Jr., Zeng, Z., Danziger, A., Lynch, J. J., Regan, C. P., Fay, J. F., Tang, Y. S., Li, C., Pudvah, N. T., White, R. B., Bell, I. M., Gallicchio, S. N., Graham, S. L., Selnick, H. G., Vacca, J. P., Kane, S. A. (2010) Pharmacological properties of MK-3207, a potent and orally active calcitonin gene-related peptide receptor antagonist. *Journal of Pharmacology and Experimental Therapeutics.* **333**(1): 152–160.
- Scheidelaar, S., Koorengel, M. C., Van Walree, C. A., Dominguez, J. J., Dorr, J. M. & Killian, J. A. (2016). Effect of Polymer Composition and pH on Membrane Solubilization by Styrene-Maleic Acid Copolymers. *Biophys J*, **111**, 1974-1986.
- Schmied, W. H., Elsässer, S. J., Uttamapinant, C., Uttamapinant, C., Chin, J. W. (2014) Efficient multisite unnatural amino acid incorporation in mammalian cells via optimized pyrrolysyl tRNA synthetase/tRNA expression and engineered eRF1. *Journal of the American Chemical Society.* **136**(44): 15577–15583.
- Schneider, E. H., Schnell, D., Strasser, A., Dove, S. and Seifert, R. (2010) Impact of the DRY Motif and the Missing “Ionic Lock ” on Constitutive Activity and G protein Coupling of the Human Histamine H 4 Receptor. *Journal of Pharmacology and Experimental Therapeutics.* **333**(2): 382–392.
- Seddon, A. M., Curnow, P. & Booth, P. J. (2004). Membrane proteins, lipids and detergents: not just a soap opera. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, **1666**, 105-117.
- Seidel, L., Coin, I. (2018) Mapping of protein interfaces in live cells using generically encoded crosslinkers. *Methods Mol Biol.* **1728**:221-235
- Serrano-Vega, M. J., Magnani, F., Shibata, Y. & Tate, C. G. (2008). Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form. *Proc Natl Acad Sci U S A*, **105**, 877-82.
- Shao, N., Singh, N. S., Slade, S. E., Jones, A. M. E., Balasubramanian, M. K. (2015) Site Specific Genetic Incorporation of Azidophenylalanine in *Schizosaccharomyces pombe*. *Scientific Reports.* **5**(1): 17196.
- Shimamura, T., Hiraki, K., Takahashi, N., Hori, T., Ago, H., Masuda, K., Takio, K., Ishiguro, M., Miyano, M. (2008) Crystal structure of squid rhodopsin with intracellularly extended cytoplasmic region. *Journal of Biological Chemistry.* **283**(26): 17753-17756
- Shoemaker, S. C., Ando, N. (2018) X-rays in the cryo-EM era: structural biology's dynamic future. *Biochemistry.* **57**(3): 277-285

- Simms, J., Uddin, R., Sakmar, T.P., Gingell, J. J., Garelja, M. L., Hay, D. L., Brimble, M. A., Harris, P. W., Reynolds, C. A., Poyner, D. R. (2018) Photoaffinity Cross-Linking and Unnatural Amino Acid Mutagenesis Reveal Insights into Calcitonin Gene-Related Peptide Binding to the Calcitonin Receptor-like Receptor/Receptor Activity-Modifying Protein 1 (CLR/RAMP1) Complex. *Biochemistry*. **57**(32): 4915–4922.
- Simms, J., Routledge, S., Uddin, R. & Poyner, D. 2019. The Structure of the CGRP and Related Receptors. In: BRAIN, S. D. & GEPPEPPI, P. (eds.) *Calcitonin Gene-Related Peptide (CGRP) Mechanisms: Focus on Migraine*. Cham: Springer International Publishing.
- Siu, F. Y., He, M., De Graaf, C., Han, G. W., Yang, D., Zhang, Z., Zhou, C., Xu, Q., Wacker, D., Joseph, J. S., Liu, W., Lau, J., Cherezov, V., Katritch, V., Wang, M.-W. & Stevens, R. C. (2013). Structure of the human glucagon class B G-protein-coupled receptor. *Nature*, **499**, 444-449.
- Song, W., Yen, H. Y., Robinson, C. V. & Sansom, M. S. P. (2019). State-dependent Lipid Interactions with the A2a Receptor Revealed by MD Simulations Using In Vivo-Mimetic Membranes. *Structure*, **27**, 392-403 e3.
- Spriestersbach, A., Kubicek, J., Schäfer, F., Block, H. & Maertens, B. 2015. Chapter One - Purification of His-Tagged Proteins. In: LORSCH, J. R. (ed.) *Methods Enzymol.* Academic Press.
- Sriskanda, V., Moyer, R. W. & Shuman, S. (2001). NAD<sup>+</sup>-dependent DNA Ligase Encoded by a Eukaryotic Virus. *Journal of Biological Chemistry*, **276**, 36100-36109.
- Strange, P. G. (2008). Agonist binding, agonist affinity and agonist efficacy at G protein-coupled receptors. *Br J Pharmacol*, **153**, 1353-63.
- Stoddart, L. A., Kilpatrick, L. E., Briddon, S. J., Hill, S. J. (2015) Probing the pharmacology of G protein-coupled receptors with fluorescent ligands. *Neuropharmacology*, **98**: 48–57.
- Stoddart, L. A., White, C. W., Nguyen, K., Hill, S. J., Pflieger, K. D. (2016) Fluorescence- and bioluminescence-based approaches to study GPCR ligand binding. *British Journal of Pharmacology*, **173** (20): 3028–3037.
- Stroud, Z., Hall, S. C. L. & Dafforn, T. R. (2018). Purification of membrane proteins free from conventional detergents: SMA, new polymers, new opportunities and new insights. *Methods*, **147**, 106-117.
- Swainsbury, D. J. K., Scheidelaar, S., Foster, N., Van Grondelle, R., Killian, J. A. & Jones, M. R. (2017). The effectiveness of styrene-maleic acid (SMA) copolymers for solubilisation of integral membrane proteins from SMA-accessible and SMA-resistant membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, **1859**, 2133-2143.
- Sun, B., Bachhawat, P., Chu, M. L., Wood, M., Ceska, T., Sands, Z. A. and Mercier, J. (2017) Crystal structure of the adenosine A<sub>2A</sub> receptor bound to an antagonist reveals a potential allosteric pocket. *PNAS*. **114**(8): 2–7.
- Tehan, B. G., Bortolato, A., Blaney, F. E., Weir, M. P. and Mason, J. S. (2014) ‘Pharmacology & Therapeutics Unifying Family A GPCR Theories of Activation’. *Pharmacology and Therapeutics*. Elsevier Inc., **143**(1): 51–60.
- ter Haar, E., Koth, C. M., Abdul-Manan, N., Swenson, L., Coll, J. T., Lippke, J. A., Lepre, C. A., Garcia-Guzman, M., Moore, J. M. (2010) Crystal structure of the ectodomain complex of the CGRP receptor, a class-V GPCR, reveals the site of drug antagonism. *Structure*. **18**(9): 1083-93

- ThermoFisher Scientific. Nunc MaxiSorp flat-bottom. [online] Last accessed: 25.11.2019 <<https://www.thermofisher.com/order/catalog/product/44-2404-21#/44-2404-21>>
- ThermoFisher Scientific. Adsorption Immunoassay Plates. [online] Last accessed: 06.08.2020 <[210](https://www.thermofisher.com/uk/en/home/life-science/protein-biology/protein-assays-analysis/elisa/elisa-microplates-plasticware/adsorption-immunoassay-plates.html#:~:text=The%20MaxiSorp%20surface%20is%20a,with%20polar%20or%20hydrophilic%20groups.&text=It%20has%20high%20affinity%20to,protein%20antigens%20including%20virus%20antigens.></a></p>
<p>Thomas, P., Smart, T. G. (2005) HEK293 cell line: A vehicle for the expression of recombinant proteins. <i>Journal of Pharmacological and Toxicological Methods</i>. <b>51</b>(3): 187–200.</p>
<p>Thorsen, T. S., Matt, R., Weis, W. I. & Kobilka, B. K. (2014). Modified T4 Lysozyme Fusion Proteins Facilitate G Protein-Coupled Receptor Crystallogenesis. <i>Structure</i>, <b>22</b>, 1657–64.</p>
<p>Tringali, G. & Navarra, P. (2019). Anti-CGRP and anti-CGRP receptor monoclonal antibodies as antimigraine agents. Potential differences in safety profile postulated on a pathophysiological basis. <i>Peptides</i>, <b>116</b>, 16–21.</p>
<p>Tsai, C.-J., Marino, J., Adaixo, R., Pamula, F., Muehle, J., Maeda, S., Flock, T., Taylor, N. M., Mohammed, I., Matile, H., Dawson, R. J., Deupi, X., Stahlberg, H. & Schertler, G. (2019). Cryo-EM structure of the rhodopsin-Gai-<math>\beta\gamma</math> complex reveals binding of the rhodopsin C-terminal tail to the g<math>\beta</math> subunit. <i>eLife</i>, <b>8</b>, e46041.</p>
<p>Tschammer, N. (2016) Allosteric Modulators of the Class A G Protein Coupled Receptors BT - Protein Targeting Compounds: Prediction, Selection and Activity of Specific Inhibitors. <i>Protein targeting compounds. Advances in experimental medicine and biology</i>. <b>917</b>: 185–207.</p>
<p>Tsukigawa, K., Liao, L., Nakamura, H., Fang, J., Greish, K., Otagiri, M., Maeda, H. (2015) Synthesis and therapeutic effect of styrene-maleic acid copolymer-conjugated pirarubicin. <i>Cancer Sci</i>. <b>106</b>(3): 270–278</p>
<p>Uddin, R., Simms, J. & Poyner, D. (2018). Functional characterisation of G protein-coupled receptors. <i>Methods</i>, <b>147</b>, 213–220.</p>
<p>Ukena, D., Daly, J. W., Kirk, K. L., Jacobson, K. A. (1986) Functionalized congeners of 1,3-dipropyl-8-phenylxanthine: Potent antagonists for adenosine receptors that modulate membrane adenylate cyclase in pheochromocytoma cells, platelets and fat cells. <i>Life Sciences</i>, <b>38</b> (9): 797–807.</p>
<p>Upton, R. N. & Mould, D. R. (2014). Basic concepts in population modeling, simulation, and model-based drug development: part 3-introduction to pharmacodynamic modeling methods. <i>CPT Pharmacometrics Syst Pharmacol</i>, <b>3</b>, e88.</p>
<p>Urizar, E. Claeysen, S. Deupi, X. Govaerts, C. Costagliola, S. Vassart, G. Pardo, L. (2005) An activation switch in the rhodopsin family of G protein coupled receptors: the thyrotropin receptor. <i>J Biol Chem</i>. <b>280</b>(17): 17135–41</p>
<p>Vallance, P. & Smart, T. G. (2006). The future of pharmacology. <i>Br J Pharmacol</i>, <b>147 Suppl 1</b>, S304–7. Pors, K. (2011). Drug Discovery into the 21st Century. <i>Drug Discovery and Development - Present and Future</i>. 69–96</p>
<p>Vecchio, E.A. White, P.J. May, L.T. (2017) Targeting Adenosine Receptors for the Treatment of Cardiac Fibrosis. <i>Front Pharmacol</i>. <b>8</b>: 243</p>
</div>
<div data-bbox=)

- Venkatakrishnan, A. J., Ma, A. K., Fonseca, R., Latorraca, N. R., Kelley, B., Betz, R. M., Asawa, C., Kobilka, B. K., Dror, R. O. (2019) Diverse GPCRs exhibit conserved water networks for stabilization and activation. *Proceedings of the National Academy of Sciences of the United States of America*. **116**(8): 3288–3293
- Venkatakrishnan, A.J. Deupi, X. Lebon, G. Tate, C.G. Schertler, G.F. Babu, M.M. (2013) Molecular signatures of G protein-coupled receptors. *Nature*. **494**(7436): 185-194
- Vilardaga, J. P., Romero, G., Friedman, P. A. & Gardella, T. J. (2011). Molecular basis of parathyroid hormone receptor signaling and trafficking: a family B GPCR paradigm. *Cell Mol Life Sci*, **68**, 1-13.
- Violin, J. D., Crombie, A. L., Soergel, D. G. & Lark, M. W. (2014). Biased ligands at G-protein-coupled receptors: promise and progress. *Trends Pharmacol Sci*, **35**, 308-16.
- Vogl, T., Gebbie, L., Palfreyman, R. W. & Speight, R. (2018). Effect of Plasmid Design and Type of Integration Event on Recombinant Protein Expression in *Pichia pastoris*. *Appl Environ Microbiol*, **84**.
- Wang, H., Wang, J. (2016) How cryo-electron microscopy and x-ray crystallography complement each other. *Protein Science*. **26**(1): 32-39
- Wang, Q., Parrish, A. R., Wang, L. (2009) Expanding the genetic code for biological studies. *Chem Biol*. **16**(3): 323-336
- Wang, Z., Sun, H., Yao, X., Li, D., Xu, L., Li, Y., Tian, S. & Hou, T. (2016). Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: the prediction accuracy of sampling power and scoring power. *Physical Chemistry Chemical Physics*, **18**, 12964-12975.
- Walker, C. S., Raddant, A. C., Woolley, M. J., Russo, A. F., Hay, D. L. (2018) CGRP receptor antagonist activity of olcegepant depends on the signalling pathway measured. *Cephalalgia*. **38**(3): 437-451
- Warfvinge, K., Edvinsson, L., Pickering, D. S., Sheykhzade, M. (2019) The presence of calcitonin gene-related peptide and its receptors in rat, pig and human brain: species differences in calcitonin gene-related peptide pharmacology. *Pharmacology* **104**: 332-341
- Watkins, H. A., Rathbone, D. L., Barwell, J., Hay, D. L., Poyner, D. R. (2013) Structure-activity relationships for  $\alpha$ -calcitonin gene-related peptide. *British journal of pharmacology*. **170**(7): 1308-1322
- Weston, C., Winfield, I., Harris, M., Hodgson, R., Shah, A., Dowell, S. J., Mobarec, J. C., Woodlock, D. A., Reynolds, C. A., Poyner, D. R., Watkins, H. A. & Ladds, G. (2016). Receptor Activity-modifying Protein-directed G Protein Signaling Specificity for the Calcitonin Gene-related Peptide Family of Receptors. *J Biol Chem*, **291**, 21925-21944.
- Wesolowski, J., Alzogaray, V., Reyelt, J., Unger, M., Juarez, K., Urrutia, M., Cauerhiff, A., Danquah, W., Rissiek, B., Scheuplein, F., Schwarz, N., Adriouch, S., Boyer, O., Seman, M., Licea, A., Serreze, D. V., Goldbaum, F. A., Haag, F., Koch-Nolte, F. (2009) Single domain antibodies: Promising experimental and therapeutic tools in infection and immunity. *Medical Microbiology and Immunology*. **198**(3): 157–174
- Wheatley, M., Charlton, J., Jamshad, M., Routledge, Sarah j., Bailey, S., La-Borde, Penelope j., Azam, Maria t., Logan, Richard t., Bill, Roslyn m., Dafforn, Tim r. & Poyner, David r. (2016). GPCR–styrene maleic acid lipid particles (GPCR–SMALPs): their nature and potential. *Biochemical Society Transactions*, **44**, 619-623.

- Wheatley, M., Wootten, D., Conner, M., Simms, J., Kendrick, R., Logan, R., Poyner, D. & Barwell, J. (2012). Lifting the lid on GPCRs: the role of extracellular loops. *Br J Pharmacol*, **165**, 1688-1703.
- White, K. L., Eddy, M. T., Gao, Z. G., Han, G. W., Lian, T., Deary, A., Patel, N., Jacobson, K. A., Katritch, V. & Stevens, R. C. (2018). Structural Connection between Activation Microswitch and Allosteric Sodium Site in GPCR Signaling. *Structure*, **26**, 259-269 e5.
- Wold, E. A., Chen, J., Cunningham, K. A. & Zhou, J. (2019). Allosteric Modulation of Class A GPCRs: Targets, Agents, and Emerging Concepts. *Journal of Medicinal Chemistry*, **62**, 88-127.
- Woolley, M. J., Simms, J., Mobarec, J. C., Reynolds, C. A., Poyner, D. R., Conner, A. C. (2017) Understanding the molecular functions of the second extracellular loop (ECL2) of the calcitonin gene-related peptide (CGRP) receptor using a comprehensive mutagenesis approach. *Molecular and Cellular Endocrinology*. **454**: 39–49.
- Wootten, D., Christopoulos, A., Marti-Solano, M., Babu, M. M. & Sexton, P. M. (2018). Mechanisms of signalling and biased agonism in G protein-coupled receptors. *Nat Rev Mol Cell Biol*, **19**, 638-653.
- Wootten, D., Lindmark, H., Kadmiel, M., Willcockson, H., Caron, K. M., Barwell, J., Drmota, T. & Poyner, D. R. (2013). Receptor activity modifying proteins (RAMPs) interact with the VPAC2 receptor and CRF1 receptors and modulate their function. *Br J Pharmacol*, **168**, 822-34.
- Wu, L., Wu, B. (2015) Analysis of protein ligand-receptor binding by photoaffinity cross-linking. *Current protocols in protein science*. **79**(1): doi: 10.1002/0471140864.ps1926s79.#
- Wu, C. H., Liu, I. J., Lu, R. M. & Wu, H. C. (2016). Advancement and applications of peptide phage display technology in biomedical science. *J Biomed Sci*, **23**, 8.
- Xian, S., Parayath, N. N., Nehoff, H., Giles, N. M., Greish, K. (2015) The use of styrene maleic acid nanomicelles encapsulating the synthetic cannabinoid analog WIN55,212-2 for the treatment of cancer. *Anticancer Research*. **35**(9): 4707-4712
- Xue, M., Cheng, L., Faustino, I., Guo, W. & Marrink, S. J. (2018). Molecular Mechanism of Lipid Nanodisk Formation by Styrene-Maleic Acid Copolymers. *Biophys J*, **115**, 494-502.
- Yan, J., Li, G., Hu, Y., Ou, W., Wan, Y. (2014) Construction of a synthetic phage-displayed Nanobody library with CDR3 regions randomized by trinucleotide cassettes for diagnostic applications. *Journal of Translational Medicine*. **12**(1): 1–12.
- Yang, L., Jin, M., Du, P., Chen, G., Zhang, C., Wang, J., Jin, F., Shao, H., She, Y., Wang, S. & Zheng, L. (2015). Study on Enhancement Principle and Stabilization for the Luminol-H<sub>2</sub>O<sub>2</sub>-HRP Chemiluminescence System. *PLoS One*, **10**, e0131193.
- Yen, H. Y., Hoi, K. K., Liko, I., Hedger, G., Horrell, M. R., Song, W., Wu, D., Heine, P., Warne, T., Lee, Y., Carpenter, B., Pluckthun, A., Tate, C. G., Sansom, M. S. P., Robinson, C. V. (2018) PtdIns(4,5)P<sub>2</sub> stabilizes active states of GPCRs and enhances selectivity of G-protein coupling. *Nature*. **559**(7714): 423-427
- Yin, H. & Flynn, A. D. (2016). Drugging Membrane Protein Interactions. *Annu Rev Biomed Eng*, **18**, 51-76.
- Zhang, H., Qiao, A., Yang, D., Yang, L., Dai, A., De Graaf, C., Reedtz-Runge, S., Dharmarajan, V., Zhang, H., Han, G. W., Grant, T. D., Sierra, R. G., Weierstall, U., Nelson, G., Liu, W., Wu, Y., Ma, L., Cai, X., Lin, G., Wu, X., Geng, Z., Dong, Y., Song, G., Griffin, P. R., Lau, J., Cherezov, V., Yang, H., Hanson, M. A., Stevens, R. C., Zhao, Q., Jiang, H., Wang, M.-W. & Wu, B. (2017). Structure of the full-length glucagon class B G-protein-coupled receptor. *Nature*, **546**, 259-264.

Zhang, H., Qiao, A., Yang, L., Van Eps, N., Frederiksen, K., Yang, D., Dai, A., Cai, X., Zhang, H., Yi, C., Cao, C., He, L., Yang, H., Lau, J., Ernst, O. P., Hanson, M. A., Stevens, R. C., Wang, M. W., Reedtz-Runge, S., Jiang, H., Zhao, Q., Wu, B. (2018) Structure of the glucagon receptor in complex with a glucagon analogue. *Nature*. **553**(7686): 106-110

Zhao, P., Liang, Y. L., Belousoff, M. J., Deganutti, G., Fletcher, M. M., Willard, F. S., Bell, M. G., Christe, M. E., Sloop, K. W., Inoue, A., Truong, T. T., Clydesdale, L., Furness, S. G. B., Christopolous, A., Wang, M. W., Miller, L. J., Reynolds, C. A., Danev, R., Sexton, P. M., Wootten, D. (2020) Activation of the GLP-1 receptor by a non-peptidic agonist. *Nature*. **577**(7790): 432-436.

Zheng, H. Loh, H.H. Law, P-Y. (2010) Agonist-selective signalling of G Protein-Coupled Receptor: Mechanisms and Implications. *IUBMB Life*. **62**(2): 112-119

Zhu, J., Pedersen, M. D., Ahmed, L. S., Abdolalizadeh, B., Grell, A. S., Berg, J. O., Thulstrup, P. W., Franzyk, H., Edvinsson, L., Sams, A., Sheykhzade, M., Hansen, P. R. (2020) Fluorescent analogues of human  $\alpha$ -calcitonin gene-related peptide with potent vasodilator activity. *Int J Mol Sci*. **21**(4): pii: E1343

## Appendix:

### Sequencing primers for pcDNA3.1- T8-HA-CLR amber stop codon mutants:

T7 primer

Forward: 5' TAATACGACTCACTATAGGGAACCC 3'

BGH primer

Reverse: 5' TAGAAGGCACAGTCGAGGCTG 3'

### Oligonucleotide primers used to generate T8-HA-CLR amber stop codon mutants:

hCLR cDNA sequence (Residues highlighted in blue were mutated with the 'tag' codon):

```
atggagaaaaagtgtaccctgtatcttctggttctcttgcctttttttatgattcttgtt
M E K K C T L Y F L V L L P F F M I L V
acagcagaattagaagagagtcctgaggactcaattcagttgggagttactagaaataaa
T A E L E E S P E D S I Q L G V T R N K
atcatgacagctcaatatgaatgttaccaaagattatgcaagacccattcaacaagca
I M T A Q Y E C Y Q K I M Q D P I Q Q A
gaaggcgtttactgcaacagaacctgggatggatggctctgctggaacgatggtgcagca
E G V Y C N R T W D G W L C W N D V A A
ggaactgaatcaatgcagctctgacctgattactttcaggactttgatccatcagaaaaa
G T E S M Q L C P D Y F Q D F D P S E K
gttacaagatctgtgaccaagatggaaactggttagacatccagcaagcaacagaaca
V T K I C D Q D G N W F R H P A S N R T
tgacaaattataccagtgtaatgttaacaccacgagaaagtgaagactgcactaaat
W T N Y T Q C N V N T H E K V K T A L N
ttgttttaccgtaccataaattggacacggattgtctattgcatcactgcttatctcgctt
L F Y L T I I G H G L S I A S L L I S L
ggcatattcttttatttcaagagcctaagttgcaaaggattaccttacacaaaaatctg
G I F F Y F K S L S C Q R I T L H K N L
ttcttctcatttgtttgtaactctgttgaacaatcattcacctcactgcagtgcccaac
F F S F V C N S V V T I I H L T A V A N
aaccaggccttagtagccacaaatcctgttagttgcaaagtgtcccagttcattcatctt
N Q A L V A T N P V S C K V S Q F I H L
tacctgatgggctgtaattacttttggatgctctgtgaaggcatttacctacacacactc
Y L M G C N Y F W M L C E G I Y L H T L
attgtggtggccgtggttgcagagaagcaacatttaattgtggtattatcttctggctgg
I V V A V F A E K Q H L M W Y Y F L G W
ggatttccactgattcctgcttgtatacatgccattgctagaagcttatattacaatgac
G F P L I P A C I H A I A R S L Y Y N D
aattgctggatcagttctgatacccatctcctctacattatccatggcccaatttgtgct
N C W I S S D T H L L Y I I H G P I C A
gctttactggtgaatcttttttcttggtaaatattgtacgcgcttctcatcaccaagtta
A L L V N L F F L L N I V R V L I T K L
aaagttacacaccaagcgaatccaatctgtacatgaaagctgtgagagctactcttatc
K V T H Q A E S N L Y M K A V R A T L I
ttgggtgccattgcttggcattgaatttgtgctgattccatggcgacctgaaggaaagatt
L V P L L G I E F V L I P W R P E G K I
gcagaggaggtatatgactacatcatgcacatccttatgcacttccagggctcttttggctc
A E E V Y D Y I M H I L M H F Q G L L V
tctaccattttctgcttctttaaattggagaggttcaagcaattctgagaagaaactggaat
S T I F C F F N G E V Q A I L R R N W N
caatacaaaatccaatttggaaacagcttttccaactcagaagctcttcgtagtgcgtctc
Q Y K I Q F G N S F S N S E A L R S A S
tacacagtgtaacaatcagtgatgggtccaggttatagtcactgctcctagtgaacac
Y T V S T I S D G P G Y S H D C P S E H
```

ttaaatggaaaaagcatccatgatattgaaaatggttctctttaaaccagaaaatttatat  
L N G K S I H D I E N V L L K P E N L Y  
aattga  
N -

### **ECL1 mutants**

L195stop

Forward: 5' GTAACAATCATTCACTAGACTGCAGTGGCCAACAA 3'

Reverse: 3' TTGTTGGCCACTGCAGTCTAGTGAATGATTGTTAC 5'

T196stop

Forward: 5' CATTACCTCTAGGCAGTGGCCAACAACCAGGC 3'

Reverse: 3' GCCTGGTTGTTGGCCACTGCCTAGAGGTGAATG 5'

A197stop

Forward: 5' CATTACCTCACTTAGGTGGCCAACAACCAGGC 3'

Reverse: 3' GCCTGGTTGTTGGCCACCTAAGTGAGGTGAATG 5'

V198stop

Forward: 5' CATTACCTCACTGCATAGGCCAACAACCAGGC 3'

Reverse: 3' GCCTGGTTGTTGGCCTATGCAGTGAGGTGAATG 5'

A199stop

Forward: 5' CATTACCTCACTGCAGTGTAGAACAACCAGGC 3'

Reverse: 3' GCCTGGTTGTTCTACTGCAGTGAGGTGAATG 5'

N200stop

Forward: 5' CATTACCTCACTGCAGTGGCCTAGAACCAGGC 3'

Reverse: 3' GCCTGGTTCTAGGCCACTGCAGTGAGGTGAATG 5'

N201stop

Forward: 5' CACTGCAGTGGCCAACACTAGCAGGCCTTAGTAGC 3'

Reverse: 3' GCTACTAAGGCCTGCTAGTTGGCCACTGCAGTG 5'

Q202stop

Forward: 5' CACTGCAGTGGCCAACAACACTGAGCCTTAGTAGC 3'

Reverse: 3' GCTACTAAGGCTCAGTTGTTGGCCACTGCAGTG 5'

A203stop

Forward: 5' GGCCAACAACCAGTGATTAGTAGCCACAAATCCTG 3'

Reverse: 3' CAGGATTTGTGGCTACTAATCACTGGTTGTTGGCC 5'

L204stop

Forward: 5' GCCAACAACCAGGCCTAGGTAGCCACAAATCCTG 3'

Reverse: 3' CAGGATTTGTGGCTACCTAGGCCTGGTTGTTGGC 5'

A206stop

Forward: 5' GGCCAACAACCAGGCCTTAGTATAGACAAATCCTG 3'

Reverse: 3' CAGGATTTGTCTATACTAAGGCCTGGTTGTTGGCC 5'

T207stop

Forward: 5' CCAGGCCTTAGTAGCCTAGAATCCTGTTAGTTGCA 3'

Reverse: 3' TGCAACTAACAGGATTCTAGGCTACTAAGGCCTGG 5'

N208stop

Forward: 5' CCAGGCCTTAGTAGCCACATAGCCTGTTAGTTGCA 3'

Reverse: 3' TGCAACTAACAGGCTATGTGGCTACTAAGGCCTGG 5'

P209stop

Forward: 5' CCAGGCCTTAGTAGCCACAAATTAGGTTAGTTGCA 3'

Reverse: 3' TGCAACTAACCTAATTTGTGGCTACTAAGGCCTGG 5'

V210stop

Forward: 5' GTAGCCACAAATCCTTAGAGTTGCAAAGTGTCCCA 3'

Reverse: 3' TGGGACACTTTGCAACTCTAAGGATTTGTGGCTAC 5'

S211stop

Forward: 5' GTAGCCACAAATCCTGTTTAGTGCAAAGTGTCCCA 3'

Reverse: 3' TGGGACACTTTGCACTAAACAGGATTTGTGGCTAC 5'

### **ECL2 mutant**

L291stop

Forward: 5' CATCTCTAGTACATTATCCATGGCCCAATTTGTGC 3'

Reverse: 3' GCACAAATTGGGCCATGGATAATGTACTAGAGATG 5'

### **ECL3 mutants**

V350stop

Forward: 5' CTTGGCATTGAATTTTAGCTGATTCCATGGCGA 3'

Reverse: 3' TCGCCATGGAATCAGCTAAAATTCAATGCCAAG 5'

L351stop

Forward: 5' CTTGGCATTGAATTTGTGTAGATTCCATGGCGA 3'

Reverse: 3' TCGCCATGGAATCTACACAAATTCAATGCCAAG 5'

I352stop

Forward: 5' CTTGGCATTGAATTTGTGCTGTAGCCATGGCGA 3'

Reverse: 3' TCGCCATGGCTACAGCACAAATTCAATGCCAAG 5'

P353stop

Forward: 5' GAATTTGTGCTGATTTAGTGGCGACCTGAAGGA 3'

Reverse: 3' TCCTTCAGGTCGCCACTAAATCAGCACAAATTC 5'

W354stop

Forward: 5' GAATTTGTGCTGATTCCATAGCGACCTGAAGGA 3'

Reverse: 3' TCCTTCAGGTCGCTATGGAATCAGCACAAATTC 5'

R355stop

Forward: 5' GAATTTGTGCTGATTCCATGGTAGCCTGAAGGA 3'

Reverse: 3' TCCTTCAGGCTACCATGGAATCAGCACAAATTC 5'

P356stop

Forward: 5' GCTGATTCCATGGCGATAGGAAGGAAAGATTGC 3'

Reverse: 3' GCAATCTTTCCTTCCTATCGCCATGGAATCAGC 5'

E357stop

Forward: 5' GCTGATTCCATGGCGACCTTAGGGAAAGATTGC 3'

Reverse: 3' GCAATCTTTCCTTAAGGTCGCCATGGAATCAGC 5'

G358stop

Forward: 5' CCATGGCGACCTGAATAGAAGATTGCAGAGGAG 3'

Reverse: 3' CTCCTCTGCAATCTTCTATTCAGGTCGCCATGG 5'

K359stop

Forward: 5' CCATGGCGACCTGAAGGATAGATTGCAGAGGAG 3'

Reverse: 3' CTCCTCTGCAATCTATCCTTCAGGTCGCCATGG 5'

I360stop

Forward: 5' CCATGGCGACCTGAAGGAAAGTAGGCAGAGGAG 3'

Reverse: 3' CTCCTCTGCCTACTTTCCTTCAGGTCGCCATGG 5'