

DOCTOR OF PHILOSOPHY

Investigation of the activation mechanism  
and structural characterization of the  
CGRP receptor

Gabriel Kuteyi

2013

Aston University

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

If you have discovered material in AURA 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

# **INVESTIGATION OF THE ACTIVATION MECHANISM AND STRUCTURAL CHARACTERIZATION OF THE CGRP RECEPTOR**

**GABRIEL OLAWALE KUTEYI**

Doctor of Philosophy

**ASTON UNIVERSITY**

**July 2013**

© Gabriel Olawale Kuteyi, 2013

Gabriel Olawale Kuteyi 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 rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledgement.

Aston University

## **Investigation of the activation mechanism and structural characterization of the CGRP receptor**

Gabriel Olawale Kuteyi

Doctor of Philosophy

2013

### **Thesis Summary**

The calcitonin gene-related peptide (CGRP) receptor is an unusual G protein-coupled receptor (GPCR) in that it comprises the calcitonin receptor-like receptor (CLR), receptor activity modifying protein 1 (RAMP1) and the receptor component protein (RCP). The RAMP1 has two other homologues – RAMP2 and RAMP3. The endogenous ligand for this receptor is CGRP, a 37 amino acid neuropeptide that act as a vasodilator. This peptide has been implicated in the aetiology of health conditions such as inflammation, Reynaud's disease and migraine. A clear understanding of the mode of activation of this receptor could be key in developing therapeutic agents for associated health conditions. Although the crystal structure of the N-terminal extracellular domain (ECD) of this receptor (in complex with an antagonist) has been published, the details of receptor-agonist interactions at this domain, and so ultimately the mechanism of receptor activation, are still unclear. Also, the C-terminus of the CLR (in the CGRP receptor), especially around the presumed helix 8 (H8) region, has not been well studied for its role in receptor signalling. This research project investigated these questions.

In this study, certain residues making up the putative N-terminal ligand-binding core of the CLR (in the CGRP receptor) were mapped out and found to be crucial for receptor signalling. They included W69 and D70 of the WDG motif in family B GPCRs, as well as Y91, F92, D94 and F95 in loop 2 of CLR N-terminus. Also, F163 at the cytoplasmic end of TM1 and certain residues spanning H8 and associated C-terminal region of CLR were found to be required for CGRP receptor signalling. These residues were investigated by site-directed mutagenesis where they were mutated to alanine (or other residues in specific cases) and the effect of the mutations on receptor pharmacology assessed by evaluating cAMP production, cell surface expression, total cell expression and  $\alpha$ CGRP-mediated receptor internalization. Moreover, the N-terminal ECDs of the CLR and RAMPs (RAMP1, RAMP2 and RAMP3) were produced in a yeast host strain (*Pichia pastoris*) for the purpose of structural interaction study by surface plasmon resonance (SPR). Following expression and purification, these receptor proteins were found to individually retain their secondary structures when analysed by circular dichroism (CD). Results were analysed and interpreted with the knowledge of the secretin family receptor paradigm.

The research described in this thesis has produced novel data that contributes to a clearer understanding of CGRP receptor pharmacology. The study on CLR and RAMPs ECDs could be a useful tool in determining novel interacting GPCR partners of RAMPs.

**Keywords:** G protein-coupled receptor, calcitonin receptor-like receptor, receptor activity modifying protein, *Pichia pastoris*, extracellular domain, site-directed mutagenesis.

Dedicated to the Glory of God

## **Acknowledgements**

I give all the glory to God Almighty for seeing me through the course of this work.

I would like to specially thank my supervisor, Prof David Poyner for his tremendous help and guidance throughout the course of this project. I also give special gratitude to my Associate Supervisor, Prof Roslyn Bill for her unmeasurable help and support during the course of this research. I would like to acknowledge Dr James Barwell for his advice and contribution to the success of this research. I also acknowledge Drs Nagamani Bora and Matt Conner for teaching me some laboratory techniques and I appreciate the support received from Dr Mohammed Jamshad, Dr Sarah Routledge Mrs Rosemary Parslow and my colleagues with whom I worked in the same laboratory.

My unreserved gratitude goes to my loving parents, Joseph and Mary Kuteyi for always being there for me; and my dearest siblings for their prayers and support.

I thank my friends who have been very supportive throughout the course of this journey. I really appreciate.

## Contents

	<b>Page</b>
Abbreviations	11
List of Figures	16
List of Tables	19
<b>Chapter 1 Introduction</b>	<b>20</b>
1.1 Membrane Proteins	20
1.1.1 The “Receptor Concept”	23
1.2 G-protein Coupled Receptors	28
1.2.1 The GPCR Superfamily	29
1.2.1.1 Rhodopsin-like (family A) GPCRs	30
1.2.1.2 Secretin-like (family B) GPCRs	31
1.2.1.3 Glutamate (family C) GPCRs	31
1.2.1.4 Frizzled/Taste2 (family D)	33
1.2.1.5 Adhesion (family E) GPCRs	34
1.2.2 Activation and signalling cascade	35
1.2.3 GPCR endocytosis	37
1.2.4 Advances in GPCRs crystal structures: the bovine rhodopsin era and beyond	41
1.2.5 GPCR 8 <sup>th</sup> helix: structural and functional variations	46
1.3 The secretin family receptors	50
1.3.1 Structural theme	50
1.3.2 Peptides of the secretin family	54
1.3.3 Mechanism of ligand binding and activation of secretin family receptors: the ‘two-domain’ model	55
1.3.4 The receptor regions in ligand binding and activation	57
1.3.4.1 N-terminal extracellular domain (ECD)	57
1.3.4.2 The transmembrane (TM) domain	57
1.3.4.3 Extracellular loop (ECL) region	59
1.3.4.4 Intracellular loops (ICLs)	60
1.4 Calcitonin family of peptides	62
1.4.1 CGRP	64

1.4.1.1	Tissue distribution and Physiological role	64
1.4.1.2	Structure of CGRP	65
1.5	The CGRP receptor complex	66
1.5.1	CLR	66
1.5.2	RAMP1	69
1.5.2.1	Physiological and pharmacological roles	69
1.5.2.2	Structure of RAMP1	71
1.5.2.3	RAMP1 analogues: RAMP2 and RAMP3	72
1.5.3	RCP	75
1.6	Expression Systems for Recombinant Protein Production	75
1.6.1	The <i>Pichia pastoris</i> Expression system	77
1.7	Mutagenesis	80
1.8	Aims and Objectives	81
<b>Chapter 2</b>	<b>Materials and Methods</b>	<b>82</b>
2.1	Generation and analysis of CLR site-directed mutants	82
2.1.1	Materials	82
2.1.1.1	Equipment	82
2.1.1.2	Media and stock solution	82
2.1.1.2.1	Cell culture media	82
2.1.1.2.2	cAMP assay media	82
2.1.1.2.3	cAMP assay buffer	83
2.1.1.2.4	Binding protein	83
2.1.1.2.5	Activated charcoal	83
2.1.1.3	Reagents	83
2.1.1.3.1	Human $\alpha$ CGRP	83
2.1.1.3.2	Antibodies	84
2.1.1.3.3	Trypsin EDTA	84
2.1.2	Methods	84
2.1.2.1	Identifying targeted sites for mutation	84
2.1.2.2	Expression constructs	84

2.1.2.3	Primer design	85
2.1.2.4	Site-directed mutagenesis	85
2.1.2.5	Mutant DNA sequencing	85
2.1.2.6	Cell culture and transfection	85
2.1.2.7	cAMP standard curve	86
2.1.2.8	Assessment of cAMP production	87
2.1.2.9	Enzyme-linked immunosorbent assay (ELISA)	89
2.1.2.10	CGRP receptor cell surface expression after $\alpha$ CGRP-dependent internalization	90
2.1.2.11	Total CLR expression	90
2.1.2.12	Crude membrane preparation	90
2.2	eCLR and eRAMPs production	91
2.2.1	Materials	91
2.2.1.1	Equipment	91
2.2.1.2	Media and stock solution	91
2.2.1.2.1	Luria-Bertani (LB) broth	91
2.2.1.2.2	BMGY (Buffered glycerol-complex medium)	92
2.2.1.2.3	BMMY (Buffered methanol-complex medium)	92
2.2.1.2.4	YPD (Yeast peptone dextrose)	92
2.2.1.2.5	YPDS Geneticin <sup>®</sup> /Zeocin plates	92
2.2.1.2.6	10x Yeast nitrogen base (YNB)	93
2.2.1.2.7	500x Biotin (0.02%)	93
2.2.1.2.8	10x Glucose (20%)	93
2.2.1.2.9	10x Methanol (5%)	93
2.2.1.2.10	10x Glycerol (10%)	93
2.2.1.2.11	1M Potassium phosphate buffer, pH 6.0	93
2.2.1.2.12	1x Tris-buffered saline with Tween 20 (TBST) pH 7.4	94
2.2.1.3	Reagents	94
2.2.1.3.1	Restriction enzymes	94
2.2.1.3.2	Ligation enzyme	94
2.2.1.3.3	SDS polyacrylamide gels	94

2.2.1.3.4	Antibiotics	95
2.2.1.3.5	Antibodies	96
2.2.1.3.6	DNA and Protein markers	96
2.2.2	Methods	96
2.2.2.1	Molecular biology	96
2.2.2.1.1	CLR and RAMPs gene amplification by PCR	96
2.2.2.1.2	Cloning	98
2.2.2.1.3	DNA sequencing	99
2.2.2.1.4	<i>E. coli</i> transformation	99
2.2.2.1.5	Miniprep and Maxiprep	100
2.2.2.2	Making <i>P. pastoris</i> competent cells	100
2.2.2.3	<i>P. pastoris</i> transformation	101
2.2.2.4	Gene expression	101
2.2.2.4.1	24-well plate (screening) format	101
2.2.2.4.2	Shake flask method	102
2.2.2.5	Protein identification	102
2.2.2.5.1	SDS PAGE	102
2.2.2.5.2	Silver staining	103
2.2.2.5.3	Western blot	103
2.2.2.6	Purification	103
2.2.2.7	Protein quantification	104
2.2.2.8	Protein dialysis	105
2.2.2.9	Gel filtration	106
2.2.2.10	Deglycosylation of proteins	106
2.3	Biophysical characterization	107
2.3.1	Materials	107
2.3.1.1	Equipment	107
2.3.2	Methods	107
2.3.2.1	Circular dichroism (CD)	107
2.3.2.2	Analytical ultracentrifugation (AUC)	108
2.3.2.3	CLR/RAMP interaction by SPR	108

<b>Chapter 3: Identification of residues within the CLR N-terminus required for CGRP receptor signalling</b>	109
3.1. Introduction	109
3.2. Method	111
3.2.1 <i>In silico</i> alanine mutagenesis	112
3.3. Results	111
3.3.1 Stimulation of cAMP production	113
3.3.2 Cell surface expression	119
3.3.3 $\alpha$ CGRP-mediated receptor internalization	121
3.3.4 Inhibition of $^{125}\text{I}$ -hCGRP radioligand binding assay	122
3.3.5 Summary of the N-terminal mutations	123
3.3.6 <i>In silico</i> alanine mutation of residues	124
3.4. Discussion	126
<b>Chapter 4: CGRP receptor pharmacology: multiple residues within Helix 8 and its associated C-terminal region play significant roles in receptor signalling</b>	140
4.1. Introduction	140
4.2. Results	143
4.2.1 Stimulation of cAMP production	143
4.2.2 Cell surface expression	148
4.2.3 $\alpha$ CGRP-mediated receptor internalization	149
4.2.5. Summary of these mutations	150
4.3. Discussion	152
<b>Chapter 5: Production of the N-terminal extracellular domains of CLR and RAMP1, -2 and -3 in <i>Pichia pastoris</i></b>	163
5.1. Introduction	163
5.2. Results	165
5.2.1 Defining the sequence of CLR and RAMP N-terminal ECDs for recombinant expression	165
5.2.2 Construction of CLR/RAMP expression vector	167
5.2.3 Transformation	168

5.2.4	Screening (small scale expression) of transformants	169
5.2.5	Troubleshooting the expression and Western blot analysis procedures	170
5.2.6	Troubleshooting by purification of samples	173
5.2.7	Large scale expression	176
5.2.8	Wild type versus protease-deficient strains in CLR and RAMPs production	180
5.2.9	Deglycosylation of eCLR and eRAMP3 revealed clear protein bands	181
5.2.10	Circular dichroism showed proteins were folded	182
5.3	Discussion	185
<b>Chapter 6: General discussion and Conclusion</b>		190
<b>References</b>		196
<b>Appendices</b>		212

## Abbreviations

A2aR	Adenosine 2a receptor
AM	Adrenomedullin
Amy	Amylin
AUC	Analytical ultracentrifugation
$\beta_1$ -AR	Beta1 adrenergic receptor
$\beta_2$ -AR	Beta2 adrenergic receptor
BCA	Bicinchoninic acid
BIBN4096BS	N-[2-[[5-amino-1-[[4-(4-pyriinyl)-1-piperazinyl]carbonyl]pentyl]amino-1-[3,5-dibromo-4-hydroxyphenyl)methyl]-2-oxoethyl-4-(1,4-dihydro-2-oxo-3(2H)-quinazoliny]
BMGY	Buffered complex glycerol medium
BMMY	Buffered complex methanol medium
Bq	Becquerel
BSA	Bovine serum albumin
cAMP	Cyclic adenosine 3', 5' monophosphate
CaSR	Calcium sensing receptor
cDNA	Complimentary deoxyribonucleic acid
CGRP	Calcitonin gene-related peptide
CLR	Calcitonin receptor like receptor
CNS	Central nervous system
CRF	Corticotropin-releasing factor
CRFR	Corticotropin-releasing factor receptor
CRSP	Calcitonin receptor stimulating peptide
CT	Calcitonin
CTR	Calcitonin receptor
Da	Dalton
dd	double distilled
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid

dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
e	Extracellular domain
<i>E. coli</i>	<i>Escherichia coli</i>
EC <sub>50</sub>	Half maximum effective concentration of agonist
ECD	Extracellular domain
ECL	Extracellular loop
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
E <sub>max</sub>	Maximum response
ER	Endoplasmic reticulum
GABA	Gamma ( $\gamma$ ) aminobutyric acid
GCR1	G-protein coupled receptor 1
GFP	Green fluorescent protein
GHRHR	Growth hormone releasing hormone receptor
GIP	Gastric inhibitory polypeptide
GIPR	Gastric inhibitory polypeptide receptor
GLP	Glucagon-like peptide
GLP1R	Glucagon-like peptide 1 receptor
GLP2R	Glucagon-like peptide 2 receptor
GLR	Glucagon receptor
GPCR	Guanine nucleotide-binding protein coupled receptor
G-protein	Guanine nucleotide-binding protein
Gs	G-protein stimulatory
GTP	Guanosine triphosphate
h	Hour
HA	Haemagglutinin
HEK	Human embryonic kidney cells
IBMX	Isobutylmethylxanthine
ICL	Intracellular loop
J	Juxtamembrane

K <sub>d</sub>	Concentration of ligand required to bind half of the receptors
kDa	Kilo Dalton
L	Litre
LB	Luria-Bertani
M	Molar concentration
mg	Milligram
min	Minute
MK-0974	N-[3R,6S)-6-(2,3-difluorophenyl)-2-oxo-1-(2,2,2-trifluoroethyl)azepan-3-yl]-4-(2-oxo-2,3-dihydro-1H-imidazo[4,5-b]pyridine-1-yl)piperidine-1-carboxamide]
ml	Millilitre
mM	Millimolar
mol	Mole
Ni-NTA	Nickel nitrilotriacetic acid
nM	Nanomolar
NMR	Nuclear magnetic resonance
OD	Optical density
OPD	<i>o</i> -Phenylenediamine dihydrochloride
PACAP	Pituitary adenylyl cyclase-activating protein
PACAPR1	Pituitary adenylyl cyclase-activating protein receptor
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank
pEC <sub>50</sub>	Negative logarithm of EC <sub>50</sub>
PEI	Polyethylenimine
pH	Negative logarithm of the hydrogen ion concentration
<i>P. pastoris</i>	<i>Pichia pastoris</i>
PKA	Protein kinase A
PKC	Protein kinase C
PLC	Phospholipase C
PTH	Parathyroid hormone

PTH1R	Parathyroid hormone receptor 1
PTH2R	Parathyroid hormone receptor 2
PTHrP	Parathyroid hormone-related peptide
RAMP	Receptor activity modifying protein
RCP	Receptor component protein
RMSD	Root mean square deviation
rpm	Revolutions per minute
s	Second
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SCR	Short consensus repeat
SCT	Secretin
SCTR	Secretin receptor
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulphate
SEM	standard error mean
SPR	Surface plasmon resonance
TBST	Tris buffered saline Tween 20
TEMED	N,N,N',N'-tetramethyl- ethane-1,2-diamine
TEV	Tobacco etch virus
TIP39	Tuberoinfundibular peptide of 39 residues
TM	Transmembrane domain
Tris	Tris(hydroxymethyl)aminoethane
Tween 20	Polyoxyethylene sorbitan monolaurate
VPAC	Vasoactive intestinal peptide receptor
v/v	Volume/volume
WT	Wild type
w/v	Weight/volume
YNB	Yeast nitrogen base
YPD	Yeast peptone dextrose
YPDS	Yeast peptone dextrose sorbitol

**Table A1:** Standard amino acid abbreviations used interchangeably throughout this thesis.

Amino acid	3-letter code	1-letter code	Amino acid	3-letter code	1-letter code
Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Methionine	Met	M
Aspartic acid	Asp	D	Phenylalanine	Phe	F
Cysteine	Cys	C	Proline	Pro	P
Glutamic acid	Glu	E	Serine	Ser	S
Glutamine	Gln	Q	Threonine	Thr	T
Glycine	Gly	G	Tryptophan	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Ile	I	Valine	Val	V

## List of Figures

### Chapter 1

Fig 1.1: Diagrammatic representation of membrane proteins	21
Fig 1.2: Receptor states and their stabilization by ligands	26
Fig 1.3: Basic structure of a GPCR	29
Fig 1.4: Structural representation of rhodopsin, secretin and glutamate families of GPCRs	33
Fig 1.5: GPCR signalling cascade	37
Fig 1.6: Schematic representation of $\beta$ -arrestin-dependent GPCR internalization	40
Fig 1.7: Structural alignment of bovine rhodopsin human PAR1 human CXCR4 and rat NTSR1 showing the different structural variations at H8 relative to rhodopsin	49
Fig 1.8: Sequence alignment and structural features of the N-terminal ECD of human family B GPCRs	51
Fig 1.9: Common structural fold of family B GPCRs	52
Fig 1.10: Sequence alignment of human calcitonin family peptides	63
Fig 1.11: Structure of CGRP showing the distribution of structural features across the peptide sequence	66
Fig 1.12: Crystal structure of CLR ECD showing the architecture of the common secretin family fold	68
Fig 1.13: Ribbon representation of RAMP1 structure.	72
Fig 1.14: Sequence alignment of human RAMPs	73
Fig 1.15: Crystal structure of the ECD of RAMP2	74
Fig 1.16: Representative model for the CGRP system	75
Fig 1.17: Methanol metabolism in <i>P. Pastoris</i>	78

### Chapter 2

Fig 2.1: Standard cAMP curve	87
Fig 2.2: Dose-response curve of mock transfected COS7 cells	89
Fig 2.3: Vector map of the <i>P. pastoris</i> vector employed in this study	99
Fig 2.4: A representation of the BSA standard curve. Graph of BSA concentration against absorbance at 570 nm	105

### Chapter 3

Fig 3.1: Representation of residues investigated in this chapter	110
Fig 3.2: Representative dose-response curves of mutant receptors showing no measurable response to 100nM $\alpha$ CGRP	116
Fig 3.3: Representative dose-response curve of mutant receptors that showed significant decrease in $\alpha$ CGRP potency compared to WT	118
Fig 3.4: Representative dose-response curves of the mutant receptors with significant decrease in both $\alpha$ CGRP potency and mean $E_{max}$ compared to WT	119
Fig 3.5: Cell surface expression of wild type and mutant receptors	120
Fig 3.6: Graph of $\alpha$ CGRP-mediated receptor internalization of wild type and mutant receptors	122
Fig 3.7: Crystal structure of CLR ECD (PDB 3N7S) showing some interacting residues	127
Fig 3.8: (a) Sequence alignment of human $\alpha$ CGRP and AMY. (b) A view of the CTR/AMY model showing interacting partners W76 of CTR and N22* of AMY	129
Fig 3.9: CLR ECD structure showing Y91 on loop 2 packed against I41 on $Ca1$	131
Fig 3.10: The CTR/AMY model showing some interacting partners between CTR and AMY	132
Fig 3.11: Helical wheel plot of CGRP and AMY $\alpha$ -helical region spanning positions 8 – 18	133
Fig 3.12: CLR ECD structure showing K103 (green) sandwiched by W75 (red) and W111	136
Fig 3.13: CLR ECD structure revealing the two hydrophobic clusters that are proposed to form the ligand-binding groove	137

### Chapter 4

Fig 4.1: Sequence alignment of helix 8 and associated C-terminal residues in human family B GPCRs	141
Fig 4.2: Schematic representation of CLR showing the residues investigated in this chapter	142
Fig 4.3: Representative dose-response curves of the mutant residues with significant decrease in $\alpha$ CGRP potency as well as decrease in mean $E_{max}$ or increase in mean basal activity compared to WT	145
Fig 4.4: Representative dose-response curve of the mutant residue with significant increase in $\alpha$ CGRP potency as well as increase in mean basal activity compared to WT	146
Fig 4.5: Representative dose-response curve of the mutant residues with (A) significant decrease in mean $E_{max}$ or (B) significant increase in mean basal activity compared to WT	147
Fig 4.6: Cell surface expression of wild type and mutant receptors	148
Fig 4.7: Graph of $\alpha$ CGRP-mediated receptor internalization of wild type and mutant receptors	150
Fig 4.8: Model structure of CLR as postulated by Vohra <i>et al</i> (2013)	153
Fig 4.9: Model structure of the CLR showing the hydrophobic residues V391, L395, W399 and Y402	154

Fig 4.10: Helical wheel plots of CLR and rhodopsin 8 <sup>th</sup> helices	155
Fig 4.11: Model structure of the CLR showing (a) the positional orientation of the neighbour basic residues R396 and R397; and (b) the proximity of R397 to F166 and K167	157
Fig 4.12: Model structure of the active form of CLR showing residues A393 and I394 in H8, and L169 in ICL1. These residues are thought to putatively interact during receptor activation	159
 <b>Chapter 5</b>	
Fig 5.1: Amino acid sequence of the N-terminal extracellular domains of CLR and RAMPs	166
Fig 5.2: (a) Schematic representation of the <i>P. pastoris</i> expression vector for recombinant production of the N-terminal ECD of CLR, RAMP1, RAMP2 and RAMP3	167
Fig 5.3: (a) Agarose gel picture of the receptor genes PCR products. (b) Digestion of the pPIC9K plasmid with BamH1 and Spe1 restriction endonucleases	168
Fig 5.4: Colony PCR to confirm integration following transformation	169
Fig 5.5: Western blot analysis of the 10xHis-tagged CLR, RAMP2 and RAMP3 transformants in expression medium using the manual method	169
Fig 5.6: Western blot analysis of the 10xHis-tagged transformants in expression medium using the CCD camera system	170
Fig 5.7: Western blot analysis of the 10xHis-tagged transformants in expression medium with different percentages of SDS gel	172
Fig 5.8: Western blot analysis of expressed transformants and the GFP	173
Fig 5.9: (a) Western blot analysis of purified 10xHis-tagged RAMP3 protein sample	175
Fig 5.10: Western blot analysis showing deglycosylated and untreated protein samples	176
Fig 5.11: Silver staining for the purified CLR, RAMP3 and RAMP2 samples	177
Fig 5.12: Western blot analysis of the purified CLR, RAMP3 and RAMP2 samples	178
Fig 5.13: Analytical ultracentrifugation analysis of protein samples	179
Fig 5.14: Silver-stained and Western-blotted SDS PAGE pictures of eCLR and eRAMPs produced in the wild type (a) and protease-deficient (b) strains of <i>P. pastoris</i>	180
Fig 5.15: Silver staining (a) and Western blot (b) analysis of glycosylated and deglycosylated (Endo H-treated) eCLR and eRAMPs	181
Fig 5.16: Far UV CD spectra of CLR and RAMPs	183
Fig 5.17: Surface plasmon resonance profiles for the (a) immobilization of CLR and (b) binding of RAMP2 to CLR	184

## List of Tables

Table A1: Standard amino acid abbreviations used interchangeably throughout this thesis	15
<b>Chapter 1</b>	
Table 1.1: Descriptive features of arrestin family members	38
Table 1.2: A summary of the receptors, major physiological roles and sequence sizes of secretin family peptides	54
Table 1.3: Pharmacological profile of receptors for the human calcitonin family of peptides	64
Table 1.4: Comparison of expression systems for recombinant protein production	77
<b>Chapter 2</b>	
Table 2.1: Composition of the polyacrylamide gels used in this project	95
Table 2.2: Table showing the PCR primers	97
<b>Chapter 3</b>	
Table 3.1: Mean pEC <sub>50</sub> of WT and mutant receptors including percentage fold difference between WT and mutants	114
Table 3.2: Mean E <sub>max</sub> and basal activity values of WT and mutant receptors	115
Table 3.3: Summary of the criteria employed in probing the N-terminal mutant receptors	123
Table 3.4: The difference in free energy of formation ( $\Delta\Delta G_f$ ) of CLR structure and energy of interaction between CLR and RAMP1 following <i>in silico</i> alanine mutagenesis	123
<b>Chapter 4</b>	
Table 4.1: Mean pEC <sub>50</sub> , E <sub>max</sub> and basal activity values of WT and mutant receptors	144
Table 4.2: Summary of the criteria employed in probing the mutant receptors	151
<b>Chapter 5</b>	
Table 5.1: Some key features of the ectodomains of CLR and RAMPs	166

## Chapter 1 Introduction

### 1.1 Membrane Proteins

The cell is the structural and functional unit of life. It is made up of biomolecules and organelles with which it performs its biological activities. These cellular constituents are enclosed within a phospholipid bilayer that serves as a barrier between the cell's internal and external environment. This bilayer is called the cell membrane. It regulates the flux of materials in and out of the cell due to its selective permeability. The fatty acyl chains of the phospholipid form a hydrophobic core that is about 25-30 Å in width and the polar head groups, a hydrophilic region each measuring about 10-15 Å (Smith and Veenstra, 2003). The presence of these two (hydrophilic and lipophilic) regions gives these membranes their "amphipathic" characteristic (Childs, 2003). There are proteins located at/associated with cell membranes that give them diverse functional characteristics. These proteins are termed "membrane proteins" (Lodish *et al.*, 2000). Biological membranes show a large amount of homogeneity in structure but differ in functions depending on the proportion of proteins to lipid, type of proteins associated with them and the type of phospholipid they contain (Childs, 2003). Their functions are usually determined by the type of the cell/organelle in which they are located (Lodish *et al.*, 2000). For example, the myelin sheath, which insulates nerve fibres, has a protein to lipid ratio of about 1:4 while the mitochondria inner membrane, on the other hand, has a higher proportion of proteins than lipid (protein:lipid = 3:1) for the purpose of ATP synthesis, importation of newly synthesized proteins and transfer of small molecules (Yeagle and Lee, 2002; Childs, 2003). That of the plasma membrane is 1:1 (Petty, 1993).

Membrane proteins (MPs) fall into two large groups – the *peripheral* and *integral* membrane proteins (Fig. 1.1) - based on their mode of interaction with the lipid bilayer. The peripheral or extrinsic membrane proteins are attached to the surfaces of the membrane, either directly to the polar head groups of the bilayer or to integral membrane protein(s) by electrostatic, hydrophobic or other non-covalent interaction. They can therefore be dissociated from the membrane upon treatment with relatively high ionic-strength buffers like that of carbonate at high pH (Smith and Veenstra, 2003;

Nelson and Cox, 2005). These water-soluble proteins play a role in the regulation of membrane-bound enzymes and in limiting the mobility of integral membrane proteins by leashing them to intracellular structures (Nelson and Cox, 2005). Examples include the cytoskeletal proteins – microfilament, actins and tubulins – and cytochrome c (Petty, 1993). The integral or intrinsic membrane proteins (IMPs), on the other hand, have a significant proportion of their mass buried within the membrane (Lodish *et al.*, 2000; Smith and Veenstra, 2003). These proteins' hydrophobic and hydrophilic regions interact more firmly with the fatty acyl hydrophobic core and the polar head groups of the membrane respectively, and can therefore only be solubilised by amphipathic agents such as detergents e.g. Triton X-100 (Smith and Veenstra, 2003; Petty, 1993). Most IMPs span the entire membrane with a portion of their structure at both phases of the membrane allowing communication between the exterior and interior of the cell/organelle and are in this case also referred to as *transmembrane proteins* e.g. glycophorin and insulin receptor (Lodish *et al.*, 2000; Smith and Veenstra, 2003).



**Fig 1.1:** Diagrammatic representation of membrane proteins. The different types of membrane proteins as well as the components of the phospholipid bilayer are shown as labelled. *Mol Cell Bio.* 2000. *WH Freeman & Co*

The membrane-spanning segments of IMPs are dominated by lipophilic amino acid residues like Leu, Ile, and Val as well as aromatic residues like Trp, Phe and Tyr. Their lipophilicity makes them highly compatible with the hydrophobic core of the membrane lipid bilayer (Muller *et al.*, 2008). So far,

only two major secondary structures – the *alpha* ( $\alpha$ ) *helix* and the *beta* ( $\beta$ ) *barrel* – have been reported for these proteins (Sternberg, 1996; Lodish *et al.*, 2000; Childs, 2003; Pifat-Mrzljak, 2007). The  $\alpha$  helix is formed from the folding of transmembrane segments made up of long stretches of non-polar side chains of the amino acids. These side chains form van der Waals interaction with the membrane lipid hydrophobic core thereby shielding the carbonyl and the imino (NH) groups. A group of positively charged Lys and Arg believed to interact with the negatively charged phospholipid head groups prevents the helix from slipping across the hydrophobic lipid bilayer (Sternberg, 1996; Lodish *et al.*, 2000). Most IMPs are composed of one or more transmembrane  $\alpha$  helix with each protein having a specific number of transmembrane  $\alpha$  helices. For example, receptor activity modifying proteins (RAMPs), glycoporphin, bacteriorhodopsin and the bacterial photosynthetic reaction centre (PRC) have 1, 2, 7 and 11 membrane-spanning  $\alpha$  helices respectively (McLathie *et al.*, 1998; Sternberg, 1996; Lodish *et al.*, 2000). Up to 19 membrane-spanning  $\alpha$  helices have also been reported for the voltage-dependent  $\text{Ca}^{2+}$  channel (Muller *et al.*, 2008).  $\beta$  barrels on the other hand are large antiparallel sheets that form cylindrical structures in which the strands have uncharged amino acids that protrude out of the outer surface of the barrel at every second position in the sequence (Sternberg, 1996). A prominent example is the class of transmembrane proteins called “porins” that have primarily different structure from other IMPs. A number of these proteins are found in the outer membrane of gram-negative bacteria like *Echerichia coli* [*E. coli*] (Lodish *et al.*, 2000). Despite the successes recorded in the past in the study of membrane proteins, a large part of the knowledge of their structures, and how this influences processes such as membrane insertion and folding mechanisms, is still considered rudimentary (Muller at al., 2008). Although models have been proposed to unravel these mechanisms (White *et al.*, 2001; Engelman, 2003), they have their limiting factors and cannot make a clear representation of an *in-vivo* mechanism. Further studies are still therefore required to better understand many structural features of these proteins.

The function of membrane proteins in a living organism cannot be overemphasized. Membrane proteins help in stabilising the structure of the cell by maintaining polarity, shape and size of the cell. A group of them function in energy transduction using electrochemical gradient to generate high

energy compounds like ATP. They also serve as transporters building and/or maintaining concentration gradients of biomolecules such as electrolytes, water and metabolic cofactors. Membrane proteins such as kinases and proteases also perform some enzyme activities. Molecule recognition in the immune system and cell adhesion in the formation of tight junctions both involve certain groups of membrane proteins (Muller *et al.*, 2008). This is just to mention a few of the countless roles played by these versatile biological molecules. It is no surprise that about 30% of all active genes encode membrane proteins (von Heijne, 2007). Of all the functions attributed to membrane proteins (especially the integral membrane proteins), a very prominent one is their role as *receptors* for extracellular ligands that bind to their extracellular or transmembrane domain (Muller *et al.*, 2008).

### **1.1.1 The “Receptor Concept”**

Contemporary scientists share the notion that hormones and drugs produce their effects by specific interactions with “receptors” in a manner similar to the interaction between enzymes and substrates. The insights came from early works of scientists who explored a number of investigations to make their points evident (Limbird, 1996). Claude Bernard (1813 – 1878) with his investigation of the selectivity and specificity of drug action pioneered the idea of specific drug interaction (Limbird, 1996). Although Bernard never used the word “receptor”, he was particular about the locus of drug action and by so doing brought to light a first method to establish the specificity of drug action. In learning the mode of action of the arrow poison curare, he discovered they were effective when administered by an arrow but harmless while administered orally. He noticed that the orally administered curare was unaltered by saliva, bile, pancreatic or gastric juice but was unabsorbed by the gastrointestinal tract (GIT), and could not reach the neuromuscular junction – its site of action (Limbird, 1996; Scheidlin, 2001). His findings suggested that the poison requires access to its target organ system to elicit its effect. This was suggested as a prerequisite for drug action.

In 1878, the English physiologist, John N. Langley first proposed the existence of physiological substances that form complexes with the pilocarpine and atropine. It was later in 1905 that he used the word “receptive substance” to describe the effects of nicotine and curare on skeletal muscle (Hollinger, 2003; Rang, 2006). A.V. Hill later used the law of mass action – derived in 1918 by Langmuir, an eminent physical chemist, to describe adsorption of gases - to study this interaction quantitatively and opined that it obeys the law. The equation describing this interaction was then tagged the Hill-Langmuir equation and this formed the basis of the “receptor theory” (Limbird, 1996; Rang, 2006). The formulation of this theory became the first significant achievement of the pharmacology discipline in terms of drug-receptor interaction. Paul Ehrlich, a German contemporary of Langley, baptized the receptive substance described by Langley as “receptor” from his interest in immunology and chemotherapy while working with bacterial toxins (Hollinger, 2003; Seth and Seth, 2009). Ehrlich’s idea was that bacterial toxins interact with nutrient-capturing structures of cells (sidechains) thereby starving the cells. He also found out that these bacterial sidechains are distinct from that of the host cells. This led him to the discovery of Salvarsan, the first potent treatment for syphilis (Rang, 2006).

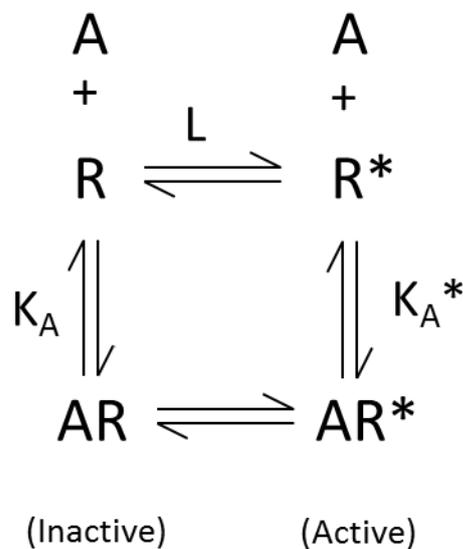
After this period, the receptor concept was now in the limelight and formed the foundation for many other scientists to build on. Notable was the work of A.J. Clark and J.H. Gaddum who worked extensively on the actions of acetylcholine and atropine on frog’s isolated heart as well as that of adrenaline and ergotamine on rabbit uterus. They derived the equations to describe the log concentration-effect curve, which is now a cornerstone of pharmacology (Rang, 2006). More work was done to seal the root of the receptor concept into the ground. The first was the study of competitive antagonism. It was found that while some drugs act as agonists, others act as antagonists. Another was direct measurement of drug binding, which made possible the techniques of receptor isolation and cloning. Most of the initial works were performed on the  $\beta$ -adrenergic receptor (Rang, 2006; Seth and Seth, 2009). Though Clark and Gaddum failed to mention drug antagonism from competition between agonist and antagonist for the same receptor in their earlier work, Gaddum later (in 1937) derived an equation to describe the binding of two drugs to the same receptor. Schild (1947)

developed this idea and came up with the “dose ratio” – “the factor by which the agonist concentration must be increased in order to produce the same level of equilibrium occupancy as the concentration of antagonist is increased” (Rang 2006).

Further modifications to Clark and Gaddum’s works by Ariens (1954), Stephenson (1956) and Furchgott (1966), added some other interesting parameters to the drug-receptor relation. At this time, the idea of drug “affinity” and “efficacy” or “intrinsic activity” sprang up (Rang, 2006; Seth and Seth, 2009). While affinity was defined as the ability to combine with receptors, efficacy represented the ability to produce response. These two terms were then used to theoretically distinguish between agonists and antagonists. While both agonists and antagonists are capable of showing affinity (i.e. interactions) for a receptor, only an agonist could possess efficacy or intrinsic activity (i.e. produce desired response) following the interaction (Seth and Seth, 2009). The antagonist in this respect prevents a response by blocking the interaction of the agonist with the receptor. Receptor ligand interaction is likened to the lock and key model postulated for enzymes and substrates, which is based on biochemical structure. Antagonists are therefore usually, but not always, similar in structure to the agonists and are bound by the receptor on this basis making the receptor “unavailable” for its ligand (agonist). The key difference here is that while the agonist binds and activates the receptor, the antagonist only binds but does not activate the receptor. Moreover, agonists and antagonists do not necessarily bind to the receptor at the same site, but their sites of binding to the receptor may overlap. There are drugs that specifically act as antagonists. An example is the opiate antagonist, naloxone, which blocks the binding of morphine (the agonist) to the opiate receptor (Seth and Seth, 2009). Agonists could be full/pure, partial or inverse. The full agonist binds to the receptor producing the desired effect. Most agonists fall into this category. Partial agonists share the characteristics of pure agonist and antagonist. They have high affinity for the receptor but typically produce a weaker response, compared to the full agonist, and in the process prevent the interaction of the full agonist with the receptor. An example is the  $\beta$ -adrenergic antagonist, oxprenolol. The inverse agonist produces an exactly opposite effect to that of the pure agonist. It reduces the activity of the receptor by inhibiting the receptor’s constitutive or basal activity. An example of an inverse agonist is  $\beta$ -

carbolines which produce an opposite effect to that of benzodiazepine in the CNS (Seth and Seth, 2009).

The mechanism by which receptors alternate between the inactive and active states in the presence or absence of a ligand has been described using the two-state model schematically similar to that postulated for the interaction between oxygen and haemoglobin by Monod, Wyman and Changeux (see Leff, 1995). The two-state model is presented in Fig 1.2 below.

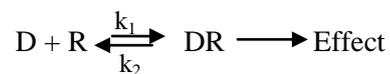


**Fig 1.2:** The two-state model for receptor activation. (*Adapted from Leff (1995)*). R and R\* represent the receptor in the inactive (or resting) and active states respectively. L is the equilibrium constant for the interconversion between the two receptor states in the absence of a ligand. A is the ligand and has affinity for both receptor states. As an agonist, it preferentially binds to R\*.  $K_A$  and  $K_A^*$  are dissociation equilibrium constants for A at the inactive and active receptor states respectively.

When A shows a higher affinity for the receptor in the inactive state (i.e. R), it is an inverse agonist. Whereas, when A displays a higher affinity for the receptor in its active state, it is an agonist. These imply that while an inverse agonist stabilizes the receptor in its inactive conformation, an agonist stabilizes the receptor's active conformation. A neutral antagonist stabilizes both receptor states (i.e. active and inactive) without altering the equilibrium.

In some cases, receptors can become active in the absence of an agonist in which case they are said to display “constitutive activity”. This is commonly observed as a result of mutation to residue(s) that helps the receptor maintain its inactive state. This mutation could be naturally occurring or experimentally induced. For instance, mutating A293 in hamster  $\alpha_{1b}$ -adrenoceptor, to any other amino acid, results in constitutive activity (Kjelsberg *et al.*, 1992). This constitutive activity is usually blocked by an inverse agonist, and by so doing, it (inverse agonist) is said to display a negative efficacy (Lambert *et al.*, 2004).

Drugs could interact with receptors reversibly by forming weak bonds (such as hydrogen, van der Waals or electrostatic) with the receptor. They could also form strong (e.g. covalent) bonds in which case they interact irreversibly, although this is rarely the case. The reversible drug-receptor complexes obey the law of mass action. The degree of response of a drug depends on the number of occupied receptors (Seth and Seth, 2009). Considering the following equation for instance,



where D is the drug, R the free receptor, and  $k_1$  and  $k_2$  are the association and dissociation rate constants respectively. The equilibrium constant  $K = k_1/k_2$ . It is a feature of both drug and receptor and it is the concentration of the drug required to occupy 50% of the receptor’s binding site at equilibrium (Seth and Seth, 2009).

Finally, receptors can be a single unit of homomeric molecules or a combination of heteromeric subunits with each subunit having equal or varying contribution towards the receptor specificity. Despite the diversity and complexity in the structure and function of these receptors, Triggle (2000) classified them into four principal families of chemically sensitive pharmacological receptors (Triggle, 2000). These are:

1. Ion channels: integral MP comprising subunits, each with membrane inserts forming different pores of the channel.
2. Enzyme-associated: one-transmembrane protein with kinase activity

3. Nuclear receptors: non-membrane intracellular proteins that regulate transcription.
4. G protein-coupled: seven-transmembrane receptor proteins that couple to proteins of the G protein family.

## 1.2 G protein-Coupled Receptors

G-protein coupled receptors (GPCRs) are seven-transmembrane (7-TM)  $\alpha$ -helical receptor proteins (Fig 1.3) that transduce signals via heterotrimeric guanine nucleotide-binding proteins (G proteins) coupled on their cytoplasmic surface (Neves *et al.*, 2002; Park *et al.*, 2009). The heterotrimeric G protein comprises the  $\alpha$ ,  $\beta$  and  $\gamma$  subunits ( $G\alpha\beta\gamma$ ) and, at resting state, are normally bound to guanosine diphosphate (GDP) at the  $\alpha$  subunit. The overall signal transduction involves the exchange of bound GDP for GTP on the  $\alpha$  subunit. When activated, for instance following the binding of an extracellular stimulus, GPCRs undergo conformational changes that allow G protein binding and subsequent release of GDP from  $G\alpha$ . On association of GTP with  $G\alpha$  subunit, conformational changes within the  $G\alpha$  subunit allow the release of bound trimeric G protein from the receptor and the dissociation of this protein into the active  $G\alpha$  (bound to GTP) and  $G\beta\gamma$ . These subunits then interact with downstream effector proteins, for example adenylyl cyclase, which produces the second messenger, cyclic AMP (Kristiansen, 2004). The cycle is terminated by the hydrolysis of GTP to GDP by  $G\alpha$  subunit which paves way for the reassociation of the  $\alpha$  and  $\beta\gamma$  subunits (Oldham and Hamm, 2008).

GPCRs constitute a remarkably large and ubiquitous family of membrane receptors through which the majority of extracellular signals including photons, neurotransmitters, ions, nucleotides, peptides and lipids mediate intracellular responses (Lundstrom and Chiu, 2006; Kroeze and Roth, 2006). They comprise a large proportion (~3.7%) of the human genome, making them the biggest single family in the genome, and their genes show over 60% sequence identity to one another (Offermanns and Rosenthal, 2008). They are found in a wide range of organisms including the invertebrates e.g. nematodes, vertebrates e.g. fish and mammals (Kroeze and Roth, 2006). For these reasons, especially

in their remarkable role in signal transduction coupled with their easy accessibility by hydrophilic hormones and drugs (since they are expressed on the plasma membrane), GPCRs are now therapeutic targets for up to 50% of clinically marketed drugs (Fredriksson *et al.*, 2005; Lundstrom and Chiu, 2006; Insel *et al.*, 2007; Daulat *et al.*, 2008; Park *et al.*, 2009). GPCRs have been reported to be the therapeutic target for almost all known antipsychotic drugs (Kroeze and Roth, 2006). In 2003, drugs targeting these receptor proteins made about 23% of total sales of prescription drugs (Lundstrom and Chiu, 2006). These drugs include Salmeterol, “an antiasthmatic  $\beta_2$  adrenergic agonist”, Losartan, “an antihypertensive angiotensin II receptor antagonist”, and the prostate cancer drug, Leuprolide, targeted at the gonadotropin-releasing hormone (GnRH) receptor (Kroeze and Roth, 2006). The physiological importance of native GPCRs accounts for their pharmacological relevance.



**Fig 1.3:** Basic structure of a GPCR. The TM helices are numbered 1 – 7 and the intracellular 8<sup>th</sup> helix is labelled 8. The extracellular loops (ECL) and intracellular loops (ICL) are indicated and numbered accordingly. *Pharmacol Ther.* 2004; 103(1):21-80

### 1.2.1 The GPCR Superfamily

There are about 850 members of the GPCR superfamily, which are mostly found within the first three classes of this superfamily of receptor proteins. They share a common backbone structure of an amino terminal region (or N-terminus), a 7-TM helical domain, and a carboxyl terminal region (or C-

terminus). The main structural distinction among these receptors is in their N-termini, which form the orthosteric site (primary binding sites for endogenous agonists) for many receptors (Kristiansen, 2004; Park *et al.*, 2009). The most widely accepted classification system for GPCRs is that introduced by Fredriksson *et al.* (2003), which is called the GRAFS classification system. According to this system, GPCRs are classified, mainly based on sequence homology, into 5 classes namely; glutamate, rhodopsin, adhesion, frizzled/tatse2, and secretin (Fredriksson *et al.*, 2003). A sum-up of these classes, with their alternative family names with which they are sometimes described, are presented below.

#### **1.2.1.1 Rhodopsin-like (family A) GPCRs**

The rhodopsin family (or family A) of GPCRs is so called because members of this family show structural and functional homology with rhodopsin, an important light-sensing receptor protein that plays a crucial role in vision (Fredriksson *et al.*, 2003; Schioth and Fredriksson, 2005; Kobilka, 2007). It was the first characterized representative member of the family. This rhodopsin- or  $\beta_2$  adrenergic receptor-like family constitutes the largest member of the GPCR family making close to 90% (701 of 800) of all GPCRs (Fredriksson *et al.*, 2003; Lundstrom and Chiu, 2006; Kobilka, 2007). They serve as receptors for the widest range exogenous and endogenous ligands including odorants, biogenic amines (adrenergic, muscarinic, dopaminergic etc), neuropeptides and peptidergic hormones, nucleotides, lipids and lipids derivatives, and photons (for rhodopsin) among others (Armbruster and Roth, 2004; Lundstrom and Chiu, 2006). They have a characteristic relatively short N-terminus, albeit with a few exceptions like the luteinizing hormone (LH) receptor and follicle-stimulating hormone (FSH) receptor. They also possess conserved disulphide bridge (between Cys110 and Cys187), a “E/DRY” motif (Glu/Asp-Arg-Tyr) on TM3 on the cytosolic phase and an “NPXXY” motif (Asn-Pro-X-X-Tyr, where X could be any amino acid) on TM7 within the lipid bilayer (Lundstrom and Chiu, 2006; Krauss, 2008; Flanagan, 2005; He *et al.*, 2001, Insel *et al.*, 2007). While the DRY motif is thought to be involved in the maintenance of a stable active/inactive conformation by the receptor, the

NPXXY motif connects the 7<sup>th</sup> TM to an 8<sup>th</sup> cytosolic helix and both roles are believed to be important in signal transduction (Flanagan, 2005; He *et al.*, 2001).

### **1.2.1.2 Secretin-like (family B) GPCRs**

The secretin/glucagon-like family of GPCRs are a much smaller group (15 members) compared to the family A (Fredriksson *et al.*, 2003; Schiöth and Fredriksson, 2005), but with diverse structure and function (Harmar, 2001). The secretin receptor was the first family B GPCR to be cloned (Ishihara *et al.*, 1991) and this probably explains why this sub-family was named the “secretin family” GPCRs. They usually bind peptide hormones of higher molecular weight such as secretin, calcitonin, glucagon and parathyroid hormones. These peptides range in length from about 27 to 52 amino acids (Insel, 2007; May *et al.*, 2007). Receptors of this family have a large N-terminal domain that is bigger than that of the rhodopsin family (see Fig 1.4), and many disulphide bonds (Krauss, 2008). Orthosteric binding to these receptors involves the N-terminus and the TM domain as well as the loops making it a complex mechanism (Hoare, 2005). A number of these members form heterodimers with receptor-activity-modifying proteins (RAMPs) - a family of single-span TM protein that modifies the receptors' ligand binding and signalling characteristics (McLathie *et al.*, 1998; Krauss, 2008). This is specifically seen in the calcitonin receptor-like receptor (CLR) from which three receptor phenotypes are derived following heterodimerization with a RAMP e.g. the CGRP is an heterodimer of the CLR and RAMP1 (McLathie *et al.*, 1998). The complexity in the orthosteric binding of these receptors coupled with the rather allosteric interaction of the RAMPs, where this exists, has made it more difficult to understand the pharmacology of these receptors.

### **1.2.1.3 Glutamate (family C) GPCRs**

The glutamate receptor family (or family C) is another small family of GPCRs with 15 major group members (Fredriksson *et al.*, 2003). They are receptors mainly to low molecular weight

neurotransmitters (e.g. glutamate and Gamma ( $\gamma$ ) aminobutyric acid (GABA)) (Insel *et al.*, 2007). Prominent in this family include metabotropic glutamate and GABA, calcium and vomeronasal (Krauss, 2008). They possess a very large (300-600 amino acids) bilobate N-terminal domain, commonly referred to as “Venus flytrap”, with which they bind their primary ligands. They have been reported to show homo- and heterodimerization (Devi, 2005; May, 2007; Krauss, 2008).

(a)



(b)



(c)



**Fig 1.4:** Structural representation of rhodopsin, secretin and glutamate families of GPCRs. (a) rhodopsin family (b) secretin family (c) glutamate family. *Endocrine Reviews* 21(1): 90-113

#### 1.2.1.4 Frizzled/Taste2 GPCRs

The Frizzled/Taste 2 receptors are a newly organized family comprising the frizzled and taste 2 subfamilies, which were formerly separate. The frizzled was first observed in *Drosophila* and was

found to be activated by the glycoprotein, Wnt, which was “curled” and “twisted” (hence the name frizzled) (Fredriksson *et al.*, 2003; Schioth and Fredriksson, 2005; Schulte and Bryja, 2007). The frizzled receptor group regulates proliferation, embryogenesis and tissue homeostasis. Their N-terminus is about 200 residues long (Schioth and Fredriksson, 2005). The taste2 (TAS2) receptors on the other hand function as bitter taste receptors as they have been found to be expressed in the tongue and palate (Schioth and Fredriksson, 2005). While the frizzled receptors are found in a wide range of organisms, from *Caenorhabditis elegans* to mammals, the taste receptors have only been found in mouse and human and this is a major reason why there are still controversies as regards their evolution (Schioth and Fredriksson, 2005). However, some few important common features, which have been observed for both groups from test sequences in humans, include the IFL, SFL and SxKTL motifs in the 2<sup>nd</sup>, 5<sup>th</sup> and 7<sup>th</sup> TMs respectively (Rognan *et al.*, 2006; Schioth and Fredriksson, 2005).

#### **1.2.1.5 Adhesion GPCRs**

This class is commonly known as the “Adhesion” family and is the second largest class of all known GPCRs, although their evolution is not clearly understood (Fredriksson *et al.*, 2003; Rognan *et al.*, 2006). Before now, this family has been given different names owing to one characteristic or the other that they possess. These names include EGF-TM7 (because they have the epidermal growth factor, EGF, domains in their N-termini); LN-TM7 (due to their long N-termini dominated by the amino acids Serine and Threonine); and B2 (owing to their similarity with the secretin receptor) (Schioth and Fredriksson, 2005; Rognan *et al.*, 2006). The Ser and Thr residues form O- and N-glycosylation sites giving the N-terminal domain a highly glycosylated, mucin-like characteristic. These residues contribute to the receptors’ specificity (Rognan *et al.*, 2006). These receptors are peculiarly known to associate with other membrane-bound proteins like lectin, globulin and cadherin, which are thought to also contribute to the receptors’ specificity (Armbruster and Roth, 2004; Schioth and Fredriksson, 2005; Rognan *et al.*, 2006). Adhesion receptors are found in virtually all organ

system of the body including the central nervous, immune and reproductive systems among others and are therefore believed to have a diverse physiological function (Rognan *et al.*, 2006; Schiöth and Fredriksson, 2005).

It is important to add that a good percentage of these GPCRs are yet to be characterized for their physiological role and are, by so doing, referred to as “orphan GPCRs” (Kobilka, 2007; Insel *et al.*, 2007; Lundstrom and Chiu, 2006; Armbruster and Roth, 2004). The majority of them are potential drug targets and are subject of ongoing research.

### **1.2.2 Activation and signalling cascade**

The activation of GPCRs is another broad subject area due to the diversity in structure and function of the various classes of this receptor. However, it has been reported that the conformational changes that accompany their activation (as observed in rhodopsin – a well-studied representative member of this superfamily of receptors) could be ascribed to other rhodopsin-like GPCRs, albeit with some degree of variations (Kobilka, 2007; Gether *et al.*, 2002; McDonough, 2003; Seifert and Wieland, 2005). The orthosteric site location of a GPCR determines a great deal in its mode of activation. Based on the location of this site, GPCRs can be broadly categorised into two. The first group have their ligand-binding pocket formed mainly within the 7-TM domain (most common with members of family A) while in some others, this is shared by both the N-terminal extracellular domain and the TM bundle (the two-domain model proposed for family B GPCRs). In the latter, the orthosteric pocket formation is favoured by the presence of Cys residues that form disulfide bonds thereby enhancing proper folding of this region. Many of these receptors’ extracellular domains have been found to bind ligands in the absence of the TM domain, based on evidence from crystallographic studies (Park *et al.*, 2007; Koth *et al.* 2010). This is specifically the case for family B GPCRs.

The successful crystallization of the full structure of rhodopsin (Palczewski *et al.*, 2000), as well as the more recent bovine opsin (ligand-free rhodopsin) (Park *et al.*, 2008) and the constitutively active rhodopsin in complex with the C-terminus of transducin (a G protein)  $\alpha$ -subunit (Standfuss *et al.*,

2011), has been a plus in the in-depth molecular study of its activation. This is usually used as a prototype of GPCR activation especially for those that have their orthosteric binding site within the TM domain (Park *et al.*, 2009); although this has now been overtaken by yet another milestone in GPCR structural studies with the release of an active  $\beta_2$ -AR-Gs complex structure by Brian Kobilka and co-workers (2011). These crystal structures are further discussed in section 1.2.4 below. The basis of rhodopsin activation is the “rotational” and “tilt” movement observed for the TM6 relative to TM3 (Kobilka, 2007; Huber and Sakmar, 2009). At the inactive state, TM3 associates with TM7 and TM6 on the extracellular and cytosolic faces of the membrane respectively and this interaction maintains the inactive conformation of the receptor (Huber and Sakmar, 2009). On binding the ligand, 11-*cis* retinal (an inverse agonist) that is rapidly converted to *all-trans* retinal (a full agonist) during light absorption, there is a conformational change in the ligand-binding pocket on TM7, which weakens the association between TM3 and 7, as well as TM3 and 6 (Huber and Sakmar, 2009). It is this change in conformation between TM3 & 7 that gives rise to the movement of TM6 generating a G-protein binding site that binds the G-protein ( $G_{\alpha\beta\gamma}$ ), which then dissociates (on binding to GTP) into its component  $G_{\alpha}$ -GTP and  $G_{\beta\gamma}$  subunits (Huber *et al.*, 2008; Jang *et al.*, 2001, Seifert and Wieland, 2005). The  $G_{\alpha}$ -GTP carries out onward signalling usually the activation of the enzyme adenylate cyclase that converts ATP (adenosine triphosphate) to cAMP (cyclic adenosine monophosphate). cAMP activates another protein kinase, protein kinase A, to produce the required effect (Pochet and Donato, 2000; Ali and Haribabu, 2006).



**Fig 1.5:** GPCR signalling. G = G  $\gamma$ -protein (subtypes s & q); PLC = Phospholipase C; PIP2 = Phosphatidylinositol 4, 5-bisphosphate; IP3 = Inositol 1, 4, 5-triphosphate; Inositol 1, 4-bisphosphate; IP1 = Inositol monophosphate; E.R. = Endoplasmic reticulum; PDE = Phosphodiesterase. *Genetic Engineering & Biotechnology News*. 2009; 29 (4)

### 1.2.3 GPCR endocytosis

Endocytosis of GPCRs involves the process of channelling GPCRs to distinct endocytic sites at the plasma membrane from where they are internalized into intracellular compartments to be degraded or recycled back to the plasma membrane (Moore *et al.*, 2007). This pathway is usually initiated by the phosphorylation of an active (i.e. ligand-bound) GPCR by a G protein-coupled receptor kinase (GRK) usually at the third intracellular loop and C-terminal region (Ferguson, 2001; van Koppen and Jakobs, 2004). GRKs are a 7-membered (GRK1 - 7) family of protein kinases with significant sequence homology (Stoffel *et al.*, 1997) that specifically phosphorylate serine/threonine residues at the amino acid side chain hydroxyl group (Moore *et al.*, 2007). Another group of protein kinases, called the second messenger-dependent protein kinases e.g. cAMP-dependent protein kinase (PKA) and protein kinase C (PKC) also perform the phosphorylation task and are capable of phosphorylating both the agonist- and non-agonist-bound receptors (Hausdorff *et al.*, 1989). These cAMP-dependent kinases are thought to be the main phosphorylating machinery in the presence of low agonist concentration (Ferguson, 2001).

The phosphorylation of GPCRs initiates the binding of cytosolic cofactor proteins called arrestins (Benovic *et al.*, 1987). Arrestins comprises 4 family members, and are further divided into two sub-families (see Table 1.1 for details). The most common of these are the  $\beta$ -arrestins found across various systems in mammals. The binding of  $\beta$ -arrestin to a phosphorylated GPCR is believed to be driven by the disruption of a polar core within  $\beta$ -arrestin by the negatively charged phosphate groups of the phosphorylated GPCRs thereby converting  $\beta$ -arrestin to a high-affinity receptor-binding state (van Koppen and Jakobs, 2004). The binding of  $\beta$ -arrestin normally results in the uncoupling of bound G-protein and/or prevention of G-protein coupling due to steric hindrance (Ferguson, 2001; van Koppen and Jakobs, 2004). The process of  $\beta$ -arrestin binding and subsequent uncoupling of G-protein is normally referred to as receptor desensitization. This process is one of the regulatory mechanisms deployed by the cell to disrupt GPCR signalling.

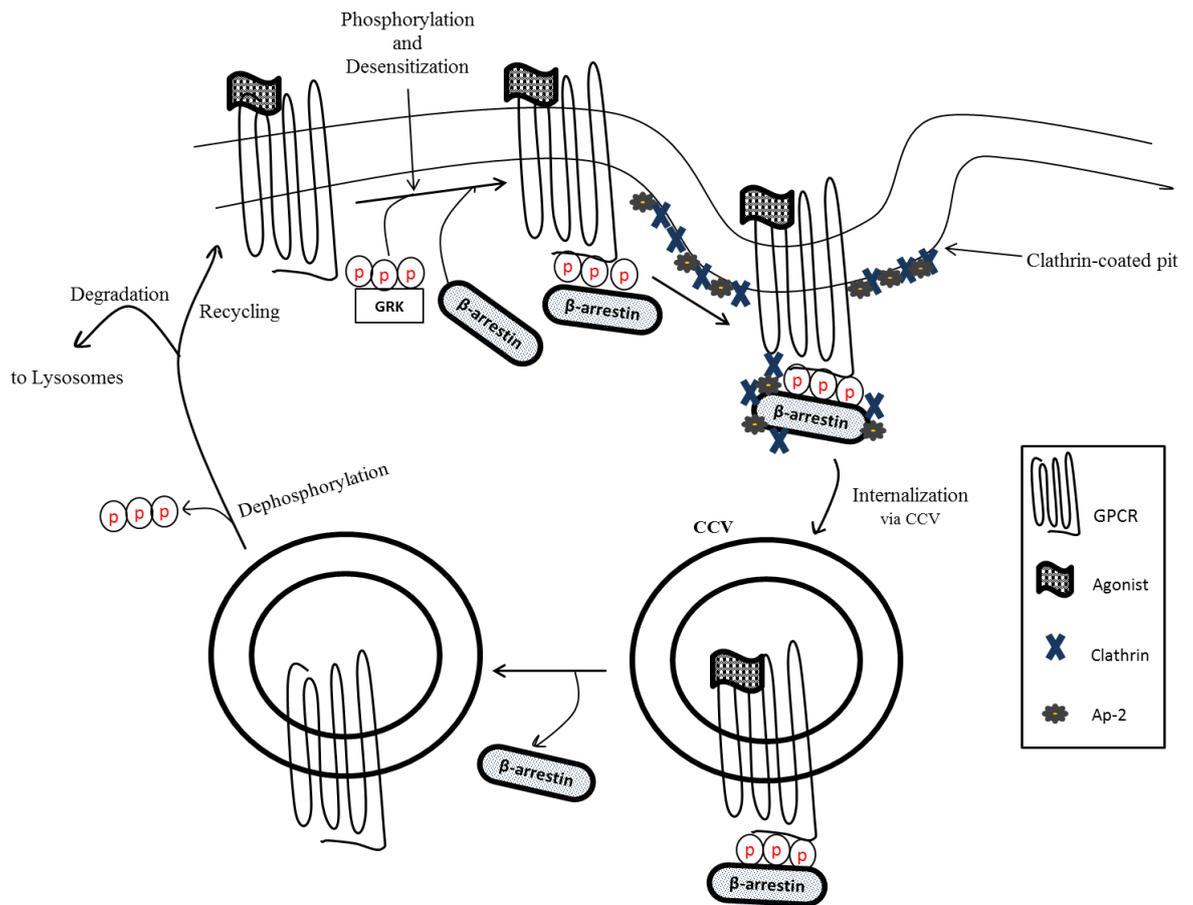
**Table 1.1:** Descriptive features of arrestin family members. *Modified from Ferguson (2001).*

Name	Other names	No of amino acids	Subfamily	Tissue distribution	Target receptors
Arrestin 1	S antigen, visual arrestin	404	Visual	Mostly light-sensing tissues e.g. retinal, pinealocytes, rod inner and outer segment. Also in cerebellum and primary blood leukocytes	Rho > $\beta_2$ AR > m2 mAChR
Arrestin 2	$\beta$ -arrestin 1	418	Non-visual	Ubiquitous. Present in most tissues except those mentioned above	$\beta_2$ AR > m2 mAChR >>Rho
Arrestin 3	$\beta$ -arrestin	410	Non-visual	Same as Arrestin 2, but abundance varies with tissue	$\beta_2$ AR, m2 mAChR >>Rho
Arrestin 4	Cone arrestin, C- or X-arrestin	388	Visual	Mainly pinealocytes, pituitary and lung	ND

ND = not determined

Upon  $\beta$ -arrestin binding, the interaction of  $\beta$ -arrestin with certain endocytic proteins, notably clathrin and clathrin adaptor (AP2) is triggered and this targets the receptor to clathrin-coated pits (CCP),

which then form clathrin-coated vesicles (CCV) that translocate the receptor, in the internalization process, to intracellular compartments (Ferguson, 2001; Moore *et al.*, 2007). The first site of deposition of the internalized receptors is the mildly acidic 'early endosome' located at the periphery of the cell and receives most vesicles from the cell membrane (Mellman, 1996). At the early endosome, internalized receptors are sorted for either degradation or recycling. Receptors for recycling can be translocated to the cell membrane at this stage while the ones marked for degradation are transported to the lysosomes (Ferguson, 2001). Enroute to the lysosome, the receptors are first delivered to the late endosome where some final sorting is carried out (Stoorvogel *et al.*, 1991). In the lysosomes, a large vacuole-resembling compartment and the last compartment of the endocytosis pathway, breaking down of unwanted receptors and other macromolecules delivered for the endosome is performed (Gruenberg and Maxfield, 1995).



**Fig 1.6:** Schematic representation of  $\beta$ -arrestin-dependent GPCR internalization. Prolonged receptor stimulation following agonist binding induces receptor phosphorylation by GRK (G protein-coupled receptor kinase) and/or other kinases (not shown). Cytosolic  $\beta$ -arrestin then binds to phosphorylated receptor and also interacts with clathrin and clathrin adaptor complex (AP-2) as well as other cytosolic components (not shown). This results in the immobilization of the GPCR in clathrin-coated pit, which forms clathrin-coated vesicle (CCV) translocating the receptor into the cytosol for processing.  $\beta$ -arrestin is then released and the receptor dephosphorylated to be recycled or degraded in the lysosome.

It is important to note that there is evidence suggesting that not all GPCRs necessarily undergo internalization in a  $\beta$ -arrestin- and clathrin-dependent manner (Ferguson, 2001). For instance, experiments assessing the extent of internalization in COS7 and HEK 293 cells and that involving dominant-negative mutants of  $\beta$ -arrestin and dynamin have suggested that the angiotensin II type 1A receptor ( $AT_{1A}R$ ) and the m2 muscarinic acetylcholine receptor (mAChR) internalize via a mechanism independent of  $\beta$ -arrestin and clathrin (see Zhang *et al.*, 1996). Aside the clathrin dependent pathway,

the other most common pathway is that involving the cholesterol-binding protein, caveolin (Parton and Simons, 2007).

#### **1.2.4 Advances in GPCRs crystal structures: the bovine rhodopsin era and beyond**

The importance of GPCRs cannot be overemphasized. The structures of this ubiquitous family of integral membrane proteins are as important as the proteins themselves, especially for drug discovery. Attempts to grow crystals of the transmembrane domain of GPCRs have been greeted with a great deal of difficulty. Structural determination of proteins, employing crystallographic technique, generally requires that they are present in a reasonably large quantity. GPCRs, however, are naturally not readily available in organisms where they are found, and therefore require to be produced using recombinant techniques to obtain large quantity (Lunstrom, 2006). The use of GPCRs produced by heterologous protein expression for structural studies has faced a number of obstacles. This is specifically due to the highly hydrophobic transmembrane region of these receptor proteins. The use of detergents to solubilized this region of the receptor to mimic the native phospholipid membrane has experienced limited success as the proteins tend to be highly unstable and fail to maintain their native form (Grisshamer *et al.*, 2006; Lundstrom and Chiu, 2006; Kobilka, 2007; McCusker *et al.*, 2007).

Despite the undisputable physiological and pharmacological significance of GPCRs, in the long years after GPCR discovery and efforts to break the deadlock in determining their structure, the first structure of a GPCR did not show up until August 2000 when Palczewski *et al.* (2000) released the crystal structure of bovine rhodopsin covalently bound to retinal (Palczewski *et al.*, 2000). The structure became the first ever high resolution structure of a GPCR to be determined. The structure showed 11-*cis* retinal bound to Lys296 in TM 7 at one end while the other end a  $\beta$ -ionone ring is submerged in a hydrophobic core (formed by Trp265, Phe212 and Tyr268) between TM 5 and 6. The interaction between Trp265 and the  $\beta$ -ionone ring causes conformational changes between the active and inactive states in the Trp265. This particular conformational change is termed the “toggle switch”, which is deemed important for the activation and inactivation of rhodopsin (Topiol and Sabio, 2009).

There is more evidence supporting the involvement of these residues in the relative movement of the TM bundle in rhodopsin activation. In the work of Lin and Sakmar (1996) for instance, using tryptophan UV-absorbance, an elimination of spectral differences that distinguishes rhodopsin from metarhodopsin II was obtained on mutating the Trp residues in TM 3 and 6 (Lin and Sakmar, 1996). Another feature in this structure is the ionic lock, which is thought to keep the receptor in its inactive state (Palczewski *et al.*, 2000). The ionic lock is an ionic interaction between the charged residues Arg135<sup>3.50</sup> and Glu247<sup>6.30</sup> at the intracellular end of TM3 and TM6 respectively. Arg135<sup>3.50</sup> is part of the E/DRY motif of the rhodopsin family receptors and has up to 98.1% conservation across the family (Smith, 2010). This ionic lock is believed to stabilize the inactive state of rhodopsin and upon rhodopsin activation, it is disrupted mainly with contributions from Tyr223, Met257 and Tyr306 (see Smith, 2010).

There were a number of shortcomings identified about using this structure as a template for constructing homology models for other GPCRs. These limitations as highlighted in reviews by Parrill (2008) and Topiol and Sabio (2009) included the unusual covalently bound retinal and the scepticism about whether other GPCRs would share the same binding-site geometry in relation to the movement and rotation of TM bundle (Parrill, 2008; Topiol and Sabio, 2009). Other limitations were the very low ( $\leq 20\%$ ) overall sequence identity between rhodopsin and other GPCRs, especially of families B and C; the small size of the binding cavity of rhodopsin compared to that obtainable for other GPCRs; the blocking of the binding site by ECL2 (as seen in rhodopsin); and the problem of deciding what oligomerization to consider in modelling a target GPCR (see Topiol and Sabio, 2009). Despite these shortcomings, there is much interest in modelling GPCRs using the rhodopsin as template. For instance, the structures of agonist-bound and antagonist-bound forms of human dopamine D3, among others, were modelled using rhodopsin as structural template (Bissantz *et al.*, 2003).

The crystal structure of rhodopsin was the only GPCR structure until 2007 when the crystal structures of  $\beta_2$ -adrenergic receptor ( $\beta_2$ -AR) were deposited (Cherezov *et al.*, 2007; Rasmussen *et al.*, 2007). The structure was a complex of the receptor with an inverse agonist, carazolol. This work employed two

different strategies both aimed at offering stability to the receptor's flexible ICL3 to aid crystallization. The first involved complex formation of the ICL3 with a monoclonal antibody. The crystals for the structure were grown with lipidic micelles producing a 3.4 Å resolution. The second strategy replaced the 'notorious' ICL3 with T4-lysozyme and the crystals grown with lipidic cubic phase with a resultant 2.4 Å resolution (Cherezov *et al.*, 2007; Rasmussen *et al.*, 2007). Despite the presence of a 'pioneer' GPCR crystal structure, this structure was nevertheless greeted with great enthusiasm as it was regarded as the first ligand-mediated GPCR crystal structure and was believed to bring with it answers to questions raised by the rhodopsin structure (Topiol and Sabio, 2009). On releasing this structure, it was imperative that comparisons were made to the rhodopsin structure. The overall architecture shares resemblance with that of bovine rhodopsin and the ligand binding position of carazolol to  $\beta_2$ -AR tallies with that of retinal to rhopsin. Moreover, the TM domain shares structural conservation of at least 13 amino acids in each helix (Parrill, 2008). Some distinguishing features are however noteworthy. For instance, the nature of ligand-receptor interaction differs. While retinal is covalently bound to rhodopsin, carazolol is bound by hydrogen bonding. Also, the ECL2 of rhodopsin contained a  $\beta$ -sheet, while that of  $\beta_2$ -AR was composed of an  $\alpha$ -helix with two disulphide bridges, one within the loop and the other connected to TM3 (Cherezov *et al.*, 2007; Topiol and Sabio, 2009). This feature douses the concern about the blockade of the binding site by ECL2 of rhodopsin as these bridges prevent similar encounter in  $\beta_2$ -AR. One controversial issue with the  $\beta_2$ -AR/rhodopsin structures was the "ionic lock" hypothesized for rhodopsin. Although the Trp, Arg and Tyr residues occupying positions 130-132 respectively in  $\beta_2$ -AR were reported to form the "D(E)RY" motif indicated in the ionic lock for rhodopsin, the ionic interaction (which represents the ionic lock) between the equivalent residues in  $\beta_2$ -AR was absent in the solved  $\beta_2$ -AR crystal structures (Rasmussen *et al.*, 2007).

In the year following the release of the carazolol-bound  $\beta_2$ -AR, another structure of the  $\beta_2$ -AR was deposited to protein data bank (PDB) by Hanson *et al.* (2008). Although this time bound to a partial inverse agonist, timolol, this was more or less a repetition of the success achieved for the previously repeated  $\beta_2$ -AR structures. Here, as before, the ICL3 was replaced by T4L and as expected, the

structure resembles that of the carazolol bound (Hanson *et al.*, 2008). The only observable difference was in the binding site conformation which was due to the presence of a different agonist. Also, two cholesterol binding sites, not involved in crystal packing, were revealed in this structure as opposed to the initial structure (Topiol and Sabio, 2009).

In another display of technical brilliance, the crystal structure of the  $\beta_1$ -adrenergic receptor ( $\beta_1$ -AR) became available through Serrano-Vega *et al.* (2008). This structure was a complex of an antagonist, cyanopindolol, and turkey  $\beta_1$ -AR. The work undertook a completely different approach in stabilizing this agonist-bound GPCR for easier crystallization. Here, the authors made six mutations (R68S, M90V, Y227A, A282L, F327A and F338M), which conferred thermostability on the receptor (Serrano-Vega *et al.*, 2008). It is important to mention however that the authors excised some residues from ICL3 and C-terminus, albeit with no extraneous protein introduced. This structure, as expected, shares a large degree of similarity with the  $\beta_2$ -AR. However, the basal activity observed for the  $\beta_2$ -AR structure was absent in  $\beta_1$ -AR. This has been suggested to be due to antagonist bound to  $\beta_1$ -AR (Topiol and Sabio, 2009). Also, the ionic lock was observed for the  $\beta_1$ -AR with different residues indicated (Serrano-Vega *et al.*, 2008).

2008 saw the release of the 3<sup>rd</sup> GPCR crystal structure in a row on the release of the human  $A_{2A}$  adenosine receptor bound to a high affinity antagonist, ZM241385, by Jaakola *et al.* (2008). The T4L replacement of ICL3 was employed in this work. This was another welcomed development in the realm of GPCR crystal structures. The ECL2 in this structure is random coiled (Jaakola *et al.*, 2008). This differs from the earlier reported  $\beta$ -sheet and  $\alpha$ -helix for rhodopsin and  $\beta_2$ -AR respectively. Moreover, in comparison with either rhodopsin or  $\beta_2$ -AR, there are significant shifts observed in the helices and a notable difference in the binding of the antagonist (Topiol and Sabio, 2009). The “ionic lock” paradigm was observed for this receptor structure (Jaakola *et al.* (2008) indicating agreement, in this wise, with rhodopsin and  $\beta_1$ -AR. This mixture of similarities and dissimilarities further reiterate the difficulty in using a GPCR structure as a template for the structural homology modelling of the other, either within or outside the same family.

The boom of GPCR structures was put to a short hold until 2010 when the crystal structures of the CXCR4 chemokine receptor (Wu *et al.*, 2010) and dopamine D3 receptor (Chien *et al.*, 2010) were released. CXCR4 was the first peptide GPCR to be crystalized. It showed a more open conformation at the extracellular end and so may be more relevant to the study of family B GPCRs. The CXCR4 structures were crystallized bound to a small molecule antagonist, IT1t, or a cyclic peptide inhibitor, CVX15, both at resolutions of 2.5Å and 2.9Å respectively. The dopamine D3 receptor structure was in complex with a selective antagonist, eticlopride. Both receptor structures employed the methods of thermostability (with mutations at L125<sup>3,41</sup> (CXCR4) and L119<sup>3,41</sup> (dopamine D3), both to Trp) and T4L fusion of ICL3. There are observed structural differences for these receptor structures compared to previous receptor structures. The extracellular end of TM1 in CXCR4 receptor, for instance, shifts towards the central axis of the receptor by 9Å (relative to  $\beta_2$ -AR) and more than 3Å (relative to A<sub>2A</sub>) while the TM2 makes a tighter (~120°) helical turn at Pro92<sup>2,58</sup> of its extracellular end with a consequential one-residue gap in sequence alignment with other receptors (Wu *et al.*, 2010). In the case of dopamine D3, some shifts are observed in TM 6 & 7 relative to  $\beta_2$ -AR and unlike  $\beta_2$ -AR, the “ionic lock” is observed (Chien *et al.*, 2010). The presence of this ionic lock tends to refute the idea that its absence in the  $\beta_2$ -AR might have been caused by the presence of the T4L.

Between the years 2011 and 2012, there was a surge in the solution of GPCR structures with up to 14 structures deposited to PDB. These included the active forms of human  $\beta_2$ -AR and bovine rhodopsin (see Venkatakrishnan *et al.*, 2013). All these structures are from within the rhodopsin family and were all crystallized via X-ray crystallographic technique except for the human CXCR1 chemokine receptor which was the first structure to be solved by nuclear magnetic resonance (NMR) spectroscopy. The most fascinating of these structures, which forms another remarkable achievement in the study of GPCRs, is the structure of the active state ternary complex of  $\beta_2$ -AR with the heterotrimeric G protein (Gs) (Rasmussen *et al.*, 2011). The conformational changes observed in this structure are comparable to that of rhodopsin. The most striking movement observed within the TM bundle, when compared to the inactive conformation, was a 14Å outward movement of TM6 taking the C $\alpha$  carbon of Glu268 as reference point. Also observed was an  $\alpha$  helix formed within ICL2 of the

active receptor which was absent in the inactive state; although this helical conformation has also been observed in ICL2 of inactive avian  $\beta_1$ -AR and therefore might not be specifically characteristic to the active state. The ICL2 however contributes to the  $\beta_2$ -AR-Gs interface (see Rasmussen *et al.*, 2011). The receptor in this structure interacts only with the  $\alpha$  subunit of Gs with no direct interaction with the  $\beta$  and  $\gamma$  subunits observed (Rasmussen *et al.*, 2011).

Most recently, in April 2013, the first non-rhodopsin family GPCR, the human smoothened (SMO) receptor (belonging to the frizzled family) was released (Wang *et al.*, 2013). Although this receptor share very low (less than 10%) sequence homology with family A GPCRs and lack most of the conserved family A motifs, it exhibits the conventional GPCR 7TM fold with connecting intracellular and extracellular loops as well as an intracellular membrane-parallel helix 8 (Wang *et al.*, 2013). Moreover, some intracellular structural features of the rhodopsin family were observed to be preserved in this receptor. For example, a helical turn was observed within the short intracellular loop (ICL) 1 of this receptor as seen for family A receptors. However, one striking difference observed within the TM bundle is the absence of any proline in the TM6. In other words, this receptor is devoid of the common kink (caused by the highly conserved P<sup>6.50</sup>) in TM6 of family A GPCRs and its TM helix 6 is therefore straighter than those of family A GPCRs (see Wang *et al.*, 2013).

### **1.2.5 GPCR 8<sup>th</sup> helix: structural and functional variations**

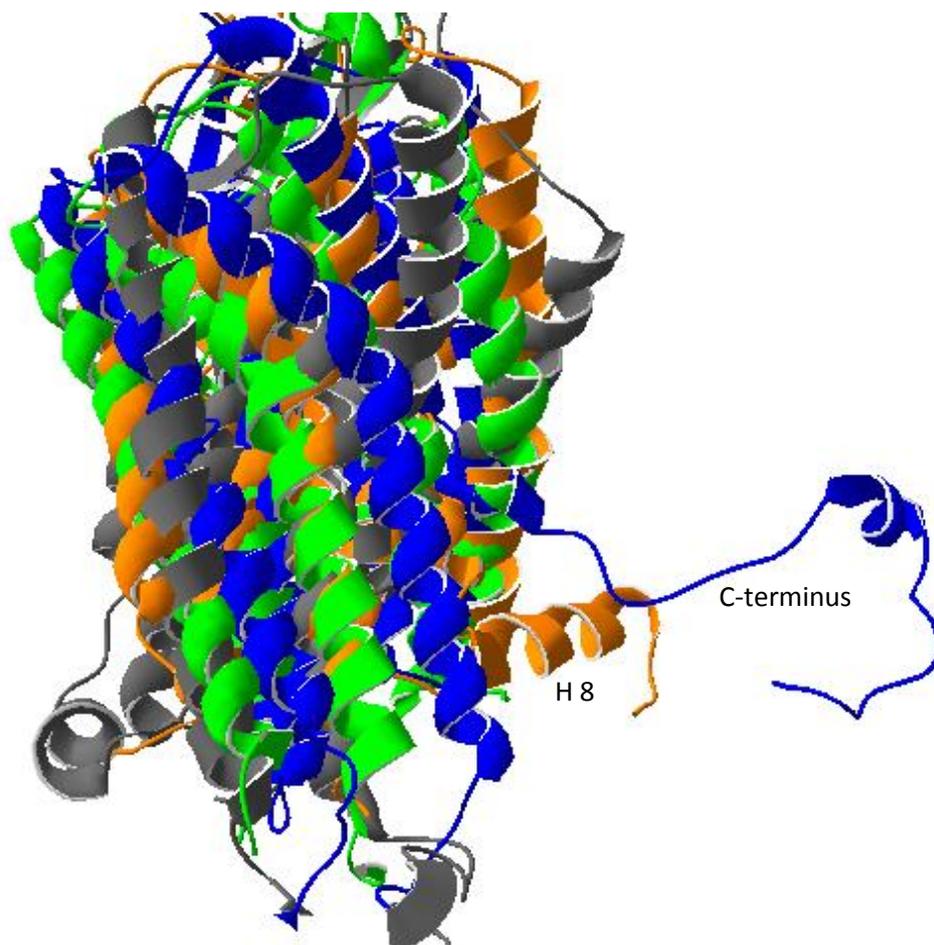
The release of X-ray crystal and NMR structures for family A GPCRs have shown that virtually all possess an amphipathic helix at the proximal end of the carboxyl tail usually oriented parallel to the membrane bilayer and perpendicular to the TM bundle. This helix (tagged helix 8 or the 8<sup>th</sup> helix or simply H8) is typically 3 helical turns in length and attached to TM7 by a short linker (Venkatakrisnan *et al.*, 2013). In many cases, but not all, this helix ends with one or two cysteine residue(s) which serve as potential palmitoylation sites for anchoring the helix to the membrane bilayer (Palczewski *et al.*, 2000; Venkatakrisnan *et al.*, 2013). While these (potential palmitoylation) sites, for instance, are present in bovine rhodopsin, the human  $\beta_2$ -AR and turkey  $\beta_1$ -AR, they are

absent in human adenosine A<sub>2A</sub> and Histamine H<sub>1</sub> receptors (Palczewski *et al.*, 2000; Cherezov *et al.*, 2007; Warne *et al.*, 2008; Jaakola *et al.*, 2008; Shimamura *et al.*, 2011). One common motif observed in the primary sequence of helix 8 in family A GPCRs is the F[R/K]xx[F/L]xxx[L/F] motif. This conserved motif is important in stabilizing the helix by forming various interactions with other structural elements within the receptor. Very recently, the structure of the human smoothed (SMO) receptor was deposited to the protein data bank (PDB) (Wang *et al.*, 2013). This receptor, which belongs to the frizzled family GPCRs and the first non-class A GPCR with its structure solved, also has a helix 8 parallel to the membrane bilayer (Wang *et al.*, 2013).

Conversely, helix 8 has been reported absent or unstructured in the human CXCR4 chemokine receptor, proteinase-activated receptor (PAR1) and rat neurotensin receptor (NTSR1) (Wu *et al.*, 2010; Zhang *et al.*, 2012; White *et al.*, 2012; Venkatakrisnan *et al.*, 2013). What is rather surprising but interesting is that while helix 8 is reportedly absent in CXCR4, it is present in the closely related human chemokine CXCR1 receptor (Park *et al.*, 2012). This however is not without debates. Firstly, CXCR1 structure was solved using solid state NMR in liquid crystalline phospholipid bilayers and in physiological conditions unlike the CXCR4 structure developed by X-ray crystallography and, like many other GPCR structures, with amino acid modifications. Moreover, the CXCR4-T4L does not bind G-protein (Wu *et al.*, 2010). Considering these, it is therefore possible that the absence of helix 8 in the CXCR4 structure is due to the experimental method used in generating this structure. It is also possible, as suggested by the authors (of CXCR4 crystal structure), that this helix is formed under certain (physiological) conditions which was not met under the experimental method employed (Wu *et al.*, 2010). On the other hand, the F[R/K]xx[F/L]xxx[L/F] motif is partially conserved in CXCR4, in which case it has an FKxxAxxxL motif where the Phe/Leu (at the mid-region of the motif) important for stabilizing the helix has been replaced by an Ala (Wu *et al.*, 2010). So, it cannot be ruled out that this receptor lacks the tendency to form helix 8 *in vivo*.

In the case of PAR1, based on phylogenetic analysis, this receptor is more distant, from other family A GPCRs with solved crystal structures. For instance, the NP<sup>7.50</sup>xxY motif is DP<sup>7.50</sup>xxY in PAR1 (Zhang *et al.*, 2012). This distance in relationship probably accounts for the absence of helix 8 in this

receptor. Moreover, the highly conserved phenylalanine residue at the proximal end of helix 8 is not conserved in this receptor. The equivalent of this is a cysteine residue and has been found to play a different, less important role, to that of F313 in rhodopsin in Gt activation (see Fritze *et al.*, 2003). The rat NTSR1 which reportedly also lack the 8<sup>th</sup> helix has its TM7 helix extended by 3 helical turns further down into the cytoplasmic face in a virtually perpendicular (not parallel) orientation to the membrane bilayer (Fig 1.7) (White *et al.*, 2012). Interestingly, this receptor has the F[R/K]xx[F/L]xxx[L/F] sequence (F<sup>376</sup>RxxFxxxL in PAR1) at a similar position (i.e. just under 2 helical turns beyond the NPxxY motif) as found in rhodopsin and  $\beta_2$ -AR. In agreement with the authors, it is therefore possible that this variation from other receptors is due to crystallization artefact (White *et al.*, 2012).



**Fig 1.7:** Structural alignment of bovine rhodopsin (Orange), human PAR1 (Green), human CXCR4 (Blue) and rat NTSR1 (Grey) showing the different structural variations at H8 relative to rhodopsin. Alignment was done using SPDB viewer bioinformatics software. H8 and C-terminus of rhodopsin and CXCR4 respectively are indicated. PDB codes: rhodopsin – 1F88; PAR1 – 3VW7; CXCR4 – 3ODU; NTSR1 – 4GRV.

The major form of stabilizing/functional interaction made by helix 8 is that involving the ICL1 and this is the most common among GPCRs with solved structures to date (Palczewski *et al.*, 2000; Venkatakrisnan *et al.*, 2013). In the bovine rhodopsin structure for instance, helix 8 interacts with ICL1 at H65. Similar interactions between helix 8 and ICL1 have also been observed for several other GPCRs with solved structure like the human adenosine  $A_{2A}$ , Histamine  $H_1$  and spingosine-1 phosphate (S1P<sub>1</sub>) receptors (Jaakola *et al.*, 2008; Shimamura *et al.*, 2011; Hanson *et al.*, 2012). However, there are variations to these in few other receptors. In squid rhodopsin, H8 interacts with a 9<sup>th</sup> helix beyond helix 8 in the C-terminus and are joined by a loop structure with a short  $3_{10}$  helix buried in the hydrophobic membrane region. This interaction stabilizes H8 and in turn helps to restrict

the motional freedom of H9 (Murakami and Kouyama, 2008). Also in CXCR1, H8 is orientated at an angle different from that observed in most other receptors and does not appear to interact with ICL1 (Park *et al.*, 2012). In the recent human smoothed receptor structure, helix 8 has an interface with helix I that is dissimilar from that observed in the rhodopsin family (Wang *et al.*, 2013).

In addition, there are conformational changes reportedly observed for H8 important for arrestin coupling (Kirchberg *et al.*, 2011) and this helix is partially unwound in the active  $\beta_2$ -AR (Hulme, 2013). In the  $\kappa$ -OR receptor, H8-H8 interaction was a major interaction in receptor dimerization, although it is not known if this dimerization is obtainable *in vivo* (Wu *et al.*, 2012). Overall, while there might appear to be some consistency in the characteristic features observed for helix 8 across receptors with solved structures, there is also a long list of variations; and with many more structures yet to be deposited, there could be more to be known than have been known about the diverse structural and functional roles of helix 8.

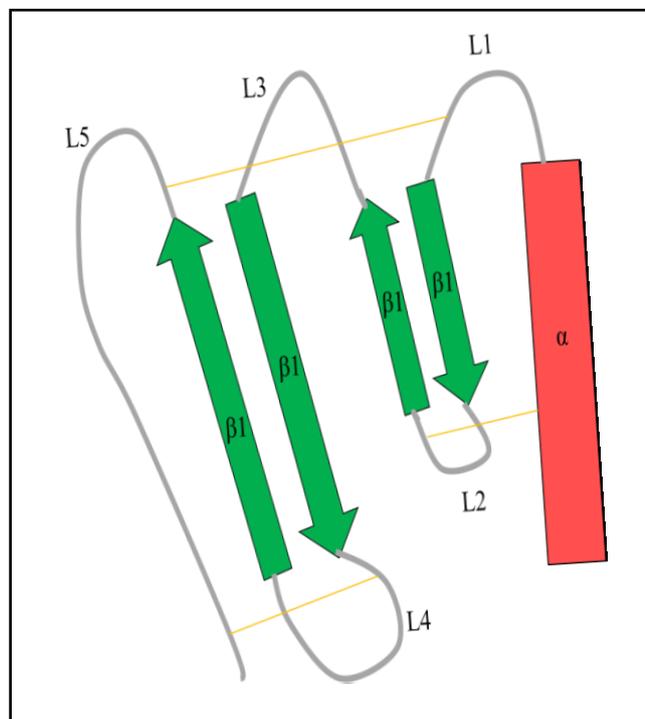
### **1.3 The secretin family receptors**

#### **1.3.1 Structural theme**

The secretin-like (family B) receptors, like the rhodopsin-like (or family A) GPCRs, are believed to exhibit the 7TM architecture (Sheikh *et al.*, 1999; Kirkpatrick *et al.*, 2012) and they share over 30% sequence homology among one another within the same family. Despite their overall similar architecture with family A, signature sequences found in family A GPCRs are largely thought to be absent in family B (Hamar, 2001; Wheatley *et al.*, 2011; Miller *et al.*, 2011). While a complete knowledge of their full structure is limited due to the lack of a crystal structure for an intact receptor within this family, the huge successes recorded within the past decade in obtaining crystal/NMR structures (Grace *et al.*, 2004, 2007; Parthier *et al.*, 2007; Sun *et al.*, 2007; Pioszak and Xu, 2008; Runge *et al.*, 2008; ter Haar *et al.*, 2010; Underwood *et al.*, 2010; Kumar *et al.*, 2011) for the N-terminal extracellular domains (ECDs) for several of these family members have provided good understanding of this part of the receptor structure.



The large and structurally-conserved N-terminus ECD of these receptors share the “secretin receptor recognition fold” which comprises two antiparallel  $\beta$ -sheets, 3 disulphide bonds, an  $\alpha$ -helix at the extreme N-terminal (Fig 1.9) which varies in length and some loop regions predominantly stabilized by the disulphide bonds (Parthier *et al.*, 2009). This domain also contains a highly conserved hydrophobic ligand-binding groove localized between the  $\alpha$ -helix and conserved loop region (Parthier *et al.*, 2009). Although some of these structures have their endogenous peptide ligands bound (Grace *et al.*, 2007; Parthier *et al.*, 2007; Runge *et al.*, 2008), there are still concerns over the correct orientation of loop regions because of the absence of the TM domain (Miller *et al.*, 2012). The role(s) of this region in ligand binding and receptor activation is discussed in section 1.3.4.1.



**Fig 1.9:** A rough sketch of the common structural fold of family B GPCRs. The N-terminal  $\alpha$ -helix is shown in red,  $\beta$  strands  $\beta 1 - \beta 4$  in green and loop regions in grey (numbered L1 – L5). Disulphide bridges which stabilize the structure are shown in yellow.

There is minimal information about the TM domain and associated loops orientation of this receptor family as a structure of this region for any of the family member is yet to be deposited to protein data bank (PDB). However, using various alignment and predictive techniques, there have been several

attempts to understanding the arrangement and orientation of the TM helices and associated loop structures (Donnelly, 1997; Scheikh *et al.*, 1999; Chungunov *et al.*, 2010; Dong *et al.*, 2011; Kirkpatrick *et al.*, 2012; Vohra *et al.*, 2013). One emerging proposition in recent structural models of receptors in this family is the presence of the TM3-TM6 ionic lock which has been observed for rhodopsin and some other family A GPCRs where they are believed to maintain receptor structure in the inactive state (Topiol and Sabio, 2009). The lock is formed by an ionic interaction between R<sup>3.50</sup> of the DRY motif at the base of TM3 and E<sup>6.30</sup> at the cytoplasmic face of TM6 (Ballesteros *et al.*, 2001). A similar interaction has been proposed in GLP1R, VPAC1R and CLR model structures; although it is suggested they might play a slightly different role in structure and function compared to the rhodopsin family (Kirkpatrick *et al.*, 2012; Chungunov *et al.*, 2010; Vohra *et al.*, 2013). While this proposition in 3 different family B receptors suggests some consistency and plausibility of this in other family members, same cannot yet be said of the residues involved in forming this lock. Despite the suggestion of the YLH motif as being the family B equivalent of the family A DRY motif, the former does not appear to directly contribute to the formation of this proposed ionic interaction. Rather, the highly conserved R173, H177 and E233 (a turn above the YLH motif; numbering according to CLR) have been suggested to directly contribute to the formation of this lock which possibly involves contribution from another highly conserved residue T338 and believed to help stabilize the receptors in their inactive conformation (Vohra *et al.*, 2013). There is further evidence supporting similar roles for corresponding residues in other family B GPCRs. In the PTHR1 for instance, natural H223R and T410P mutations in patients with Jansen's metaphysealchondrodysplasia – an uncommon type of short-limbed dwarfism – resulted in constitutively active receptors (Schipani *et al.*, 1996). Moreover, mutagenesis studies on equivalent residues in GIPR and VPAC1R also significantly caused constitutive activity (Tseng and Lin, 1997; Gaudin *et al.*, 1998) suggesting that these residues likely play a role in maintaining receptors inactive conformation. Also, the basic residues R188 and R190 (equivalent of R173 in CLR) in GLP1R and VPAC1R respectively have been indicated in a putative TM2-TM3-TM6 interaction suggested to be crucial in Secretin-like family structure and function (Kirkpatrick *et al.*, 2012; Chungunov *et al.*, 2010). This region of the receptor will further be discussed for their role in receptor activation in section 1.3.4.2.

### 1.3.2 Peptides of the secretin family

The secretin family peptides comprise endogenous ligand peptides usually 27 – 52 amino acid in length (Insel, 2007). They commonly share a helix-loop-helix architecture and usually have an  $\alpha$  helical backbone with which they normally bind the N-terminal ECD of their receptors (Mierke and Pellegrini, 1999; Jin *et al.*, 2000; Parthier *et al.*, 2009; Watkins *et al.*, 2012). Their various receptors, which make up the secretin family GPCRs, and physiological roles are summarised in Table 1.2.

**Table 1.2:** A summary of the receptors, major physiological roles and sequence sizes of secretin family peptides. Adapted from Hoare (2005). CGRP, calcitonin gene-related peptide; CLR, calcitonin receptor like receptor; RAMP, receptor activity modifying protein; CRF, corticotropin-releasing factor; GHRH, growth hormone-releasing hormone; GIP, glucose-dependent insulintropic peptide; GLP, glucagon-like peptide; PTH, parathyroid hormone; TIP39, Tuberoinfundibular peptide of 39 residues; VIP, vasoactive intestinal peptide; PACAP, pituitary adenylate cyclase-activating polypeptide; CTR, calcitonin receptor.

Ligand	Size (a.a)	Principal receptor	Major physiological roles
CGRP	37	CLR + RAMP1	Vasodilation
Adrenomedullin	52	CLR + RAMP2 CLR + RAMP3	Vasodilation
CRF	41	CRF receptor 1	Adrenocorticotropin hormone (ACTH) release, central stress responses
Urocortin (UCN)	40	CRF receptor 2	Central stress responses (UCN1); cardiac contractility (UCN2); hearing (UCN3)
GHRH	44	GHRH receptor	Stimulation of growth hormone release
GIP	42	GIP receptor	Secretion of insulin
Glucagon	29	Glucagon receptor	Regulation of blood glucose
GLP-1	36/37	GLP-1 receptor	Regulation of insulin and glucagon secretion
GLP-2	33	GLP-2 receptor	Gut mucosal growth
PTH1	84	PTH receptor 1	Ca <sup>2+</sup> homeostasis
TIP39	39	PTH receptor 2	Hypothalamic secretions, nociception
Secretin	27	Secretin receptor	Pancreas secretions
VIP	28	VPAC1R; VPAC2R	Vasodilation, vascular and neuroendocrine functions
PACAP	38	PACAPR1 VPAC1R; VPAC2R	Neurotransmission, neuronal modulation and regulation
Calcitonin	32	Calcitonin receptor	Ca <sup>2+</sup> homeostasis
Amylin	37	CTR + RAMP1 CTR + RAMP3	Reduces feeding by slowing gastric emptying. Partly in bone metabolism

The secretin peptides family also possess a characteristic helix N-cap which, in principle, helps protect against helix unwinding of peptide helix thereby stabilizing the peptide ligands. This N-capping motif is thought to represent a specific fold that aids interaction of the ligand N-terminus with the receptor juxta-membrane domain (the two-domain model, discussed below) and helps initiate receptor activation (Neumann *et al.*, 2008). This motif however is absent in the calcitonin subfamily of peptides. Here, the equivalent of this motif is a disulphide bridge between the invariantly conserved C1 and C7 (numbered for calcitonin) and is believed to perform a similar function to the N-capping motif of other peptides of the secretin family (Neumann *et al.*, 2008).

### **1.3.3 Mechanism of ligand binding and activation of secretin family receptors: the ‘two-domain’ model**

The “two-domain” model describes ligand binding and activation of family B GPCRs as process involving two steps. The first involves an affinity binding of the C-terminal region of the peptide ligand to the N-terminus of the receptor. The second step, involves the interaction between the N-terminus of the ligand and the juxta-membrane (J) domain of the receptor (Hoare *et al.*, 2005). While the former is significant for affinity-driven, specific ligand recognition, the latter interaction is primarily responsible for receptor activation. Even though no structure currently exists for an intact family B receptor, studies employing mutagenesis, photoaffinity labelling, molecular modelling and, on receptor N-terminal fragments, NMR and X-ray crystallography have all provided substantial evidence supporting this model. Early studies in this respect include those of Stroop *et al.* (1995), Holtman *et al.* (1995) and Bergwitz *et al.* (1996) where hybrid peptide ligands and receptor chimeras were generated to test ligand-receptor interaction within the family B ligands and receptors.

In the study by Bergwitz *et al.* for instance, peptide ligand hybrids made up of the N-terminal portion of salmon CT and C-terminal portion of bovine PTH and vice versa (i.e. sCT/bPTH and bPTH/sCT); and receptor chimeras formed by N-terminus of rat PTH and J-domain plus C-terminus of porcine CTR and vice versa (i.e. rPTH/pCTR and pCTR/rPTH) were generated. The CT/PTH and PTH/CT

hybrids respectively activated the PTH/CTR and CTR/PTH receptors, but not the wild type (Bergwitz *et al.*, 1996). The study showed that the two composite ligands, notwithstanding their widely varying primary sequence, displayed an identical pattern of ligand-receptor interaction. It also indicated that both CT and PTH exhibit similar architecture involving two functional receptor-specific domains (Bergwitz *et al.*, 1996).

Today, NMR and X-ray crystal structures of ligand-free and agonist-bound N-terminal extracellular domains of several family GPCRs have shown that the C-terminus together with the mid-region of ligands in this family recognise and bind to the N-terminus of their corresponding receptors (Parthier *et al.*, 2009). This pattern of interaction, which sees the predominantly-helical ligand within a conserved hydrophobic core of the ECD, has been tagged the “hot-dog-in-a-bun” mode of ligand binding (Pioszak and Xu, 2008).

Although available evidence supporting the interaction of the ligand N-terminus with the receptor J-domain are only predictive (e.g. Runge *et al.*, 2003; Dong *et al.*, 2011, Coopman *et al.*, 2011; Kirkpatrick *et al.*, 2012), there is more concrete evidence indicating the J-domain as being primarily responsible for receptor activation. In a study by Shimizu *et al.* (2001), N-terminally truncated PTH receptor was activated with a maximum response similar to wild type following agonist treatment, although this was unsurprisingly with significantly reduced potency (Shimizu *et al.*, 2001). Also, mutation to the J-domain could result in constitutive activation of receptors (Hjorth *et al.*, 1998). As earlier mentioned, the helix N-capping of secretin family peptides, and the C1-C7 disulphide bridge in the calcitonin peptides subfamily, is thought to be the primary signature that initiates receptor activation further supporting this model (Neumann *et al.*, 2008).

### **1.3.4 The receptor regions in ligand binding and activation**

#### **1.3.4.1 N-terminal extracellular domain (ECD)**

One fact that has been gathered over the years about the N-terminus of secretin family receptors is their ability to bind endogenous peptide agonists even in the absence of the TM bundle, albeit with low affinity (Perrin *et al.*, 2003; Koth *et al.*, 2010). Available ECD structures revealed that this domain binds a large portion (the  $\alpha$ -helical mid region and C-terminus) of their peptide ligands (Parthier *et al.*, 2009), showing that they play a crucial role in ligand binding. The ligand helix lies along a binding core formed by a hydrophobic cluster of amino acids. This cluster is made of residues of the WDG motif within the highly conserved  $\beta$ -hair pin structure and the conserved large loop structure between the second antiparallel  $\beta$ -sheets as well as residues from the N-terminal  $\alpha$ -helix. There is evidence from both structural and mutagenesis studies supporting the importance of these residues in ligand binding and receptor signalling (Parthier *et al.*, 2007; Pioszak and Xu, 2008; Pioszak *et al.*, 2009; Barwell *et al.*, 2010; Kusano *et al.*, 2011; Kumar *et al.*, 2011). While the role of the ECD in ligand binding cannot be disputed, it is not clear what role they play in receptor activation. As will be discussed in subsequent sections, it is widely believed that the activation of this family of receptors involves, though is not restricted to, the TM bundle and extracellular loops (Dong *et al.*, 2011; Kirkpatrick *et al.*, 2012; Wheatley *et al.*, 2011).

#### **1.3.4.2 The transmembrane (TM) domain**

The conformational changes that accompany the activation of family B GPCRs cannot yet be ascertained due to a lack of a crystal/NMR structure for any member of this family. However, there are some findings based on predictive studies which, with the knowledge of the well-studied family A GPCRs, could in part plausibly explain the activation mechanism. Overall, the TM helices 3 and 6 appear to play a central role in the inactive-active interchange of these receptors. The latter seems to be more structurally designed to be pivotal in these movements that drive receptor activation. A major characteristic feature is the proposed kink which is suggested to be caused by the highly conserved

proline in the mid region of TM6. This has been shown to be crucial for G-protein coupling (Conner *et al.*, 2005; Bailey and Hay, 2007) and in maintaining receptor structure (Knudsen *et al.*, 2001). Inactive-active movement observed for TM6 in  $\beta_2$ AR (family A) is also believed to be present in PTH1R of family B (Sheikh *et al.*, 1999). As earlier mentioned, highly conserved Arg and His around the cytoplasmic face of TM2 and a Glu at the base of adjacent TM3 have been suggested to form the family B equivalent of the TM3-TM6 ionic lock in family A. What is not clear however is if there is a corresponding TM6 residue in family B that is directly involved in this interaction as suggested by Kirkpatrick *et al.* (2012) and Vohra *et al.* (2013). It is believed that upon ligand binding, this ionic lock breaks to create a binding pocket for G $\alpha$ . Other networks of interactions like the TM2-TM3-TM7 have also been proposed in the GLP1 and VPAC1 receptors (Kirkpatrick *et al.*, 2012; Chungunov *et al.*, 2010). The interaction network between N229 (TM3) and Q380 (TM7) in VPAC1R for instance seems to maintain TM7 in conformation required for G-protein activation (Chungunov *et al.*, 2010).

There is a large network of polar interactions in the rhodopsin family GPCRs which are believed to be important for maintaining receptor conformation in the inactive and active states (Gether *et al.*, 2002). Although the conserved polar residues in family A receptors are absent in family B, they have their own distinct polar residues within the TM region which could play similar role as those in family A (Wooten *et al.*, 2013). Kirkpatrick *et al.* (2012) have reported that a good number of these polar residues may form hydrogen bonds and salt bridges that keeps receptor in either an active or inactive conformation and plays a role in conformational switches accompanying receptor activation. The TM bundle also plays a role in ligand binding and receptor activation by forming a binding pocket at its extracellular end for the N-terminus of the peptide according to the two state model (Hoare, 2005) and as observed, for instance, in the GLP1 receptor-ligand model structure (Coopman *et al.*, 2011). This role is normally played together with the extracellular (EC) loops which are discussed below.

### 1.3.4.3 Extracellular loop (ECL) region

One bane of successful crystallization of membrane proteins, including GPCRs, is the constant movement of loop structures (Bill *et al.*, 2011). Even when successfully crystallized, the correct positioning/orientation of loop structures still remains in question. Modelling the loop structures therefore from family A crystal structure is very challenging and it is difficult to draw conclusions from structures obtained in this respect. One interesting thing however is that there have been studies on structural modelling where the role of loop structures have been supported by mutagenesis and photo affinity cross-linking data on intact receptors (see Wheatley *et al.*, 2011). There are a few conserved residues within the three ECL 1-3 of the family B GPCRs. ECL2 has the highest number of conserved residues. Although the exact boundaries of the ECLs cannot be currently ascertained, ECL2 probably starts with conserved basic residue (R/K) and it is possible that these residues interact with membrane phospholipids as suggested in family A GPCRs (Hawtin *et al.*, 2006). A conserved Cys also exists in the mid-region of ECL2. There are results from mutagenesis studies supporting the presence of a disulphide bond between Cys residues in ECL2 and at the top of TM3 in GLP1R (Mann *et al.*, 2010; Koole *et al.*, 2012). This locus is suggested to contribute to the ligand binding pocket within the juxtamembrane domain. Also, residues Met and Tyr at positions 204 and 205 respectively in ECL1 of rGLP1R have been suggested to be important in the formation of binding site for ligand N-termini following a double mutation of these residues to Ala (Lopez de Maturana *et al.*, 2004). These residues did not affect receptor signalling when individually mutated to Ala in the same study and it is consistent with that obtained for equivalent residues of CLR (Barwell *et al.*, 2011), although the equivalent of M204 in the glucagon receptor (an Arg) has been proposed as required in glucagon binding and receptor activation (Unson *et al.*, 2002) and Y205 has been reported as a point of interaction for a photoactive analogue of GLP1 (Chen *et al.*, 2010). It is not particularly clear whether these residues might contribute to ligand binding and/or receptor activation in other family members.

Following several attempts to predict the exact mode of interactions of seretin family peptides to their receptors (e.g. Monaghan *et al.*, 2008; Dong *et al.*, 2011, Coopman *et al.*, 2011; Kirkpatrick *et al.*, 2012), two forms of interaction between the ligand N-terminus and receptor juxta-domain have been

postulated. The N-terminus either penetrates the TM bundle in a perpendicular orientation or lies in a parallel manner across this domain. This might explain the different outcomes in the studies by Lopez de Maturana *et al.* (2004) and Barwell *et al.* (2011) on ECL1 of GLP1 and CGRP receptors respectively. While only a double Ala mutation of M204 and Y205 significantly affected ligand binding and receptor activation in the former, up to 9 residues affected ligand binding and/or receptor activation in the latter. Moreover, the ECL3 of CLR appears to be highly involved in the activation of the adrenomedulin (AM) 1 and 2 receptors but shows less involvement in the CGRP receptor (Kuwasako *et al.*, 2012; Barwell *et al.*, 2011). This in addition suggests that the ECLs might be a key definer of receptor activation following ligand binding. While the exact mechanism involved in the EC loops' contribution to ligand binding and receptor activation is not particularly known, at least it is certain that they play a crucial role.

#### **1.3.4.4 Intracellular loops (ICLs)**

ICL1 and 2, the membrane associated regions of ICL3 and the proximal segment of the C-terminus has variously been reported to participate in G-protein recognition and binding (Nabhan *et al.*, 1995; Wess, 1997; Conklin and Bourne, 1993; Rasmussen *et al.*, 2011). For instance, alternative splicing in ICL1 of CRF1 and calcitonin receptors have been shown to influence G-protein coupling (Nabhan *et al.*, 1995; Nussenzueig *et al.*, 1994). The highly conserved KL dipeptide, part of the RKLH motif in ICL1 of family B GPCRs, has been reported as crucial for G-protein coupling in human GLP1, VPAC2, secretin and CGRP receptors (Mathi *et al.*, 1997; Hilaiet *et al.*, 2001; Chan *et al.*, 2001; Conner *et al.*, 2006b); although in the CGRP receptor, where the equivalent dipeptide is SL (in CLR), only the L was investigated. Moreover, K167 of the KSLS (equivalent of RKLH) motif in CLR was observed to interact with G<sub>β</sub> in the CLR/G-protein complex model by Vohra *et al.* (2013). Overall, considering earlier discussions involving this loop, the mechanism underlining the role played by ICL1 in receptor activation appears to revolve around its interaction with helix 8 and the

interactions suggested for the invariantly conserved Arg (R173 in CLR) as mentioned in previous sections.

It has been suggested that the movement of ICL2 may create a binding pocket for G-protein binding in CLR (Conner *et al.* 2006). This may be related to the observation made for this loop in the  $\beta_2$ AR-Gsa complex where ICL2 assumes an  $\alpha$ -helical conformation in contrast to the extended loop structure of this loop in the inactive form (Rasmussen *et al.*, 2011); although the authors noted that this might not be a feature peculiar to the active receptor as the former conformation is observed in the inactive avian  $\beta_1$ AR. Moreover, F139 at the beginning of ICL2 helix was observed to sit in a pocket formed by a network of mostly hydrophobic residues in  $\beta_1$ ,  $\beta_3$  and  $\alpha 5$ -helix of G-protein (Rasmussen *et al.* 2011). This residue is said to be conserved in GPCRs that couple to Gs but are variable in other non-Gs-coupling receptors. However, it is not known whether this is extended to other non-family A. So even though CLR for instance, which also couple to Gs, has a Phe close to the mid-region of ICL2, it is not clear whether they play a similar role.

ICL3 is thought to contain the major determinants for specific G-protein coupling in family B GPCRs as in family A (Pisegna *et al.*, 1996; Christopoulos *et al.*, 2003). This is probably owing to its structural association with the TM6, so that it is in a good position to interact with the G-protein. A study using splice variants in the rPACAP receptor has indicated ICL3 as important for PACAP-dependent cAMP stimulation (Pisegna and Wank, 1996). One important motif in this loop is the KxxK motif which is shared by both families A and B (Vohra *et al.*, 2013). The two (first and last) basic residues in this motif have been investigated in human secretin (Chan *et al.*, 2001; Garcia *et al.*, 2012), VPAC1 & 2 (Langer and Robberecht, 2005; Langer *et al.*, 2005), CRF1 (Pun *et al.*, 2012) and rat GLP1 receptors, and all have generated varying results. In CLR where the second basic residue is an Arg (R314 precisely), R314 (but not K311 of the same motif) at the ICL3-TM6 junction disrupts G-protein coupling when mutated to an Ala and has been suggested to stabilize ICL3/TM6 association by possibly interacting with phosphate head groups of membrane phospholipids (Conner *et al.*, 2006b). Whereas in hVPAC2 receptor, individual Ala mutation of these basic residues (i.e. R325A and R328A) both significantly reduced receptor potency and cAMP levels (Langer *et al.*, 2005); only

R341A (but not R338A) significantly reduced IP<sub>3</sub> levels in the closely related hVPAC1. On the other hand, in hCRF1R, both K311A and K314A individually showed significant increase in Gs coupling with coupling to Gq significantly impaired (Pun *et al.*, 2012). These, among other things, suggest that while the second basic residue (of the KxxK motif) appears to be of higher importance probably owing to its localisation at the ICL3-TM6 junction, the importance (and plausibly the role) of both residues vary from receptor to receptor and could be designed for specific interaction with G-protein isoforms.

Overall, the intracellular loops have not been extensively studied across family B GPCRs especially in relation to structure. So while a good amount of evidence exists to show they are important, there is not enough evidence to highlight the specific structural mechanism underlining their roles in receptor signalling.

#### **1.4 Calcitonin family of peptides**

The calcitonin family of peptides majorly comprises calcitonin (CT), calcitonin gene-related peptide (CGRP), adrenomedulin (AM) and Amylin (Amy). While these peptide ligands show low sequence homology, they exhibit more similarity in their structures (Watkins *et al.*, 2012).

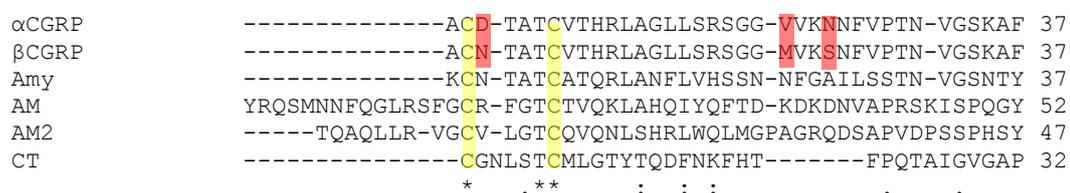
CT is a 32 amino acid peptide and the shortest in length among the family. It is secreted by thyroid C cells. It plays a crucial role in the modulation of calcium homeostasis through the inhibition of osteoclast-mediated bone resorption (Sexton *et al.*, 1999). This peptide is found in humans and many other vertebrates. It is interesting to note that the salmon CT is a more potent agonist of the human calcitonin receptor (hCTR) than the human CT (Dong *et al.*, 2004). Owing to its prominent role in calcium homeostasis and bone formation, this peptide is a well explored therapeutic target. The drug Miacalcin, for instance, is an sCT and it used in the treatment of bone disorders like Paget's disease.

AM is a 52 amino acid peptide and the longest within the family (Kitamura *et al.*, 1993). It is located on chromosome 11 (Hinson *et al.*, 2000). It has a homologue, AM2 (47 amino acid long), in human

(Roh *et al.*, 2004). In fish, 5 forms are present namely; AM1, AM2, AM3, AM4 and AM5 (Ogoshi *et al.*, 2006). AM is a potent vasodilator and have been suggested to be involved in the process of homeostasis and regulation of blood pressure (Nishio *et al.*, 1997). It has also recently been suggested to play a role in cancer as high levels have been observed in many cancer cells (see Hay *et al.*, 2011).

Amy is a 37 amino acid peptide secreted by  $\beta$ -cells of the pancreas usually together with insulin. It was first isolated in the pancreas of patients with type 2 diabetes (Cooper *et al.*, 1987). It mainly helps control feeding by inhibiting gastric emptying, gastric acid secretion and post-prandial glucagon secretion (Höppener *et al.*, 2000).

CGRP is also a 37 amino acid peptide and is a potent vasodilator that acts near its site of release. It mediates effects mainly through a complex of the calcitonin receptor-like receptor (CLR) and the receptor activity modifying protein 1 (RAMP1) (Poyner *et al.*, 2002, Taylor *et al.*, 2006). CGRP is further discussed in section 1.4.1 below.



**Fig 1.10:** Sequence alignment of human calcitonin family peptides. Invariantly conserved residues are marked with asterisk (\*), residues with similar side-chain properties are marked with colon (:), while residues with similar shape are marked with dot (.). Disulphide-forming cysteine residues are highlighted yellow while residues distinguishing  $\alpha$ CGRP from  $\beta$ CGRP are highlighted in red. Alignment was conducted using ClustalW2. The number of residues for each peptide is indicated at the end of the respective sequence.

The receptors for the calcitonin family of peptides come from the secretin family of GPCRs. These are the calcitonin receptor (CTR) or calcitonin receptor like receptor (CLR). These receptors function either as a monomer (in the case of CTR) or in complex with the RAMPs (CTR and CLR). The pharmacological profiles of these receptors are presented in Table 1.3.

**Table 1.3:** Pharmacological profile of receptors for the human calcitonin family of peptides. (Hay *et al.*, 2006)



## 1.4.1 CGRP

### 1.4.1.1 Tissue distribution and Physiological role

CGRP is a 37 amino acid peptide with its gene located on chromosome 11. It comprises two homologues ( $\alpha$  and  $\beta$ ) which differ by three amino acids in humans (Fig 1.10) and they share a much pharmacological similarity (Barwell *et al.*, 2010; Moore *et al.*, 2010). It is produced as a result of differential splicing of RNA transcripts from the CT gene (Amara *et al.*, 1982). This splicing is tissue-specific and therefore plays a significant role in distribution of CGRP, especially  $\alpha$ CGRP. For instance, while the CT mRNA is predominant in the thyroid, that of  $\alpha$ CGRP predominates the hypothalamus (Lou and Gagel, 1998). This peptide is widely distributed throughout the nervous system and in the cardiovascular system.

CGRP has been suggested to be actively involved in the control of blood flow. In knockout models where  $\alpha$ CGRP or both  $\alpha$ CGRP and CT were deleted, an increase in blood pressure was observed (Kurihara *et al.*, 2003; Li *et al.*, 2004). Its over-secretion therefore causes facial flushing, oedema and inflammation. Plasma high levels of CGRP have been associated with vascular diseases such as Raynaud's disease (characterized by the discolouration of the fingers, toes and occasionally other

areas) as a result of vasoconstriction in the extremities. It has also been suggested to play a role in bone formation (Wedemeyer *et al.*, 2007). Increased levels of CGRP has also been observed in migraine attack (Nichols *et al.*, 2010), a chronic neurovascular disorder that affects about 12% of the general population (Durham, 2004). Evidence supporting the involvement of this peptide in migraine was reported by Lassen *et al.* (2008) where the intravenous (IV) infusion of  $\alpha$ CGRP caused migraine in patients (Lassen *et al.*, 2008). Targeting drugs at the CGRP receptor to curtail CGRP's excesses has therefore been suggested as effective therapy for migraine (Durham, 2004; Zhang *et al.*, 2007; Nichols *et al.*, 2010).

A recent review by Hay *et al.* (2011) has presented evidence supporting a link between CGRP and cancer progression, but not initiation. Several studies have observed an increased level of CGRP in both plasma and tumours from certain cancers (see Hay *et al.*, 2011). Although this requires further investigations, there are already suggestive opinions making a case for CGRP receptor antagonists (especially those that simultaneously block the  $AM_1$  and  $AM_2$  receptors) in cancer therapy.

Antagonists binding to the CGRP receptor have been developed with the aim of treating disease conditions, especially migraine. The first CGRP antagonist was the CGRP<sub>8-37</sub>, i.e. CGRP devoid of the first 7 amino acid residues (Chiba *et al.*, 1989). This truncated peptide binds CGRP with high affinity but unfortunately, it was of no pharmacological use due to its short half-life *in vivo*. It has however been useful in the characterization of the CGRP receptor. Doods *et al.* (2000) developed a high affinity and highly selective antagonist, BIBN4096BS with about 150x higher affinity than CGRP<sub>8-37</sub>. This compound was efficacious in the treatment of acute migraine (Olesen *et al.*, 2004). Moreover, an orally active antagonist, MK-074 has also been developed for this receptor for the treatment of migraine (Salvatore *et al.*, 2007) and this made it to phase 3 clinical trials.

#### **1.4.1.2 Structure of CGRP**

Although no crystal/NMR structure has been deposited for CGRP in the protein data bank, its structure has been investigated by NMR and molecular modelling techniques (Lynch and Kaiser,

1988; Breeze *et al.*, 1991; Conner *et al.*, 2002). The first of four domains of the CGRP structure comprises the first 7 amino acid residues, the absence of which produces an antagonist (CGRP<sub>8-37</sub>). A disulphide bridge between residues 2 and 7 is crucial for activation of the receptor by this peptide (Conner *et al.*, 2002; Barwell *et al.*, 2010). Affinity of this peptide is conferred in part by the second domain made up of an amphipathic  $\alpha$ -helix between 8<sup>th</sup> and 18<sup>th</sup> residues. This characteristic is specifically driven by Arg residues on positions 11 and 18 sitting on the hydrophilic face of the helix. Residues 19-27 make up the third domain beginning with a  $\beta$  or  $\gamma$  turn and act as a hinge with no stringent constraint on its composition. The last ten residues make up the C-terminal domain with two turn regions (centred on Pro-28 and Gly-33). This is thought to be crucial in high-affinity binding (Conner *et al.*, 2002). A more recent review of this peptide has given more insight into the roles played by various residues making the different regions of the peptide (Watkins *et al.*, 2012).



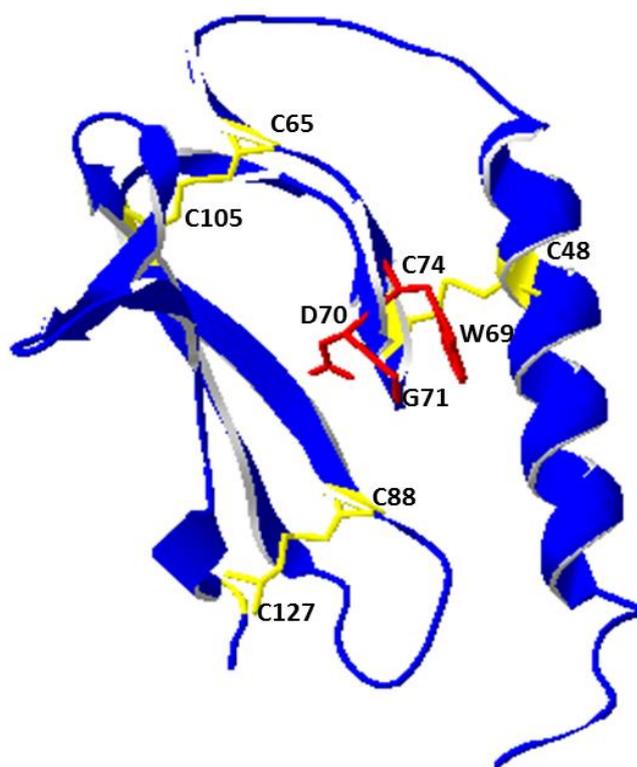
**Fig 1.11:** Structure of CGRP showing the distribution of structural features across the peptide sequence. The residues implicated in making receptor contact are shaded grey (Conner *et al.*, 2002; Watkins *et al.*, 2012)

## 1.5 The CGRP receptor complex

### 1.5.1 CLR

The calcitonin receptor-like receptor (CLR) is a 7 transmembrane (TM) receptor protein of the family B which is approximately 460 amino acids in length. Earlier reports (e.g. Conner *et al.*, 2007) on the presumptive structure of the CLR N-terminal region were made based on the published NMR

structure of the mouse corticotropin-releasing factor receptor type 2 (CRF-R2 $\beta$ ) N-terminal extracellular domain (Grace *et al.*, 2004) – a prototype of the family B GPCRs. An accurate structure became possible following the release of the crystal structure of the ECD of the CGRP receptor (ter Haar *et al.*, 2010). The structure reveals an N-terminal  $\alpha$ -helix (spanning residues Gly35 to Met53) that packs against a core of two antiparallel  $\beta$ -sheets. The length of the helix is similar to those of PTH and GIP receptors. The helix is followed by a long loop of irregular structure between residues Asp55 and Tyr64, leading to a long ‘finger-like motif’ from Cys65 to Gly81 which is highly conserved in all ectodomain structures. The N-terminus is stabilised by three disulphide bonds (Cys48 – Cys74; Cys65 – Cys105; and Cys88 – Cys127), each performing distinctive stabilising function. The configuration observed by the two  $\beta$  sheets, linked by the second disulphide bond, resembles the ‘short consensus repeats’ (SCR) revealed in the CRF-R2 $\beta$  N-terminus (Parthier *et al.*, 2009) between residues 39 and 133 (Grace *et al.*, 2004). Sequence homology has clearly revealed that residues involved in intramolecular interactions that contribute to structural stability, in addition to the disulphide bond system, are conserved in many other family members. For instance, Trp75 and Trp111, the aromatic indol ring systems of which sandwich the basic residue Arg-103, are highly conserved (ter Haar *et al.*, 2010).



**Fig 1.12:** Crystal structure of CLR ECD (ter Haar *et al.*, 2010) showing the architecture of the common secretin family fold. Disulphide bridges stabilizing the ECD are shown as yellow sticks. Residues making the WDG motif are shown in red.

There have been molecular structural models proposed for the TM bundle of this receptor (Vohra *et al.*, 2007; 2013; Conner *et al.*, 2005). CLR, like other family B GPCRs, is believed to exhibit the conventional 7 TM helical fold already shown for family A GPCRs. Existing model structures of CLR suggest that this receptor also has a kink within TM6, caused by P343, as commonly observed in family A GPCR structures, although it is thought to be roughly 2 turns above the position of the equivalent Pro residue (P<sup>6.50</sup>) in family A (Vohra *et al.*, 2013). Mutagenesis studies have shown that this residue is important for receptor activation and the kink introduced by it may drive structural activation as in family A GPCRs (reviewed in Barwell *et al.*, 2013). The most recent of the CLR model structures, Vohra *et al.* (2013), was based on structural homology using an intermediate GPCR, GCR1 – the most studied plant GPCR which shares homology with both families A and B GPCRs. In this modelling, key molecular signatures of the CLR (as a representative family B GPCR) were

associated with family A (Table 2, Vohra *et al.*, 2013). While some of these motifs appear to play similar roles to their family A counterparts, the same does not look plausible for others. (see Vohra *et al.*, 2013).

A putative 8<sup>th</sup> helix, anchored via a palmitoylation motif to the membrane in rhodopsin, has also been suggested for family B GPCRs but this currently remains postulative (Conner *et al.*, 2008; Vohra *et al.*, 2013). Although the palmitoylation motif is said to be absent in the CLR, a biophysical study (Conner *et al.*, 2008) on a synthetic CLR C-terminus reported that there is a possibility for the formation of a lipid anchor on this region. This has been attributed to Trp399 of the CLR at the distal end of H8. The EFxxxL<sup>8.54</sup> motif in H8 of family A is also believed to be exhibited by CLR where it is present as EVxxxL<sup>8.54</sup>.

## 1.5.2 RAMP1

### 1.5.2.1 Physiological and pharmacological roles

RAMP1 is the first member of a 3-member small (148 – 175 amino acids) family of proteins that serve as accessory proteins for some family B GPCRs. They were first identified as partners of the CLR (McLathie *et al.*, 1998). Each member possesses a single transmembrane  $\alpha$ -helix, an extracellular amino terminus, and a short intracellular C-terminus. They have different tissue distribution (Hay *et al.*, 2006) and share less than 30% sequence identity with one another (Kusano *et al.*, 2008). The human RAMP1 gene is present on chromosome 2 (CLR is also present on chromosome 2 but not particularly close to RAMP1) at 2q36 (Poyner *et al.*, 2002). RAMP1, together with RAMP2 and 3, are relatively ubiquitously distributed with at least 1 RAMP expressed in every tissue (Sexton *et al.*, 2001). RAMP1 is specifically found in many body tissues including brain, uterus and pancreas, among others (see Hay *et al.*, 2006).

RAMP1 associates with the CLR to form the receptor for CGRP in a 1:1 stoichiometry (Hay *et al.*, 2006; Kusano *et al.*, 2008; ter Haar *et al.*, 2010). This stoichiometric ratio is however under debate as

there are reports suggesting a monomer of RAMP1 interacting with a CLR dimer (Heroux *et al.*, 2007). A clearer insight into the structure of RAMP1 and the stoichiometry of its interaction with CLR was boosted by the emergence of the crystal structure of the human RAMP1 ECD I association with CLR ECD (ter Haar *et al.*, 2010). The most notable role of RAMP1 comes from its relation with CLR and this involves its (RAMP1) role in regulating the transport (i.e. acting as a chaperone) and pharmacological phenotype of CLR (McLathie *et al.*, 1998). RAMP1 is required for the translocation of CLR from the endoplasmic reticulum (ER) to the cell surface as CLR expressed in the absence of RAMP1 (and other RAMPs) is retained intracellular (McLathie *et al.*, 1998). The N-terminus appears to be most crucial in CLR/RAMP1 interaction. The importance of this domain has been reported in several studies (Fraser *et al.*, 1999; Udawela *et al.*, 2006b). Studies involving the use of chimeras have suggested the TM and C-terminus to be less crucial for receptor function (Udawela *et al.*, 2006a), but the TM is important for the formation of functional receptor at the cell surface (Fitzsimmons *et al.*, 2003). The C-terminus of RAMP1 contains an ER retention motif (QSKRT) adjacent to the plasma membrane (Steiner *et al.*, 2002). This probably plays a role in the overall regulation of receptor function as its deletion causes RAMP1 to be translocated to the cell surface, although this was only tested in COS 7 cells.

RAMP1 tends to modulate the pharmacology of CLR directly/indirectly altering its (CLR) structure or by contributing to an interface for binding agonists or antagonists. Mutagenesis studies and information from the crystal structures of the ECD of RAMPs 1 and 2 (in complex with CLR ECD) have shown that while two or more residues may show similar effect on receptor pharmacology, their mechanism may differ. For instance, while F93A and F101A both reduced CGRP receptor cell surface expression in an earlier mutagenesis study (Kuwasako *et al.*, 2003), a later structural study showed that they are unlikely to act via the same mechanism (Kusano *et al.*, 2008). While F101 is believed to contribute to the CLR binding interface, it was unlikely that F93 does the same (Kuwasako *et al.*, 2003; Kusano *et al.*, 2008). The contribution of RAMP1 to an interface for binding agonists and antagonists in the CGRP receptor is a feature of this protein in CGRP receptor pharmacology. One very important residue that has been implicated in this respect is W84, mutation

of which significantly reduced CGRP potency and to a lesser extent CGRP receptor cell surface expression (Moore *et al.*, 2010). This residue has been indicated in peptide and non-peptide binding (Moore *et al.*, 2010).

### 1.5.2.2 Structure of RAMP1

The structure of RAMP1 was first predicted by an *ab initio* model designed by Simms *et al.* (2006). The secondary structure was predicted from two prediction routines. Their model predicted RAMP1 to be composed of an N-terminal region of three alpha helices (helices 1, 2 and 3) and a relatively short TM domain spanning residues 118 – 139. The disulphide forming cysteine residues were tested using site directed mutagenesis (Simms *et al.*, 2006).

Today, a crystal structure of RAMP1 ECD exists in the protein data bank in its monomeric state (PDB code: 2YX8) as well as in complex with the CLR ECD (PDB code: 3N7S). The overall architecture of this structure agrees with that already predicted by Simms *et al.* (2006). The ECD RAMP1 is a three-helix bundle –  $\alpha 1$ ,  $\alpha 2$  and  $\alpha 3$  - stabilized by three disulphide bonds (Cys40 – Cys72; Cys27 – Cys82; Cys57 – Cys104). There is also a short helical structure between the  $\alpha 1$  and  $\alpha 2$  helices. The  $\alpha 2$  helix is aligned in an antiparallel position to the  $\alpha 1$  and  $\alpha 3$ , and the  $\alpha 1$  is has a kink on Leu-39 (Kunaso *et al.*, 2008). This released structure has appreciable similarity with that earlier predicted for an *ab-initio* model of RAMP1 (Simms *et al.*, 2006) especially in the helical composition. However, there are clear significant differences especially in the predicted residues marking the helices and those involved in certain intramolecular interactions. This thus emphasizes an importance of obtaining crystal structures for these proteins. The biochemical roles of RAMP1 can be summarized as cell-surface targeting, direct ligand binding and indirect modulation of the CLR conformation (Mallee *et al.*, 2002).



**Fig 1.13:** Ribbon representation of RAMP1 structure. The three disulphide bonds are shown in yellow.  $\alpha 1$ ,  $\alpha 2$  and  $\alpha 3$  are respectively the first, second and third helices and are shown in red. *Protein Sci.* 17(11):1907-14

### 1.5.2.3 RAMP1 analogues: RAMP2 and RAMP3

RAMP2 and RAMP3, the two homologues of RAMP1, are present on chromosome 17 at 17q12-21.2 and chromosome 7 at 7p13-12 respectively in human (Derst *et al.*, 2000). The associations of RAMP2 and RAMP3 with the CLR form receptors for adrenomedullin (AM) with distinct affinities (Table 1.3) (McLatchie *et al.*, 1998; Poyner *et al.*, 2002). Despite their low sequence identity, sequence alignment shows that key residues involved in intramolecular hydrophobic interaction and disulphide linkages in the RAMP1 are also contained in the RAMP2 and RAMP3 isoforms (Fig 1.15). One main exception is the absence of the disulphide bond corresponding to Cys27 – Cys82 in RAMP2, leaving it with two disulphide bonds compared to three observed for RAMP1 and predicted RAMP3. The disulphide bonds in RAMP2 however have been reported to be more crucial in receptor cell surface expression compared to the other two RAMPs (Kuwasako *et al.*, 2003). Unlike in RAMP1, RAMP2 and RAMP3 possess N-glycosylation sites which, for instance in the case of RAMP2, has been implicated as influential in receptor trafficking (Flahaut *et al.*, 2002). While RAMP3 has four potential N-





**Fig 1.15:** Crystal structure of the ECD of RAMP2 (Kusano *et al.*, 2012). The disulphide bridges stabilizing this domain are shown in yellow. Alpha helices 1, 2 and 3 are labelled  $\alpha 1$ ,  $\alpha 2$  and  $\alpha 3$  respectively.

RAMP3 appears to be very closely related to RAMP1 especially from sequence homology (see Fig 1.14) and possibly in functional mechanism. The RAMP3 equivalent of the highly important W74 in RAMP1 (E74 in RAMP3), for instance, has been mutated have been found to play a similarly crucial role as RAMP1 (Hay *et al.*, 2006b). The little information available for the RAMP3 structure is extrapolated from RAMP1 and is therefore predictive. Bailey *et al* (2010) constructed a homology model for RAMP3 from its closest associate, RAMP1. The structure is predicted to be stabilized by disulphide bonds formed by Cys residues equivalent to the disulphide-forming Cys residues in RAMP1. Also in this structure, helix 1 is slightly unwound and the C-terminus is orientated in a slightly different manner, as a result of the presence of a kink, when compared to RAMP1.

### 1.5.3 RCP

Another addition to the complexity of the CGRP system is the receptor component protein (RCP) discovered during expression systems targeting the CGRP receptor. It is a 148-amino acid cytoplasmic protein found to co-immunoprecipitate with CLR in cell culture and tissues. It has been found to be involved in the activation of the CGRP receptor (Tolun *et al.*, 2007). An earlier study has strongly suggested that direct interaction between RCP and CLR is required for CGRP receptor activation and signalling. In the study, signalling of endogenous CLR was inhibited in cells from which RCP has been co-immunoprecipitated with an interacting intracellular CLR domain expressed as soluble fusion protein. CLR trafficking was however not inhibited, suggesting that RCP may not play a chaperone role for CLR (reviewed in Egea and Dickerson, 2012).



**Fig 1.16:** Representative model for the CGRP system. AC is adenylate cyclase. *Protein Expr Purif* 52: 167-174

### 1.6 Expression Systems for Recombinant Protein Production

The use of proteins for scientific research requires that they are present in considerably large amount. This is paramount when they are required for most biophysical studies. Isolating proteins from their native source, or their synthetic production, is cost intensive and sometimes raises ethical issues, especially when a human source is required. The emergence of recombinant DNA technology, in the early 1970s (Jackson *et al.*, 1972), has been a very useful tool in overcoming this hurdle. The

recombinant DNA technology involves the splicing of relatively short DNA molecule which can be replicated in a host cell. The basic steps in recombinant DNA technology, as reported by Reddi *et al.* (2000), are summarized as;

- Isolation and specific cleavage of DNA
- Ligation of DNA fragment into a cloning vehicle, called *a vector*
- Transformation and selection
- Physical mapping and DNA sequencing to confirm cloned gene
- Expression of the cloned gene.

Since proteins are made from DNA, recombinant proteins are therefore proteins produced from recombinant DNA. The last of these steps, the gene expression, requires a host cell into which the cloned DNA-harbouring vector is introduced and the protein purified.

A number of host cells are used to obtain recombinant proteins. This could either be a prokaryote or eukaryote. The most commonly used are the *Escherichia coli* (*E. coli* - a prokaryote), yeast, insect and mammalian cells (all eukaryotes). They have all been reported in the isolation of various GPCRs (Fraser, 2006). These systems differ in different ways with each having their advantages and disadvantages. The *E. coli* and yeast expression systems both have the advantages of rapid cell growth, minimal complexity and cost effectiveness. These are not attributed to the insect and mammalian cell lines. However, the insect and mammalian cell lines have a special advantage of posttranslational modification (e.g. folding of the protein) - a feature absent in the *E. coli* cells but present in the yeast cells (Higgins and Cregg, 1998). These features form the major criteria used to descriptively classify these expression systems. Of these systems, the yeast expression system is most favourable as it shares the desirable features of the prokaryotic and eukaryotic expression system.

The yeast expression system consists of two major cell types – *Saccharomyces cerevisiae* and *Pichia pastoris*. *S. cerevisiae* was the first to be discovered and has been widely studied with its entire genome sequenced (Higgins and Gregg, 1998). *P. pastoris*, though less characterized compared to the former, has out-shone the *S. cerevisiae* in its use for recombinant protein production. One main

advantage it has over the *S. cerevisiae* is its preference for respiratory growth which facilitates its culturing resulting in higher cell density compared with the fermentative *S. cerevisiae* (Higgins and Cregg, 1998; Burrowes *et al.*, 2005). The *P. pastoris* expression system, employed in this study, will further be discussed in some details.

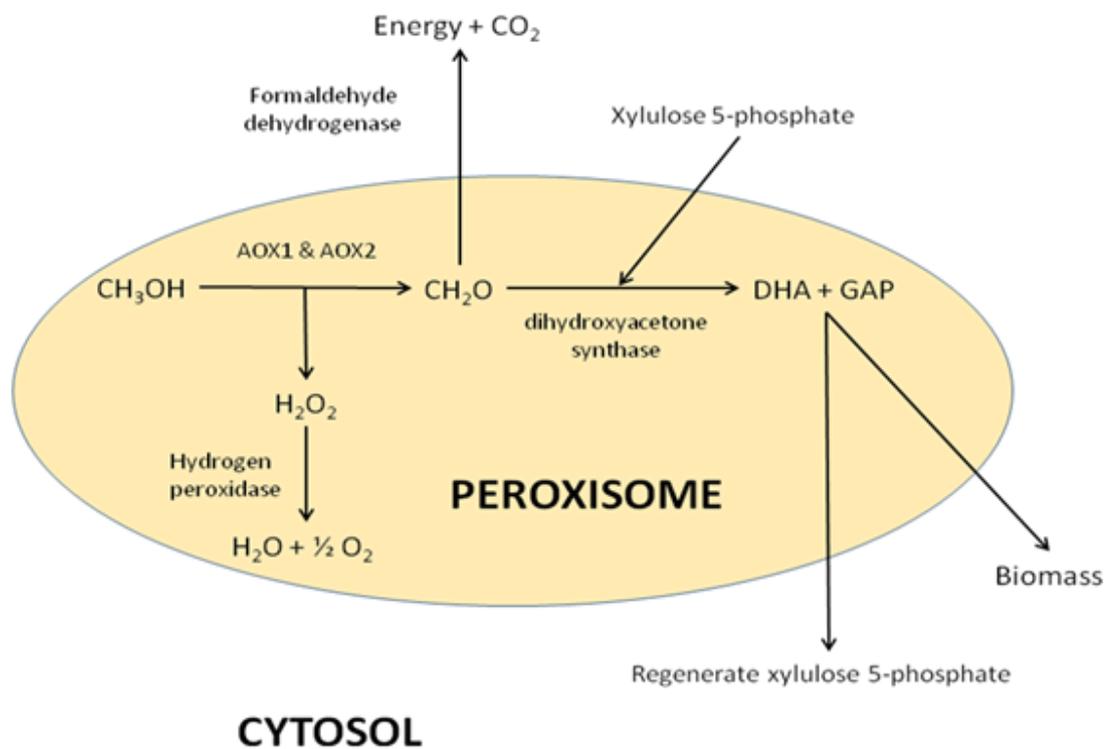
**Table 1.4:** Comparison of expression systems for recombinant protein production. Adapted from GenWay Biotech Inc. (2013).

Characteristics	<i>E. coli</i>	Yeats	Insect cell	Mammalian cell
Cell growth	Fast	Fast	Slow	Slow
Level of expression	High	Low – High	Low – High	Low – Moderate
Yield (mg/L culture)	50 – 500	10 – 200	10 – 200	0.1 – 100
Extracellular expression	Secretion to Periplasm	Secretion to medium	Secretion to medium	Secretion to medium
Protein folding	Folding usually required	Folding may be required	Proper folding	Proper folding
N-linked glycosylation	None	High mannose	Simple, No Sialic acid	Complex
O-linked glycosylation	No	Yes	Yes	Yes
Phosphorylation, Acetylation and Acylation	No	Yes	Yes	Yes
Cost	Low	Low	Mid-way	High

### 1.6.1 The *Pichia pastoris* Expression system

*Pichia pastoris* is an ascomycetous budding yeast that most commonly exist in a vegetative state (Higgins and Cregg, 1998). It is methylotropic with the ability to utilize methanol as a sole carbon source. The methanol is usually required in low concentration to prevent toxicity to the cells. Although methanol metabolism (illustrated in Fig 1.17) by this organism produces hydrogen peroxide (a toxic compound) as by-product, this is detoxified to water and oxygen by endogenous hydrogen peroxidase. This happens within the peroxosome – another feature that protects the organism from the reach of this toxic compound (Burrowes *et al.*, 2005). One striking feature that makes *P. pastoris* a

very productive expression system is the presence of the alcohol oxidase 1 (*AOX 1*) gene that encodes the alcohol oxidase enzyme required for the utilization of methanol for growth and also for the overexpression of heterologous genes introduced downstream in a *Pichia* expression vector. This allows for up to 100-fold recombinant protein production in this organism compared to the traditionally used *Saccharomyces* (Higgins and Cregg, 1998; Burrowes *et al.*, 2005).



**Fig 1.17:** Methanol metabolism in *P. pastoris*. Adapted from Lin-Cereghino and Cregg (2000). AOX, alcohol oxidase; DHA, dihydroxyacetone; GAP, glyceraldehyde 3-phosphate.

Moreover, certain features give *Pichia* an edge over other expression systems in its use for the production of GPCRs. These are its characteristic easy manipulation and relative cost effectiveness, in addition to its ability to glycosylate the protein produced (Singh *et al.*, 2008). An addition to this is its adaptability to large-scale culture in bioreactors (Singh *et al.*, 2008) and this is much more productive

compared to the conventional shake flask system (Burrowes *et al.*, 2005). Bioreactors allow a defined regulation of growth parameters, mainly air, pH and carbon source, and this gives room for growth to an ultra-high density therefore maximising recombinant protein production (Singh *et al.*, 2008). It is important to note that this organism also allow for the secretion of proteins into the growth media, a feature which favours its choice for this research work. This property is conferred by the presence of a *S. cerevisiae*  $\alpha$ -factor signal peptide gene in some of its vectors. This is coupled with the fact that this organism secretes a rather negligible amount of endogenous protein into the culture media, making the secreted recombinant protein the vast majority of total protein in the media (Aoki *et al.*, 2003). The soluble recombinant protein(s) secreted in this way do not require solubilisation by detergents, making it more suitable for biophysical characterization.

The *Pichia* vector used in cloning the receptor genes of interest is the pPIC9K-MepNet modified by Andre *et al.* (2006) from the invitrogen pPIC9K vector. The expression cassette is made up of an N-terminal  $\alpha$ -factor signal sequence followed by a Flag-tag and a decahistidine-tag. Two tobacco etch virus (TEV) protease sites flank the multiple cloning site and a biotinylation-tag from *Propionibacterium shermanii* follows the second TEV site at the C-terminal end of the protein to be expressed (Andre *et al.*, 2006). This vector, specifically designed by this group in the Membrane Protein Network (MepNet) to increase GPCR production in *Pichia*, has been reported to have added advantage over the original vector (Andre *et al.*, 2006; Zeder-Lutz *et al.*, 2006). The host strains employed for the expression of these receptor proteins are the wild type X33, and the mutant strains, GS115 and SMD1168. The GS115, like the wild type, has a methanol utilization plus (Mut<sup>+</sup>) phenotype but has a mutation in the *his4* (histidinol dehydrogenase) gene making transformants selective on a minimal (histidine-deficient) media plate. This is an added selectivity to the antibiotic resistance. The SMD1168 strain (*his4 pep4*) has both features described for GS115 but in addition is deficient in the proteinase A encoded by the *pep4* gene. Although this strain has the proteinase B, the latter requires proteinase A for activation. The proteinase B has half of the activated proteinase activity prior to activation, this is however considerably reduced (Higgins and Cregg, 1998). These

features described form the basis for the rationale for choosing these strains, from which the best expression strain would be chosen following screening.

## 1.7 Mutagenesis

The technique of site-directed mutagenesis (a form of mutagenesis) was first invented by Michael Smith in 1978 (Hutchison *et al.*, 1978). It is now widely employed, especially in pharmacology, for molecular receptor-ligand interaction mechanisms, which in turn provides a good knowledge of designing drugs for these receptors. In a simpler term, this technique helps determine the role of amino acid residues in the binding of ligands and activation of the receptor. In a broader perspective, it is also employed in introducing a desired residue(s) into a peptide/protein for several intentions. The most common of this is alanine mutagenesis. In principle, it involves making a point-substitution mutation by substituting an amino acid moiety with alanine (for alanine mutagenesis) at a specific site on known sequence (Hulme *et al.*, 1999). The effects of such mutation(s) can be interpreted based on four basic outcomes and these include changes in expression, basal activity, agonist affinity binding and signalling efficacy (see Hulme *et al.*, 1999). This mutation is usually incorporated by designing an oligonucleotide primer (containing the new nucleotide base around the centre of the primer) complementary to the sequence base pairs around the site of mutation. The primer is then used to synthesize a new plasmid vector (from a template vector harbouring the gene of interest) which now contains the mutated gene. The template gene is eliminated by restriction enzyme that targets methylated sites on DNA, which is not found in the newly synthesized DNA. This is transformed into *E. coli* competent cells to repair nicks and isolated for DNA sequencing to confirm the mutation (Campbell and Farrell, 2006). The mutated construct can now be transfected into the host cell for direct functional studies of the protein or could be isolated for onward biophysical studies.

In this study, the mutagenesis technique would mainly be employed in the *in vivo* study of the activation mechanism of the CGRP receptor using parameters such as cAMP production and cell surface expression levels. The host cell lines to be employed would be the mammalian cell line COS-

7, derived from kidney cells of African green monkey. The mammalian cells are used in this wise because of the evolutionary differences and distribution of the G-proteins required in GPCR signalling among organisms.

## **1.8 Aims and Objectives**

Understanding the molecular mechanism/basis of the CGRP receptor-ligand interaction, hence, designing exogenous molecule targeting this receptor for therapeutic purpose has been hampered by the absence of crystal structures for this receptor and accompanying accessory proteins. This research work is aimed at using mutagenesis to investigate several residues at the N- and C-termini of the CLR for their role in receptor signalling. The former, involving the N-terminus, is in the quest to determine and investigate certain residues within a putative ligand-binding core of CLR that are crucial for CGRP binding and hence, receptor activation. The sites have been predicted following information from the crystal structures of CLR and some family B GPCR ectodomains - the CRFR, PTHR and GIPR. The latter involves using alanine scanning mutagenesis to investigate several residues spanning H8 of CLR as well as the associated C-terminal region.

This research project is also aimed at expressing, purifying and characterizing soluble ectodomains of CLR and the RAMPs (i.e. RAMP1, RAMP2 and RAMP3). The purpose of this is to produce receptor proteins with the plausibility of studying their interactions (i.e. CLR/RAMP interaction) using surface plasmon resonance (SPR). The ultimate goal of this is to develop a system that could be a tool in determining novel partners for RAMP proteins.

## **Chapter 2      Materials and Methods**

### **2.1      Generation and analysis of CLR site-directed mutants**

#### **2.1.1    Materials**

##### **2.1.1.1    Equipment**

Invitrogen Countess<sup>TM</sup> Automated Cell Counter (Life Technologies, UK)

Microflow Advanced BioSafety Cabinet Family B

Thermo Scientific Multiskan GO Plate Reader (Thermo Scientific, UK)

Sigma 2-6E bench centrifuge (Sigma-Aldrich, UK)

##### **2.1.1.2    Media and stock solution**

###### **2.1.1.2.1      Cell culture media**

500 ml DMEM 4.5g/L Glucose with L-Glutamine (Lonza, UK) was supplemented with 50 ml heat-inactivated fetal calf serum (PAA, UK) and 5 ml penicillin/streptomycin. The media solution was stored at 4°C. The fetal calf serum was stored as 50 ml aliquots in 50 ml universal tubes at -20°C.

###### **2.1.1.2.2      cAMP assay media**

40 µl of 500 mM IBMX in DMSO and 20 mg BSA were added to 20 ml serum-free DMEM or 1x HBSS (Hank's Balanced Salt Solution; Life Technologies, UK) and mixed properly. This was prepared freshly before each assay. This volume is designed for one 48 well plate. Stock solution of 500 mM IBMX in DMSO was made by dissolving 100 mg IBMX in 900 µl DMSO. This was stored as 80 µl aliquots in 0.5 ml Eppendorf at -20°C.

#### **2.1.1.2.3 cAMP assay buffer**

This was composed of 20 mM HEPES pH 7.5 and 5 mM EDTA. To prepare, 2.38 g HEPES acid (or 2.6 g HEPES Na salt) and 0.93 disodium EDTA were dissolved in 500 ml ddH<sub>2</sub>O and the pH adjusted to 7.5 with NaOH (or HCl). This was stored at 4°C.

#### **2.1.1.2.4 Binding protein**

This was composed of 0.02% w/v cAMP-dependent protein kinase A (PKA) in 1 mM sodium citrate pH 6.5 with 2 mM dithiothreitol (DTT). To make, 10.5 mg citric acid and 15.42 mg DTT were dissolved in 50 ml ddH<sub>2</sub>O. This was allowed to cool on ice after which 10 mg 3'5'-cyclic AMP-dependent protein kinase (Sigma-Aldrich, UK) was added. This was stored as 2.5 ml aliquots at -20°C.

#### **2.1.1.2.5 Activated charcoal**

The activated charcoal stock solution was made up of 5% w/v activated charcoal and 0.2% w/v BSA in cAMP assay buffer. The solution was prepared by dissolving 400mg BSA in 200 ml assay buffer with gentle stirring. 10 g activated charcoal (100-400 mesh) was slowly added and allowed to stir a little longer. The solution was kept at 4°C overnight to allow for equilibration before the first use.

### **2.1.1.3 Reagents**

#### **2.1.1.3.1 Human $\alpha$ CGRP**

264  $\mu$ l of 1 mM acetic acid was directly added to 1 mg human  $\alpha$ CGRP (Merk4Biosciences, UK) in the original container to give a 1 mM stock concentration. This was stored as 5  $\mu$ l aliquots at -20°C.

### **2.1.1.3.2 Antibodies**

The antibodies used in this section were employed to probe for the HA tag on the CLR. The primary antibody used is mouse anti-HA antibody (Sigma-Aldrich, UK) while the secondary is anti-mouse horseradish peroxidase conjugated antibody (Cell signalling Technology, UK).

### **2.1.1.3.3 Trypsin EDTA**

Trypsin EDTA was employed to dissociate the adherent COS 7 cells from their culture vessel surface (trypsinization). This is required for passaging or plating the cells. Trypsin EDTA was purchased as 100 ml stock solution from PAA, UK. These were stored in original container at 20°C. Once thawed however and in constant use, the working solution was stored at 4°C.

## **2.1.2 Methods**

The methods employed here for generation and analysis of the mutants have been previously published (e.g. Conner *et al.*, 2006; Barwell *et al.*, 2011).

### **2.1.2.1 Identifying targeted sites for mutation**

Various residues of the CLR were selected for investigation by site-directed mutagenesis. The residues were selected on the presumption that they play a role in receptor function. This was adjudged based on the CLR N-terminal and transmembrane structures using the Swiss-PDB viewer. The regions from which the residues were selected included the ligand-binding N-terminal region, the intracellular loop 1 (ICL1) and its associated transmembrane region as well as the helix 8 (H8) and its associated C-terminal region. The regions were defined as specified on UniProt/Swiss-Prot and were compared to the literature. Sequence alignments were performed using ClustalW<sup>TM</sup> alignment tool.

### **2.1.2.2 Expression constructs**

The human CLR and RAMP1 constructs were kindly provided by Dr James Barwell (LHS, Aston University). The CLR cDNA was in pcDNA3.1- mammalian vector with a T8 signal peptide and an

N-terminal heamagglutinin (H8) tag. The RAMP1 cDNA was also incorporated pcDNA3.1- vector with a CD33 signal peptide and N-terminal *myc* epitope tag. The introduction of the tags has been shown not to affect the pharmacology of the receptor (McLatchie *et al.*, 1998, Fraser *et al.*, 1999). The translated sequence of the T8-HA CLR and CD33-*myc* RAMP1 used for all site-directed mutagenesis investigations are given in the appendix section.

### **2.1.2.3 Primer design**

Mutagenesis primers were designed with the aid of a primer designing tool, PrimerX™. The forward and reverse oligonucleotide primers were designed to incorporate an alanine (or another amino acid as desired) as a point mutation in to the wild type receptor. Primers were synthesized desalted by Life Technologies, UK.

### **2.1.2.4 Site-directed mutagenesis**

Site directed mutagenesis was carried out on selected regions of the full length CLR subcloned in the mammalian vector pcDNA3.1(-). Mutations were generated using the QuikChange II Site-Directed mutagenesis kit (Agilent Technologies, UK) according to manufacturer's instructions.

### **2.1.2.5 Mutant DNA sequencing**

Mutant plasmids were sequenced at the Functional Genomics Laboratory (University of Birmingham, UK) to confirm mutation. Sequencing reaction was set up according to the laboratory's recommendation. The T7, TM2, TM4 and BGH oligonucleotide primers (see appendix section for details) for the pcDNA3.1(-) plasmid were used in sequencing the CLR-coding region of the plasmid construct.

### **2.1.2.6 Cell culture and transfection**

COS 7 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% (v/v) fetal calf serum and 1% (v/v) penicillin/streptomycin in a tissue culture treated 75 cm<sup>2</sup> cell

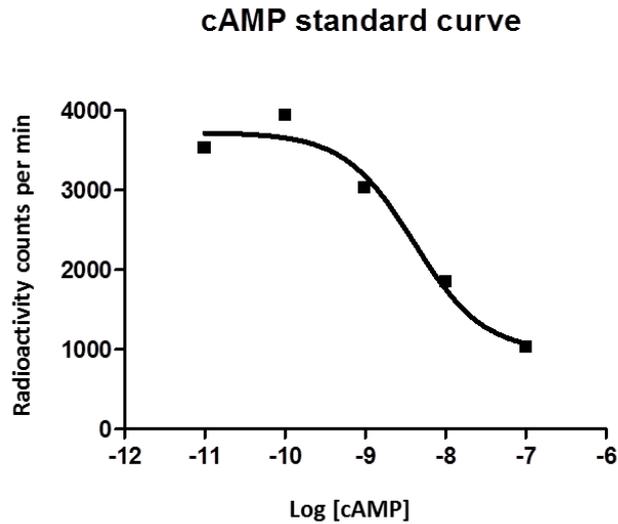
culture flask and incubated at 37°C with 5% CO<sub>2</sub>. Cells were allowed to grow to 70-90% confluency before plating them for transfection.

To transfect, cells were plated in a 96- or 48-well plate or seeded in a 10 mm petri dish depending on the intended assay. 30,000 cells in 200 µl volume were seeded per well of a 48-well plate and this was used as standard to calculate the number of cells seeded in a 96-well plate or 10 mm petri dish as a function of their surface areas. Cells were transiently transfected 24 h after plating with equal amounts of the HA CLR (wild type or mutant) and *myc* RAMP1 vectors. For a 48-well plate, 1 µg total DNA (for instance, 0.5 µg HA CLR and 0.5 µg *myc* RAMP1) was used per well. For a well of 96-well plate and a petri dish, 0.5 µg and 10 µg total DNA were used respectively. For a negative control (Fig 2.1), cells were also transfected with HA pcDNA3.1(-) (i.e. empty vector without CLR) and *myc* RAMP1. This was to check for any CGRP receptor activity in cells not transfected with HA CLR. The transfection mix contained 4.5 µl of 10 mM polyethyleneimine (PEI), 1 µg of DNA, 40 µl of 5% glucose and DMEM to a total volume of 200 µl per well of a 48-well plate (100 µl per well and 8ml for 96-well plate and 10 mm petri dish respectively). The volume was scaled up to accommodate the required number of wells. The transfection mix was set up as follows: CLR and RAMP1 vectors were added to appropriate amount of 5% glucose, mixed and incubated for 10 mins. Corresponding volume of PEI was then added, mixed and incubated for 20-30 mins, after which the volume was made up with DMEM. The plates were brought out of the incubator and old growth media replaced with the transfection mix, the plates were agitated and returned to the incubator. Cells were assayed ~48 h after transfection.

#### **2.1.2.7 cAMP standard curve**

A typical cAMP competition curve was generated using a cAMP standard, which was assessed by the radio receptor assay (a competition assay between cAMP and [<sup>3</sup>H]cAMP) described in section 2.1.2.8 below. This was important in order to determine the linear portion of the cAMP stimulation curve. cAMP standard were diluted to concentrations ranging from 100 nM to 10 pM. Here, the cAMP standards replaced the cell extracts under the competition assay step described in section 2.1.2.8. The

radioactivity counts per min were plotted against log concentration of cAMP using GraphPad Prism 4 (GraphPad Software Inc., San Diego, USA).



**Fig 2.1:** Standard cAMP curve. Standard curve was generated from a cAMP/[<sup>3</sup>H]cAMP competition assay. cAMP standard was diluted to concentrations of 100 nM – 10 pM. Sigmoidal curves were fitted using GraphPad prism 4.

#### **2.1.2.8 Assessment of cAMP production**

Following ~48 hr incubation post transfection from above, the growth media (or transfection mix) was replaced by 100 µl assay media and the cells were incubated at 37°C for 30 mins for stimulation. During the incubation, a 10 µl aliquot of 1 mM αCGRP stock (already thawed on ice) was diluted in assay media to concentrations ranging from 300 nM to 30 pM. 50 µl αCGRP from each dilution was added to the assay media in each well (i.e. each well corresponding to a particular αCGRP concentration and represents an assay point) to make the final concentrations 100 nM to 10 pM. The control assay point contained 50 µl assay media instead of αCGRP. Each assay point was made in duplicate for each condition (wild type or mutant). Cells were then incubated at 37°C for 15 mins after which the media were thoroughly aspirated and 100 µl of ice cold ethanol added to each well. The plates were stored at -20°C for at least 15 mins.

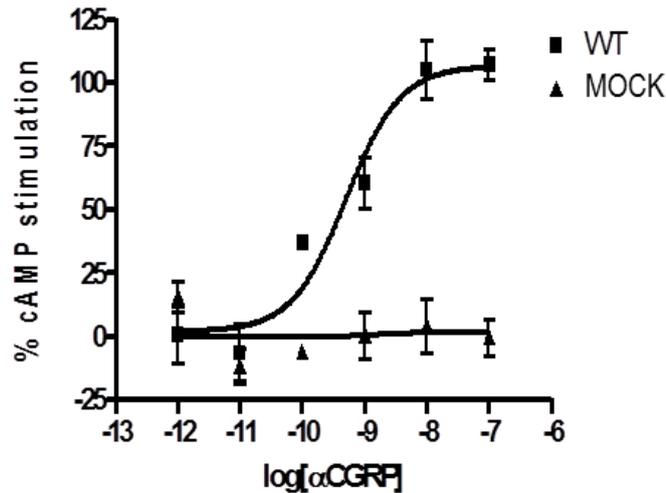
To measure cAMP production, the ethanol was allowed to evaporate by placing plates in a fume hood. The binding protein was allowed to thaw and an equal volume of assay buffer added. 100 µl of assay buffer was added to each well and the plate gently shaken for ~5 mins to re-suspend sample. 50 µl of cell extracts were transferred to 1.5 ml Eppendorf tubes. 2 µl <sup>3</sup>H cAMP (74kBq) was diluted in 4 ml assay buffer and 50 µl added to each Eppendorf tube. 100 µl of the diluted binding protein was then added to each tube and mixed thoroughly by inverting 4 - 6 times. The tubes were incubated for 2 – 24 hrs.

Following incubation, the activated charcoal was stirred gently for few mins and 100 µl was added to each tube and mixed. The tubes were centrifuged (14000g, 4°C, 5 mins) for charcoal pellet to form. 185 µl of supernatant was carefully transferred to a scintillation vial and 4 ml of high performance ScintiSafe 2 scintillation fluid was added. The vial was capped and shaken to mix content. The samples were counted using the Packard 1600TR liquid scintillation analyser.

Raw cAMP data generated were fitted to non-linear regression concentration-response curves using GraphPad Prism 4. Below is the sigmoidal concentration-response equation with which the raw cAMP data were fitted:

$$Y = \text{Bottom} + (\text{Top} - \text{Bottom}) / (1 + 10^{((\log EC_{50} - X))})$$

where Y is the response running from the bottom through to the top of the sigmoidal curve. Bottom and Top represent the lowest and highest plateau values respectively on the Y axis. The Hill slope was assumed to be 1.



**Fig 2.2:** Dose-response curve of mock transfected COS7 cells. Following 48 h incubation after transfection, mock (HA pcDNA3.1(-) and *myc* RAMP1) transfected cells were challenged with 100nM – 10pM  $\alpha$ CGRP with a control assay point containing no  $\alpha$ CGRP. Sigmoidal curves were fitted using GraphPad prism 4. Each curve is representative of one of at least three independent experiments. Each point on the curve represents duplicate assay data with standard error bars.

### 2.1.2.9 Enzyme-linked immunosorbent assay (ELISA)

This assay probes for the HA-tagged CLR to determine the cell surface expression of wild type and mutant CGRP receptor. COS 7 cells in a 48-well plate were transiently transfected (with wild type HA CLR/*myc* RAMP1, mutant HA CLR/*myc* RAMP1 and empty pcDNA3.1(-)/*myc* RAMP1) to give three different conditions in each experiment. The empty pcDNA3.1(-)/*myc* RAMP1-transfected cells (negative control) controlled for non-specific antibody binding. Assay points for each condition were in triplicates or more and at least three independent experiments were performed. Following transfection and ~48 hr incubation period, the growth media was aspirated and cells fixed with 150  $\mu$ l 3.7% formaldehyde for 15 min after which the cells were washed three times with 250 $\mu$ l PBS. The cells were blocked with 200  $\mu$ l 2% BSA in PBS for 45 min and were then treated with 150  $\mu$ l of primary antibody diluted 1:2000 in 1% BSA in PBS for 1 h. The cells were washed three times with PBS and treated with 150  $\mu$ l secondary antibody diluted 1:2000 in 1% BSA in PBS for another 1 h. The cells were washed again three times with PBS and then developed with 150  $\mu$ l SIGMAFAST™

OPD tablets dissolved in water according to manufacturer's instructions. The reaction was terminated with 100  $\mu$ l 1M H<sub>2</sub>SO<sub>4</sub>. The reaction product was measured at 490 nm using the Thermo Scientific Multiskan GO plate reader.

#### **2.1.2.10 CGRP receptor cell surface expression after $\alpha$ CGRP-dependent internalization**

The procedure followed here is as described above except that the cells were treated with 100 nM  $\alpha$ CGRP in full medium for 1 h at 37°C before fixing with 3.7% formaldehyde.

#### **2.1.2.11 Total CLR expression**

This was conducted for mutants with significantly different cell surface expression values to that of the wild type. The procedure again is as described for the assessment of cell surface expression by ELISA except that after fixing with 3.7% formaldehyde, the cells were permeabilized by treating with 0.1% Triton-X 100 in PBS for 1 h.

#### **2.1.2.12 Crude membrane preparation**

COS 7 cells were seeded into 10 cm tissue culture dishes and was transfected after ~24 hr. Following ~48 hr incubation post-transfection, each culture dish was washed with 3 ml ice cold PBS. 2 ml ice cold homogenization buffer (20 mM HEPES, 1 mM EGTA, 10 mM MgCl<sub>2</sub> and SIGMAFAST™ Protease Inhibitor cocktail tablet EDTA-free, pH 7.5) was then added and the cells were scrapped into a sterile 50 ml falcon tube on ice. The cells were subjected to three 15 sec burst with an Ultra-Turrax T25 tissue homogenizer at maximum capacity, with the cell suspension placed on ice for 45 sec between burst. The homogenate was centrifuged at 20,000 g for 25 min at 4°C. The pellet was resuspended in 2 ml binding buffer (20 mM HEPES, 2 mM MgCl<sub>2</sub> pH 7.5). 20  $\mu$ l aliquot of the resuspension was transferred into 0.5 ml eppendorf tube for Western blot analysis. The remainder was stored at -20°C.

## **2.2 eCLR and eRAMPs production**

### **2.2.1 Materials**

All reagents/chemicals, unless otherwise stated, were supplied by Sigma-Aldrich, Fisher Scientific or Invitrogen (UK). Enzymes and primers were purchased from New England Biolabs, and Invitrogen respectively. The supplier of each equipment used is stated in parenthesis as it is first mentioned in the course of this report. The reagents used were certified at analytical grade for scientific research purpose.

#### **2.2.1.1 Equipment**

ÄKTA purifier (GE Healthcare Life Sciences, UK)

Bio-Rad Power Pac 1000 and Power Pac 300 (Bio-Rad, UK)

Centrifuges: AccuSpin™ Micro R (Fisher Scientific, UK), Beckman Coulter Avanti™ J-20 XP and Allegra™ 25R (Beckman Coulter Inc., UK)

GeneAmp® PCR System 9700 (Applied Biosystems, UK)

Jenway Genova Spectrophotometer (Bibby Scientific, UK)

Uvitec CCD Camera (Uvitec, UK)

#### **2.2.1.2 Media and stock solution**

##### **2.2.1.2.1 Luria-Bertani (LB) broth**

10 g of LB powder was dissolved in 500 ml distilled H<sub>2</sub>O. 10 g agar granules were added if making up agar plates. The solution was sterilised by autoclaving at 121°C for 20 min and was allowed to cool to room temperature. LB broth was stored at room temperature while agar plates were stored at 4°C.

#### **2.2.1.2.2 BMGY (Buffered glycerol-complex medium)**

10 g yeast extract and 20 g peptone were dissolved in 700 ml dH<sub>2</sub>O. It was autoclaved at 121°C for 20 min and allowed to cool to room temperature. 100 ml 1 M potassium phosphate pH 6.0, 100 ml 10x yeast nitrogen base (YNB), 2 ml 500x biotin and 100 ml 10x glycerol (all sterile; see below) were added to make a final volume of 1 L. The media was stored at 4°C for at most 2 months.

#### **2.2.1.2.3 BMMY (Buffered methanol-complex medium)**

10 g yeast extract and 20 g peptone were dissolved in 700 ml dH<sub>2</sub>O. It was autoclaved at 121°C for 20 min and allowed to cool to room temperature. 100 ml 1 M potassium phosphate pH 6.0, 100 ml 10x YNB, 2 ml 500x biotin and 100 ml 10x methanol were added to make a final volume of 1 L. The media was stored at 4°C for at most 2 months.

#### **2.2.1.2.4 YPD (Yeast peptone dextrose)**

10 g yeast extract and 20 g peptone were dissolved in 900 ml dH<sub>2</sub>O. 20 g of agar granules were added if making agar YPD agar plates. This was autoclaved and 100 ml of sterile 10x glucose (dextrose) added after cooling. The media was stored at 4°C.

#### **2.2.1.2.5 YPDS (Yeast peptone dextrose with sorbitol) Geneticin<sup>®</sup>/Zeocin plates**

10 g yeast extract, 20 g peptone and 182 g sorbitol were dissolved in 900 ml dH<sub>2</sub>O. 20 g agar granules were then added and the solution made sterile by autoclaving. It was allowed to cool to around 55°C with constant gentle stirring after which 100 ml 10x glucose was added. Appropriate amount of Geneticin or Zeocin was added to give the desired final antibiotic concentration. The solution was poured into sterile 10 cm polystyrene petri dish to about two-third of dish volume and allowed to solidify. The plates were stored at 4°C.

#### **2.2.1.2.6 10x Yeast nitrogen base (YNB)**

134 g of yeast nitrogen base (with ammonium sulphate and without amino acids) was dissolved in ddH<sub>2</sub>O to a total volume of 1 L. Solution was heated slightly to allow YNB dissolve. This was filter-sterilised and stored at 4°C.

#### **2.2.1.2.7 500x Biotin (0.02%)**

20 mg biotin was dissolved in 100 ml ddH<sub>2</sub>O and filter-sterilized and was stored at 4°C.

#### **2.2.1.2.8 10x Glucose (20%)**

200 g D-glucose was dissolved in 1 L ddH<sub>2</sub>O and was autoclaved at 121°C for 20 min. This was stored at 4°C after cooling to room temperature.

#### **2.2.1.2.9 10x Methanol (5%)**

5 ml absolute methanol was mixed with 95 ml ddH<sub>2</sub>O and was filter-sterilized. This was stored at 4°C for no later than 6 weeks.

#### **2.2.1.2.10 10x Glycerol (10%)**

100 ml of glycerol was mixed 900 ml ddH<sub>2</sub>O. The solution was filter-sterilized and stored at room temperature.

#### **2.2.1.2.11 1M Potassium phosphate buffer, pH 6.0**

132 ml 1 M K<sub>2</sub>HPO<sub>4</sub> was mixed with 868 ml 1 M KH<sub>2</sub>PO<sub>4</sub> and the pH adjusted to 6.0 with phosphoric acid. The buffer was sterilized by autoclaving and stored at room temperature. To make 1 M K<sub>2</sub>HPO<sub>4</sub>, 174.2 g of this salt was dissolved in ddH<sub>2</sub>O to a total volume of 1 L. 1 M KH<sub>2</sub>PO<sub>4</sub> was made by dissolving the salt in ddH<sub>2</sub>O to 1 L total volume.

#### **2.2.1.2.12 1x Tris-buffered saline with Tween 20 (TBST) pH 7.4**

10 ml 1 M Tris base pH 7.3 was mixed with 40 ml 2.5 M NaCl and the volume made up to 1 L. 1-2 ml Tween-20 was then added and mixed thoroughly. The solution was stored at room temperature. To make 1 M Tris base stock solution, 121.1 g Tris base was dissolved in 700 ml ddH<sub>2</sub>O and pH adjusted to 7.3 with concentrated HCl. The volume was made up to 1 L. 2.5 M NaCl was made by dissolving 146.1 g of NaCl in 1 L ddH<sub>2</sub>O. These were stored at room temperature.

#### **2.2.1.3 Reagents**

##### **2.2.1.3.1 Restriction enzymes<sup>4</sup>**

The restriction enzymes (restriction endonucleases) used in this project were *Bam*HI, *Spe*I and *Pme*I (New England BioLabs UK Ltd). Reaction conditions were set up as recommended by the manufacturer. Details of these enzymes can be sought from supplier's website (<http://www.neb.uk.com/tools/index.aspx?req=enzyme finder>).

##### **2.2.1.3.2 Ligation enzyme**

T4 DNA ligase, an enzyme which catalyses the formation of a phosphodiester bond between DNA strands, was employed to ligate digested plasmid vectors and PCR products. It was used according to supplier's instructions. It was purchased from New England BioLabs UK Ltd.

##### **2.2.1.3.3 SDS polyacrylamide gels**

The resolving gels used included 10%, 12% and 15% gels and are appropriately indicated where used. The stacking gel was usually 4%. The composition of these gels is presented in Table 2.1.

**Table 2.1:** Composition of the polyacrylamide gels used in this project.

Reagents	Volume per gel (ml)			
	Stacking gel	Resolving gel		
	4%	10%	12%	15%
Polyacrylamide 30%	0.3	1.9	2.3	2.9
ddH <sub>2</sub> O	1.5	2.2	1.8	1.2
1.5M Tris pH 8.8	-	1.5	1.5	1.5
1M Tris pH 6.8	0.6	-	-	-
10% SDS	0.02	0.06	0.06	0.06
10% Ammonium persulfate (APS)	0.01	0.02	0.02	0.02
TEMED	0.0025	0.0045	0.0045	0.0045

#### 2.2.1.3.4 Antibiotics

Antibiotics were added to agar plates and media for selectivity in the growth of transformed cells. The vectors contain sequences that confer antibiotic resistance on transformed cells (i.e. cells that have taken up the plasmid DNA).

**Ampicillin:** 500 mg ampicillin was dissolved in 10 ml sterile distilled water to give a 50 mg/ml stock and was stored as 1 ml aliquots in Eppendorf tubes at -20°C.

**Kanamycin:** Kanamycin was purchased as a 50 mg/ml solution from Sigma-Aldrich, UK (catalogue No: K0254). It was stored at 4°C.

**Geneticin (G418):** 250 mg G418 was dissolved in 5 ml distilled water to give a 50 mg/ml working stock. This was filter-sterilised using a microfilter. This was prepared freshly when required.

**Zeocin:** 500 mg zeocin was dissolved in 10 ml sterile distilled water to give a 50 mg/ml stock and was stored as 1 ml aliquots at -20°C. The tubes were wrapped with foil to avoid light as zeocin is light-sensitive.

#### **2.2.1.3.5 Antibodies**

The antibodies were employed in Western blot to probe the 10x His tag on the eCLR and eRAMPs. The primary antibody used is the albumin-free 6x His Monoclonal antibody (Clontech, UK. Cat No: 631212) and the secondary antibody is Goat Anti-mouse IgG (Fab Specific) peroxidase conjugate (Sigma-Aldrich, UK. Cat No: A3682).

#### **2.2.1.3.6 DNA and Protein markers**

The DNA marker used for all agarose electrophoresis was the 1 Kb Plus DNA Ladder (100bp – 12Kb) (Life Technologies, UK. Cat. No: 10787-018). The protein markers were PageRuler™ Prestained Protein Ladder (10-170 kDa) (Fermentas, UK. Cat. No: SM0671) and ProtoMetrics - Recombinant Protein Markers (10-225 kDa) (Geneflow, UK. Cat. No: L1-0100). The latter of the protein markers was detectable on Western blot.

### **2.2.2 Methods**

#### **2.2.2.1 Molecular biology**

##### **2.2.2.1.1 CLR and RAMPs gene amplification by PCR**

The genes encoding the n-terminal extracellular domains of human *CLR*, *RAMP1*, *RAMP2* and *RAMP3*, excluding the signal peptide, were amplified by polymerase chain reaction (PCR) using the GoTaq® (Promega, UK). The forward and reverse primers (Table 2.2) were designed to incorporate the *BamHI* and *SpeI* restriction sites respectively into the constructs.

**Table 2.2:** Table showing the PCR primers. The *Bam*HI and *Spe*I recognition sequence in the forward and reverse primers respectively are highlighted bold.

CLR	Forward: 5' GGGGG <b>GGATCC</b> GAAT TAGAAGAGAGTCCT 3' Reverse: 5' GGGGG <b>ACTAGT</b> TTACTCGTGGGTGTTAACATT 3'
RAMP1	Forward: 5' CT <b>GGATCC</b> GCCTGCCAGGAGGCTAAC 3' Reverse: 5' C <b>CACTAGT</b> TTAGCTGCCGGGCGGGTCCCCG 3'
RAMP2	Forward: 5' GGGGG <b>GGATCC</b> GAATTAGAAGAGAGAC 3' Reverse: 5' GGGGG <b>ACTAGT</b> TTACTCGTGGGTGTTAATG 3'
RAMP3	Forward: 5' GGGGG <b>GGATCC</b> GGCTGCAACGAGACAGGC 3' Reverse: 5' GGGGG <b>ACTAGT</b> TTAGAACCTCGTCTGGGGGGT 3'

The reaction was set up in sterile PCR tubes on ice as follows:

- 10 µl 5x GoTaq polymerase buffer
- 1 µl forward primer (10 pmoles)
- 1 µl reverse primer (10 pmoles)
- 1 µl DNA template (100 ng) – not included in negative control
- 1 µl dNTPs (10 mM)
- 35 µl sterile dH<sub>2</sub>O
- 1 µl GoTaq polymerase

Total reaction volume = 50 µl.

The PCR programme was set up as follows:

1. 95°C for 2 min
  2. 95°C for 30 sec
  3. 55°C for 30 sec
  4. 72°C for 2 min
  5. 72°C for 5 min
  6. 4°C until reaction tubes are collected.
- } 30 cycles

The template genes, harboured in pcDNA3.1, were kindly provided by Dr James Barwell. The PCR products were analysed on 1% agarose gel with the TrackIt™ 1 Kb Plus DNA Ladder (Invitrogen UK). Gels were run at 85V for 1h 30mins.

#### **2.2.2.1.2 Cloning**

The amplified gene products were subcloned into a modified *Pichia pastoris* expression vector named pPICK9K\_MepNet (Fig 2.2) from the Membrane Protein Consortium (kindly provided by Dr Sarah Routledge, Aston University). The amplified gene products and the vector were sequentially digested with the BamHI and SpeI restriction endonucleases, according to supplier's instructions, generating sticky ends. The digested products were ligated with T4 DNA ligase for 16 hr at room temperature. The ligated products were transformed with *E. coli* competent cells (described in section 2.2.2.1.4) followed by isolation of positive clones by selecting on Luria broth (LB) agar plates containing a 50µg/ml final concentration of kanamycin. The clones were grown overnight at 37°C with shaking at 200 rpm. Plasmid isolation was performed using Plasmid Miniprep Kit (Sigma-Aldrich, UK). Isolated plasmids were sequenced using the  $\alpha$ -Factor and 3' AOX1 sequencing primers to confirm insertion.



**Fig 2.3:** Vector map of the *P. pastoris* vector employed in this study. AOX1: alcohol oxidase 1. Courtesy of Dr Mohammed Jamshad (University of Birmingham).

#### **2.2.2.1.3 DNA sequencing**

The cloned genes were sequenced at the Functional Genomics laboratory, University of Birmingham. The primers used were the  $\alpha$ -factor and 3'AOX1 sequencing primers. The sequencing reaction mix was set up according to the laboratory's instruction.

#### **2.2.2.1.4 *E. coli* transformation**

*E. coli* cells to be transformed were brought out of the -80°C freezer and allowed to thaw on ice. Sterile 1.5 ml eppendorf tubes and pipette tips to be used were pre-chilled in a -20°C freezer. While placed on ice, 40  $\mu$ l of the *E. coli* cells was aliquoted into each tube, each for individual construct plasmid. 1-5  $\mu$ g of the plasmid to be transformed was added to the respective labelled tube and allowed to incubate on ice for 30 min. The reaction mix was heat-shocked by incubating it at 42°C for exactly 45 sec. The cells were immediately transferred back on ice to recover for 2 min after which 0.5 ml LB was added and the cells incubated at 37°C for 1 h. A 150  $\mu$ l aliquot of the cells were plated on an antibiotic (kanamycin or ampicillin)-treated LB plate and incubated at 37°C for 12 to 16 h.

Depending on the purpose of the transformation, colonies were picked and grown in a 1-2 ml (miniprep) or 200-250 ml (maxiprep) LB culture overnight.

#### **2.2.2.1.5 Miniprep and Maxiprep**

Overnight cultures from earlier described were harvested for extraction of plasmid DNA by miniprep or maxiprep as required. The miniprep and maxiprep were performed using GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich, UK) and Marligen PowerPrep™ HP Plasmid Purification Kit (Insight Biotechnology Limited, UK) respectively according to manufacturer's instruction. While the miniprep is performed to extract plasmid DNA of up to 15 µg, the maxiprep is carried out to yield a desired amount of up to 500 µg plasmid DNA. Final elutions for miniprep extractions were done with 30 µl of sterile distilled H<sub>2</sub>O. Dilution of DNA pellets from maxiprep extraction was done with varying volume of sterile distilled H<sub>2</sub>O depending on the desired concentration.

#### **2.2.2.2 Making *P. pastoris* competent cells**

1 µl of stock *P. pastoris* (X33 or SMD1163) cells was streaked on a YPD plate and incubated at 30°C until colonies were seen. 5 ml of YPD was inoculated with a colony from the plate and grown overnight at 37°C with shaking at 200 rpm. The overnight culture was diluted to an OD<sub>600</sub> of 0.15 – 0.20 in 50 ml YPD in a 250-ml shake flask. The cells were grown at 30°C with shaking at 200 rpm to an OD<sub>600</sub> of 1.0 – 1.2. The culture was centrifuged at 5000 rpm for 5 min and the supernatant taken off. The pellet was re-suspended in 9 ml BEDS solution (10 mM bicine-NaOH pH 8.3, 5% v/v dimethyl sulphoxide (DMSO), 3% v/v ethylene glycol and 1M sorbitol) supplemented with 1 ml 0.1 M dithiothreitol (DTT) and was incubated for 5 min at 30°C with shaking at 100 rpm. The culture was again centrifuged at 5000 rpm for 5 min at room temperature and re-suspended in 250 µl of BEDS solution without DTT. Aliquots of 40 µl were transferred to cell storage vials and stored at -80°C for up to 6 months.

### **2.2.2.3 *P. pastoris* transformation**

Isolated individual plasmids containing the genes for the ectodomains of the CLR, RAMPs 1, 2 and 3 (further referred to in this report as *eCLR*, *eRAMP1*, *eRAMP2* and *eRAMP3* respectively) were linearized with *PmeI* restriction endonuclease at 37°C for 16hrs to ensure maximum linearization. The linearized plasmids were analysed on 0.8% agarose gel to confirm linearization. This was followed by transformation with *P. pastoris* wild type (X33) or protease-deficient (SMD1163) host strain by electroporation (at 1800v) using an electroporation device according to manufacturer's instruction. The cells were quickly revived with 500µl of ice cold 1M sorbitol and 500µl ice cold YPD (1% yeast extract, 2% peptone, 2% dextrose) and the cell pellets were resuspended. The resuspended cells representing each construct were spread on YPDS (1% yeast extract, 2% peptone, 2% dextrose, 2% agar, 1M sorbitol) plates with varying final concentrations of G418 (0, 0.1, 0.5, 1.0 and 2.0 mg/ml) and incubated at 30°C for 96 hr with the plates checked after every 24hrs. After the 96-hr incubation period, a few colonies from the 1.0mg/ml G418 plates were restreaked on fresh 0.1 and 0.5mg/ml G418 plates and incubated to obtain bigger, more resistant colonies.

### **2.2.2.4 Gene expression**

#### **2.2.2.4.1 24-well plate (screening) format**

This was performed to check for protein-producing clones. Selected clones from the G418-YPDS plates were cultured in BMGY (1% yeast extract, 2% peptone, 1.34% yeast nitrogen base without amino acids, 0.4µg/ml biotin, 1% glycerol, 0.1 M potassium-phosphate buffer pH 6) at 30°C for 24hrs with shaking at 250 rpm. 3 ml aliquots of BMMY (same as BMGY with 5% methanol replacing glycerol) medium in a 24-well plate were inoculated with the 24-hr cultures to a starting OD<sub>600</sub> of 1. This was incubated at 30°C, 250 rpm. The AOX1 promoter was induced 24hrs after inoculation by the addition of methanol (1% final concentration) to each culture well. Prior to induction, a 100 µl aliquot was collected from each well and centrifuged at 10,000g for 10mins. The pellets and supernatants were separated and both stored at -80°C for SDS PAGE and Western blot analyses. A 100 µl aliquot

of cells were harvested from each culture well 24 and 48 hr post-induction. Harvested cells were cooled on ice and stored at -80°C for further analyses.

#### **2.2.2.4.2 Shake flask method**

Here, the actual protein production was carried out in contrast to the 24-well expression format which was specifically performed to determine expressing clones. The clones found to best express the proteins of interest were cultured in 25 ml BMGY in a 250 ml shake flask for 24 hr to grow a biomass. All incubations were performed at 30°C and 250 rpm. The 25 ml cell culture was harvested by centrifuging at 5000g and the supernatant discarded. The cell pellets were resuspended in 100 ml BMMY in a 500 ml shake flask and incubated for 72 hr. Inductions (addition of 5% methanol to a 1% final concentration) were carried out 24 and 48 hr post-inoculation. The cells were then harvested 24 hr after the second induction by centrifuging at 5000 rpm for 25 min at 4°C in a sterile 50 ml falcon tube. The supernatant were analysed for the presence of desired proteins by Western blot before purification.

#### **2.2.2.5 Protein identification**

##### **2.2.2.5.1 SDS PAGE**

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE) lower (or resolving) and upper (or stacking) gels were prepared as shown in Table 2.1. Each gel was set using a 10-well comb and the number of gels prepared was dependent on the number of samples to be analysed at the time of experiment. Samples to be analysed were treated with 1x sample loading buffer (50 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 1%  $\beta$ -mercaptoethanol and 0.02% bromophenol blue) and incubated at 85 - 90°C for 5 – 10 min. 5 - 20  $\mu$ l of protein samples were loaded onto each well depending on the concentration of the protein. The gel plates were set up in an electrophoretic tank, samples loaded in the presence of 1x running buffer (3.03g Tris, 18.8g glycine and 1g SDS in 1L ddH<sub>2</sub>O) and run at 200 volts for 1 hr using the Bio-Rad Power Pac 1000 or Power Pac 300 as power source. The gels were collected and prepared for silver staining or western blotting.

#### **2.2.2.5.2 Silver staining**

Silver staining was performed using the SilverQuest™ Silver Staining Kit (Life Technologies, UK) according to manufacturer's instruction.

#### **2.2.2.5.3 Western blot**

The gel from SDS-PAGE was placed in the transfer buffer (3.03g Tris, 14.4g glycine and 200 ml methanol in 1L ddH<sub>2</sub>O). For each gel, four filter papers and one Whatman nitrocellulose paper were cut to rectangles approximately 10 x 7 cm and 9 x 6 cm in sizes respectively. These papers were placed in the transfer buffer to equilibrate for 15 – 20 min. The transfer sandwich was set up in the following order in the direction of transfer from cathode to anode:

- o fibre pad
- o two chromatography (filter) papers
- o the SDS PAGE gel
- o nitrocellulose transfer membrane
- o another two chromatography (filter) papers, and
- o the fibre pad.

The set-up was performed carefully in a bowl containing transfer buffer to avoid air bubbles. The cassette was then closed and fixed into the transfer cell alongside an ice cooling unit and the transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol) introduced to fill the cell. Transfer was conducted at 150 volts for 45 min.

After the transfer, the nitrocellulose membrane, which now contained the protein, was blocked with 5% Marvel dried skimmed milk in TBST (1 M Tris pH 7.3, 2.5 M NaCl, 0.1-0.2% Tween-20) for 1 hr at room temperature or for 24 hrs at 4 °C. This is to prevent unspecific interaction between the membrane and the antibody used for the detection of the protein. The blocking solution was then replaced by a solution of primary antibody in 5% milk/TBS (1:5000) and was incubated for 1 hr at room temperature with gentle shaking. This antibody, in principle, binds to the “unblocked” protein of

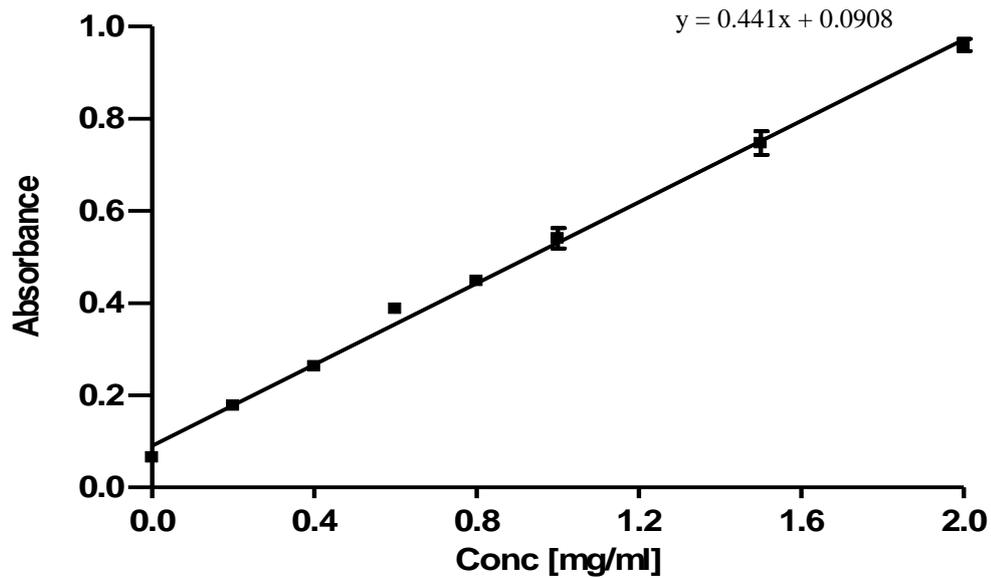
target and could then be probed by a secondary antibody (from same animal origin), which is in turn detected by chemiluminescence. The membrane was washed with TBST for 5 min thrice to remove unbound primary antibody and was treated with a solution of secondary antibody in 5% milk/TBS (1:5000) for 1 hr at room temperature with gentle shaking. The membrane was washed again, thrice for 5 min each with TBST. Protein was detected using EZ-ECL™ chemiluminescence solution (Geneflow, UK. Cat. No: 20-500-120) according to manufacturer's instruction.

#### **2.2.2.6 Purification**

The receptor proteins were aimed to be purified by selectively targeting the N-terminal 10x His tag on the proteins. Ni-NTA agarose resin was used. 2 ml of 25% Ni-NTA agarose resin was added to 30 ml supernatant and incubated at 4°C for 1 hr or overnight with constant gentle shaking. The mixture was loaded onto 5 ml Qiagen gravity flow column. The resin was allowed to settle after which the column bottom cap was removed to allow flow under gravity. The column was washed at least twice with 4 ml wash buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl and 20 mM imidazole, pH 8) and the protein eluted thrice with 4 ml elution buffer (50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl and 250 mM imidazole, pH 8). All fractions were collected on ice. 50 µl aliquots were collected from each fraction for silver stain and Western blot analyses. The collected fractions were stored at 4°C.

#### **2.2.2.7 Protein quantification**

Protein concentration was determined using 1 mg/ml BSA as standard in a flat bottom 96 well plate. A standard curve was generated by plating BSA standard in triplicate wells at concentrations of 0.0, 0.2, 0.4, 0.6, 0.8 and 1.0 mg/ml in a 10 µl total volume with PBS [8.0 g/l sodium chloride, 0.2 g/l potassium chloride, 1.15 g/l di-sodium hydrogen phosphate (Na<sub>2</sub>HPO<sub>4</sub>), 0.2 g/l potassium dihydrogen orthophosphate (Fisher Scientific, UK)]. Samples were diluted 20- and 40-fold to a total volume of 10 µl and were plated in triplicate. 200 µl BCA reagent (50:1 (v/v) mixture of BCA solution and 4% (w/v) copper II sulphate solution) was added to all triplicate wells of standards and samples. The plate was incubated at 37°C for 30 min and the absorbance values measured at 570 nm using a Biotek EL800 microplate reader. A typical standard curve generated is shown in Fig 2.3.



**Fig 2.4:** A representation of the BSA standard curve. Graph of BSA concentration against absorbance at 570 nm.

#### 2.2.2.8 Protein dialysis

The purified receptor protein was dialysed into dialysis buffer (50 mM sodium phosphate and 150 mM NaCl, pH 7.4). Dialysis membrane was cut to a length enough to accommodate sample volume. It was folded twice and clipped tightly at one end and the protein sample introduced. The other end was then folded and clipped tightly allowing a little free space within the membrane. The membrane end to be folded was briefly dipped in the buffer to soften it for easy folding. The membrane was put in a beaker with one end taped to the beam of the beaker and the dialysis buffer was poured in to cover all part of membrane containing the sample. The dialysis system was first incubated for 4 hrs at room temperature with the dialysis buffer replaced after every two hrs. After the second buffer replacement, the system was left to incubate for 24hrs at 4°C. The buffer was gently stirred using a magnetic stirrer during the entire incubation period with enough space allowed between the clip and magnetic stirrer to avoid any contact as this could cause high agitation which might rupture the membrane. The dialysed sample was concentrated using vivaspin 20 (Sartorius Ltd, UK. Cat. No: VS2011), quantified and stored at 4 °C.

### **2.2.2.9 Gel filtration**

This was performed to analyse the homogeneity of the eCLR and eRAMP3 protein samples. The Superdex™ 200 column was used while connected to an ÄKTA purifier. The column was washed with 2x column volumes dH<sub>2</sub>O which has been filter-sterilized and degassed (dissolved air removed) using Millipore® Stericup™ vacuum filter units and a vacuum pump. The column was equilibrated with 1.5 column volumes of 50 mM sodium phosphate buffer pH 7.5 containing 150 mM NaCl (filter-sterilized and degassed). 0.5 ml of protein sample (0.08 – 1.2 mg/ml concentration range) was injected into the purification system and the phosphate buffer run through at 0.5 ml per min to elute the protein. 2 ml elution fractions were collected. The elution profile was determined by plotting the elution volume against UV absorbance (mAU). The fractions were analysed on SDS PAGE and detected by silver staining.

### **2.2.2.10 Deglycosylation of proteins**

Removal of high mannose N-glycans from CLR and RAMP receptor protein samples was performed using Endoglycosidase H (Endo H; New England Biolabs, UK. Cat No: P0702S). Following dialysis and concentration of protein samples as earlier described above, 10 µg of protein sample was denatured by adding 1 x Glycoprotein denaturing buffer (0.5% SDS, 40 mM DTT) and incubating at 100°C for 10mins. 1 x G5 reaction buffer (50 mM sodium citrate pH 5.5 at 25°C) was added and 2 µl Endo H then introduced into the reaction mix. The reaction was incubated for 1 h at 37°C. The enzyme was heat-inactivated at 75°C for 10 min. The reaction products were analysed by SDS-PAGE.

## **2.3 Biophysical characterization**

### **2.3.1 Materials**

#### **2.3.1.1 Equipment**

Jasco J-715 Spectropolarimeter (School of Biosciences, University of Birmingham)

Beckman Optima™ XL-I Analytical Ultracentrifuge (School of Biosciences, University of Birmingham)

Reichert SR7000DC Dual Channel SPR Instrument (ARCHA Advanced Research Facility Unit, Aston University)

### **2.3.2 Methods**

#### **2.3.2.1 Circular dichroism (CD)**

CLR and RAMPs samples that have been dialysed and concentrated were diluted to within 0.5 – 0.7 mg/ml with 50 mM sodium phosphate buffer pH 8.0. The CD spectra were read using the Jasco J715 spectropolarimeter with the parameters set as follow;

- Sensitivity - 100 mdeg
- Data pitch - 0.5nm
- Scanning mode - continuous
- Response - 1 sec
- Band width - 2.0 nm
- No of scans - 8

1mm cuvettes (Starna/Optiglass , Hainault, U.K.) were used and the spectra were recorded from 260 nm to 180 nm. The CD spectra were then buffered corrected against 50 mM sodium phosphate buffer pH 8.0. The spectra were truncated at low wavelength where the high voltage of the detector indicated

no linear correlation between the signal and protein concentration. The data were plotted using Microsoft Excel spreadsheet.

### **2.3.2.2 Analytical ultracentrifugation (AUC)**

The samples to be used for AUC were diluted to an OD at 280 nm of 0.3 – 0.5. The analysis was carried out using the Beckman Optima™ XL-I Analytical Ultracentrifuge at the University of Birmingham (Birmingham, UK). The equipment was allowed to cool to 4°C. The cells were carefully loaded with 500µl of protein sample and one was loaded with an equal volume of 50 mM sodium phosphate buffer as reference. The cells were weighted to ensure their weights were virtually the same. The cells were loaded on to the ultracentrifuge and the samples were run overnight at maximum speed. The wavelength range was between 220 and 310 nm. The flow rate was set 2ml/min. The data were analysed using the Sedfit and Sednterp analytical programmes with the help of Mrs Rosemary Parslow (University of Birmingham).

### **2.3.2.3 CLR/RAMP interaction by SPR**

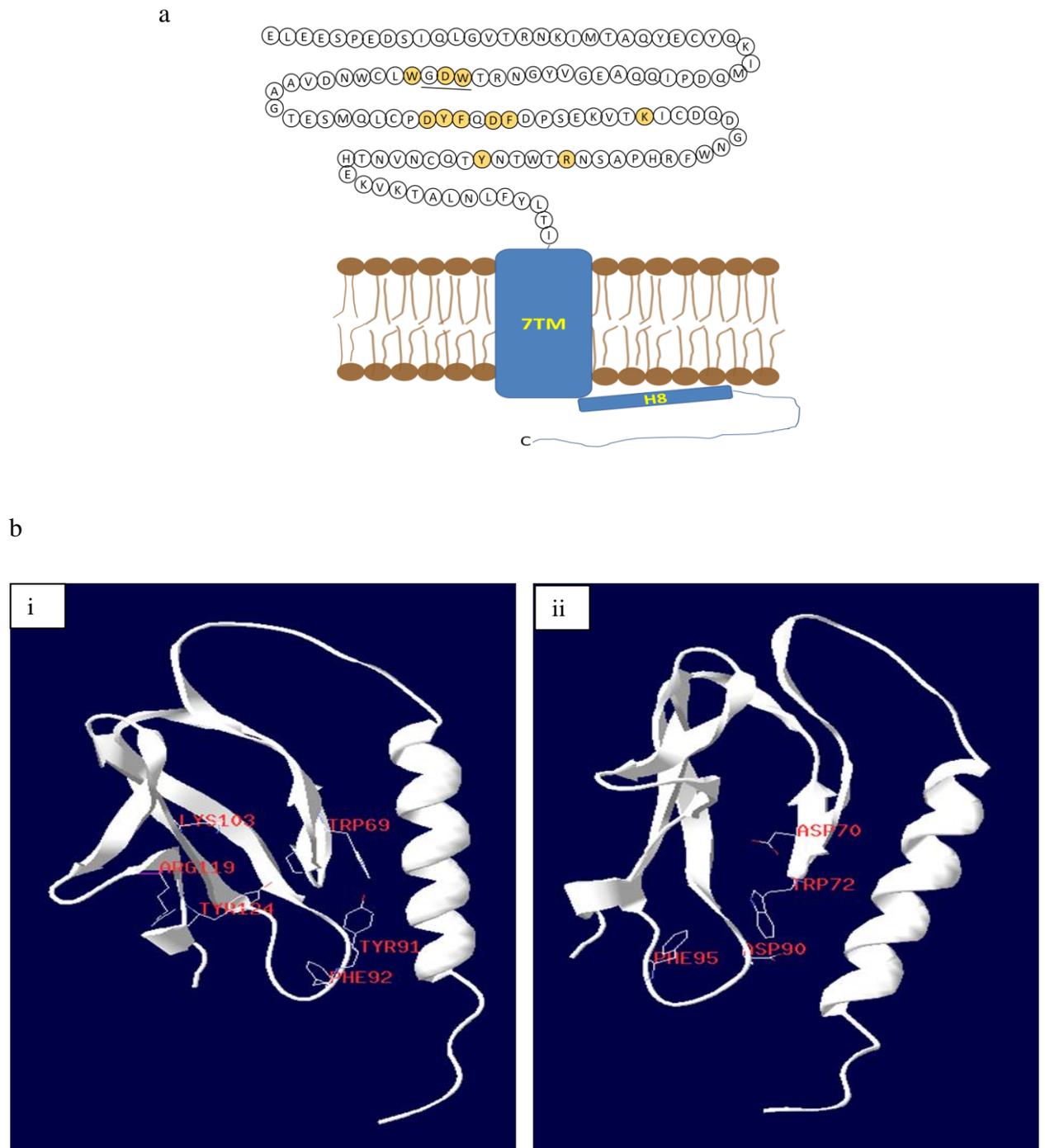
CLR/RAMP interaction was investigated by surface plasmon resonance (SPR), using the SR7000DC Dual Channel Surface Plasmon Resonance System. Prior to the experiment, C-terminally biotinylated CLR (to be immobilized) was dialysed onto 10 mM Sodium acetate buffer (pH 5.2) and was concentrated to 0.1-0.2 mg/ml (~1 ml final volume) using the Vivaspin column. RAMP1, 2 and 3 (the analytes) were diluted into HBS running buffer (10mM HEPES pH 7.4, 150 mM NaCl, 0.005% Tween 20). All buffers were filtered through 0.2 µm filters and degassed by filtering under vacuum. Samples were degassed using a syringe and parafilm. These were done at room temperature. CLR was immobilised onto Streptavidin/NeutrAvidin sensorchip (Reichert Technologies, USA) resulting in around 300 response units (RU). 100 µl of a RAMP was injected for ~4mins. Experiments were performed at 25°C and at a flow rate of 20 µl/min. The binding profiles of these proteins were monitored on a sensogram.

## **Chapter 3: Identification of residues within the CLR N-terminus required for CGRP receptor signalling**

### **3.1. Introduction**

A review of various models proposed for the binding of peptide and non-peptide ligands to the class B GPCRs has indicated the importance of the N-terminus in these interactions (Hoare, 2005; Parthier *et al.*, 2009). The C-terminal region of  $\alpha$ CGRP has been reported to interact with the amino terminal domain of the CLR (Koller *et al.*, 2002; Banerjee *et al.*, 2006), further emphasizing the role of this domain in the receptor activation and signalling. More evidence is drawn from Barwell *et al.* (2010) where important sites of ligand interaction were identified at the extreme N-terminus (residues 23 – 60) of CLR. In order to further investigate the roles played by some other residues within this pharmacologically significant N-terminal region, a number of residues were selected based on the crystal structure released by ter Haar *et al.* (2010). According to the structure, these residues were situated around and/or pointing towards the putative binding pocket for the receptor ligand and could possibly be involved in  $\alpha$ CGRP binding. The final list of residues (see Fig 3.1) was also aided by the availability of the crystal structures of some family B GPCR N-terminal domains with bound endogenous ligands. These N-terminal ECD structures included those of GIPR (Parthier *et al.*, 2007), PTH1R (Pioszak and Xu, 2008), GLP-1R (Runge *et al.*, 2008) and CRFR1 (Pioszak *et al.*, 2008). The selected residues are shown in Fig 3.1 and 3.2.

So far, there has been a great deal of consistency in the general structural architecture of the deposited crystal structures of the ECD for family B GPCRs (see Chapter 1). Most of these receptor ECDs, excluding the CGRP and AM receptors, are bound to endogenous ligands in their deposited structures. Despite their particular common fold, there is inconsistency in the way they bind their endogenous ligands (Miller *et al.*, 2012). This probably explains the basis of their ligand specificity, hence a reason to study individual receptor for their specific ligand binding and activation mechanism.



**Fig 3.1:** Representation of residues investigated in this chapter. (a) Schematic representation of CLR showing the amino acid sequence of the N-terminus. The residues investigated in this chapter are coloured yellow. The WDG motif is underlined. **NB:** only the N-terminus amino acid sequence is shown for the purpose of this figure. Amino acid sequence of the full receptor is shown in chapter 1. (b) Crystal structures of the N-terminal ECD of CLR showing selected residues investigated in this study for CLR-CGRP interaction. (i) Residues Trp(W)69, Tyr(Y)91, Phe(F)92, Lys(K)103, Arg(R)119 and Tyr(Y)124. (ii) Residues Asp(D)70, Trp(W)72, Asp(D)90 and Phe(F)95

Although the work of ter Haar and co-workers (2010) has provided a better understanding of the mechanism of CGRP receptor activation, a good number of questions are still left unanswered. The association of the receptor ECD with the antagonists in this structure cannot be considered a clear representation of the receptor conformation when bound to its endogenous ligands. Moreover, it does not serve as a tool to affirmatively define important residues required for the receptor pharmacology. Also, this particular structure is devoid of the TM and associated EC loops, which are required to bind the N-terminus of the ligand (see chapter 1 and Wheatley *et al.*, 2011) and have been shown to be important for receptor activation (see Barwell *et al.*, 2013). So it is not known what influence this region has on specific residue interactions of the ligand with receptor N-terminus, even though the CGRP receptor ECD, like other family B GPCRs, is capable of binding its endogenous ligand without the TM domain, albeit with lower affinity when compared to the full receptor (Koth *et al.*, 2010). The screening of the selected residues by alanine scan in this chapter was therefore designed to provide answers to some of these questions.

The residues selected (in Fig 3.2) were mutated to alanine and the effect of the mutation on receptor function analysed by assessing cAMP production, cell surface expression and  $\alpha$ CGRP-induced receptor internalization.

Overall, the aim was to select and investigate putative residues within/around the ligand-binding cleft of the CLR (in the CGRP receptor ECD complex) required for CGRP binding. This would be useful in predicting the binding pattern of CGRP to its receptor. Also, this would help to better understand the receptor activation mechanism, and could be a useful tool in developing drug candidates targeted at this receptor. The results obtained are presented and discussed in this chapter.

### **3.2 Method**

The general methods employed in this chapter are as described in Chapter 2. The method employed specifically in this chapter and not described under the general methods, in this case *in silico* mutagenesis, is however described below.

### 3.2.1 *In silico* alanine mutagenesis

Residues investigated in this chapter were also mutated to alanine *in silico* to check of the effect (of mutation) on CLR stability and CLR/RAMP1 interaction using the protein design program FoldX (Van Durme *et al.*, 2011) as a plug in for the YASARA (Yet Another Scientific Artificial Reality Application) molecular visualization programme (Krieger *et al.*, 2002). Briefly, FoldX is a computer algorithm primarily developed to evaluate the effect of mutations on the interactions contributing to the stability of proteins and protein complexes (Guerois *et al.*, 2002). This has now been modified for easier usability as a plug in for YASARA and estimates free energy difference between the WT and mutated protein ( $\Delta\Delta G$ ) and the change in interaction energy where the protein interacts with another in a complex, both in kcal/mol, (Van Durme *et al.*, 2011). The FoldX and YASARA (view) are both available for free download for non-commercial use (see Van Durme *et al.*, 2011). FoldX uses the following equation to calculate free energy in kcal/mol;

$$\Delta G = q(\Delta G_{\text{vdw}}) + r(\Delta G_{\text{solvH}}) + s(\Delta G_{\text{solvP}}) + t(\Delta G_{\text{wb}}) + u(\Delta G_{\text{hbond}}) + v(\Delta G_{\text{el}}) + w(\Delta G_{\text{kon}}) + x(T\Delta S_{\text{mc}}) + y(T\Delta S_{\text{sc}}) + z(\Delta G_{\text{clash}})$$

where  $\Delta G_{\text{vdw}}$  is total van der Waals forces from all atoms.  $G_{\text{solvH}}$  and  $G_{\text{solvP}}$  are respectively the difference in energy of solvation for non-polar and polar groups, when moving from unfolded to folded state.  $G_{\text{wb}}$  is the extra stabilising free energy obtained from a water molecule that makes more than one H-bond with protein.  $G_{\text{hbond}}$  is the difference in free energy of formation of H-bonds.  $G_{\text{el}}$  is the electrostatic contribution from interaction between charged groups while  $G_{\text{kon}}$  is the additional electrostatic contribution from interactions between atoms of different polypeptide chains.  $S_{\text{mc}}$  and  $S_{\text{sc}}$  are the entropy costs of fixing the backbone and a side chain respectively in a particular conformation.  $G_{\text{clash}}$  is the steric overlaps between atoms in the structure. T is temperature and  $q - z$  represent relative weights of the different energy terms used in the calculation (Schymkowitz *et al.*, 2005; Guerois *et al.*, 2002). While a positive  $\Delta G$  value implies disruption to structure stability or interaction between structures, a negative value implies the opposite.

The analysis was conducted using the crystal structure of the CGRP receptor (CLR/RAMP1) ectodomain complex in PDB (PDB code: 3N7S) as template. Residues were analysed one at a time.

Each residue to be mutated was selected and the 'mutate residue' command selected under the FoldX drop down menu. All boxes in the subsequent checkbox menu were checked to enable repair of template structure (i.e. 'RepairPDB' function, which energetically optimizes the PDB structure by eliminating small van der Waals clashes), calculate stability change and interaction energy change after which Ala was selected for the residue options menu as the desired substitution residue. Analysis parameter were as follows; temperature = 298K, pH = 7.0, ionic strength = 50 mM, van der Waals design = 2 and number of runs = 3. These are the default settings of the programme.

A difference in free energy ( $\Delta\Delta G$ ) value of around  $\pm 0.8$  kcal/mol is often considered FoldX error and thus insignificant while the threshold value above which difference in free energy values are considered significant is around  $\pm 1.6$  kcal/mol (Guerois *et al.*, 2002; Alibes *et al.*, 2010). In this current analysis therefore, average  $\Delta\Delta G$  values obtained from the 3 runs that are  $\geq 2$  kcal/mol or  $\leq -2$  kcal/mol are considered as potentially important while values between -2 and 2 (exclusive) are considered insignificant.

### **3.3 Results**

#### **3.3.1 Stimulation of cAMP production**

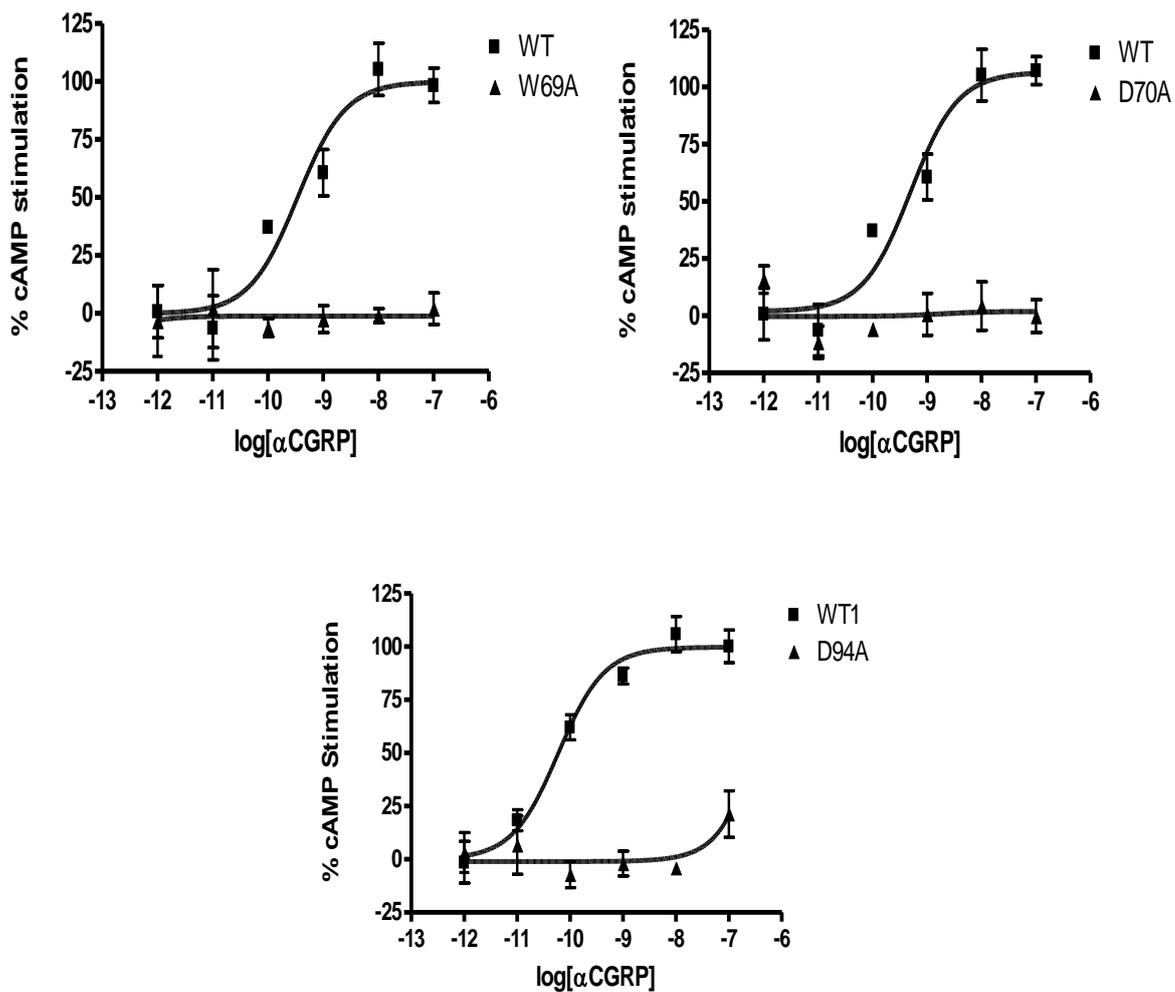
Each mutant residue was assayed for level of cAMP production in comparison to the wild type. The resulting data were fitted with sigmoidal dose-response curves and pEC<sub>50</sub> values generated (Table 3.1; Figs 3.2, 3.3 and 3.4). W69A, D70A and D94A produced no measurable cAMP response even at 100nM  $\alpha$ CGRP. The dose-response cAMP data generated did not fit into a sigmoidal dose-response curve like the WT, but instead produced a straight line (Fig 3.2) similar to that obtained for the mock receptor (i.e. HA-pcDNA3.1-/mycRAMP1; Fig 2.1). Due to the unmeasurable response observed for these mutants, the basal activity and maximum response were therefore undeterminable.

**Table 3.1:** Mean pEC<sub>50</sub> of WT and mutant receptors including percentage fold difference between WT and mutants. Values are presented as mean ± SEM. \* and \*\* represent p < 0.05 and p < 0.01 respectively. WT and mutant values were compared using unpaired *t*-test (two-tailed). Hyphen (-) indicates no measurable response. N = no of independent experiments

<b>Mutant</b>	<b>pEC<sub>50</sub> WT</b>	<b>pEC<sub>50</sub> Mutant</b>	<b>Fold decrease in αCGRP potency</b>	<b>N</b>
W69A	9.44 ± 0.18	-	-	7
D70A	9.55 ± 0.18	-	-	7
W72A	9.93 ± 0.19	8.29 ± 0.29*	44 fold	6
D90A	10.01 ± 0.04	9.59 ± 0.70*	3 fold	4
Y91A	9.94 ± 0.09	8.30 ± 0.17*	44 fold	4
F92A	9.88 ± 0.10	8.28 ± 0.07**	40 fold	5
D94A	9.80 ± 0.10	-	-	5
F95A	9.86 ± 0.17	7.59 ± 0.29*	186 fold	5
K103A	9.91 ± 0.14	8.11 ± 0.19*	63 fold	4
R119A	9.91 ± 0.14	8.16 ± 0.20*	56 fold	4
Y124A	9.76 ± 0.09	7.85 ± 0.09**	81 fold	5

**Table 3.2:** Mean  $E_{\max}$  and basal activity values of WT and mutant receptors. Values are presented as mean  $\pm$  SEM. \* and \*\* represent  $p < 0.05$  and  $p < 0.01$  respectively. WT and mutant values were compared using one way ANOVA. Hyphen (-) indicates no measurable response. N = no of independent experiments

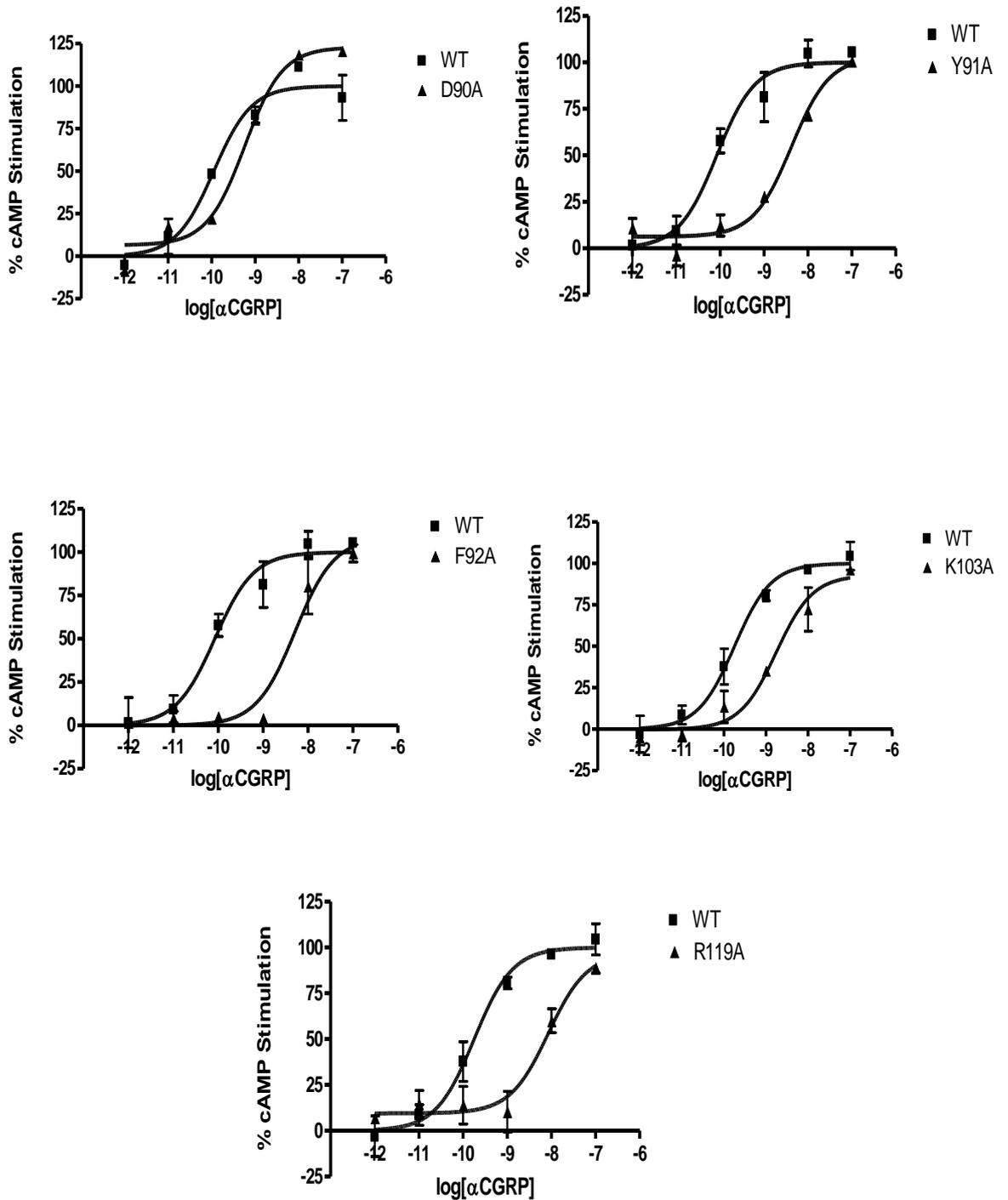
<b>Mutant</b>	<b>Mean <math>E_{\max}</math> (% WT)</b>	<b>Mean basal activity (%WT)</b>	<b>N</b>
W69A	-	-	7
D70A	-	-	7
W72A	$76.55 \pm 3.58^{**}$	$1.76 \pm 5.87$	6
D90A	$116.20 \pm 4.12$	$6.47 \pm 3.10$	4
Y91A	$101.80 \pm 9.86$	$2.42 \pm 1.88$	4
F92A	$94.87 \pm 9.16$	$0.03 \pm 2.99$	5
D94A	-	-	5
F95A	$69.80 \pm 8.56^*$	$2.39 \pm 6.40$	5
K103A	$97.62 \pm 11.10$	$1.12 \pm 3.47$	4
R119A	$95.01 \pm 6.76$	$12.93 \pm 8.24$	4
Y124A	$79.62 \pm 2.50^*$	$-2.69 \pm 9.17$	5



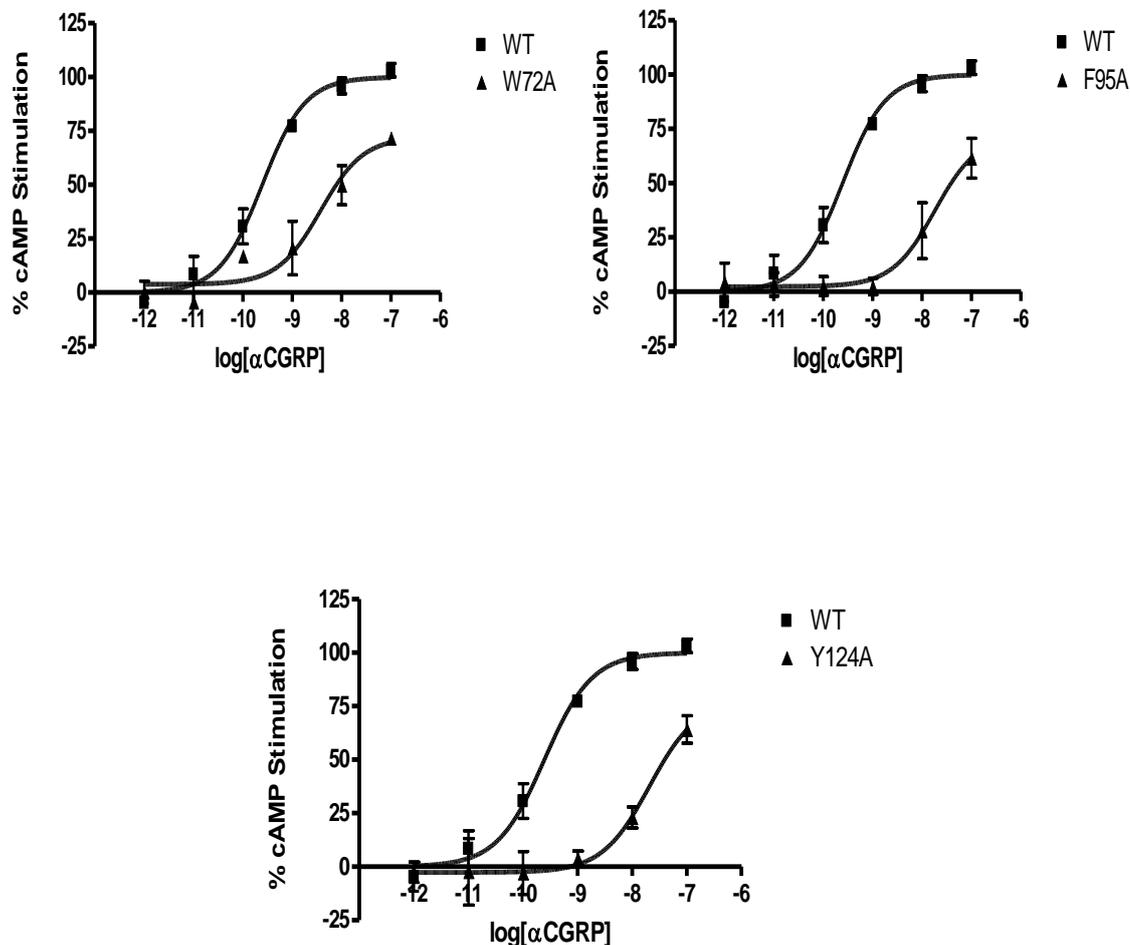
**Fig 3.2:** Representative dose-response curves of mutant receptors showing no measurable response to 100nM  $\alpha$ CGRP. WT and mutant receptors were challenged with 100nM – 10pM  $\alpha$ CGRP with a control assay point containing no  $\alpha$ CGRP. Sigmoidal curves were fitted using GraphPad prism 4. Each curve is representative of one of at least three independent experiments that best represents the mean  $pEC_{50}$ ,  $E_{max}$  or basal activity values. Each point on the curve represents duplicate assay data with standard error bars.

All other mutants showed significant decrease in  $\alpha$ CGRP potency as a measure of their  $pEC_{50}$  values compared to WT. F95A showed the largest fold decrease (186 fold) in potency while the least fold decrease (3 fold) was observed for D90A. The  $pEC_{50}$  values, with fold difference when compared to WT, are summarized in Table 3.1 above.

W72A, F95A and Y124A showed significant decrease in maximum cAMP levels (Table 3.2 and Fig 3.4), with the largest decrease (30%) observed for F95A. W72A and Y124A showed significant 24% and 20% decrease in mean  $E_{\max}$  respectively (Table 3.2). No significant shift in mean basal activity was observed for any of the mutants. The mean maximum and minimum responses are presented as percentage of WT in Table 3.2 above.



**Fig 3.3:** Representative dose-response curve of mutant receptors that showed significant decrease in  $\alpha$ CGRP potency compared to WT. WT and mutant receptors were challenged with 100nM – 10pM  $\alpha$ CGRP with a control assay point containing no  $\alpha$ CGRP. Sigmoidal curve was fitted using GraphPad prism 4. The curve is representative of one of at least three independent experiments that best represents the mean pEC<sub>50</sub>, E<sub>max</sub> or basal activity values. Each point on the curve represents duplicate assay data with standard error bars.

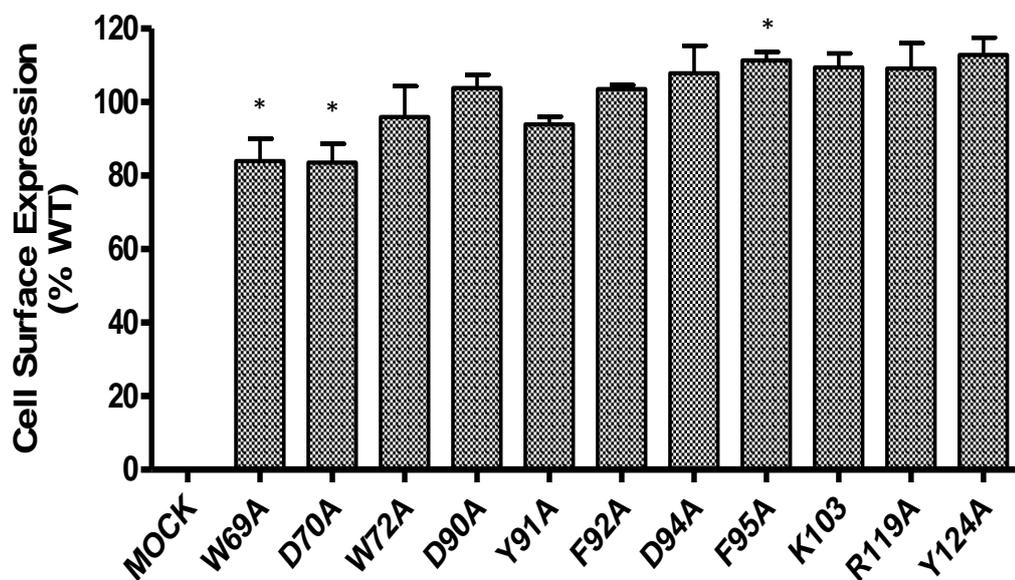


**Fig 3.4:** Representative dose-response curves of the mutant receptors with significant decrease in both  $\alpha$ CGRP potency and mean  $E_{max}$  compared to WT. WT and mutant receptors were challenged with 100nM – 10pM  $\alpha$ CGRP with a control assay point containing no  $\alpha$ CGRP. Sigmoidal curves were fitted using GraphPad prism 4. Each curve is representative of one of at least three independent experiments that best represents the mean  $pEC_{50}$ ,  $E_{max}$  or basal activity values. Each point on the curve represents duplicate assay data with standard error bars.

### 3.3.2 Cell surface expression

Cell surface expression of the mutant receptors as well as wild type was performed to determine the importance mutated residues in receptor cell surface expression. This was measured using cell-surface

ELISA. Resulting data were normalized and expressed as percentage of wild type (%WT) with the WT representing 100% cell surface expression. Reduction in cell surface expression was observed only for W69A ( $83.98 \pm 6.06$ ) and D70A ( $83.57 \pm 5.07$ ), both showing modest approximately 16% reduction. On the other hand, F95A ( $111.40 \pm 2.30$ ) caused a very subtle (11%) increase in cell surface expression. Y124A also resulted in a very mild increase in receptor cell surface expression as observed for F95A, although this was non-significant. Mutations to the other residues did not significantly affect cell surface expression. The cell surface expression data are graphically represented in Fig 3.5.

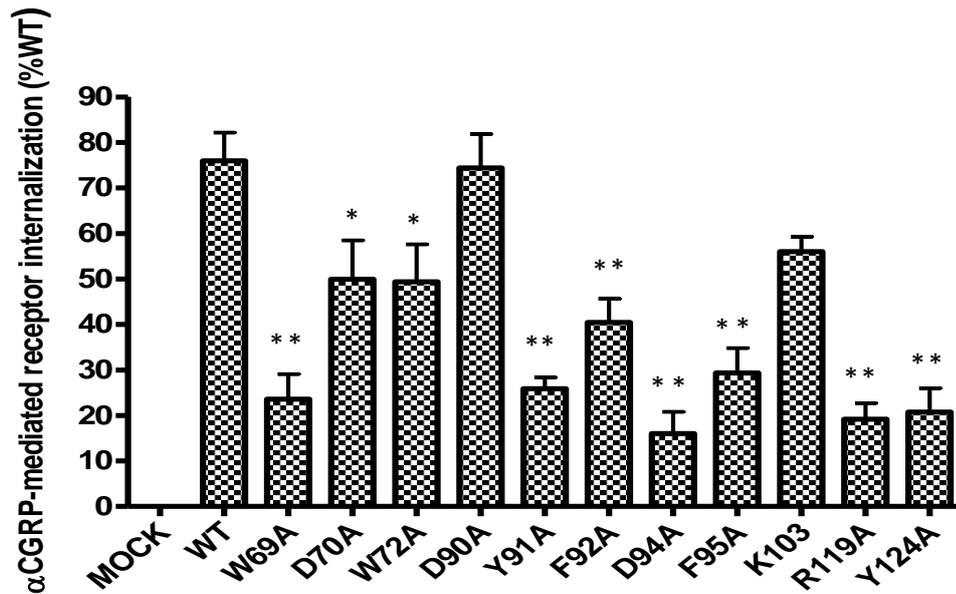


**Fig 3.5:** Cell surface expression of wild type and mutant receptors. The presence of the HA tag on WT and mutant receptors (HA-CLR/*myc*-RAMP1) was detected by ELISA. The negative control was HA-empty vector (pcDNA3.1-)/*myc*-RAMP1. Values were normalized and expressed as %WT. The WT and negative control represents 100% and 0% cell surface expression respectively. At least 3 independent experiments were performed and assay points were in triplicates or quadruplicates. Values were plotted on column graph using GraphPad prism 4 and each column represents mean  $\pm$  SEM. Mutant cell surface expression were compared to 100% using one way ANOVA. \* represents  $p < 0.05$ .

### 3.3.3 $\alpha$ CGRP-mediated receptor internalization

The mutants were analysed for their influence on  $\alpha$ CGRP-induced receptor internalization. The resulting values were compared to WT. The WT showed  $75.98 \pm 6.19\%$  internalization. Of all mutants investigated, 9 significantly reduced receptor internalization by varying degrees. 5 of these mutants reduced receptor internalization by 50% or more and they included D94A ( $16.03 \pm 4.79$ ), R119A ( $19.16 \pm 3.54\%$ ), Y124A ( $20.75 \pm 5.25\%$ ), W69A ( $23.56 \pm 5.66\%$ ) and Y91A ( $25.85 \pm 2.53\%$ ) representing approximately 60%, 57%, 55%, 52% and 50% reduction respectively when compared to the WT. Others reduced internalization by less than 50%. These included F95A ( $29.34 \pm 5.48\%$ ), F92A ( $40.45 \pm 5.21\%$ ), W72A ( $49.40 \pm 8.20\%$ ) and D70A ( $49.95 \pm 8.55\%$ ) with approximately 47%, 36%, 27% and 26% reduction respectively in comparison to WT.

Other mutants had no significant effect on  $\alpha$ CGRP-mediated receptor internalization. The values obtained for the WT and all mutants are represented in Fig 3.6.



**Fig 3.6:** Graph of  $\alpha$ CGRP-mediated receptor internalization of wild type and mutant receptors. The presence of the HA tag on WT and mutant receptors (HA-CLR/myc-RAMP1) was detected by ELISA following 1hr incubation with 100nM  $\alpha$ CGRP. Cell surface expression for the  $\alpha$ CGRP-treated mutant and WT receptors were expressed as percentage of the unstimulated WT receptor. % internalization was determined by subtracting the resulting values from 100. The WT and negative control represented 100% and 0% cell surface expression respectively. At least 3 independent experiments were performed and assay points were in triplicates or quadruplicates. Values were plotted on column graph using GraphPad prism 4 and each column represents mean  $\pm$  SEM. Mutant % internalization were compared to WT using Dunnett's multiple comparison test . \* and \*\* represent  $p < 0.05$  and  $p < 0.01$  respectively.

### 3.3.4 Inhibition of $^{125}$ I-hCGRP radioligand binding assay

Radioligand binding assay was performed for WT and mutant receptors from membrane preparations as earlier described (see chapter 2) to test for the role of mutants in ligand binding. However, it was not possible to obtain three successful independent experimental data as at the time of compiling this report due to technical issues and time constraints.

### 3.3.5 Summary of the N-terminal mutations

All mutants tested in this chapter that showed significant influence on at least one of the parameters, i.e. ability to stimulate cAMP, cell surface expression and agonist-induced internalization, are presented in Table 3.3.

**Table 3.3:** Summary of the criteria employed in probing the N-terminal mutant receptors. All values are approximate to the nearest whole number. NS = not significantly different to WT; NMR = no measurable response.

Mutant	Fold shift in $\alpha$ CGRP potency	Mean $E_{max}$ (% WT)	Cell surface expression (%WT)	$\alpha$ CGRP-induced internalization (%WT)
W69A	NMR	NMR	16% reduction	52% reduction
D70A	NMR	NMR	16% reduction	26% reduction
W72A	44 fold reduction	24% reduction	NS	27% reduction
D90A	3 fold reduction	NS	NS	NS
Y91A	44 fold reduction	NS	NS	50% reduction
F92A	40 fold reduction	NS	NS	36% reduction
D94A	NMR	NMR	NS	60% reduction
F95A	186 fold reduction	30% reduction	11% increase	36% reduction
K103A	63 fold reduction	NS	NS	NS
R119A	56 fold reduction	NS	NS	57% reduction
Y124A	81 fold reduction	20% reduction	NS	55% reduction

The result summary shows that virtually all the selected mutants show very significant variation from WT in the receptor's ability to stimulate cAMP and agonist-induced receptor internalization. It is interesting to note that some of these mutants (W69A, D70A and D94A) completely render the

receptor inactive at the maximum  $\alpha$ CGRP dose tested. Of all mutants investigated, only W69A, D70A and F95A appeared to significantly influence all the parameters tested.

Residues investigated here were also tested by *in silico* alanine mutagenesis to see the effect of mutations on both the stability of CLR and the interaction between CLR and RAMP1 in the CGRP receptor ectodomain complex; and more importantly, to compare these results to the ones earlier experimentally generated. The results are reported below.

### **3.3.6 *In silico* alanine mutation of residues**

The average values for the difference in free energy of formation of the CLR structure ( $\Delta\Delta G_f$ ), which represents its stability, and the difference in energy of interaction between CLR and RAMP1 ( $\Delta\Delta G_i$ ) for the WT and mutant structures were recorded for the 3 runs (Table 3.4). From the result, D70A caused the most instability to the CLR structure with a  $\Delta\Delta G_f$  value of 6.02 kcal/mol while R119A showed the least, insignificant effect with  $\Delta\Delta G_f$  of -0.24 kcal/mol. Other mutants which markedly affected CLR structure stability, in decreasing order of their  $\Delta\Delta G_f$  values, are Y124A (5.47 kcal/mol), W69A (4.84 kcal/mol), Y91A (4.56 kcal/mol), F95A (3.68 kcal/mol), F92A (2.70 kcal/mol) and W72A (2.34 kcal/mol). Interestingly, on the other hand, none of the mutants significantly affected the interaction between CLR and RAMP1 with their  $\Delta\Delta G_i$  values ranging only between -0.02 and 0.67 kcal/mol. W69A and R119A returned the two highest  $\Delta\Delta G_i$  values, although these were not considered important. These values are presented in Table 3.4.

**Table 3.4:** The difference in free energy of formation ( $\Delta\Delta G_f$ ) of CLR structure and energy of interaction between CLR and RAMP1 following *in silico* alanine mutagenesis. The CGRP receptor ectodomain crystal structure (PDB code: 3N7S; resolution 2.10 Å) was used as template. Values are average values from 3 runs.# indicates values considered important because they are  $\geq 2$  kcal/mol.

<b>Mutant</b>	<b>Average <math>\Delta\Delta G_f</math> (kcal/mol)</b>	<b>Average <math>\Delta\Delta G_i</math> (kcal/mol)</b>
W69A	4.84#	0.67
D70A	6.02#	0.02
W72A	2.34#	-0.02
D90A	0.23	0.01
Y91A	4.56#	0.19
F92A	2.70#	0.27
D94A	-0.18	0.0
F95A	3.68#	0.0
K103A	1.93	0.17
R119A	-0.24	0.31
Y124A	5.47#	-0.02

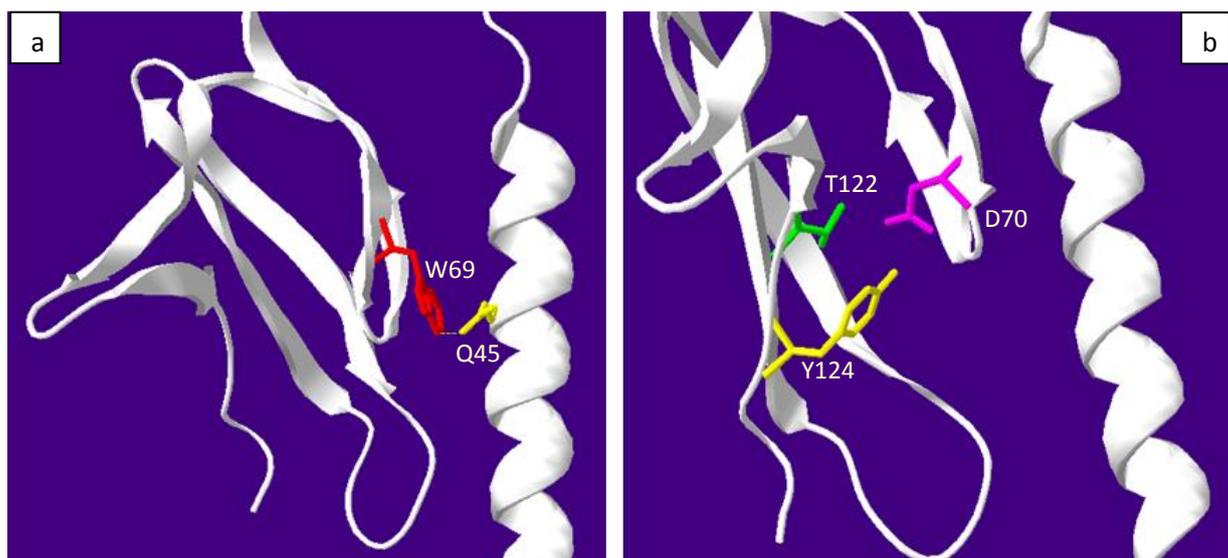
Overall, the results from the *in silico* mutagenesis show great deal of consistency with the experimental *in vitro* mutagenesis especially considering the non-pronounced effects these mutants had on CLR and RAMP1 interaction which is in agreement with the cell surface expression data. Plausible reasons for these observations and possible ligand-receptor interaction mechanism are discussed below.

### 3.4 Discussion

A clear mechanism of CGRP binding to its receptor is still unknown, and even though a structure exists for the receptor ECD, a structure for the CGRP is yet to be deposited in the protein data bank. In this chapter, some residues appearing to play very crucial role in CGRP binding to its receptor have been identified. These residues are present as part of different structural elements within the receptor ECD and would be interesting to understand the roles they play in the binding of CGRP to its receptor.

W69 and D70 are part of the WDG motif in family B GPCRs with D (D70 in CLR) invariantly conserved in all members. Both displayed a high level of importance in receptor signalling in that their mutation renders the receptor inactive. They have been thought to play a role in receptor stabilization (Kumar *et al.*, 2011). W69 forms a hydrogen bond with Q45 on the CLR  $\alpha$ -helix 1 (Ca1; Fig 3.7a). W69 together with Q45, I41 and A44 forms a hydrophobic cluster (Fig 3.13a) that stabilizes the orientation of the Ca1 towards the core of the CLR ECD in contribution to the stabilizing effect of the disulphide bond-forming residues C48 and C74 (ter Haar *et al.*, 2010). More importantly, this hydrophobic cluster is one of the few clusters that form the basis of ligand interaction within the putative binding cleft of this receptor and has been implicated for some other secretin family GPCRs where residues in corresponding positions, though less conserved, play a similar role (Parthier *et al.*, 2007; Pioszak and Xu, 2008; Pioszak *et al.*, 2009). These partner residues of W69 have been shown to significantly reduce  $\alpha$ CGRP potency and/or binding affinity (Barwell *et al.*, 2010) further substantiating the importance of this interaction, hence the importance of W69A. D70 on the other hand helps to stabilize the receptor structure by making backbone H-bonds with G71, W72 and L73. Its side chain also makes H-bonds with the side chains of T122 and Y124 (Fig 3.7b). This residue is a major stabilizing force of the  $\beta$ -hairpin structure in the CLR, like other family B GPCRs, which forms an integral part of the ligand binding cleft. The stability of this hairpin structure, which is structurally conserved in all family members with solved structures, is believed to be highly important for strengthening the hydrophobic cluster (involving W69) that holds the Ca1 in place (Pioszak *et al.*, 2007). This helix, as earlier discussed (in Chapter 1), is crucial for ligand binding. Moreover, the turn

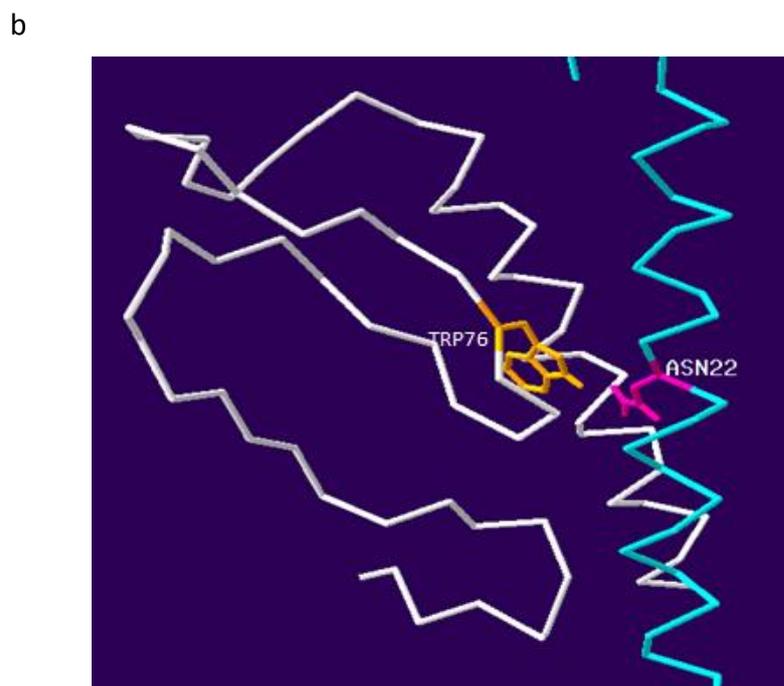
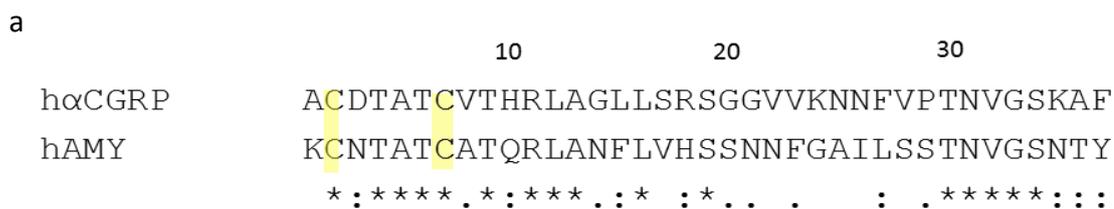
between the short antiparallel  $\beta$ -strands of this same  $\beta$ -hairpin appears to play a crucial role in holding and maintaining the orientation of W72 (Fig 3.1bii) which, in this chapter (from the results above and later discussed below), has been indicated to be very important in ligand binding. D70 mutation has been extensively explored in mouse CLR (equivalent residue is D69) where it has been mutated to Ala, Glu and Asn and also deleted as part of a W DGN ( $\Delta 68 - 71$ ) deletion mutant (see Ittner *et al.*, 2004). Interestingly, as observed for D70A in this chapter, the mouse CLR D69A mutant was inactive even at  $1\mu\text{M}$   $\alpha\text{CGRP}$  and co-immunoprecipitation with RAMP1 was significantly reduced when compared to WT receptor. D69E was only partially responsive at  $\geq 100\text{nM}$  while D69N did not affect receptor potency but reduced  $E_{\text{max}}$  by more than 70% when compared to WT (Ittner *et al.*, 2004). This further emphasizes the very important role of this residue in receptor stabilization and perhaps, in ligand binding.



**Fig 3.7:** Crystal structure of CLR ECD (PDB 3N7S) showing some interacting residues. (a) H-bond interaction between W69 (red) in the  $\beta$ -hairpin and Q45 (yellow) on CLR  $\alpha$ -helix 1 (C $\alpha$ 1). (b) D70 (purple) interacts with Y124 (yellow) and T122 (green). NB: the structure of RAMP1 has been omitted for clarity and this is same for subsequent Figures.

Although it seems very likely that W69 and D70 are involved in the stabilization of the ligand-binding groove, there is a possibility that they could work by disrupting signal transduction. The knowledge of this is limited by the absence of CGRP binding data for these mutant receptors. However, these

residues have been implicated in direct ligand interaction in some other family B GPCRs. In GIP-bound GIPR ECD structure (Parthier *et al.*, 2007) for instance, a Leu in position 26/27 was reported to participate in hydrogen bonding involving D66 (the equivalent of D70 in the GIP receptor) but it is not known whether this is a direct bond interaction or mediated by surrounding water molecules. Although the ligand-binding architecture observed for the agonist-bound ECDs of GIPR and PTH1R, when compared with those of the CRF1R and PAC1R, are thought to be closely related to the pattern proposed for the calcitonin peptide family (Parthier *et al.*, 2009), it is still difficult to extrapolate this to the CGRP receptor. In a model of amylin (AMY) bound to the calcitonin receptor, by Dr James Barwell (<http://www.pA2online.org/abstracts/Vol10Issue4abst026P.pdf>), a CGRP-binding role was observed for W69. This structure was first designed based on mutagenesis data from the lab of Dr Debbie Hay (University of Auckland, New Zealand) on the calcitonin receptor as well as AMY receptor and was modified based on additional data from this chapter. In this structure, W76 (W69 in CLR) interacts with N22\* (\* indicates residue of a ligand and hereafter is used for same purpose) of AMY (Fig 3.8b), which also interacts with Q52 (Q45 in CLR). Although W69 and Q45, as earlier mentioned above, are interaction partners within the ligand binding cleft, it is not certain whether they make similar interaction with the CGRP. AMY, for which solution NMR structures exist (Patil *et al.*, 2009; Nanga *et al.*, 2011), shares an appreciable 37% sequence identity (55% along the first 20 residues) with the CGRP (Fig 3.8a) and is considered the most relevant structure to this peptide (reviewed in Watkins *et al.*, 2012). It is therefore possible that W69 is involved in a similar interaction with V22\*/23\* of CGRP.



**Fig 3.8:** (a) Sequence alignment of human  $\alpha$ CGRP and AMY. Alignment was performed using CustalW2. “\*” indicates strictly conserved residues. “:” indicates residues with similar side-chain property and “.” indicates residues with similar shapes. The disulphide bond-forming Cys residues are highlighted in yellow. Both peptides are 37 amino acids in length. (b) A view of the CTR/AMY model showing interacting partners W76 of CTR and N22\* of AMY. The AMY structure is coloured blue. *Model was courtesy of Dr James Barwell*

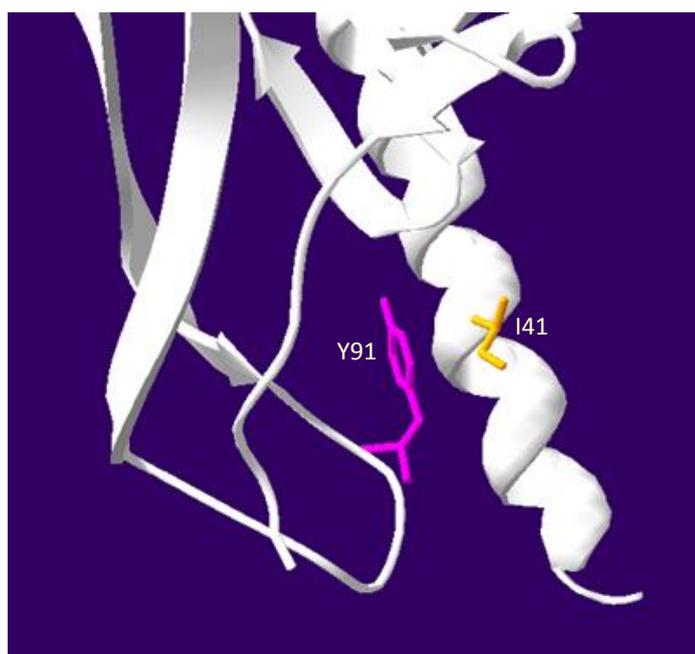
It is important to note that the total loss of response observed for W69A and D70A are most likely a cumulative effect of their role in stabilization of receptor ligand-binding pocket, ligand binding and reduced cell surface expression.

W72A caused significant 44 fold reduction in CGRP potency as seen in its decreased pEC<sub>50</sub> value when compared to WT. This is most likely a reflection of its central position in the antagonist-binding

groove where it is thought to interact with CGRP. Apart from this residue being indicated as making important interactions in the published structures of agonist-bound family B GPCR ECDs, it was also observed to be central in the binding of the antagonists, telcagepant and olcegepant, to CGRP receptor ECD (ter Haar *et al.*, 2010). These antagonists, in binding the ECD, lie across the ligand-binding cleft in such a way that they block the ligand binding groove and preventing CGRP binding. Considering the manner of binding of these antagonists and their interaction with W72, it strongly suggests that this residue is contained in a central position within the ligand-binding core of the receptor ECD and would interact with the ligand. This is supported by the effect this residue has on CGRP potency. Moreover, in a study by Yokoyama and co-workers on adrenomedulin (AM) receptor (i.e. CLR/RAMP2), W72A was found to heavily reduce affinity binding of AM as indicated by a 1,740-fold increase in  $K_D$  when compared to WT (Kusano *et al.*, 2011), suggesting that it is involved in direct binding of the ligand which also belongs to the calcitonin family of peptides and is expected to show appreciable similarities in their mode of binding. Although, in this current study, W72A did not affect receptor cell surface expression, its direct involvement in ligand binding cannot be ascertained since the  $K_D$  values for CGRP binding to these mutant receptors were not available and a role in signal transduction cannot be formally excluded.

D90, Y91, F92, D94 and F95 are part of the residues forming loop 4 of the CLR ECD. All of these residues significantly reduced CGRP potency, although the effect of D90A is very unpronounced compared to the others. Y91, F92, D94 and F95 are particularly important to this loop. Y91 contributes to the hydrophobic cluster around loop 2 which is important in forming the ligand-binding core. This residue is also packed against I41 on C $\alpha$ 1 (Fig 3.9). I41A has been shown to significantly reduce CGRP potency and specific binding and is suggested to be directly involved with CGRP binding (Barwell *et al.*, 2010). Considering these, and more importantly the loci of the residues in the receptor ECD, it is possible that Y91 interacts with CGRP within the putative  $\alpha$ -helical region (spanning positions 8-18) of CGRP, although its principal role appears to be in stabilizing the hydrophobic ligand binding cleft. F92 is another important residue as its mutation significantly reduced CGRP potency. It lies just below W72 at the tip of the  $\beta$ -hairpin structure and it is part of the

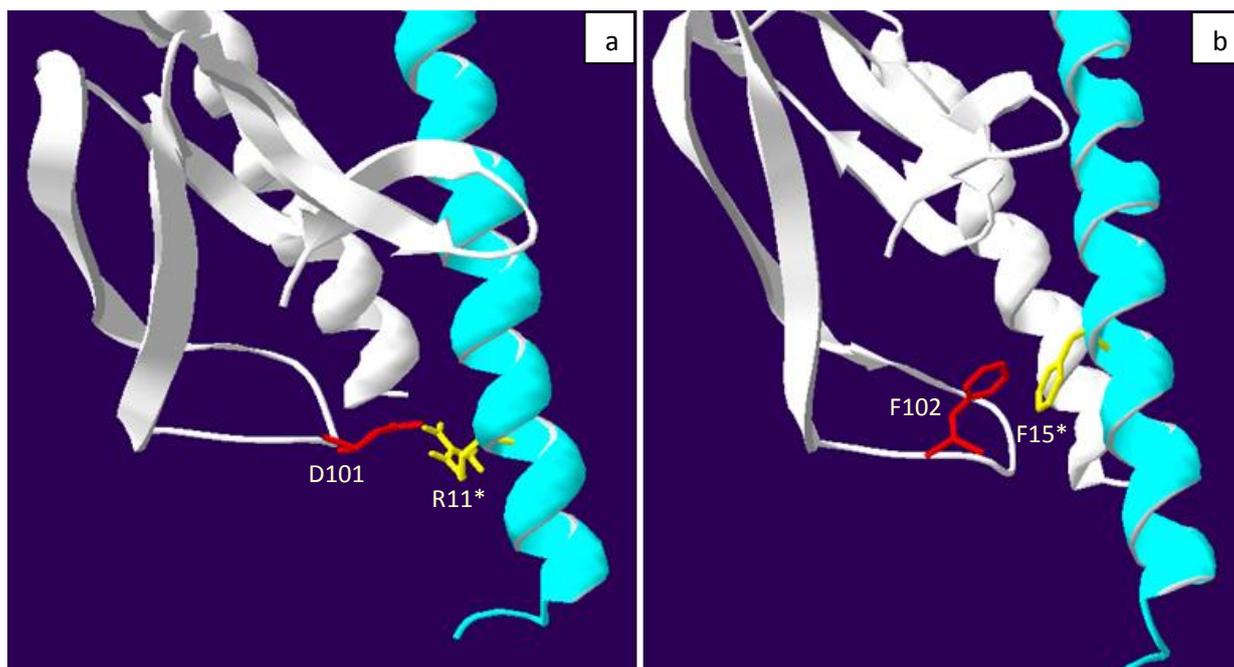
hydrophobic cluster (Fig 3.13b), between the  $\beta$ -turn and loop 2, which forms a huge portion of the ligand binding pocket. Like W72, F92 side chain interacts with the terminal pyridyl of olcegepant (see ter Haar *et al.*, 2010) suggesting a strong likelihood of CGRP interaction. In the AM receptor, F92 causes a huge 1,870-fold reduction in AM affinity binding (Kusano *et al.*, 2011). Both Y91 and F92 are conserved within human family B GPCRs with residues at these two positions coming from a small pool of hydrophobic Y, F, L, I and V (see Fig 1.5).



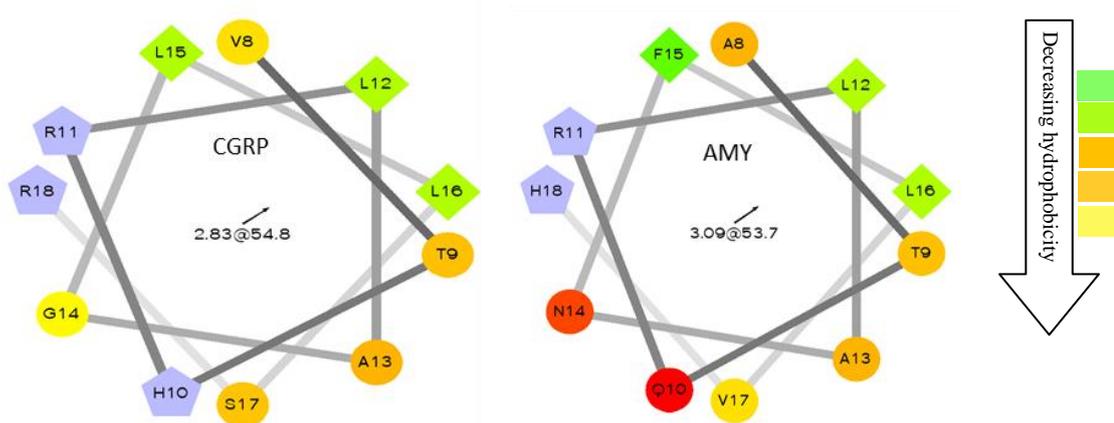
**Fig 3.9:** CLR ECD structure showing Y91 (purple) on loop 2 packed against I41 (orange) on Ca1.

The loss of response to 100nM  $\alpha$ CGRP observed upon Ala substitution of D94 indicates the extreme importance of this residue in receptor signalling, perhaps in ligand binding. D94 is not conserved in family B GPCRs and, considering its localization within the receptor structure, might be involved in conferring specificity in receptor-ligand interaction. However, it is strictly conserved in four (CLR, CTR, PTH1R and PTH2R). In PTH1R, D137 (the equivalent of D94) makes a pair of charged interactions with R20\* of PTH1 as well as hydrogen bonds with D29 and M32 and was said to be of high importance in binding affinity and specificity (Pioszak and Xu, 2008). It is therefore possible that D94 plays a similar role in the CGRP receptor. Also, the side chain carboxyl of D94 partakes in

hydrogen bonding with the terminal pyridyl of olcegepant (ter Haar *et al.*, 2010) suggesting that is also localized within the ligand binding groove of the receptor ECD. In the model structure by Dr Barwell, D101 (D94 in CLR) interacts with R11\* of AMY (Fig 3.10a). The Arg is invariantly conserved in CGRPs and AMY in human and across most species (see Fig 1.10 and Watkins *et al.*, 2012). D94 possibly interacts with same residue (R11\*) within the  $\alpha$ -helical region of CGRP. Moreover, Ala substitution of this residue (i.e. R11\*) has been found to reduce the affinity of CGRP<sub>37</sub> by up to 60 fold in SK-N-MC cells endogenously expressing the CGRP receptor (Howitt *et al.*, 2003). This indicates that this residue is important for affinity binding and hence further supports a possible interaction between this residue and D94, although other interaction partners might exist for R11\* and could be responsible for its effect.



**Fig 3.10:** The CTR/AMY model showing some interacting partners between CTR and AMY. The AMY structure is coloured blue. Residues in CTR and AMY are coloured red and yellow respectively (a) D101 and R11\*. (b) F102 and F15\*.



**Fig 3.11:** Helical wheel plot of CGRP and AMY  $\alpha$ -helical region spanning positions 8 – 18. Hydrophilic residues are presented as circles, hydrophobic residues as diamonds, potentially negatively charged as triangles, and potentially positively charged as pentagons. The most hydrophobic residue is coded green, and the amount of green decreases proportionally to the hydrophobicity, with zero hydrophobicity coded as yellow. Hydrophilic residues are coded red. The potentially charged residues are coded light blue. Arrows show the direction of the hydrophobic moment and so point toward the hydrophobic side of the helix.

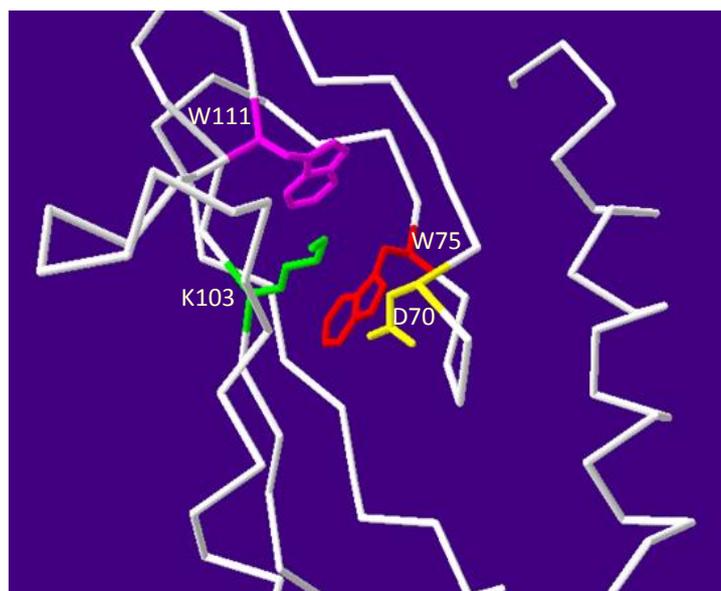
F95A showed the largest fold reduction in CGRP potency and highest percentage reduction maximum cAMP levels (Table 3.1), albeit not the largest reduction in agonist-mediated receptor internalization. F95 is part of the crucial and putative ligand-binding loop 2 of the CLR ECD and forms a part of the hydrophobic cluster (Fig 3.13b) around this region important for CGRP binding. It is only partially conserved in family B GPCRs as it is present as a Phe in 8 (including CTR, PTHRs and PAC1R but not in CRFRs or GIPR) of the 15 human family B GPCRs (see Fig 1.8). This residue has been indicated to be involved in ligand binding of PTH1 and PACAP (Pioszak and Xu, 2008; Kumar *et al.*, 2011) and, although the peptide was docked in the PACAP:PAC1R structure, this was however supported by alanine scan where F84 (the PAC1R equivalent of F95) was found to significantly reduce maximum cAMP levels. In the CTR/AMY model structure, F102 of CTR interacts with F15\* of AMY (Fig 3.10b). It is plausible that CLR makes a similar interaction with L15\* in CGRP, or more likely L16\*. Mutation to L16\* caused a 700-fold decrease in CGRP<sub>8-37</sub> potency while that of L15\* was a modest 5-fold decrease (Howitt *et al.*, 2003).

In the earlier mentioned study by Dr Hay's group, alanine mutation of the calcitonin receptor residues W79, F99, D101, F102 and Y131 (W72, F92, D94, F95 and Y124 respectively in CLR and the only residues in common with those investigated in this chapter) resulted in significant decrease in calcitonin (CT) potency (Proceedings of the British Pharmacological Society at <http://www.pA2online.org/abstracts/Vol10Issue4abst026P.pdf>). The same was observed for rAMY in presence of RAMP1 (unpublished data). This is in agreement with data obtained in this chapter and further supports the importance of these residues in receptor cAMP signalling. In addition to the similarities shared by CGRP and AMY, helical wheel plots (Fig. 3.10) show that their residues contained within the helical region appear on similar face of the helix. This suggests that they are likely to be buried in the binding pocket in the same orientation and it is therefore plausible to make a direct extrapolation from the CTR/AMY model. However, it is important to stress that as this is only a model, it requires further experimental verifications and inferences based on it are speculative. Moreover, there could be significant differences in binding conformation and specific interaction displayed by these peptides in binding their receptors.

It is again important to point out that, as earlier mentioned for the W69, D70 and W72 alanine mutants, it is difficult to specifically state whether these CLR residues are directly involved in CGRP binding or not, considering the absence of binding data. An alternative interpretation is that the residues are involved in linking agonist binding to receptor activation.

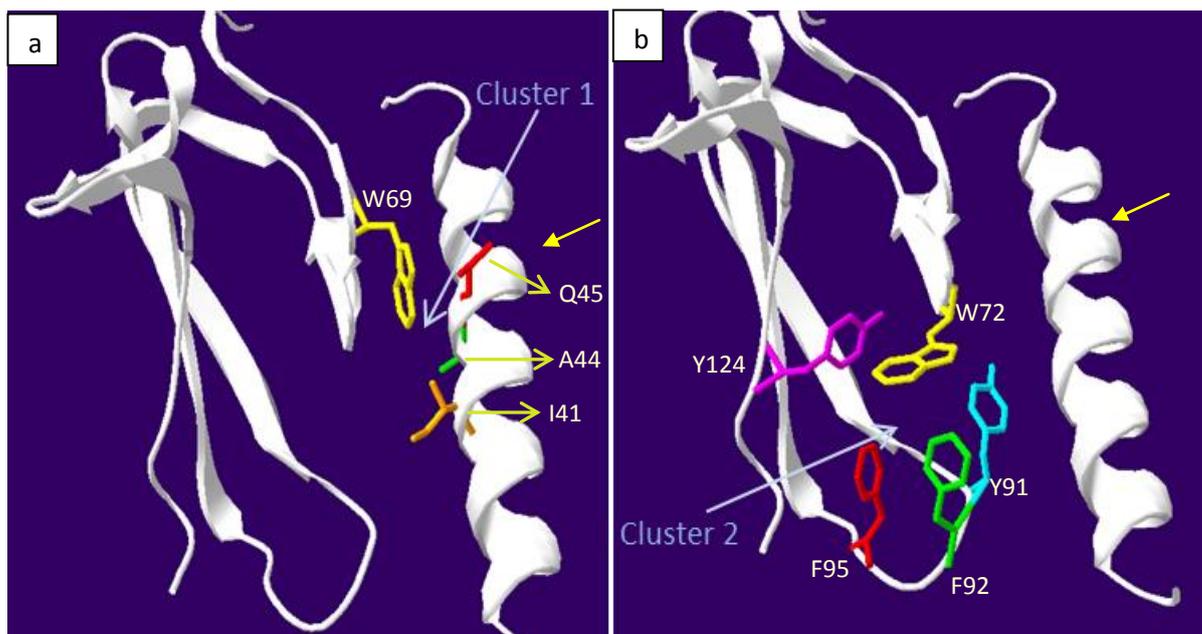
D90, apart from its mild effect on  $\alpha$ CGRP potency (Table 3.1; Fig 3.2), behaved like WT in receptor cell surface expression and internalization (Fig 3.4 and 3.5). In the CLR ECD structure (ter Haar *et al.*, 2010), D90 faces away from the putative binding cavity towards the CLR  $\alpha$ -helix 1 (C $\alpha$ 1) and does not appear to participate in ligand binding. This is an interesting result as it suggests the constitution of this core is confined among certain residues. Although this residue seems to interact with K40 on the C $\alpha$ 1, it is not known what importance this offers to receptor signalling or whether it is responsible for the effect observed for this residue.

K103A and R119 both showed ~63 and ~56 fold reduction in CGRP potency. Both basic residues lie on the same left half of the CLR ECD with each apparently acting in dissimilar manner. K103 is sandwiched between W75 and W111 (Fig 3.12) and serves as a cushion to keep these hydrophobic residues in place for structural stability. Its localization makes it very unlikely to make any contact with CGRP. Its stabilization role is consistent in other published family B GPCR ECD structures. This probably reflects the high conservation of this residue (and its sandwich partners) in all family members. In the mouse CRFR2 $\beta$  (Grace *et al.*, 2007) and human PAC1R (Kumar *et al.*, 2011) structures, it was shown to form a salt bridge with the invariantly conserved Asp (D70 in CLR) in the highly structurally conserved  $\beta$ -hairpin. Although this was not reported for the CLR, the residues appear to be in close proximity to make interaction. It is not known however if a salt bridge is formed between these residues *in vivo* during the activation process. K103 might be more involved in receptor activation rather than ligand binding, more so that it had no significant effect on receptor internalization. R119 has its side chain pointed towards the ligand binding domain and most likely interacts with CGRP. In the CTR/AMY model, R126 of CTR (equivalent of R119) interacts with S19\* of AMY. This residue is also conserved in CGRP and it is possible that it makes similar interaction with CGRP S19\*. Unlike K103, R119 is not conserved in family B GPCRs and may be one of the few residues that define ligand specificity.



**Fig 3.12:** CLR ECD structure showing K103 (green) sandwiched by W75 (red) and W111 (purple). The invariantly conserved Asp70 (yellow) in the  $\beta$ -hairpin, which may also interact with K103, is also shown.

Y124A showed significant reduction in both mean  $pEC_{50}$  and  $E_{max}$  (Table 3.1 and Fig 3.3). This residue has been indicated as part of the hydrophobic cluster that stabilizes the receptor ECD as well as the ligand-binding groove in, for instance, the CRFR1 (Pioszak *et al.*, 2008), PTH1R (Pioszak and Xu, 2008), GIPR (Parthier *et al.*, 2007) and PAC1R (where it reduces maximum cAMP levels by more than 30%; Kumar *et al.*, 2011) but hardly in ligand interaction. In the CLR ECD structure (ter Haar *et al.*, 2010), this residue contributes to, and probably stabilizes the hydrophobic cluster formed between the  $\beta$ -hairpin and loop 2 (Fig 3.13b). Considering its positioning within the ECD, it probably makes no direct interaction with CGRP. Its role in stabilisation of the CGRP binding pocket appears to be very crucial. It is packed against the pivotal W72 (discussed above) and the conserved V101. In addition, its side chain forms a hydrogen bond with the side chain of the very crucial D70 (discussed above). This interaction contributes to the orientation of D70 helping it make the necessary interactions.



**Fig 3.13:** CLR ECD structure revealing the two hydrophobic clusters that are proposed to form the ligand-binding groove. (a) Cluster 1: W69 (yellow), I41 (orange), A44 (green) and Q45 (red). (b) Cluster 2: W72 (yellow), Y91 (blue), F92 (green), F95 (red) and Y124 (purple). The yellow arrows show the direction of RAMP1 interaction.

Although there are currently no available competition ligand-binding data for these residues, there is appreciable evidence to show that almost all these residues directly participate in ligand binding. The normal cell surface expression of these mutants (except for W69A and D70A), when compared to WT, rules out the likelihood of a reduction in potency resulting from a poor availability of the receptor at the cell surface. The receptor internalization data are in general agreement with the cAMP data and, since receptor internalization is agonist-mediated, further supports the idea of these residues involved in ligand binding. It is important to point out however that despite the massive effect on the ability of the D70A and F95A mutants to stimulate cAMP when compared to most other mutants, they showed much less, but significant effect on agonist-mediated receptor internalization. A mutant like Y91A, on the other hand, with much less effect on cAMP stimulation compared to most other mutants showed a much higher effect on receptor internalization especially when compared to D70A and F95A (Table 3.3). This suggests that some of these residues, apart from their roles in ligand binding,

might influence the coupling of the receptor to  $\beta$ -arrestin. This could possibly be by slightly altering the N-terminus conformation, hence effecting the interaction of the N-terminus with the TM domain. D90A is internalized as the WT and this further suggests that it has no significant influence on ligand binding. Of all the mutants with pronounced reduction in CGRP potency, only K103A does not significantly alter receptor internalization. In other words, this residue does not necessarily affect receptor-ligand interaction but may play a role in receptor activation following ligand association. The highly conserved K103 could therefore be part of a system that dictates/controls Gs protein coupling and  $\beta$ -arrestin binding pathways.

The results from the FoldX alanine mutation analysis was in great deal of agreement with *in vivo* mutagenesis data discussed above. Firstly, as earlier mentioned, there were no pronounced disruptions to the CLR/RAMP1 interaction – an interaction important for the trafficking and cell surface expression of the CGRP receptor (McLathie *et al.*, 1998). This is consistent with the cell surface expression data where no mutant, except W69A and D70A, showed any significant decrease in receptor cell surface expression (Fig 3.5). The significant reduction in receptor cell surface expression exhibited by W69A and D70A however are mild. Interestingly, residues earlier suggested to exhibit their importance in ligand binding by rather playing a role in stabilizing the ligand-binding groove (i.e. W69, D70, Y91 and Y124) showed marked and much higher (the four highest)  $\Delta\Delta G_f$  values (D70A = 6.02 kcal/mol, Y124A = 5.47 kcal/mol, W69A = 4.84 kcal/mol and Y91A = 4.56 kcal/mol) when mutated to Ala compared to other mutants with similar effect on cAMP stimulation and receptor internalization (see Table 3.3 and 3.4). This however does not rule out the possibility of direct involvement of these residues in ligand interaction. The noticeable disruption to the stability of CLR N-terminus structure observed for F95A, F92A and W72A, as reflected by their  $\Delta\Delta G_f$  values (3.68 kcal/mol, 2.70 kcal/mol and 2.34 kcal/mol respectively), probably suggests these residues contribute, albeit to a lesser degree, to the stabilization of the ligand-binding core in addition to their role in ligand interaction. It is worth mentioning that despite the highly significant effect D94A had on cAMP stimulation and agonist-mediated receptor internalization (Table 3.3), it does not appear to affect the stability of CLR N-terminus structure. This plausibly supports the notion that D94 is strictly

involved in direct ligand interaction possibly via the putative D94 (CLR) and R11\*(CGRP) interaction (Fig 3.10a). While the FoldX algorithm has been shown to have an appreciably high success rate and in several cases has shown laudable consistency with experimental data (Guerois *et al.*, 2002; Schymkowitz *et al.*, 2005; Alibes *et al.*, 2010), it should only be considered a guide as the *in vivo* physiologic conditions cannot be completely imitated. Moreover, the template structure used here was the CGRP ectodomain structure which is devoid of the TM bundle and associated loops.

In conclusion, this study has been able to identify residues that constitute a putative binding cleft of CLR in the CGRP receptor. Based on the data obtained in this study and other available published data, it is strongly believed that CGRP binds in a somewhat similar overall architecture as observed for, most especially, the PTH1R in agreement with the two-domain model (Hoare, 2005). However, unlike other known family B ligand-ECD complexes, the CGRP is most likely bent towards the RAMP1 at the side of the  $\alpha$ -helix 1 via a kink formed by the tri-peptide Ser-Gly-Gly, which has been shown to be required for peptide potency (reviewed in Watkins *et al.*, 2012). The involvement of RAMP1 in the CGRP receptor makes it more complex compared to other family B GPCRs and even members of other families. While the structure of CGRP is still awaited, the major clusters that form the ligand-binding groove have here been drawn out (Fig 3.13). This could be helpful in designing drug candidates that target this receptor.

## **Chapter 4: CGRP receptor pharmacology: multiple residues within Helix 8 and its associated C-terminal region play significant roles in receptor signalling.**

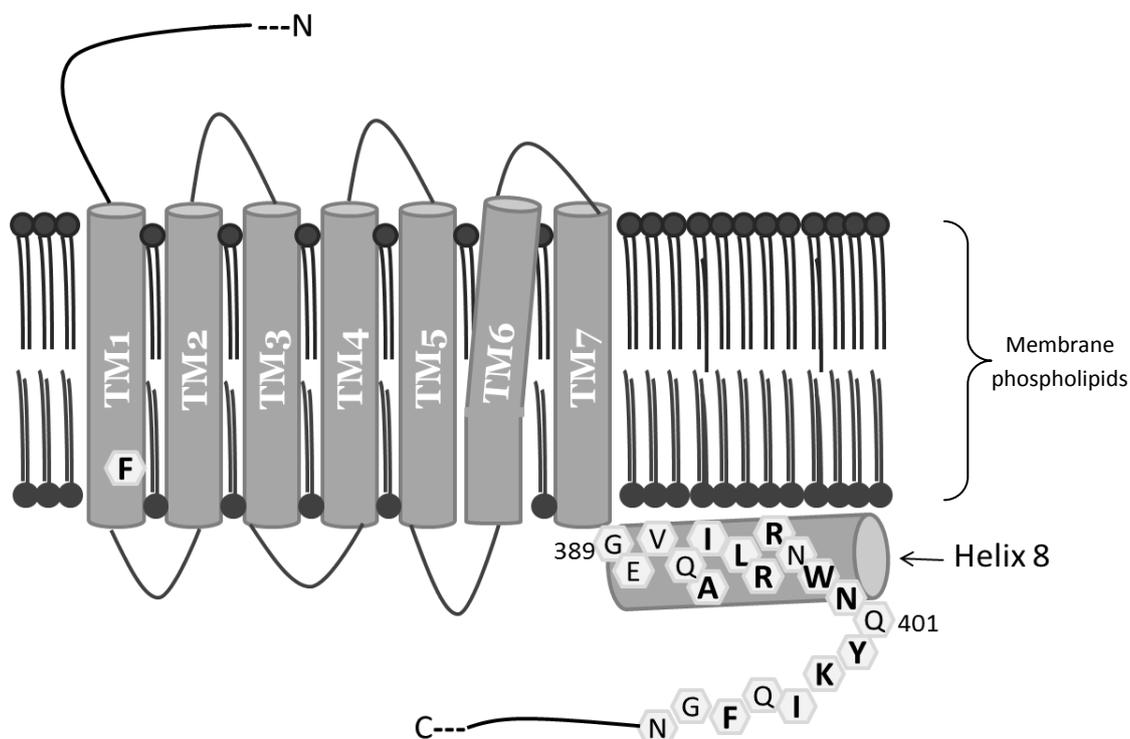
### **4.1. Introduction**

One of the contributions made by the crystal structures of some family A GPCRs in understanding the structural architecture of these receptors is the revelation of an eighth helix within the C-terminal intracellular domain of the receptors. The GPCR signal transduction cascade involves the interaction of intracellular G-proteins with the C-terminal tail of receptor (Wettschureck and Offermans, 2005). The C-terminal helix, unlike the transmembrane helices, is orientated in a parallel position to the lipid bilayer and perpendicular to the TM bundle axis. This part of the receptor has been shown to be of importance in receptor cell surface expression,  $\beta$ -arrestin binding and G-protein coupling. For instance, it has been shown to participate in conformational changes accompanying rhodopsin activation (Hoersch *et al.*, 2008) as well as  $\beta$ -arrestin-dependent receptor desensitization (Kirchberg *et al.*, 2011). It is difficult to make a direct and reliable extrapolation of these findings to family B GPCRs as they share little or no sequence similarity with family A GPCRs (Vohra *et al.*, 2013). Although no crystal structure currently exists for a family B GPCR, the presence and the importance of the helix 8 in some family B GPCRs have been postulated. Conner *et al.* (2008) showed that the C-terminal tail of the CLR contained an  $\alpha$ -helix at the origin and that it played a role in cell surface expression. More recently, Kuwasako *et al.* (2010) analysed the helix 8 of the CLR as part of the adrenomedulin 1 (AM<sub>1</sub>) receptor (i.e. CLR/RAMP2) and further stressed the importance of this region in receptor signalling. The postulated CLR helix 8 is believed to run from E389 to W399 (Conner *et al.*, 2008; Kuwasako *et al.*, 2010). Despite the previous work done on the CLR 8<sup>th</sup> helix, none has identified the key residues involved in the signal transduction pathways of the CGRP receptor (CLR/RAMP1). While Conner *et al.* (2008) studied the helix as a whole via deletions at specific residues, Kuwasako *et al.* (2010) studied individual residues in the AM<sub>1</sub> (not CGRP) receptor. In this chapter, multiple residues within this helical region of the C-terminal tail were analysed by alanine/aspartic acid scan mutagenesis for their role in receptor pharmacology.



Also investigated in this chapter is phenylalanine (F) 163 at the C-terminal end of TM1. The model structure of CLR TM bundle and helix 8 constructed by Vohra *et al* (2013) showed that helix 8 is held in close proximity to ICL1 and bottom end of TM1. This is a common feature of GPCRs seen in most crystal structures. Interaction between H8 and residues from the KSLS motif in ICL1 has been proposed. The KSLS motif is the family B equivalent of the KKLH motif in ICL1 of family A GPCRs (Vohra *et al.*, 2013). A close examination of this model structure revealed certain residues putatively involved in this interaction. One of these residues, situated at the bottom end of TM1, is F163 but it is yet to be investigated. This hydrophobic residue was mutated to alanine and the larger methionine to see what effect these have on receptor signalling.

All residues examined in this chapter, as well as their positions within the receptor are presented in Fig. 4.2. These residues' roles were probed on multiple criteria, which included their ability to stimulate cAMP production, cell surface expression and agonist ( $\alpha$ CGRP)-mediated receptor internalization.



**Fig 4.2:** Schematic representation of CLR showing the residues (highlighted in bold) investigated in this chapter.

Summarily, this chapter was aimed at investigating several residues making up the membrane-parallel helix 8 and associated residues at the distal end of this helix, as well as residue F163 at the bottom of TM helix 1 of CLR. This is in order to gain further understanding of their role in CGRP receptor signalling. The experimental data obtained are presented and discussed below.

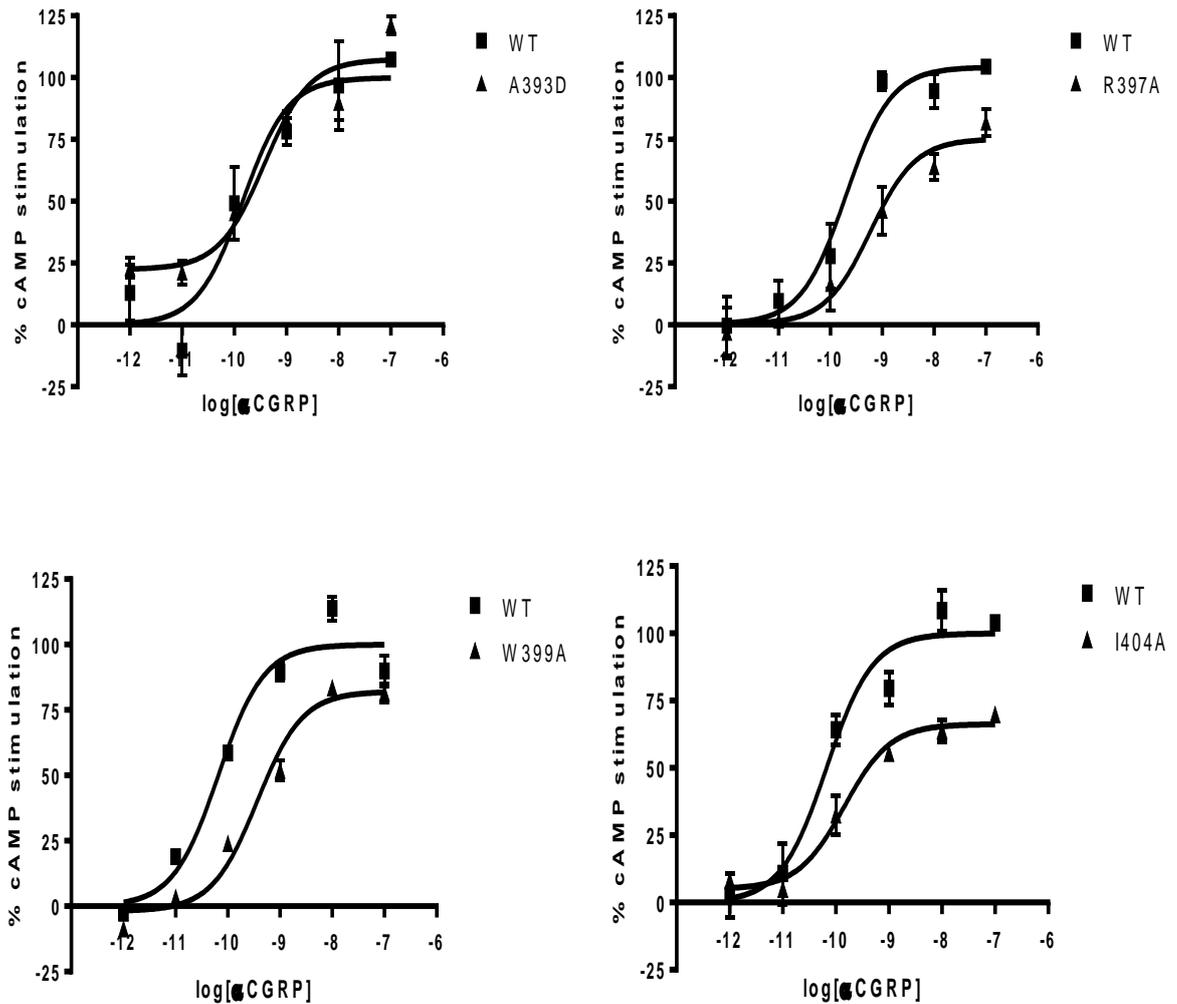
## **4.2. Results**

### **4.2.1 Stimulation of cAMP production**

Each mutant receptor was assayed for its level of cAMP production in comparison to the wild type receptor. The resulting data were fitted with sigmoidal dose-response curves and  $pEC_{50}$  values generated. Comparing the mean  $pEC_{50}$  values for CGRP of mutants with wild type (WT) revealed a modest but significant decrease in potency for A393D (~2 fold), R397A (~3 fold), W399A (~7 fold) and I404A (~3 fold). Of these residues, R397A, W399A and I404A showed decrease in mean  $E_{max}$  by approximately 22%, 13% and 33% respectively with no significant shift in mean basal activity. A393D on the other hand showed no significant decrease in mean maximum response but significant increase (~21%) in mean basal activity (see Table 4.1 and Fig. 4.3).

**Table 4.1:** Mean pEC<sub>50</sub>, E<sub>max</sub> and basal activity values of WT and mutant receptors. Values are presented as mean ± SEM. \*, \*\* and \*\*\* represent p < 0.05, p < 0.01 and p < 0.001 respectively. WT and mutant pEC<sub>50</sub> values were compared using unpaired *t*-test (two-tailed) while mutants' mean E<sub>max</sub> and basal activity were compared to WT using one way ANOVA

Mutants	pEC <sub>50</sub> WT (mean ± SEM)	pEC <sub>50</sub> Mutant (mean ± SEM)	Mean E <sub>max</sub> (% WT)	Mean basal activity (%WT)	N
F163A	9.90 ± 0.13	9.75 ± 0.22	76.15 ± 6.10**	0.88 ± 3.8	4
F163M	9.71 ± 0.16	10.34 ± 0.12*	107.60 ± 3.48	28.95 ± 7.49*	4
A393D	9.94 ± 0.08	9.65 ± 0.05*	110.2 ± 6.45	21.11 ± 4.85**	5
I394A	9.89 ± 0.04	10.04 ± 0.09	100.9 ± 15.1	0.87 ± 3.10	3
L395A	9.99 ± 0.12	10.12 ± 0.29	82.09 ± 4.56*	0.41 ± 1.88	3
R396A	9.93 ± 0.06	9.67 ± 0.20	87.11 ± 16.3	1.36 ± 2.54	3
R397A	9.80 ± 0.10	9.32 ± 0.10*	77.98 ± 10.22*	-0.62 ± 3.40	4
W399A	10.05 ± 0.07	9.22 ± 0.18*	87.07 ± 4.46*	0.18 ± 0.40	4
N400A	10.05 ± 0.08	9.89 ± 0.29	104.7 ± 5.71	35.50 ± 4.24*	4
Y402A	10.01 ± 0.08	9.70 ± 0.25	98.42 ± 8.73	26.77 ± 5.17**	4
K403A	10.24 ± 0.05	9.85 ± 0.25	90.22 ± 12.95	-2.13 ± 8.31	3
I404A	10.24 ± 0.05	9.82 ± 0.04**	67.26 ± 3.80***	3.85 ± 2.92	3
F406A	10.10 ± 0.14	10.03 ± 0.03	89.0 ± 8.89	2.21 ± 0.94	4

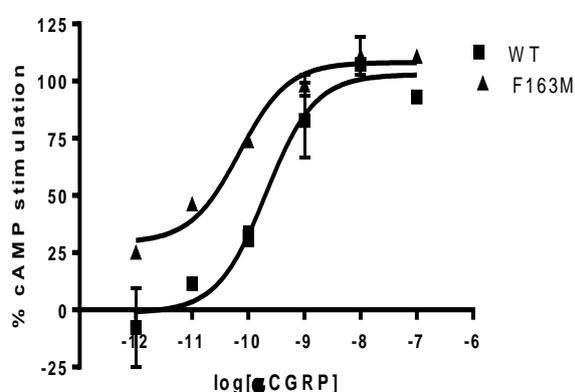


**Fig 4.3:** Representative dose-response curves of the mutant residues with significant decrease in  $\alpha$ CGRP potency as well as decrease in mean  $E_{max}$  or increase in mean basal activity compared to WT. WT and mutant receptors were challenged with 100nM – 10pM  $\alpha$ CGRP with a control assay point containing no  $\alpha$ CGRP. Sigmoidal curves were fitted using GraphPad prism 4. Each curve is representative of one of at least three independent experiments that best represents the mean  $pEC_{50}$ ,  $E_{max}$  or basal activity values. Each point on the curve represents duplicate assay data with standard error bars.

Of all mutants analysed, only F163M was found to show an increase (~5 fold) in potency for CGRP. This was in addition to a significant increase (~29%) in basal activity (Table 4.1). There was however no significant increase in mean  $E_{max}$  observed for this mutant. The dose-response curve for this mutant is presented in Fig. 4.4.

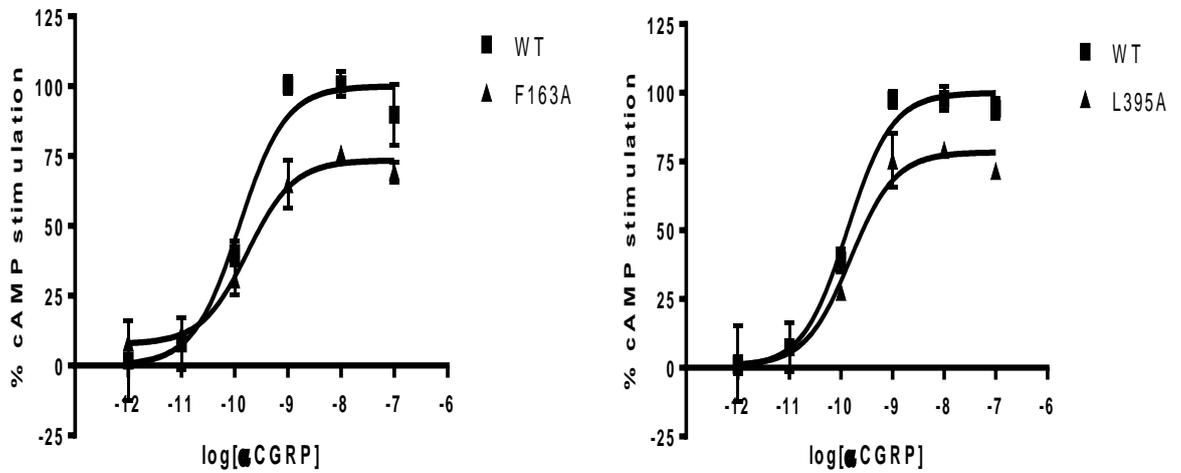
A significant decrease in mean  $E_{\max}$  or increase in mean basal activity was observed for F163A, L395A, N400A and Y402A. While F163A and L395A showed significant decrease in  $E_{\max}$  by approximately 24% and 18% respectively, N400A and Y402A respectively showed ~36% and ~27% increases in basal activity (see Table 4.1).

All other mutants behaved like the WT in their response to  $\alpha$ CGRP. A summary of the  $pEC_{50}$  values as well as the percentage shifts in mean  $E_{\max}$  and basal activities are presented in Table 4.1.

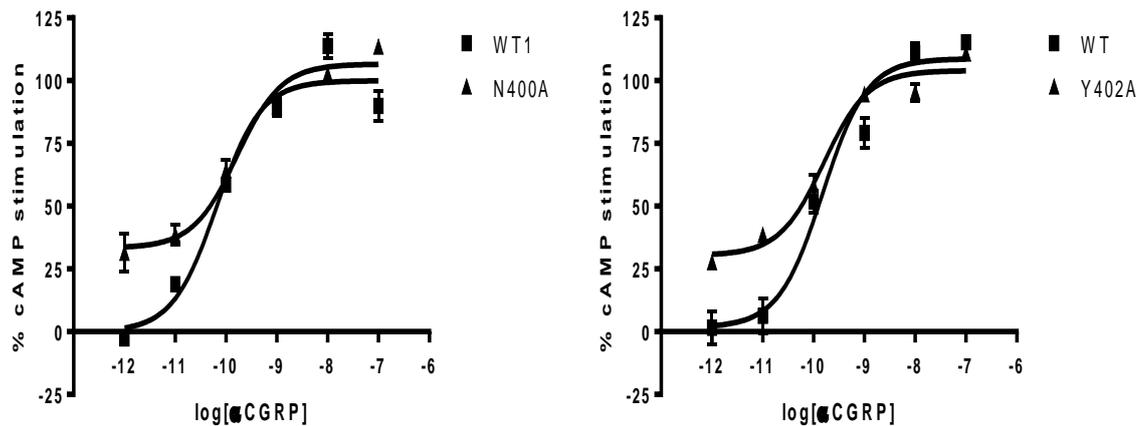


**Fig 4.4:** Representative dose-response curve of the mutant residue with significant increase in  $\alpha$ CGRP potency as well as increase in mean basal activity compared to WT. WT and mutant receptors were challenged with 100nM – 10pM  $\alpha$ CGRP with a control assay point containing no  $\alpha$ CGRP. Sigmoidal curve was fitted using GraphPad prism 4. The curve is representative of one of at least three independent experiments that best represents the mean  $pEC_{50}$ ,  $E_{\max}$  or basal activity values. Each point on the curve represents duplicate assay data with standard error bars

A.



B.

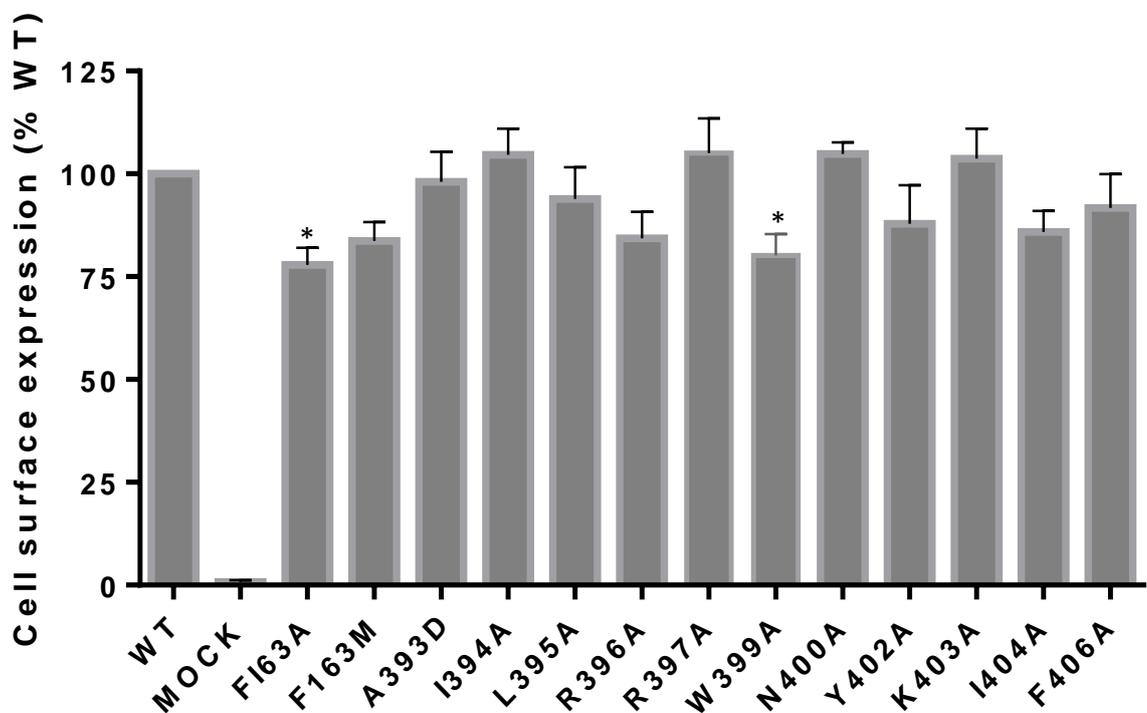


**Fig 4.5:** Representative dose-response curve of the mutant residues with (A) significant decrease in mean  $E_{max}$  or (B) significant increase in mean basal activity compared to WT. WT and mutant receptors were challenged with 100nM – 10pM  $\alpha$ CGRP with a control assay point containing no  $\alpha$ CGRP. Sigmoidal curves were fitted using GraphPad prism 4. Curves are representative of one of at least three independent experiments that best represents the mean  $pEC_{50}$ ,  $E_{max}$  or basal activity values. Each point on the curve represents duplicate assay data with standard error bars.

#### 4.2.2 Cell surface expression

Cell surface expression of the mutant receptors as well as wild type was measured employing cell-surface ELISA. Data were normalized and expressed as %WT while the WT represents 100% cell surface expression. When compared to WT, only F163A ( $77.82 \pm 4.19$ ) and W399A ( $80.15 \pm 9.84$ ) showed significant reduction in cell surface expression – both showing ~ 22% and ~20% reduction respectively. All other mutants behaved like WT in their expression at the cell surface (Fig. 4.6).

It is noteworthy that mutant residues F163M and R396A showed the highest not significant reduction in cell surface expression with each expressed 16% less than WT.



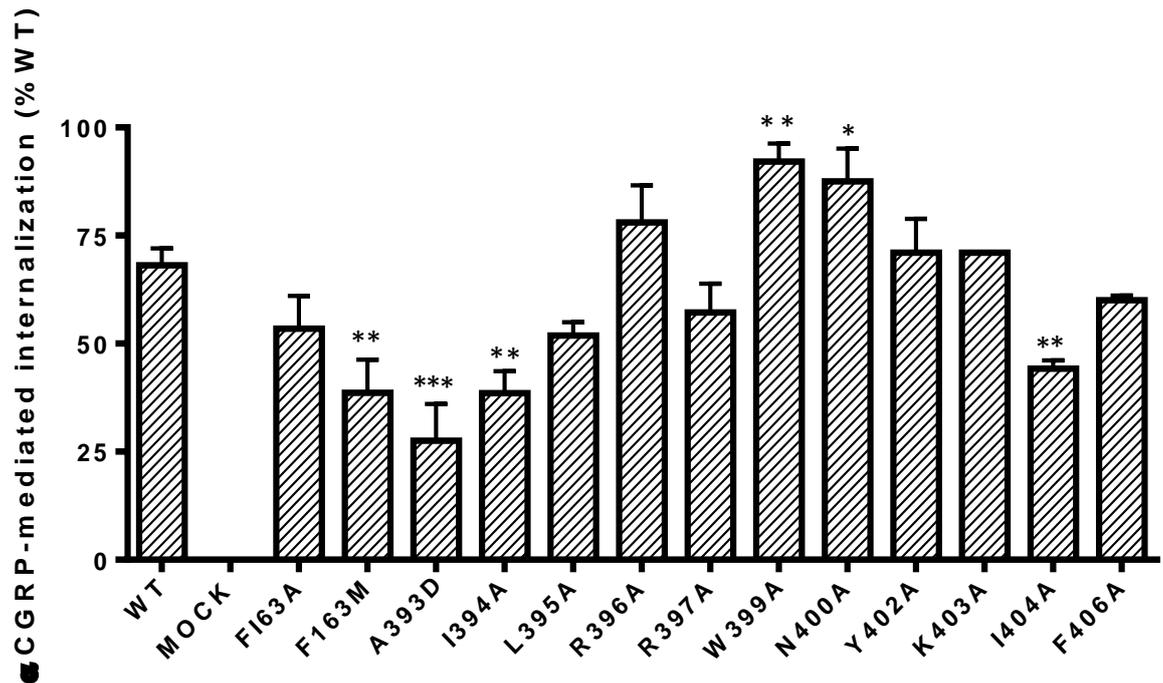
**Fig 4.6:** Cell surface expression of wild type and mutant receptors. The presence of the HA tag on WT and mutant receptors (HA-CLR/*myc*-RAMP1) was detected by ELISA. The negative control was HA-empty vector (pcDNA3.1-)/*myc*-RAMP1. Values were normalized and expressed as %WT. The WT and negative control represents 100% and 0% cell surface expression respectively. At least 3 independent experiments were performed and assay points were in triplicates or quadruplicates. Values were plotted on column graph using GraphPad Prism 4 and each column represents mean  $\pm$  SEM. Mutant cell surface expression were compared to WT using column statistics. \* represents  $p < 0.05$ .

### 4.2.3 $\alpha$ CGRP-mediated receptor internalization

The mutants were analysed for their influence on  $\alpha$ CGRP-induced receptor internalization. The resulting values were compared to WT and their significant difference statistically tested. Of all mutants, A393D showed the largest most significant reduction in receptor internalization with a % internalization value of  $27.55 \pm 8.51$  compared to WT ( $68.09 \pm 3.95$ ), which represents a 59.6% reduction. Other residues, which showed reduced internalization included F163M ( $38.64 \pm 7.62$ ), I394A ( $38.52 \pm 5.15$ ) and I404A ( $44.24 \pm 1.89$ ) representing 43.2%, 43.5% and 35% reduction respectively.

On the other hand, a significant increase in agonist-induced internalization was observed for W399A ( $92.09 \pm 4.23$ ) and N400A ( $87.56 \pm 7.59$ ) compared to  $68.09 \pm 3.95$  for WT. These represent ~35% and 29% increase in agonist-induced internalization by W399A and N400A.

Other mutants had no significant influence on receptor internalization. The values obtained for the WT and all mutants are presented in Fig. 4.7.



**Fig 4.7:** Graph of  $\alpha$ CGRP-mediated receptor internalization of wild type and mutant receptors. The presence of the HA tag on WT and mutant receptors (HA-CLR/*myc*-RAMP1) was detected by ELISA following 1h incubation with 100nM  $\alpha$ CGRP. Cell surface expression for the  $\alpha$ CGRP-treated mutant and WT receptors were expressed as percentage of the unstimulated WT receptor. % internalization was determined by subtracting the resulting values from 100. The WT and negative control represented 100% and 0% cell surface expression respectively. At least 3 independent experiments were performed and assay points were in triplicates or quadruplicates. Values were plotted on column graph using GraphPad prism 4 and each column represents mean  $\pm$  SEM. Mutant % internalization were compared to WT using Dunnett's multiple comparison test. \*, \*\* and \*\*\* represent  $p < 0.05$ ,  $p < 0.01$  and  $p < 0.001$  respectively.

#### 4.2.5. Summary of these mutations

All mutants that showed significant influence on at least one of the parameters, i.e. ability to stimulate cAMP, cell surface expression and agonist-induced internalization, tested in this chapter are presented in Table 4.2.

**Table 4.2:** Summary of the criteria employed in probing the mutant receptors. All values are approximate to the nearest whole number. ns = not significantly different to WT.

<b>Mutants</b>	<b>Fold shift in <math>\alpha</math>CGRP potency</b>	<b>Mean <math>E_{max}</math> (% WT)</b>	<b>Mean basal activity (%WT)</b>	<b>Cell surface expression (%WT)</b>	<b><math>\alpha</math>CGRP-induced internalization (% WT)</b>
F163A	ns	24% reduction	ns	22% reduction	ns
F163M	5 fold increase	ns	29% increase	ns	43% reduction
A393D	2 fold reduction	ns	21% increase	ns	60% reduction
I394A	ns	ns	ns	ns	44% reduction
L395A	ns	18% reduction	ns	ns	ns
R397A	3 fold reduction	22% reduction	ns	ns	ns
W399A	7 fold reduction	13% reduction	ns	20% reduction	35% increase
N400A	ns	ns	36% increase	ns	29% increase
Y402A	ns	ns	27% increase	ns	ns
I404A	3 fold reduction	33% reduction	ns	ns	35% reduction

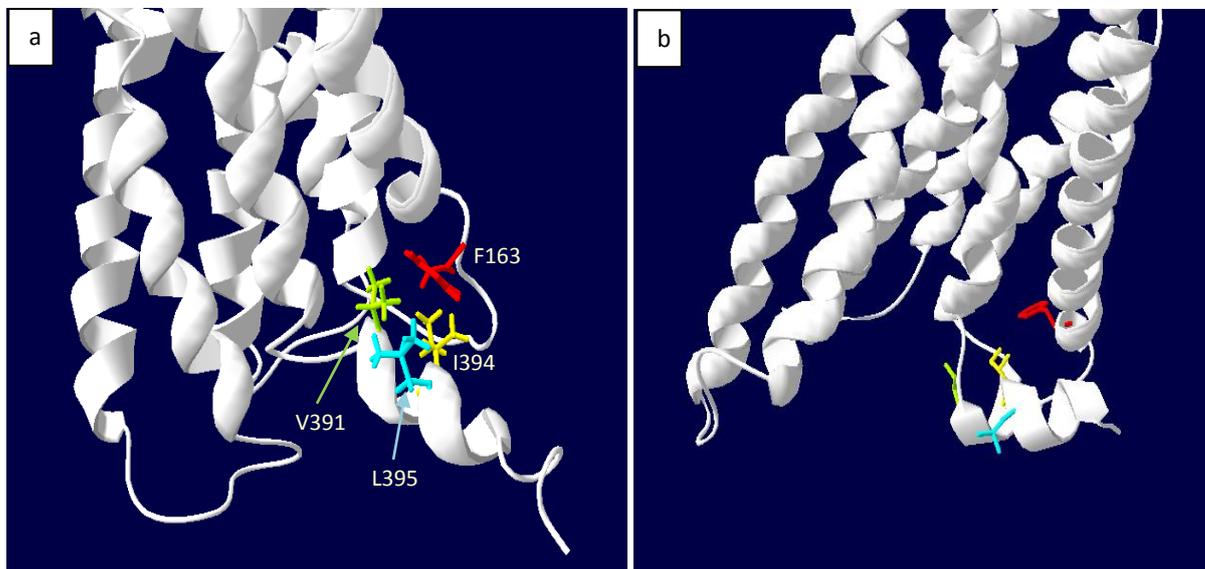
Overall, from the data obtained, mutations to residues F163, A393 and W399 appear to significantly influence most parameters tested. In all, the residues investigated also appear to almost evenly influence CGRP potency and CGRP-mediated receptor internalization. The least overall effects observed were for cell surface expression. These data are discussed below.

### 4.3. Discussion

The C-terminal domains of GPCRs are essential for receptor signalling as they are mostly responsible for interacting with cytoplasmic biomolecules (Bockaert *et al.*, 2004). Since the presence of a helix 8 in GPCRs was unambiguously revealed in the first GPCR crystal structure (Palczewski *et al.*, 2000), this helical region has been reported to play a role in receptor pharmacology in many family A GPCRs (Marin *et al.*, 2000; Swift *et al.*, 2006; Yasuda *et al.*, 2009; Parker and Parker, 2010) and also in some family B GPCRs (Couvineau *et al.*, 2003; Conner *et al.*, 2008). The residues involved in the crucial role of the region have not been well studied especially in the class B GPCRs. Some of these residues in the CGRP receptor were examined in this chapter with a few playing important roles. Also, F163 of CLR located on the bottom end of TM1 (part of the H8-interacting region postulated by Vohra *et al.* (2013)) was found to be important for receptor function in this chapter.

The F163A mutant showed significant ~24% and 22% decrease in mean  $E_{max}$  and cell surface expression respectively with no significant change in  $\alpha$ CGRP potency. Mutating this residue to a methionine (i.e. F163M) in contrast showed significant ~5fold increase in  $\alpha$ CGRP potency and ~29% increase in mean basal activity, implying a constitutive activity for this mutant receptor. These results, first of all, show the importance of this residue in receptor signalling. Using the structural model of Vohra *et al.* (2013), F163 appears to make hydrophobic interactions with V391, I394 and L395 in the inactive state (Fig. 4.8a) whereas in the active state, it is completely separated from these residues and points towards the base of TM2 (Fig 4.8b). A closer examination of the structure showed that the movement of this residue appears to have resulted from the unwinding of the base of TM1. This suggests that F163 might play a role in the maintenance of receptor inactive state as well as the structural switch between the inactive & active states of the receptor. Although one might expect F163M to foster the hydrophobic interaction, and help maintain the inactive state, the absence of an aromatic side chain in Met might explain this. In Fig 4.8a for instance, the aromatic ring of F163 is positioned in such a way that it gives room for some form of gliding movement over the putative interaction partners i.e. V391, I394 and L395. This sort of movement has been suggested for the aromatic ring of Trp residues in receptor activation, where they act to dampen such motions (Hulme,

2013). Mutation of F163 to a Met *in silico* showed that the hydrophobic tail of the Met residue faces downwards in-between the three interacting residues earlier stated. Further work is required to substantiate the exact mechanism of action of this residue.

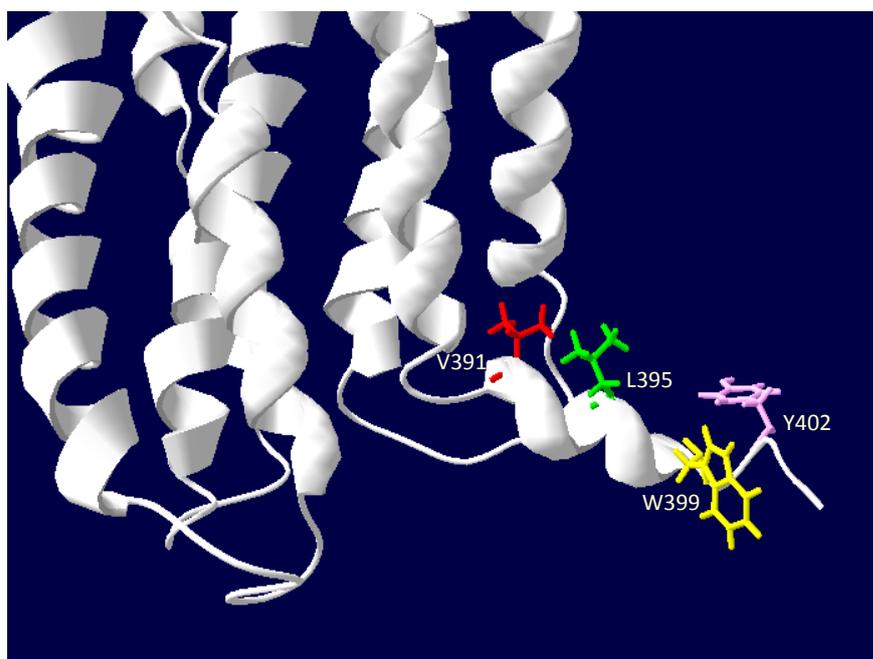


**Fig 4.8:** Model structure of CLR as postulated by Vohra *et al* (2013). (a) Inactive CLR showing putative interaction of F163 (red) in ICL1 with V391 (green), I394 (yellow) and L395 (blue) in H8. (b) Active form of CLR showing a change in receptor conformation and resultant prevention of the interaction observed in (a).

On the other hand, the larger size of the hydrophobic side chain of Met, relative to Ala, might explain the reduction in  $E_{\max}$  and cell surface expression observed for F163A but not for F163M. Met might be capable of maintaining the structural integrity in the mutant receptor in activation and cell surface expression as in the WT. It is however important to stress that other factors may play a role and that this model might not perfectly represent the actual receptor structure. This residue is yet to be investigated in any other family B GPCR.

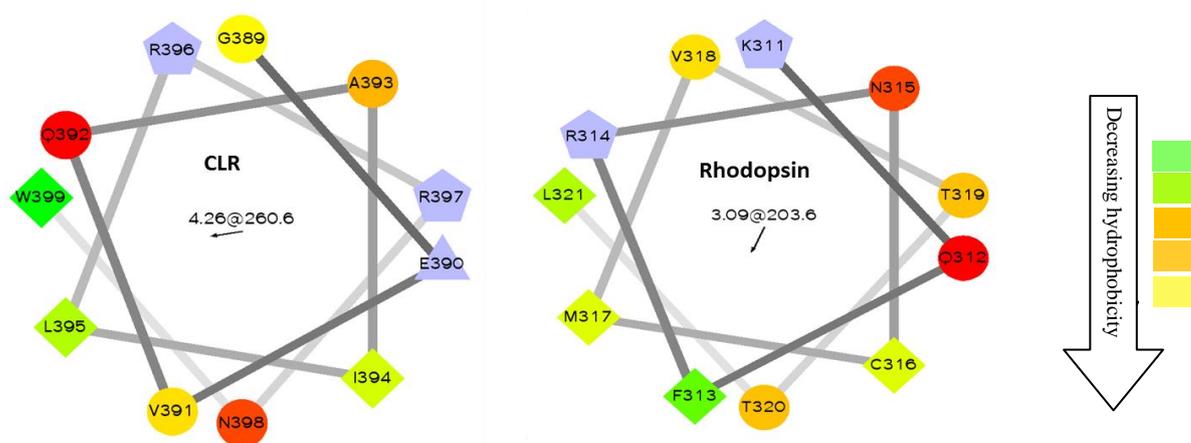
W399 appears to play crucial role in receptor pharmacology. W399A showed a significant  $\sim 7$  fold decrease in  $\alpha$ CGRP potency as well as a significant  $\sim 13\%$  decrease in maximum cAMP production suggesting a role in G-protein coupling. The small but significant reduction in cell surface expression ( $\sim 20\%$ ) for this mutant shows that it has a minor role in the receptor's cell surface expression. This is in agreement with previous studies on the CLR in the CGRP receptor (Conner *et al.*, 2008) and

AM1 receptor (Kuwasako *et al.*, 2010). Although the mechanism of action of W399 is still unknown, there has been the suggestion it acts as lipid membrane anchor (Conner *et al.*, 2008) performing similar role played by Cys in family A GPCRs. The orientation of W399 in the CLR model structure suggests it might interact with the phospholipid bilayer (Fig. 4.9), although its side chain would probably require some rotation to put it in closer proximity to the membrane as it appears to be slightly shifted sideways of the membrane core. A similar rotation has been observed in the crystal structure of olcegepant-bound CLR/RAMP1 ECD complex for the side chain of W72 within the N-terminus of the CLR where it (the W72 side chain) rotates  $\sim 70^\circ$  when compared to unliganded complex (ter Haar *et al.*, 2010). This rotation helps to form a “Trp shelf” on which the piperidine ring of olcegepant stacks. It is however not known, especially for the CLR helix 8, whether such rotation is applicable under normal physiological conditions.



**Fig 4.9:** Model structure of the CLR showing the hydrophobic residues V391 (red), L395 (green), W399 (yellow) and Y402 (purple) thought to be involved in hydrophobic interactions significant for receptor cell surface expression.

From a broader perspective, it is possible that the overall architecture of this helix was wrong since it was modelled based on homology with a family A GPCR. This, as earlier discussed (Chapter 1), could be very misleading based on variations observed between, and even within, GPCR families. On this note, a helical wheel was plotted for the helix 8 of CLR with the assumption that it forms a regular helix (Fig. 4.10). This was to have an independent picture of the orientation of residues making up the helix and observe them against the results obtained in this chapter. This plot was also conducted for rhodopsin (Fig. 4.10) to compare with the available crystal structure of this receptor. Comparison indicated that the wheel plot presented residues in virtually similar plane as the crystal structure. Worthy of mention, for instance, are residues F313, M317 and L321, which are orientated towards the membrane hydrophobic core as in the crystal structure. This comparison was also made for the  $\beta_2$ -AR (figure not shown). In the CLR, W399 appeared to be in the hydrophobic core supporting earlier suggestions of its potential interaction with the membrane bilayer - an interaction required for receptor stability and cell surface expression.

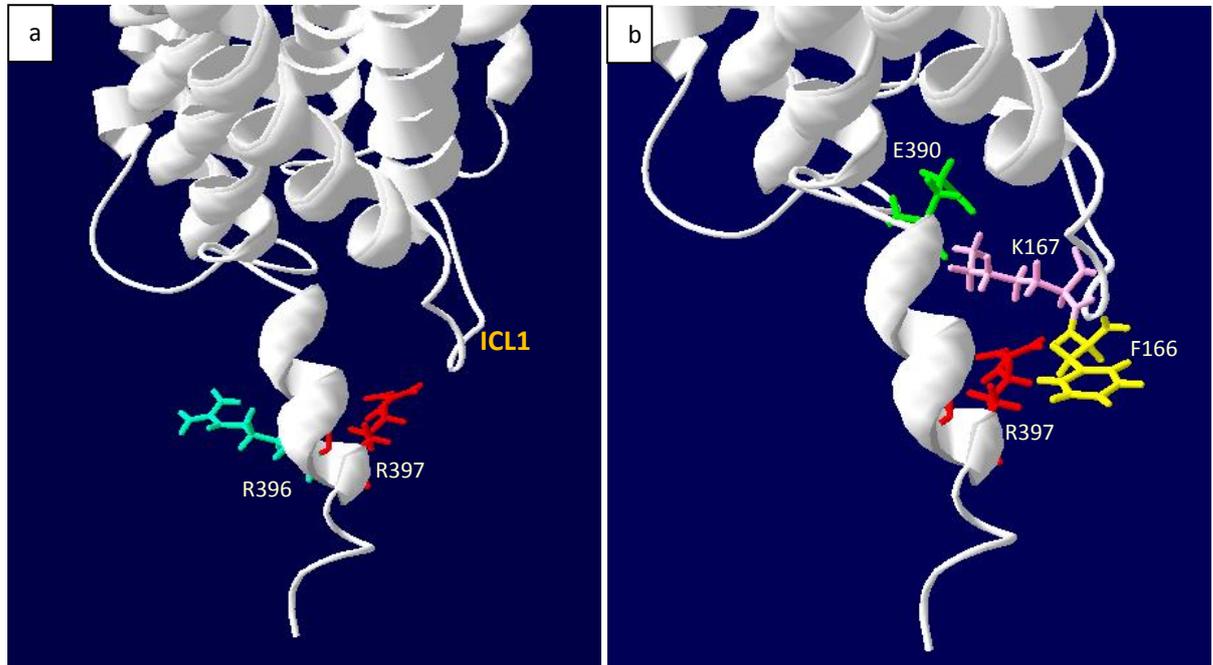


**Fig 4.10:** Helical wheel plots of CLR and rhodopsin 8<sup>th</sup> helices. Hydrophilic residues are presented as circles, hydrophobic residues as diamonds, potentially negatively charged as triangles, and potentially positively charged as pentagons. The most hydrophobic residue is coded green, and the amount of green is decreases proportionally to the hydrophobicity, with zero hydrophobicity coded as yellow. Hydrophilic residues are coded red with pure red being the most hydrophilic (uncharged) residue, and the amount of red decreases proportionally to the hydrophilicity. The potentially charged residues are coded light blue. Arrows show the direction of the hydrophobic moment and so point toward the hydrophobic side of the helix.

Y402 in this structure might also play a supportive similar role to W399 as they appear to be orientated on a similar plane (Fig. 4.9). Y402A showed a very modest ~13% decrease in receptor cell surface expression, although this was insignificant. Since W399A appear not to act alone in its role in cell surface expression (see Conner *et al.*, 2008), Y402 might be one of those hydrophobic residues (in addition to V391 and L395 suggested by Conner *et al.*, 2008), which may play a collective role in the structural and functional stability of this region and in receptor cell surface expression. These hydrophobic residues are orientated in the same plane (Fig.4.9, 4.10). Mutations at this region of some family A GPCRs have been reported to disrupt membrane trafficking (Tobin and Wheatley, 2004). The significant ~27% increase in mean basal activity observed for Y402A supports the idea of its role in receptor structure stability (especially in the inactive state) via hydrophobic interaction with the membrane phospholipid bilayer. Moreover, this hydrophobic aromatic amino acid structure (Y/F/W) is conserved in 12 of all 15 human GPCRs (see Fig. 4.1).

R397A showed significant ~3 fold and 22% reductions in  $\alpha$ CGRP potency and mean  $E_{max}$  respectively. R396A on the other hand showed no significant effect on receptor response to  $\alpha$ CGRP. This is in agreement with that reported for the AM1 receptor, although R396A was found to modestly decrease maximum response (Kuwasako *et al.*, 2010). Despite these two basic residues sitting next to each other in helix 8, only R397A appears to play important role in receptor signalling. This probably explains their orientation observed in Vohra *et al.*, (2013). While R397 points towards the ICL1, R396A points away to intracellular space (Fig. 4.11a). Although it is not clear what R397 exactly does, it appears to be in close proximity to F166 and K167 in ICL1, pointing in-between the two residues (Fig. 4.11b). According to this structure, it is possible that R397, with its positively charged side chain, pushes away the side chain of K167, in turn helping it (K167) to assume a better orientation to form ionic interaction with E390 at the proximal end of helix 8. Conversely, it is possible that a strong ionic interaction formed between K167 and E390 provides the force that pushes R397 into the position it occupies where its side chain might make hydrophobic interaction with that of F166. Although this ionic interaction between acidic and basic residues in helix 8 and ICL1 has

been reported for some family A GPCRs (e.g. see Swift *et al.*, 2006), and therefore plausible, the speculative nature of this model structure however leaves the role of R397 uncertain.



**Fig 4.11:** Model structure of the CLR showing (a) the positional orientation of the neighbour basic residues R396 (light green) and R397 (red); and (b) the proximity of R397 (red) to F166 (yellow) and K167 (pink). E390 (green) is also shown.

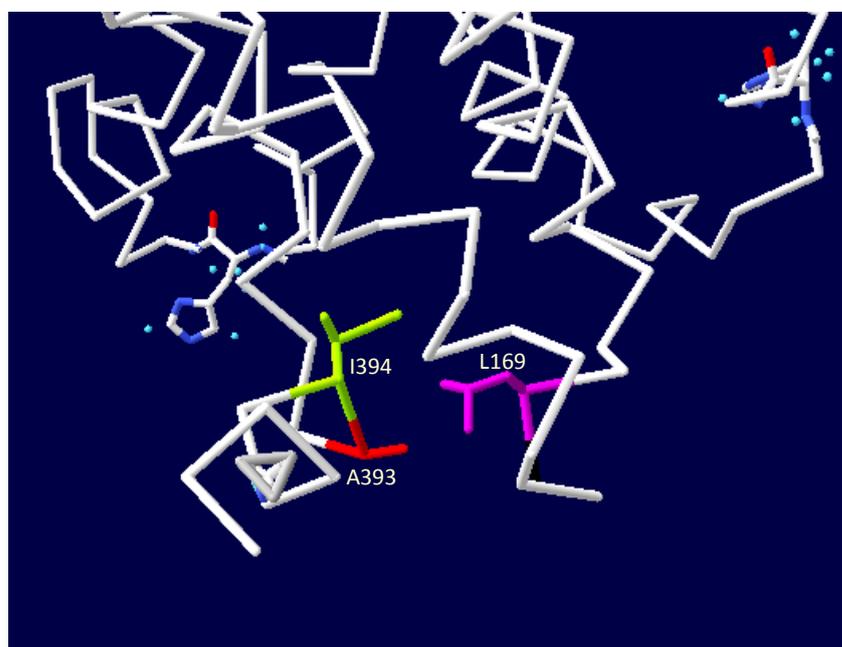
Moreover, the helical wheel plot presents a different, but possibly more plausible explanation. According to this wheel (Fig. 4.10), assuming that the hydrophobic face of the helix faces the membrane, R397 appears to be on cytoplasmic face of the helix; and considering its effect on  $pEC_{50}$  and mean  $E_{max}$  values, with no significant reduction in cell surface expression, it might be directly involved in G-protein coupling. In the helical wheel, R396 also appears to be present at the cytoplasmic face. Although these residues are adjacent to each other on a helix, their side chains most likely face opposite directions. A charge-charge repulsion might be responsible for these opposing directions towards which their side chains face. Since these two residues are completely identical, and R396 appeared to play a subtle role in receptor signalling compared to that of R397, it suggests that

their orientations within the helix are of high physiological importance. Although R396 appears to play no significant role in CGRP receptor signalling from the results gathered so far, the ~13% decrease in  $E_{\max}$  observed for this residue is comparable with the significant ~19% reduction in maximum response observed for the AM1 receptor by Kuwasako *et al* (2011). Depending on the true orientation of its side chain, it is possible that R396 forms a salt bridge with an Asp/Glu residue further down the C-terminal tail as observed in the squid rhodopsin crystal structure (Murakami and Kouyama, 2008) where Lys321 of helix 8 (in a roughly similar position as R396) forms a salt bridge with Glu351 (in the 9<sup>th</sup> helix) further down the C-tail. Unfortunately, no assertions can currently be made as to the correct orientation of these residues based on the limited available resources; this therefore re-echoes the need for a crystal structure of the full receptor, or even a family B GPCR.

Neither R396A nor R397A showed significant effect on cell surface expression and receptor internalization. This is in contrast to that obtained in AM1 receptor where R397A causes significant decrease in cell surface receptor and agonist-induced internalization. The disagreement might be due to differences in cell type, transfection, human RAMP form and/or specific agonist (for receptor endocytosis). However, double mutation to these neighbour residues (R400A/R401A) in human VPAC1 receptor showed no significant effect on receptor cell surface expression and cAMP production as well as radioligand binding (Couvineau *et al.*, 2003). It is important to state however that this hVPAC1 mutant showed a noticeably modest reduction in potency and agonist binding affinity as represented by the  $EC_{50}$  and  $K_d$  values respectively (Couvineau *et al.*, 2003). This further reflects the importance of these residues, which show a high degree of conservation within the family, in family B receptor signalling.

A393D showed a mild but significant ~2 fold decrease in receptor potency with a significant ~21% positive shift in basal activity. The slight reduction in potency observed might be an evidence for a possible hydrophobic interaction of L169 in ICL1 with A393 and I394 in helix 8 in the active CLR model structure by Vohra and coworkers (Fig. 4.12). L169A has also been reported to cause ~3 fold decrease in the  $pEC_{50}$ , suggesting this interaction might be important for receptor activation. The increased basal activity observed for this receptor apparently is not readily explainable and so must be

due to some other factor(s). There is a possibility that the negatively charged side chain of Asp, from the A393D mutation, repels the negative charge of the polar head groups of the phospholipid bilayer causing a shift in favour of the active state but not capable of increasing  $\alpha$ CGRP potency. This sort of repulsion has been observed in  $V_{1a}$  vasopressin receptor where the substitution of highly conserved Arg125 at the extracellular end of TM3 by an Asp was found to be detrimental due to charge-charge repulsion between the Asp side chain carboxyl and membrane lipid phosphate head groups. This resulted in increased solvent accessibility at the extracellular end of TM3/TM4 and altered local conformation (Hawtin *et al.*, 2006). A393 is conserved in 40% of class B GPCRs (Fig 4.1) and no charged residue is present at the same locus in other members of the family. This therefore suggests that, even though this A393 might not directly make interactions involved in receptor signalling, it may be required for maintaining the inactive-active state equilibrium. This, although, is plausible assuming the orientation of Ala in the current model is right. At the same time, the effect posed by A393D might be as a result of disruption in the integrity of the helix caused by the charged and hydrophilic Asp residue. It will therefore be interesting to see what effect a mutation to Leu will have on the receptor function.



**Fig 4.12:** Model structure of the active form of CLR showing residues A393 (red) and I394 (green) in H8, and L169 (purple) in ICL1. These residues are thought to putatively interact during receptor activation.

Of all the three residues (A393, R397 and W399) in Helix 8 whose mutation showed significant decrease in  $\alpha$ CGRP potency, A393 and R397 appear to be on the cytoplasmic side of the helix considering the helical wheel plot. Going by this, A393, like R397, might be involved in G-protein coupling. In spite of this, the most probable explanation observed for the effects of the Ala – Asp mutation is a possible disruption in the helix as earlier suggested. Also, in  $\beta_2$ -AR, unwinding of the 8<sup>th</sup> helix is one of the conformational changes observed to accompany receptor activation (Hulme, 2013). It is therefore possible that A393D causes an agonist-independent unwinding of CLR helix 8 resulting in constitutive activity observed for this mutant.

N400A showed a very significant ~36% increase in cAMP basal activity with no significant effect on  $pEC_{50}$ ,  $E_{max}$  or cell surface expression values. It is possible that mutation to this residue causes unwinding of the helix thereby shifting the receptor equilibrium towards the active state.

Although mutation to almost half of the total residues investigated in this chapter caused significant reduction in CGRP potency and/or maximum cAMP levels, it cannot be ascertained what role, if any, they play in CGRP binding and the formation of the R\* form of the receptor. This information is limited owing to the unavailability of ligand binding assay data for these mutant receptors. However, as they are on the cytoplasmic surface of the receptor, they cannot have a direct role

It is important to mention that some of these residues that appear to play a role in receptor signalling could act by interacting with receptor component protein (RCP). RCP, cytoplasmic protein found to co-immunoprecipitate with CLR in cell culture and tissues, are required for CGRP receptor signalling (Tolun *et al.*, 2007). The site of interaction of RCP at the receptor complex is still unknown, so it is difficult to speculate what effects mutants might have on receptor signalling. This makes CGRP receptor signalling more complicated.

Six of the residues investigated significantly influence  $\alpha$ -CGRP mediated receptor internalization. W399A and N400A cause significant ~35% and ~29% increase in receptor internalization respectively. This strongly suggests that W399 and N400A negatively regulate  $\alpha$ -CGRP receptor internalization. W399's effect on receptor internalization agrees with that reported for AM1

(Kuwasako *et al.*, 2010). The role of membrane anchor earlier suggested for W399 might have come to play here. The mutation of this residue to Ala may favour interaction with  $\beta$ -arrestin that drives the endocytosis pathway thereby resulting in increased receptor internalization observed for this mutant. F163M showed significant ~43% decrease in internalization. This reiterates the importance of this residue, not only in receptor activation but in endocytosis. Although not significant, Ala mutation of this residue showed a mild ~16% decrease in receptor internalization. A393D, 1394 and I404A resulted in significant ~60%, ~44% and ~35% decrease in receptor internalization respectively. Previous studies in porcine CTR and human CLR have indicated the region between positions 391 and 418 as important for CTR and CGRP receptor internalization (Findlay *et al.*, 1994; Conner *et al.*, 2008). These results therefore reveal some of the key residues responsible for the role this region plays in receptor internalization. The mechanisms involved however need to be investigated. Although this region has also been shown to affect receptor internalization in the AM1 receptor (Kuwasako *et al.*, 2010), the C-terminus of the GLP2 receptor has been suggested to have no role in receptor endocytosis (Estall *et al.*, 2005). This suggests that the role observed for this region might be receptor-specific.

It is interesting to note that almost all mutants, which showed significant effect on cAMP levels, also appeared to alter agonist-induced internalization. Although this does not necessarily suggest the dependence of receptor internalization on G-protein coupling or vice versa, it rather shows that the part of the receptor is important for both cAMP signalling and agonist-induced receptor internalization

In conclusion, these results suggest that the helix 8 of the CLR plays a significant role in cAMP response, cell surface expression and receptor internalization. Notably among the important residues within this helical region are W399, R397 and A393. I404 in the extended C-terminal region of H8 is also noteworthy. The importance of F163 is also very interesting. The investigation of these residues in this receptor is novel. As the very limited existing investigations on this receptor region in other family B GPCRs are performed employing receptor truncation, more investigations need to be conducted to see the role these residues play in other family B GPCRs. As pointed out in Chapter 1

however, there could still be huge differences observed for the presence, absence or overall conformation of the 8<sup>th</sup> helix even within the class B GPCRs. This therefore makes interpretations and extrapolations from other receptors more difficult. Above all, the availability of a crystal structure will help a great deal to better interpret some of these results.

## **Chapter 5: Production of the N-terminal extracellular domains of CLR and RAMP1, -2 and -3 in *Pichia pastoris***

### **5.1 Introduction**

Most of the work done in the past few years on the CGRP receptor has been on the receptor in its native membrane and usually in intact (mammalian) cells. Studies carried out in this way have limitations to how much information they can give about the receptor and the forms of interaction between its (the receptor) components. Conducting detailed structural investigation on this receptor, as for other GPCRs, involves carrying out biophysical studies on isolated receptors and this normally requires that the protein be produced in some larger quantity than would happen in cells where it naturally occurs. For this reason, recombinant protein production techniques are usually employed to make large amounts of GPCRs for various studies. This technique involves subcloning and expressing the GPCR gene in prokaryotic (most commonly *Escherichia coli*) or eukaryotic (most commonly yeast, insect or mammalian) cells, with the growth conditions optimized to increase the rate of the protein synthesis pathway. Each of these expression systems has its own peculiar features and the choice of expression system usually depends on the purpose of study (which includes the amount of protein desired) and affordability (Terpe, 2006), with a generally valued consideration believed to be the evolutionary closeness of the host system to the target membrane protein (Bill *et al.*, 2011). The GPCRs produced in this process are isolated from the cell membrane, which would usually contain other (unwanted) proteins, and purified. This isolation process requires the use of detergents or other lipid mimics e.g. styrene maleic acid lipid particles (SMALPs) (Knowles *et al.*, 2009) to provide an artificial amphipathic membrane-like environment for the proteins in solution.

The entire production process of GPCRs is a difficult task. The reasons for this range from insufficient yield to the inability of making a perfect environment that mimics that of the natural cellular membrane, which houses the transmembrane domain. Though some detergents have been successfully used to achieve this purpose for some GPCRs, they have also been reported to have negative impact on the stability of GPCRs (Lundstrom, 2005; Kobilka, 2007). This has particularly

been the bane of successful crystallization especially for family B GPCRs, which assume variable conformations when purified in detergents (Bill *et al.*, 2011). This problem is compounded by the fact that these receptors need to retain their functions as much as possible outside their native environment. A way of avoiding this problem has seen researchers producing and working with the truncated forms of these receptors, especially the N-terminal extracellular domain (ECD). The absence of the TM domain makes the ECD easier isolate and study. However, it should be clearly noted that although the difficulty in producing the TM domains of these proteins have led to more concentration on the ECDs, the ectodomains in themselves play a large role in ligand binding and receptor activation, and hence hold a lot of importance. There have been successes recorded for the family B GPCRs due to their large (~100 – 160 amino acids) ECD, which can interact with ligands without the transmembrane domain (Hoare, 2005; Perrin *et al.*, 2003; Koth *et al.*, 2010; Kumar *et al.*, 2011). This forms the basis of this project.

The CLR differs from other family B members by being the only member that absolutely requires RAMP1, RAMP2 or RAMP3 for its function. Therefore, any study intended on the CLR would normally involve the RAMP. Here, a strategy developed to produce the ECDs of the CLR and RAMP1, -2 and -3 (hereafter referred to as RAMPs except otherwise stated) in a eukaryotic system is reported. Expression in this system avoids the issues associated with the absence of posttranslational modification peculiar with prokaryotic expression system. In this segment of the research project, genes encoding the ECDs of the CLR and RAMP were amplified and subcloned into a *Pichia pastoris* (yeast) vector – pPIC9KMep\_Net. The yeast cells were transformed with the resulting constructs from the subcloning. The genes were expressed and the receptor proteins produced. The proteins were analysed by western blot analysis to determine if the proteins of interest had been produced. The protein samples were used for further biophysical characterization.

Overall, the proteins produced in this study were intended to be used for interaction studies using surface plasmon resonance (SPR). This is aimed at developing an SPR-based tool for investigating novel interacting partners for RAMPs.

## 5.2 Results

### 5.2.1 Defining the sequence of CLR and RAMP N-terminal ECDs for recombinant expression

The optimum sequence to express as the soluble ectodomain of the receptor proteins (CLR and RAMP1, -2 and -3, hereafter called RAMPs) were determined with the help of sequence homology with other family B GPCRs in earlier research works (Augen *et al.*, 2008) and the Uniprot<sup>TM</sup> bioinformatics database. The RAMPs ECD residues were also determined in reference to Kusano *et al* (2008) and employing the Uniprot<sup>TM</sup> with slight modifications. For instance, RAMP1 and RAMP3 were mapped a residue earlier than reported by Kusano *et al* (2008) to avoid having the disulphide-forming Cys residue as the first residue. The residues making up the CLR and RAMP1 ECDs were drawn with the assistance of Dr James Barwell. These amino acid sequences, for the proteins produced in this research work, are shown in Fig 5.1. The number of residues and some key features (e.g. glycosylation sites and disulphide bonds) of these domains are summarised in Table 5.1.

**CLR**<sub>(23-133)</sub>  
 ELEESPEDSIQLGVTRNKIMTAQYE<sup>C</sup>YQKIMQDPIQQAEGVY<sup>C</sup><sup>N</sup>R<sup>T</sup>WDGWL  
<sup>C</sup>W<sup>N</sup>DVAAGTESMQL<sup>C</sup>PDYFQDFDPSEKVTKI<sup>C</sup>DQDGNWFRHPAS<sup>N</sup>R<sup>T</sup>WTN  
 YTQ<sup>C</sup><sup>N</sup>VN<sup>T</sup>HE

**RAMP1**<sub>(26-117)</sub>  
 A<sup>C</sup>Q<sup>E</sup>A<sup>N</sup>Y<sup>G</sup>A<sup>L</sup>L<sup>R</sup>E<sup>L</sup>L<sup>R</sup>E<sup>L</sup><sup>C</sup>L<sup>T</sup>Q<sup>F</sup>Q<sup>V</sup>D<sup>M</sup>E<sup>A</sup>V<sup>G</sup>E<sup>T</sup>L<sup>W</sup><sup>C</sup>D<sup>W</sup>G<sup>R</sup>T<sup>I</sup>R<sup>S</sup>Y<sup>R</sup>E<sup>L</sup>A<sup>D</sup><sup>C</sup>T<sup>W</sup>H  
 M<sup>A</sup>E<sup>K</sup>L<sup>G</sup><sup>C</sup>F<sup>W</sup>P<sup>N</sup>A<sup>E</sup>V<sup>D</sup>R<sup>F</sup>F<sup>L</sup>A<sup>V</sup>H<sup>G</sup>R<sup>Y</sup>F<sup>R</sup>S<sup>C</sup>P<sup>I</sup>S<sup>G</sup>R<sup>A</sup>V<sup>R</sup>D<sup>P</sup>P<sup>G</sup>S

**RAMP2**<sub>(43-145)</sub>  
 Q<sup>P</sup>L<sup>P</sup>T<sup>T</sup>G<sup>T</sup>P<sup>G</sup>S<sup>E</sup>G<sup>G</sup>T<sup>V</sup>K<sup>N</sup>Y<sup>E</sup>T<sup>A</sup>V<sup>Q</sup>F<sup>C</sup>W<sup>N</sup>H<sup>Y</sup>K<sup>D</sup>Q<sup>M</sup>D<sup>P</sup>I<sup>E</sup>K<sup>D</sup>W<sup>C</sup>D<sup>W</sup>A<sup>M</sup>I<sup>S</sup>R<sup>P</sup>  
 Y<sup>S</sup>T<sup>L</sup>R<sup>D</sup><sup>C</sup>L<sup>E</sup>H<sup>F</sup>A<sup>E</sup>L<sup>F</sup>D<sup>L</sup>G<sup>F</sup>P<sup>N</sup>L<sup>A</sup>E<sup>R</sup>I<sup>I</sup>F<sup>E</sup>T<sup>H</sup>Q<sup>I</sup>H<sup>F</sup>A<sup>N</sup><sup>C</sup>S<sup>L</sup>V<sup>Q</sup>P<sup>T</sup>F<sup>S</sup>D<sup>P</sup>P<sup>E</sup>D<sup>V</sup>

**RAMP3**<sub>(27-119)</sub>  
 G<sup>C</sup><sup>N</sup>E<sup>T</sup>G<sup>M</sup>L<sup>E</sup>R<sup>L</sup>P<sup>L</sup><sup>C</sup>G<sup>K</sup>A<sup>F</sup>A<sup>D</sup>M<sup>M</sup>G<sup>K</sup>V<sup>D</sup>V<sup>W</sup>K<sup>W</sup><sup>C</sup><sup>N</sup>L<sup>S</sup>E<sup>F</sup>I<sup>V</sup>Y<sup>Y</sup>E<sup>S</sup>F<sup>T</sup><sup>N</sup><sup>C</sup>T<sup>E</sup>M<sup>E</sup>  
 A<sup>N</sup>V<sup>V</sup>G<sup>C</sup>Y<sup>W</sup>P<sup>N</sup>L<sup>A</sup>Q<sup>G</sup>F<sup>I</sup>T<sup>G</sup>I<sup>H</sup>R<sup>Q</sup>F<sup>F</sup>S<sup>N</sup><sup>C</sup>T<sup>V</sup>D<sup>R</sup>V<sup>H</sup>L<sup>E</sup>D<sup>P</sup>P<sup>E</sup>V<sup>L</sup>

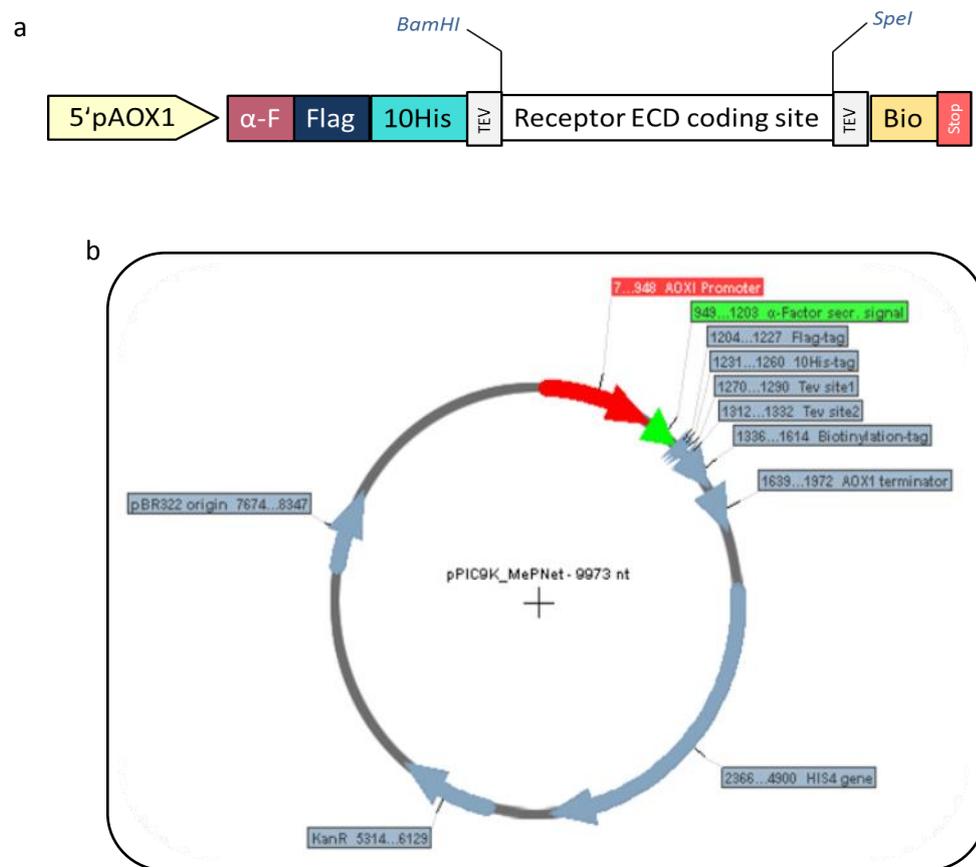
**Fig 5.1:** Amino acid sequence of the N-terminal extracellular domains of CLR and RAMPs. Putative N-glycosylation sites are highlighted red while disulphide-forming Cys residues are highlighted yellow. The RAMP3 disulphide Cys residues were determined based on sequence homology with RAMP1 and 2.

**Table 5.1:** Some key features of the ectodomains of CLR and RAMPs

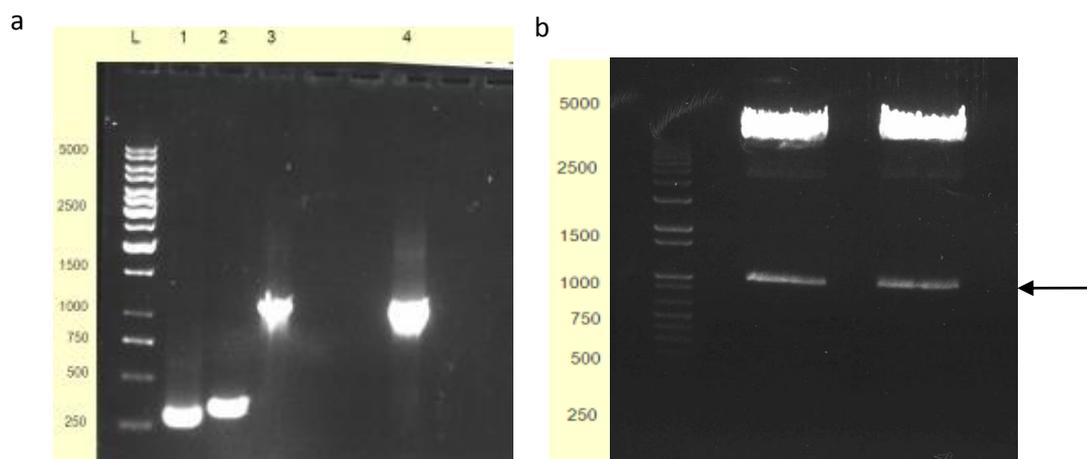
Protein	N-terminal residues positions	No of residues	Calculated MW + tags (~kDa)	No of glycosylation sites	Di-sulphide bonds
CLR	23 – 133	111	18	3	3
RAMP1	26 – 117	90	16	0	3
RAMP2	43 – 145	103	17	1	2
RAMP3	27 – 119	93	16	4	2

### 5.2.2 Construction of CLR/RAMP expression vector

The genes of interest (i.e. CLR/RAMPs) were amplified by polymerase chain reaction (PCR). The resulting PCR products (Fig 5.3a) were inserted into the pPIC9K\_MepNet expression vector (Fig 5.2) following digestion of the PCR product and the vector. The vector digestion was confirmed by running an agarose gel electrophoresis (Fig 5.3b). The inserts, indicated by the lower bands, confirms this.



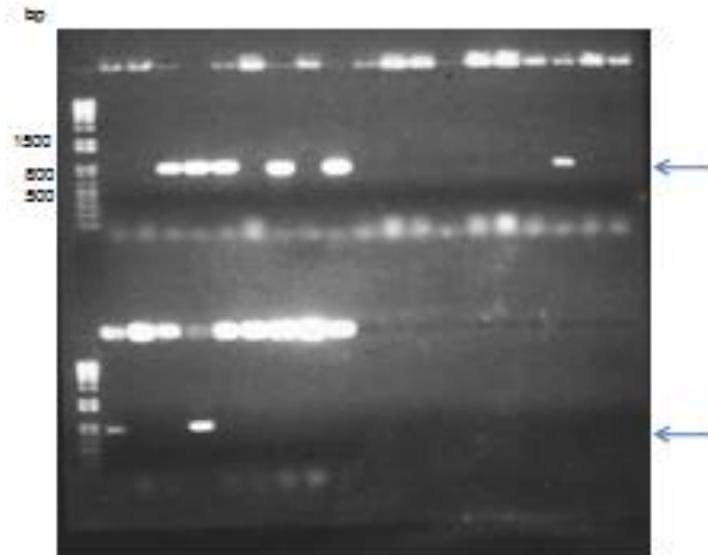
**Fig 5.2:** (a) Schematic representation of the *P. pastoris* expression vector for recombinant production of the N-terminal ECD of CLR, RAMP1, RAMP2 and RAMP3. The various regions indicated are the coding regions for; the 5' alcohol oxidase 1 gene promoter (5'pAOX1);  $\alpha$ -factor ( $\alpha$ -F) secretional signal from *Saccharomyces cerevisiae*; FLAG-tag; decahistidine (10His) tag; tobacco etch virus (TEV) protease cleavage site; biotinylation (Bio) tag from *Propionibacterium shermanii* and the terminator (Stop) domain. *Bam*HI and *Spe*I are the restriction endonucleases used to clone the receptor ECD genes. (b) Vector of pPIC9K\_MepNet constructed using the Serial Cloner v2.



**Fig 5.3:** (a) Agarose gel picture of the receptor genes PCR products; L= ladder, 1 = N-terminus of RAMP3 (277 bp); 2 = N-terminus of CLR (332 bp); 3 and 4 = full length of CLR (1134 bp). (b) Digestion of the pPIC9K plasmid with BamHI and SpeI restriction endonucleases. Both lanes represent the cut plasmid. The insert is indicated by the arrow.

### 5.2.3 Transformation

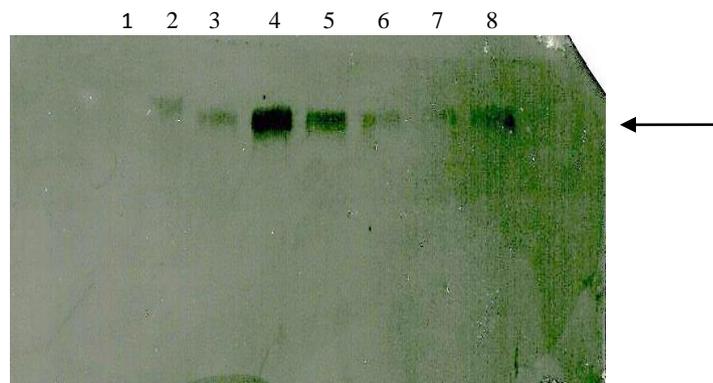
The integration of the cloned plasmid vector into the *Pichia* cells was confirmed by colony PCR. Colonies were picked at random, each representing either a CLR, RAMP2 or RAMP3 clone. The clone loaded in each lane was noted. The bands observed corresponds with the number of base pairs (~800bp) expected for the cloned genes (Fig 5.4). The 800 bp represents the number of bp CLR/RAMP2/RAMP3 in addition to the number of bp between this gene and the primer sequence on the expression vector. This indicates that the respective clones contain the desired inserts. Going by the result from the colony PCR, an integration frequency of ~40% was observed as roughly 3 out of every 8 colonies showed bands corresponding to the expected 800 bp for each receptor protein ECD construct. It is important to note however that this might not represent the actual integration frequency as the colony PCR might have failed for some samples owing to reasons such as too many cells (from transformant colony) being present in the PCR reaction mix.



**Fig 5.4:** Colony PCR to confirm integration following transformation. Each lane represents a CLR/RAMP2/RAMP3 colony picked at random. The bands indicated by the arrow show that the cloned gene is present in the respective colonies.

#### 5.2.4 Screening (small scale expression) of transformants

Having confirmed integration, the genes were expressed in a 24-well expression plate in a 3ml volume per well. A 20 $\mu$ l aliquot of the sample from each expression medium was analysed by western blot (Fig 5.5).



**Fig 5.5:** Western blot analysis of the 10xHis-tagged CLR, RAMP2 and RAMP3 transformants in expression medium using the manual method. Wells 1-3= CLR; 4-6=RAMP3; 7-8=RAMP2

All the visible bands were observed to be retained in the wells of the polyacrylamide gel. This was not as expected as it appears the proteins have not traversed the gel. The photographic film for the western blot at this stage had been fixed and developed manually, it was therefore not clear whether the result obtained was due to the technique employed in the fixing and development of the film or something else. In order to eliminate any doubts arising from the detection system employed in the western blot procedure, there became the need to use a more trusted and automated system (e.g. a CCD camera) for the detection process.

### 5.2.5 Troubleshooting the expression and Western blot analysis procedures

Since the use of photographic film was not sufficiently reliable as a detection method, a CCD camera was used to repeat the detection. No visible bands were obtained (Fig 5.6). It was not clear if the western blot had failed, not enough protein had been loaded or there was no expression. The process was therefore repeated with increased concentrations of protein sample and antibody used in the same and separate experiments. No significant changes were observed.

1 2 3 4 5 6 7 8

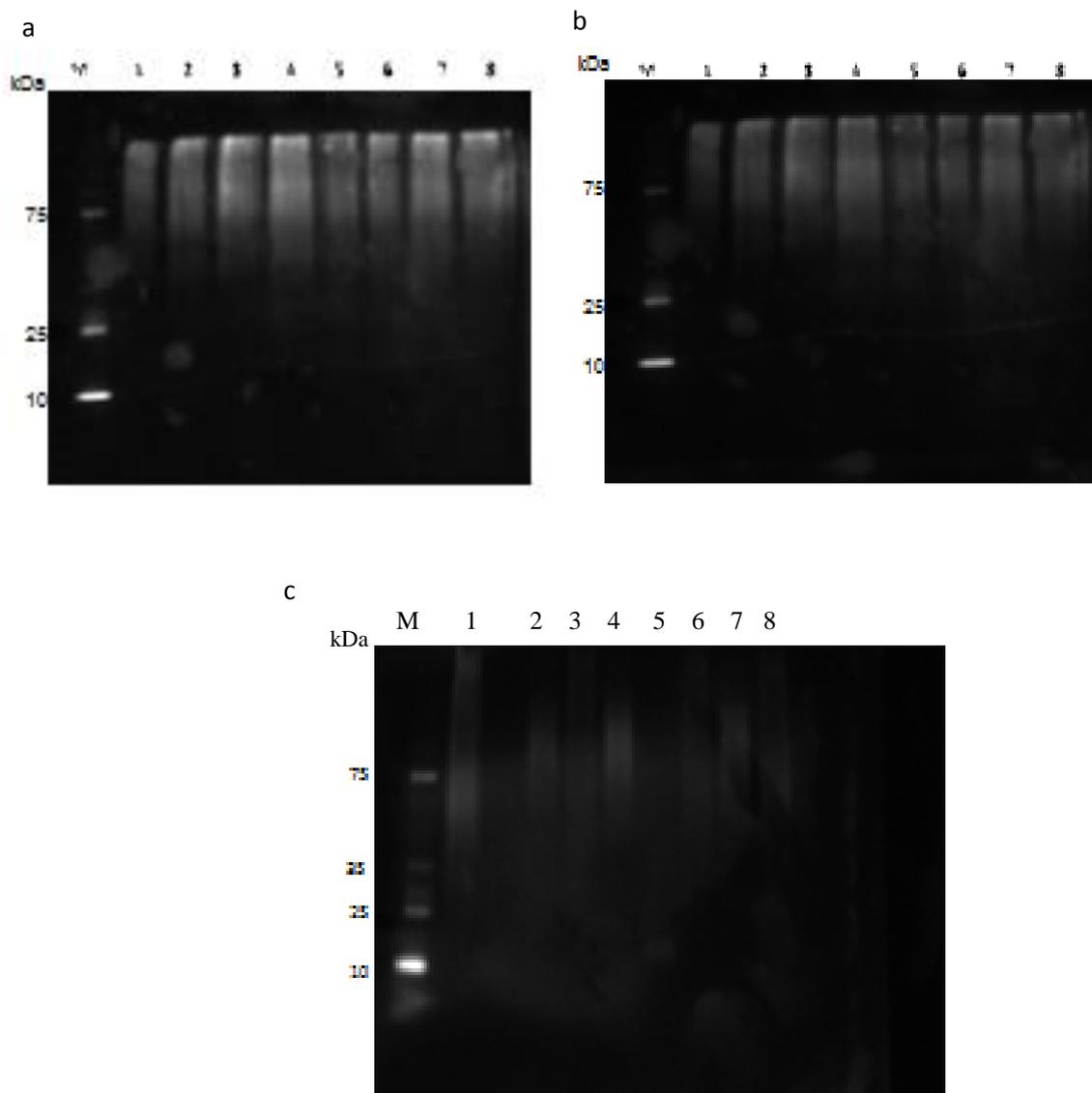


**Fig 5.6:** Western blot analysis of the 10xHis-tagged transformants in expression medium using the CCD camera system. Wells 1-3= CLR; 4-6=RAMP3; 7-8=RAMP2

At this point, it was thought that since the protein had been stored at -20 °C for ~90 days, it might have been degraded and therefore be the reason why virtually nothing appeared on the western blot image. As the proteins had not previously been expressed by this means, their susceptibility to degradation was unknown. The expression process was therefore repeated to produce fresh samples for analysis.

To improve the analysis, the percentage of polyacrylamide resolving gel used was also explored. A 10% and 15% polyacrylamide resolving gel was used together with the 12% already employed in the previous analyses. As a positive control for the sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE) and the western blot, a set of His-tagged molecular weight markers, detectable on western blot, was used.

The results confirmed that both the SDS PAGE and western blotting processes were working (Fig 5.7a & b). In addition to finding the best gel percentage for the protein analysis, a 4-20% gradient gel was used. Although the images obtained on this occasion showed the loaded samples were traversing the gel, there were no clear bands seen as the samples were all smears (Fig 5.7a – c). This process was repeated to check for consistency but the results remained similar. Thus it seemed the western blot was working but that there might be something wrong with the sample. It was also concluded at this point that the manual photographic film method of detecting the blot was not in itself responsible for the bottlenecks earlier encountered.

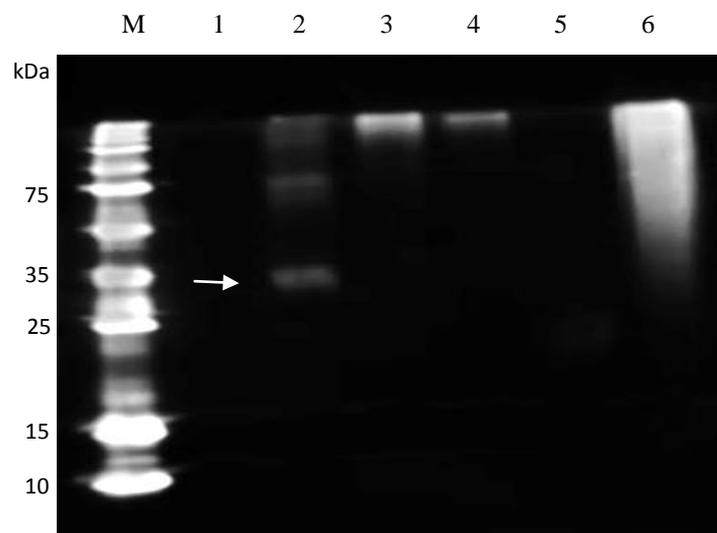


**Fig 5.7:** Western blot analysis of the 10xHis-tagged transformants in expression medium with different percentages of SDS gel. 10% (a) 15% (b) and 4-20% (c) gradient gels. M=marker; wells 1-3= CLR; 4-6=RAMP3; 7-8=RAMP2.

It was now not clear whether the protein of interest was being produced, either at the induction or some other stage. It was thought that the proteins appearing on the western blot might be a yeast protein with multiple histidine residues and which was detected in the western blot by non-specific interaction with the anti-His antibody. To check the expression process, positive and negative controls were therefore devised. For the positive control, a green fluorescent protein (GFP)-transformed *Pichia pastoris* cell sample, already successfully used to produce GFP by Holmes *et al.* (2009), was obtained

from Dr Sarah Routledge (Aston University). The negative control was an untransformed X33 strain of the *Pichia pastoris* – the same strain used to express the target proteins of interest.

Following another expression trial with these controls, it became clear that the process was working as a clear band was seen for the GFP at the correct molecular weight (Fig 5.8). The problem of smears and unclear bands however, still persisted for the CLR and RAMP3 samples. It was then thought at this point that there might be substances in the expression medium interfering with the protein and making them aggregate and/or produce smear. To rule out this factor of interference, there was therefore a need for the samples to be purified.

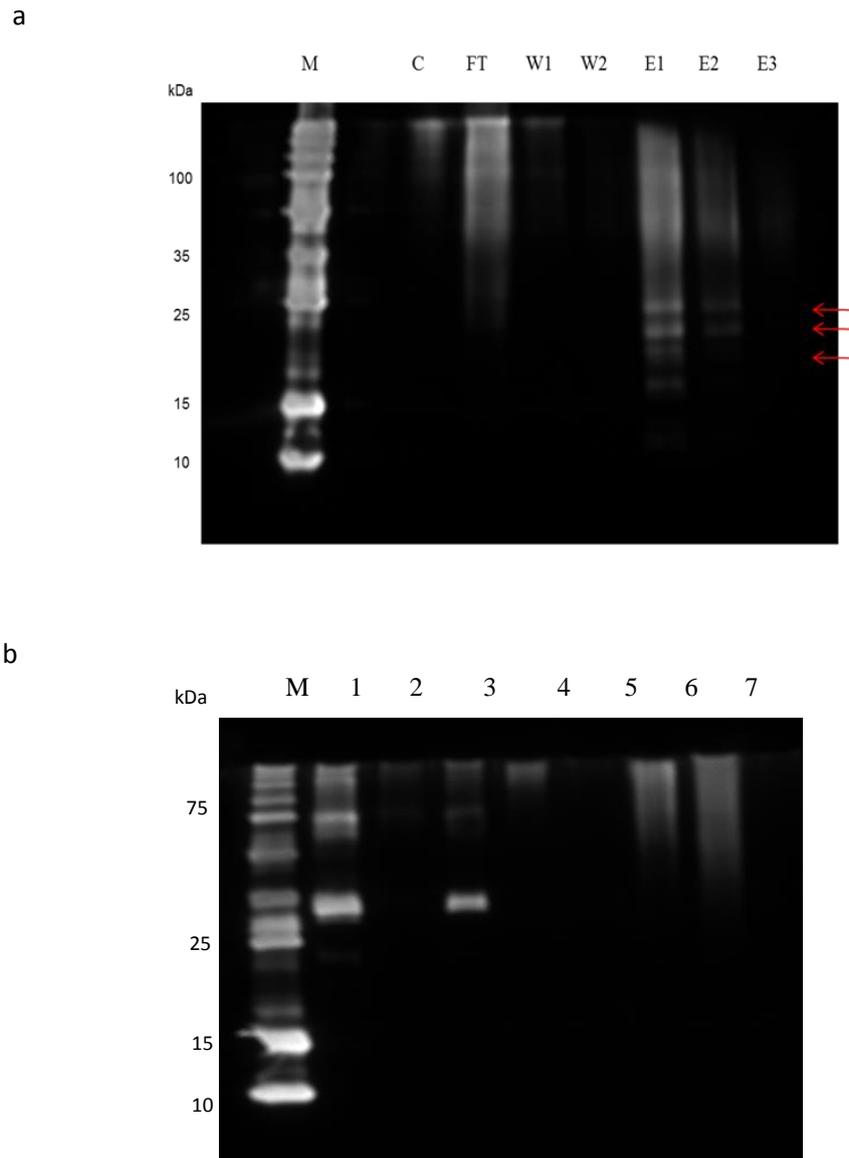


**Fig 5.8:** Western blot analysis of expressed transformants and the GFP. M=marker, Lanes 2=GFP; 3&4=CLR; 6=RAMP3. The arrow indicates the GFP band.

### 5.2.6 Troubleshooting by purification of samples

The samples were purified using the batch method, modified from Singh *et al* (2010) and optimized for this expression. For instance, the equilibration buffer contained 20 mM imidazole to minimize non-specific binding and the protein-resin mixture was incubated overnight at 4°C to maximize batch-binding. Following purification and another western blot analysis, some clear, though faint, bands

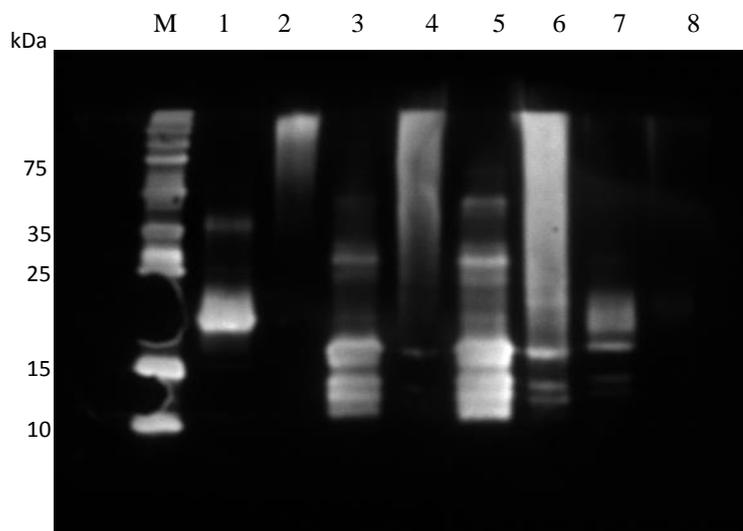
were seen for RAMP3 (Fig 5.9a) but not for CLR (Fig 5.9b). The GFP positive control showed that the purification process was working (Fig 5.9b). Although the purification process appeared to have yielded some positive result as observed for the RAMP3, the persistence of this problem for the CLR sample required the troubleshooting process to be taken a step further. The conclusion at this point was that it was possible that the smearing was due to the glycosylation of the expressed proteins. This therefore presented the need to conduct a deglycosylation reaction.



**Fig 5.9:** (a) Western blot analysis of purified 10xHis-tagged RAMP3 protein sample. M = Marker; C = crude; FT = flow through; W1 & 2 = 1st and 2nd washes respectively; E1, 2 & 3 = 1st, 2nd and 3rd elutions respectively. Arrows indicate bands most likely corresponding to the different glycosylated forms of RAMP3. (b) Western blot analysis of purified 6xHis-tagged GFP and 10xHis-tagged CLR. M=marker, Wells 1=crude GFP; 2=GFP wash; 3=GFP elution; 4=crude CLR; 5= CLR wash; 6&7=CLR elutions 1&2 respectively.

Consequently, a deglycosylation reaction was carried out to remove the sugar bound to the protein. This procedure resulted in the CLR for the first time appearing as bands of the expected molecular weight (Fig 5.10). Although some smears still appeared with the RAMP3 sample, clear bands around the right molecular weight for the monomer and dimer for this protein could be observed (Fig 5.10).

An SDS PAGE analysis and fluorescent staining was performed for these samples, the bands excised and sent for mass spectrometry analysis to confirm that they were the proteins of interest.

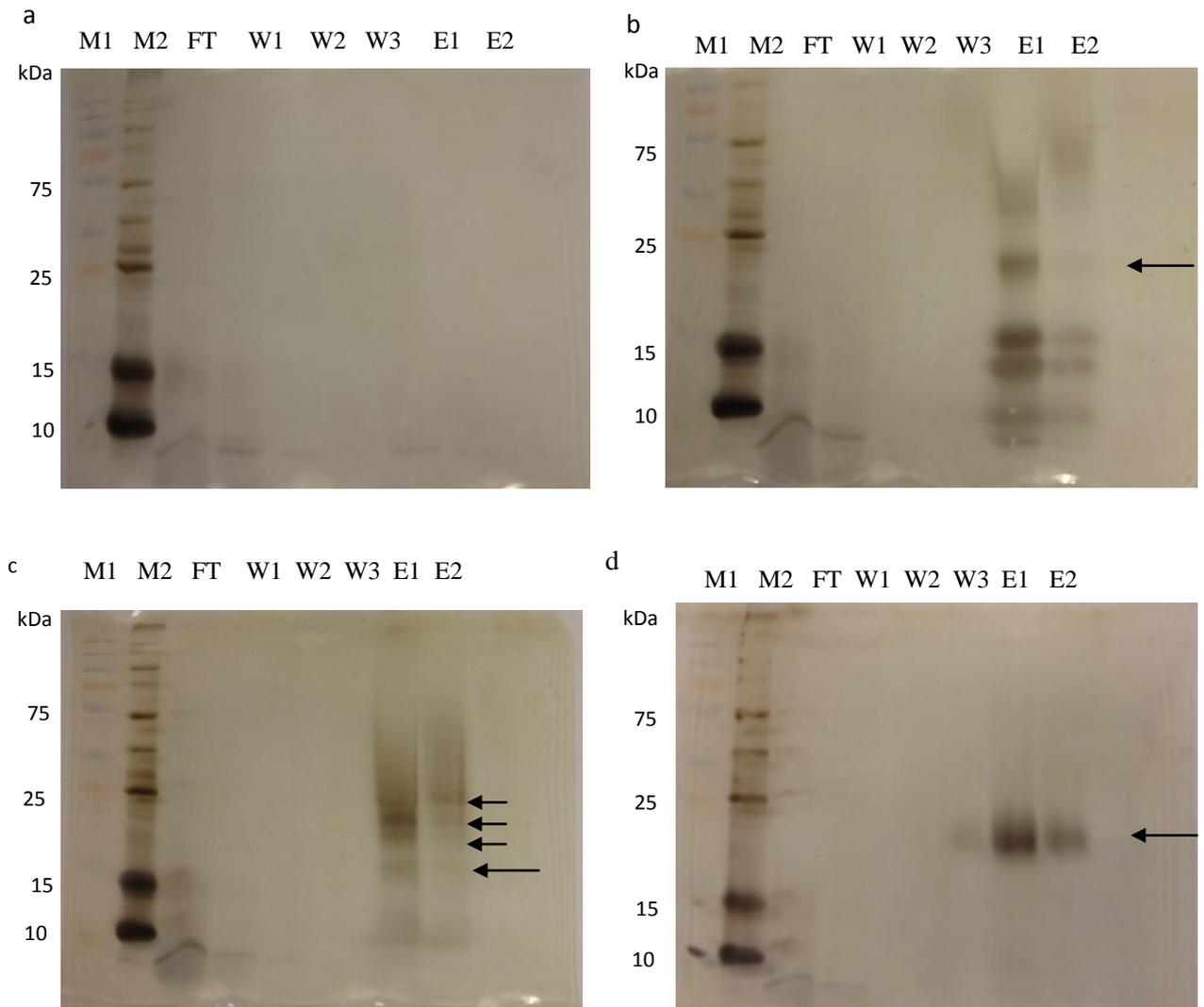


**Fig 5.10:** Western blot analysis showing deglycosylated and untreated protein samples. M=marker; Wells 1=Deglycosylated CLR; 2=Non deglycosylated CLR; 3&5=Deglycosylated RAMP3; 4&6=Non deglycosylated RAMP3; 7=Deglycosylated RAMP2; 8=Non deglycosylated RAMP2.

### 5.2.7 Large scale expression

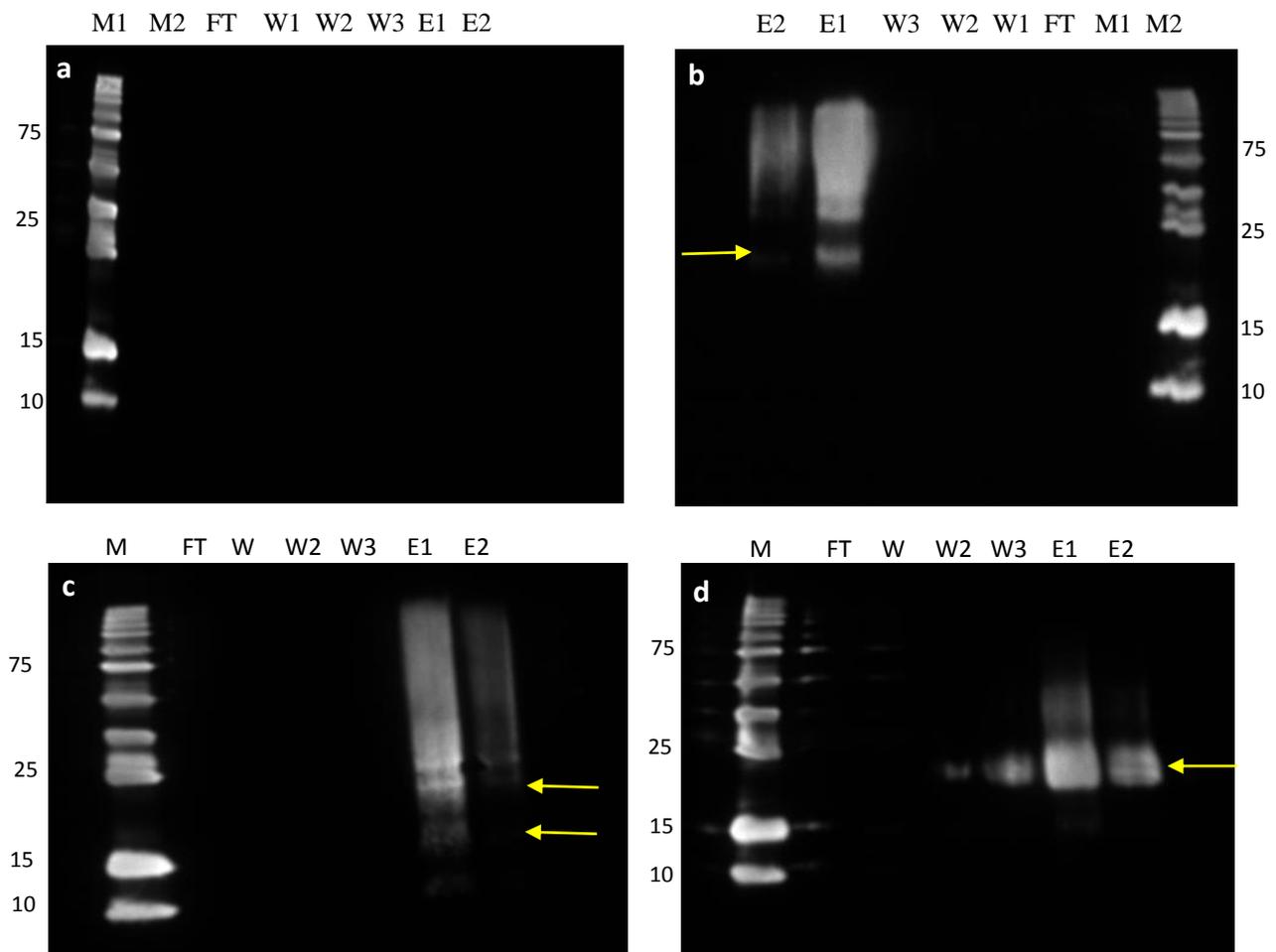
Having isolated suitable expression clones and optimised their purification, 150ml volume expression was carried out in shake flasks in order to obtain sufficient material for biophysical characterization. The protein samples were harvested and purified. Silver stain and western blot analyses were carried out and showed the samples were successfully purified as bands corresponding to the correct calculated molecular weights were observed for the proteins with no bands observed for the control condition (Fig 5.11 and 5.12). The silver-stained SDS polyacrylamide gels in Fig 5.11b, 5.11c and 5.11d show the bands (indicated by arrows) for the CLR, RAMP3 and RAMP2 respectively at the correct calculated molecular weight. Although some multiple bands were observed for the CLR and RAMP3 proteins, this was thought to be due to degradation, aggregation, oligomerization as well as glycosylation. The effect of glycosylation may be as a result of full or partial glycosylation as

reported, for instance, for CLR by Flahaut *et al.* (2003). The negative control (Fig 5.11a), as expected showed no observable band(s).



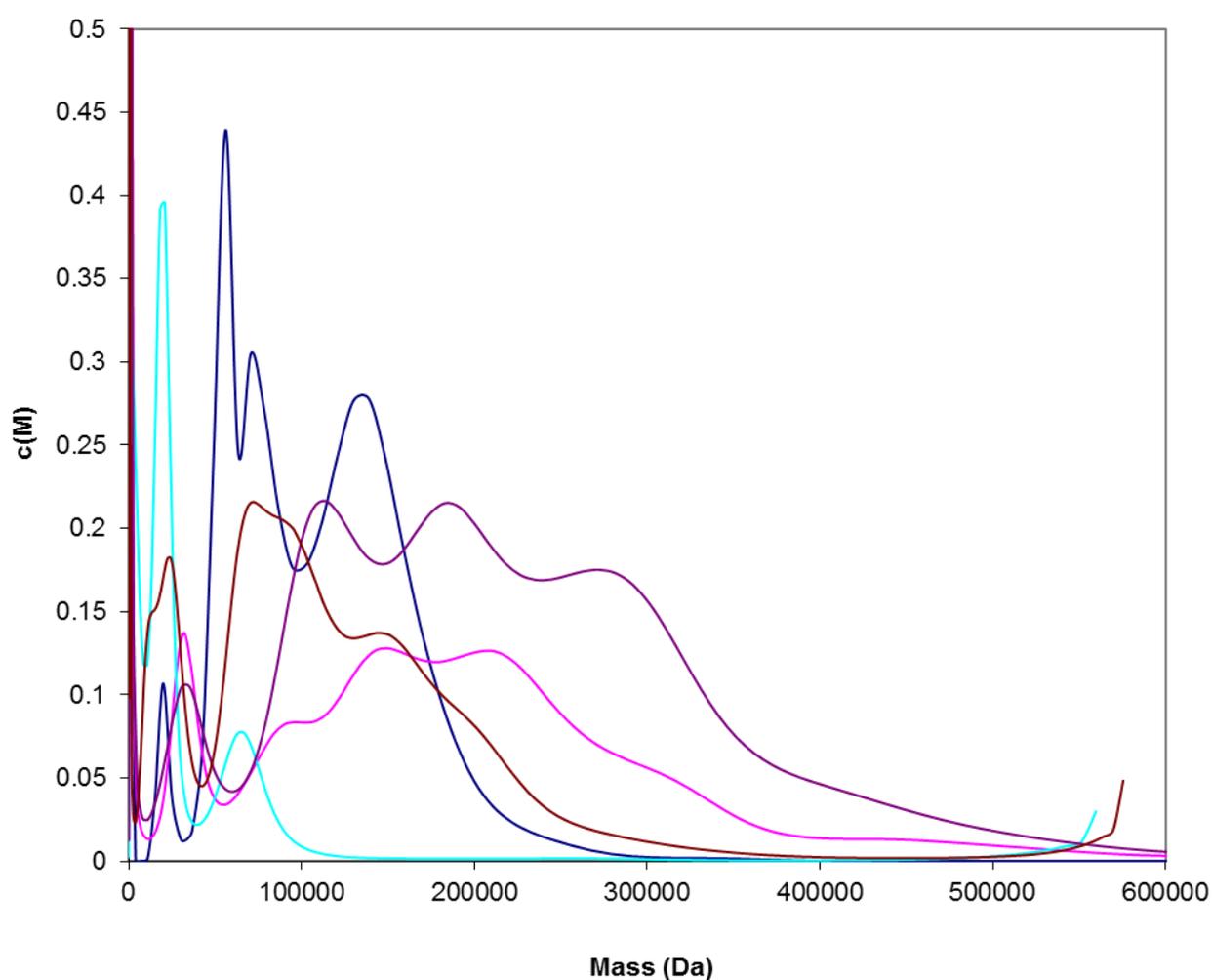
**Fig 5.11:** Silver staining for the purified CLR, RAMP3 and RAMP2 samples. (a) Untransformed X33 (negative) control; (b) CLR; (c) RAMP3; (d) RAMP2. The bands corresponding to the right molecular weight of these proteins are indicated by arrows. The shorter arrows for RAMP3 probably represent fully and partially glycosylated RAMP3. M=prestained marker; M2=His-tagged marker; FT=flow through; W1, W2 & W3=Washes 1, 2 & 3 respectively; E1&E2= Elutions 1&2 respectively.

The western blot pictures, like the silver staining, also reveal these proteins at their calculated molecular weight (Fig 5.12b – d). However, unlike the silver staining, no multiple bands were observed below the right molecular weight of the CLR. This is probably because there were no His tags present on the degraded fragments of these proteins.



**Fig 5.12:** Western blot analysis of the purified CLR, RAMP3 and RAMP2 samples. (a) Untransformed X33 control; (b) CLR; (c) RAMP3; (d) RAMP2. The bands corresponding to the right molecular weight of these proteins are indicated by the arrows. The double arrow in c probably reflects the varying glycosylation pattern in RAMP3. M=pre-stained marker; M2=His-tagged marker; FT=flow through; W1, W2 & W3=Washes 1, 2 & 3 respectively; E1&E2= Elutions 1&2 respectively. Markers are marked in kDa.

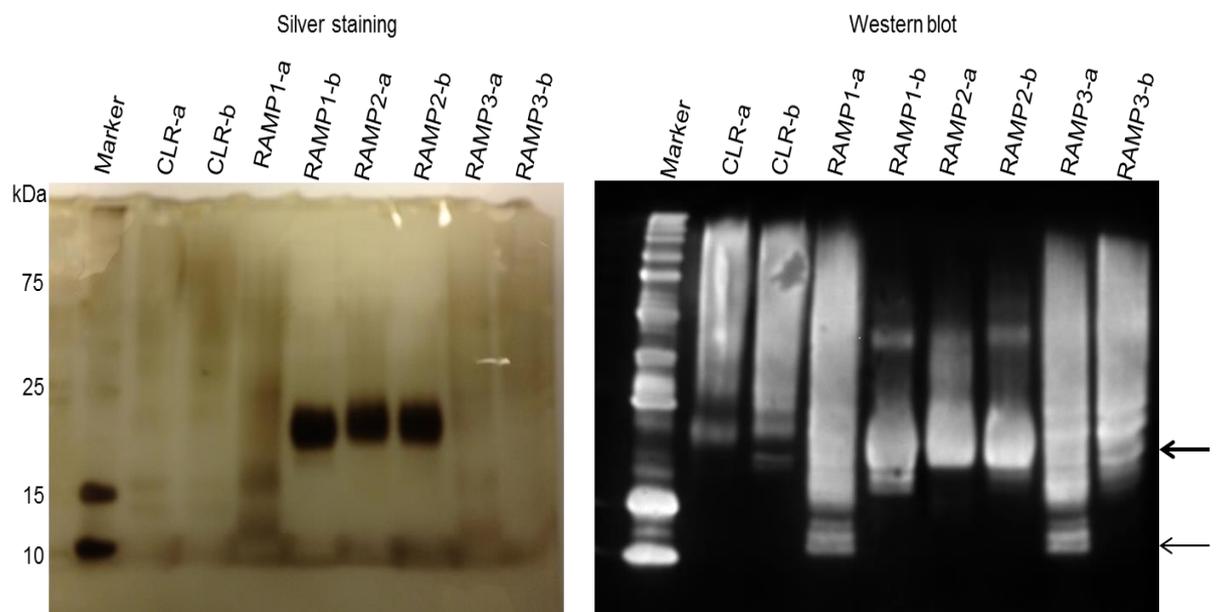
The protein samples were also subjected to analytical ultracentrifugation (AUC) to further check for oligomerization and/or aggregation. The results obtained (Fig 5.13) were consistent with those obtained for silver staining (Fig 5.11). To also check for a possible association between CLR and RAMP2 or RAMP3, 1:1 molar ratio mixture of CLR and RAMP2 or CLR and RAMP3 were incubated for 24h at 4°C and then subjected to AUC. From the AUC profile (Fig 5.13) there are several peaks, though broad, which could represent CLR/RAMP ectodomain association and/or individual ectodomain aggregation.



**Fig 5.13:** Analytical ultracentrifugation analysis of protein samples. CLR (blue), RAMP2 (cyan), RAMP3 (pink), CLR/RAMP2 (brown), and CLR/RAMP3 (purple). CLR/RAMP2 or CLR/RAMP3 were a 1:1 mixture of CLR and RAMP2 or -3 in a 0.5 ml total volume and were incubated at 4°C for ~24 h before ultracentrifuging; please see Chapter 2 for further experimental procedures.

### 5.2.8 Wild type versus protease-deficient strains in CLR and RAMPs production

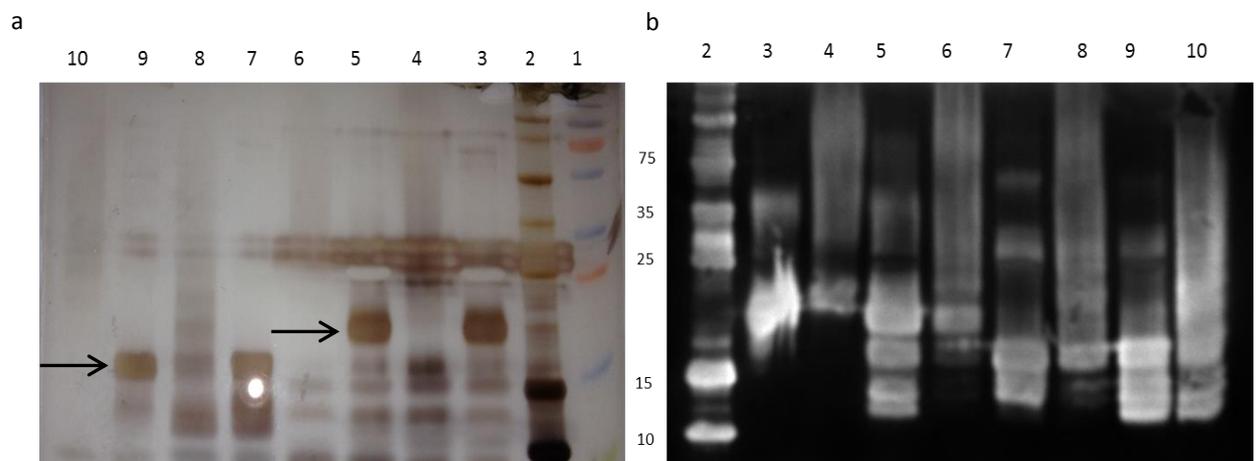
The presence of multiple bands observed for purified CLR and RAMP3 protein samples was a point of concern. The reasons for this, as earlier mentioned, have been attributed to the singular or combined effect(s) of aggregation, oligomerization and degradation. The latter is most likely responsible for the bands observed below the calculated molecular weight of these proteins. In order to check for the possibility of alleviating the problem of degradation without altering the purity of the protein (e.g. without the addition of a protease inhibitor), these protein constructs were expressed in the protease-deficient (SMD1163) strain and compared to the wild type (X33) strain. Following expression and purification of these proteins, no observable difference was noticed for the CLR and RAMP3 expressed in both strains after SDS PAGE silver staining and Western blot analyses (Fig. 5.14). However, a very significant difference, this time in protein quality, was observed for RAMP1 with the protease-deficient strain producing a clear and distinct band as opposed to the wild type (Fig. 5.14)



**Fig 5.14:** Silver-stained and Western-blotted SDS PAGE pictures of eCLR and eRAMPs produced in the wild type (a) and protease-deficient (b) strains of *P. pastoris*. The bold arrow indicates the corresponding calculated molecular weights of these proteins. The light arrow indicates likely product of degradation.

### 5.2.9 Deglycosylation of eCLR and eRAMP3 revealed clear protein bands

Protein samples of eCLR and eRAMP3 from the expression performed in section 5.8 were deglycosylated by treating with endoglycosidase H (Endo H). This was done to again test the hypothesis that deglycosylating these proteins helps their mobility and separation on SDS PAGE. The silver-stained and Western-blotted polyacrylamide gel pictures reveal clearer and more distinct bands for the eCLR and eRAMP3 fragments. This is shown in Fig 5.15.



**Fig 5.15:** Silver staining (a) and Western blot (b) analysis of glycosylated and deglycosylated (Endo H-treated) eCLR and eRAMP3. 1 & 2 – MW markers; 3 – eCLR (X33) deglycosylated; 4 – eCLR (X33) glycosylated; 5 – eCLR (SMD1163) deglycosylated; 6 – eCLR (SMD1163) glycosylated; 7 – eRAMP3 (X33) deglycosylated; 8 – eRAMP3 (X33) glycosylated; 9 – eRAMP3 (SMD1163) deglycosylated; 10 – eRAMP3 (SMD1163) glycosylated. X33 – Wild type host strain; SMD1163 – Protease-deficient strain. The arrows indicate the distinct bands for eCLR and eRAMP3 that corresponds with their calculated molecular weight in their respective lanes.

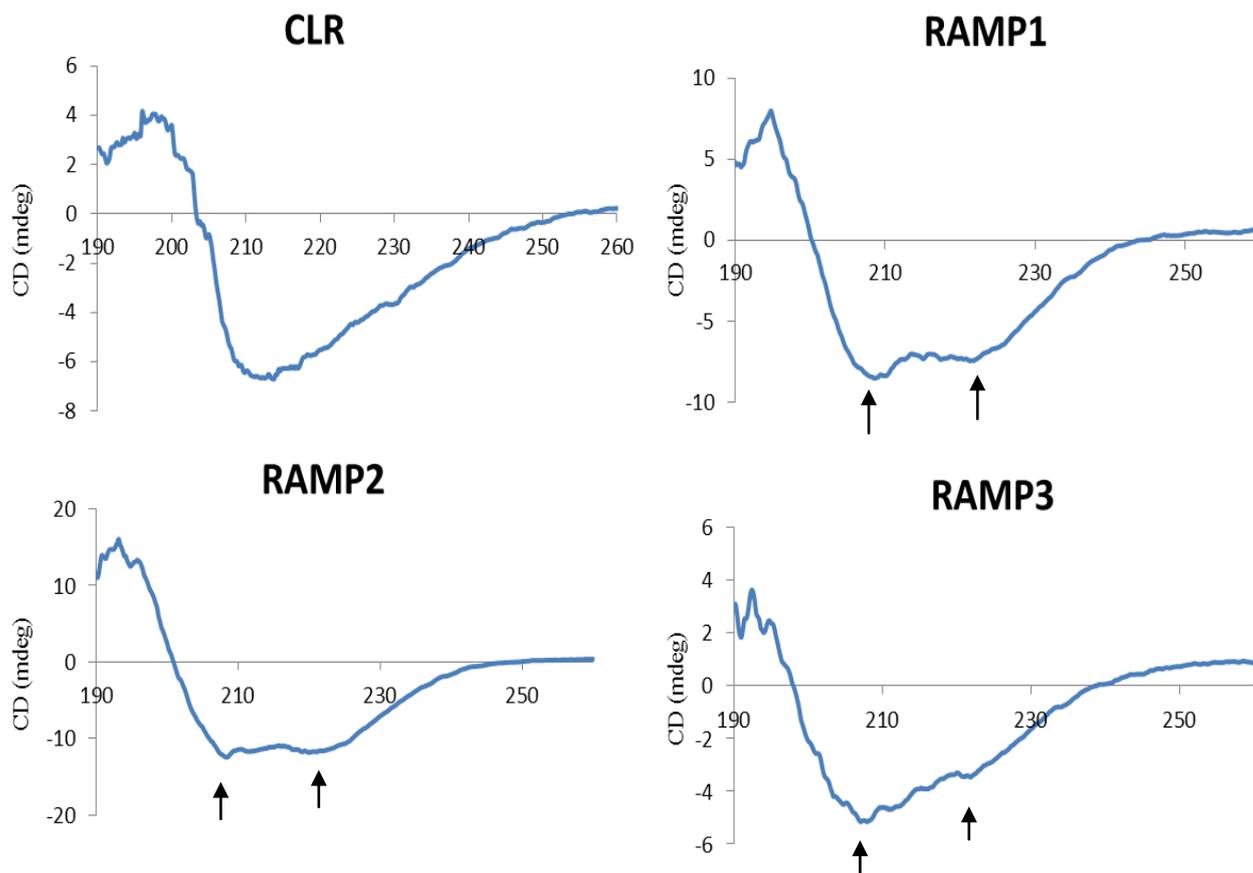
From Fig 5.15, distinct bands were observed for these natively glycosylated proteins following treatment with a deglycosylation enzyme. This is more appreciable on the silver staining gel picture (with arrow indications). The bands below those of the right molecular weight are most likely products of degradation, although the ones closest to the actual bands may represent partially glycosylated protein fragments as earlier suggested in section 5.2.7. This is therefore in correlation

with earlier experiment carried out in section 5.2.6 where deglycosylation was shown to aid distinctiveness of bands.

Moreover, it is worth mentioning that deglycosylation also helped in revealing potential oligomers for the different protein samples. This is clearly observed in the Western blot picture for lanes representing deglycosylated proteins (Fig 5.15, lanes 3, 5, 7 and 9).

### **5.2.10 Circular dichroism showed proteins were folded**

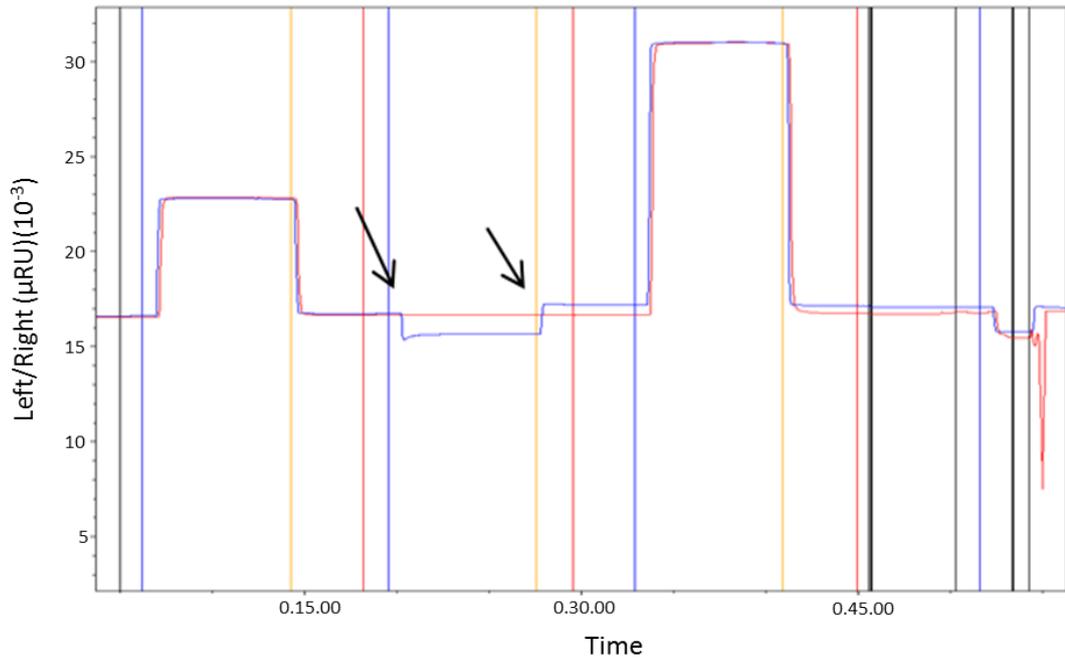
Intact unmodified protein samples of CLR and RAMPs were analysed primarily to determine if they were folded or not, i.e. if they exist in their secondary structural conformation. Since the secondary structural content of these proteins are known (ter Haar *et al.*, 2010; Kusano *et al.*, 2011), circular dichroism profiles would reveal if they are richly  $\alpha$ -helical, mainly composed of  $\beta$ -sheet or both. The results reveal that CLR contains both  $\beta$ -sheet (~45%) and  $\alpha$ -helix (~10%) while the RAMPs are predominantly  $\alpha$ -helical (over 80%) (Fig 5.16). Although the  $\alpha$ -helical profile observed for RAMP3 is less pronounced compared to RAMP1 and RAMP2, this is probably due to the quality of the protein sample. It is important to note that the secondary structure estimation obtained via this system is not always representative of the exact secondary structure component of the protein as there are limiting factors and so should serve as a guide. For instance, the technique requires the accurate protein concentration and this is difficult to obtain as most methods for determining protein concentration vary with protein types (Kelly *et al.*, 2005).



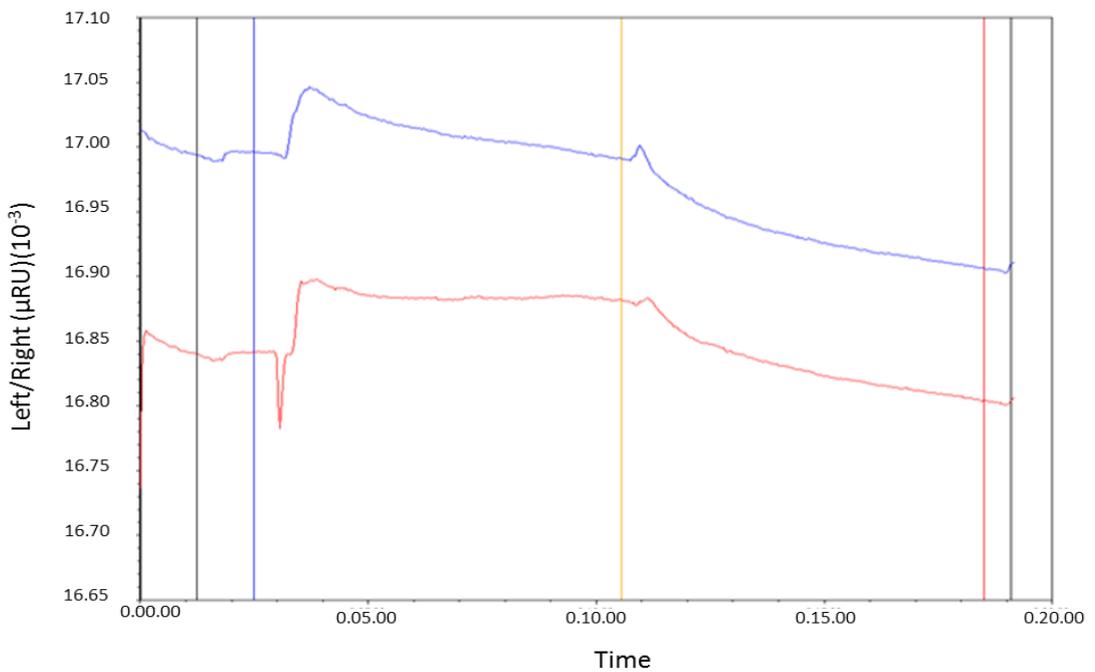
**Fig 5.16:** Far UV CD spectra of CLR and RAMPs. 1 mm cuvettes were filled with 0.7 – 1.0 mg/ml of protein samples (one for each) dialysed in 50 mM sodium phosphate buffer pH 7.5. Spectra were recorded from 260 to 190nm with 8 scans averaged for each sample. Arrows indicate troughs and shoulders at around 208 and 222 nm respectively indicating the presence of  $\alpha$ -helix.

The CLR and RAMP2 samples were tested for their interaction using the SPR. This was done to develop an SPR-based system with which novel interacting partners of RAMPs could be detected. CLR was immobilized and RAMP2 applied to the immobilized phase to investigate binding. Although the immobilization of CLR was successful (Fig 5.17a), the association of RAMP2 was inconclusive as nonspecific binding was also observed (Fig 5.17b). The nonspecific binding might be due to imperfect immobilization of the CLR or some technical reasons and therefore required troubleshooting.

a



b



**Fig 5.17:** Surface plasmon resonance sensograms for the (a) immobilization of CLR and (b) binding of RAMP2 to CLR. C-terminally biotinylated CLR ECD was immobilized on streptavidin-coated sensor chip and RAMP2 run over to investigate any specific interaction (see section 2.3.2.3 for more details on methodology). The difference in the two levels indicated by arrows in (a) is indicative of successful immobilization. The similarity observed in the pattern of the two horizontal lines indicates that there was non-specific binding.

### 5.3 Discussion

The soluble N-terminal extracellular domains of the human CLR and RAMP1, -2 & -3 have here been successfully produced, albeit with some bottlenecks, especially the heavy glycosylation of CLR and RAMP3. The recombinant expression of these receptor proteins' soluble N-termini was accomplished by the presence of an  $\alpha$ -factor secretion peptide upstream the plasmid vector cloning site. This secretion peptide was necessary as CLR requires the RAMPs for trafficking to the cell surface from the endoplasmic reticulum (McLathie *et al.*, 1998, Miret *et al.*, 2002) and vice versa, though to a certain and varying degree(s) (Hilairret *et al.*, 2001; Flahaut *et al.*, 2003). The constructs also contained a 10xHis epitope tag in-frame to the 5' end of CLR- or RAMP-encoding cDNA for purification and detection.

The results obtained at various stages of the expression and characterization processes showed the success of the strategy employed in the production of these receptor proteins. Although the expression of human CLR and RAMPs in some eukaryotic systems e.g. *S. cerevisiae* (Miret *et al.*, 2002) and *Xenopus* oocyte (Flahaut *et al.*, 2003) have been reported, none has been reported for the ectodomains of these receptor proteins in *P. pastoris*. The major problem encountered in this study was the heterogeneous glycosylation of these proteins especially for CLR and RAMP3. These N-glycosylation sites are all localized within the N-terminus of the proteins. The inability of the protein samples to efficiently traverse the gel was totally alleviated following protein deglycosylation (Fig 5.10). It is most likely that the heavily formed sugar moieties caused the forming of smears. Smearing is usually observed for glycosylated proteins expressed in eukaryotic expression systems (Miret *et al.*, 2002; Flahaut *et al.*, 2003). Apart from the mobility of these proteins, deglycosylation also helped reveal a clearer resolution and distinct bands (Fig 5.10 and 5.15). Deglycosylation helped reveal bands corresponding to dimers especially for CLR at ~37 kDa and RAMP3 at ~32 kDa, which were not clearly observed for the non-deglycosylated forms (Fig 5.10 and 5.15b). This, together with that observed for RAMP1 and RAMP2 in the western blot analysis of Fig 5.14, suggests that CLR and RAMPs form stable homodimers which are resistant to denaturing and/or reducing conditions. This finding is consistent with that observed for RAMP1 and RAMP3 by Sexton *et al.* (2001). This current

study further reveals that the dimerization could be independent of the transmembrane domains of these proteins, although the degree or nature of dimerization may vary. It is also possible that RAMP3 also forms a homotrimer as reflected by bands observed at around 50 kDa which is consistent in the results obtained in Fig 5.10 and 5.15.

It is not known what role homo-oligomerization specifically plays in overall receptor signalling. CLR and RAMPs associations have been consistently described as a 1:1 heterodimerization (McLatchie *et al.*, 1998; ter Haar *et al.*, 2010; Kusano *et al.*, 2011) meaning that it is unlikely to play a role in receptor formation. However, as suggested by Sexton *et al.* (2001), it is possible that homo-oligomerization plays a role in regulating the amount of component protein available for CLR/RAMP hetero-dimerization for functional receptor formation, or even by preventing the availability of the receptor for endocytosis pathway especially in the case of CLR. In a broader sense, this could also imply some individual functions played by CLR or RAMPs independent of the other but this remains speculative.

The multiple bands observed for RAMP3 in Fig 5.11c (bands indicated by short arrows, though faint) and in Fig 5.15a (lane 8) imply its heterogenous glycosylation and is consistent with those observed for this protein in previous studies (Flahaut *et al.*, 2003; Sexton *et al.*, 2001). Although the RAMP3 protein expressed in these previous studies was the full length, this should not affect the glycosylation pattern as the glycosylation sites of RAMP3 (as with CLR and RAMP2) are located within the N-terminus. RAMP3 has four consensus N-linked glycosylation sites at N28, N58, N71 and N103 (Fig 5.1) which are glycosylated to various degrees with N71 and N103 more efficiently glycosylated than N28 and N58 (Flahaut *et al.*, 2003). This protein therefore usually appears as 3 molecular species and here in this study, was observed at around 19, 21 and 24 kDa (Fig 5.11c). The band corresponding to the 16 kDa marker could be as a result of incomplete glycosylation while the bands below the molecular weight mark of 15 kDa, especially observed in Fig 5.15, are most likely degraded fragments of the protein. CLR on the other hand has N-linked glycosylation sites at positions 66, 118 and 128 (Muff *et al.*, 2001) but usually show uniform glycosylation and was observed as a single band for the full length receptor expressed in a eukaryotic expression system (Flahaut *et al.*, 2003). In Fig

5.11b and 5.12b, purified CLR ECD showed a distinct band at around 23 kDa which is ~ 5 kDa higher than the calculated MW as a result of glycosylation. This is therefore in agreement with the work of Flahaut *et al.* (2003). Although there were additional bands observed below the said MW mark in Fig 5.11b, these are most likely as a result of degradation and/or the loss of the epitope tag as these bands were not observed in the equivalent western blot picture (Fig 5.12b). The loss of epitope tag has also been observed for calcitonin receptor and RAMP1 expressed in *Escherichia coli* (personal communication with Dr Harriet Watkins, University of Auckland, New Zealand). This suggests that the *P. pastoris* expression system could be a good representation of the mammalian system in recombinantly producing these receptor proteins especially when considering posttranslational modifications such as glycosylation.

The role of glycosylation in calcitonin family peptide receptor pharmacological properties has not been well established. While CLR is terminally glycosylated when co-expressed with RAMP1 in mammalian cells, it shows core glycosylation in the presence of RAMP2 or -3 suggesting that glycosylation may play a role in CLR and RAMP association (Foord and Marshall. 1999), although this was thought to have been influenced by the choice of epitope tag and concluded that RAMPs made no difference to glycosylation (Hilairt *et al.*, 2001). In Schneider insect cells, the glycosylation of CLR is undisturbed by any of the RAMPs but still influenced its functioning (Aldecoa *et al.*, 2000). In CLR, single mutation of N60 or N112 to threonine had no significant effect on cell surface expression while double mutation of same residues to threonine disrupted cell surface expression and N117T substitution impaired receptor function in the presence of RAMP1 and -2 (see Buhlmann *et al.*, 2000 and Gujer *et al.*, 2001). It is still not certain if these effects were as a direct result of altered glycosylation as N117D mutation showed caused no significant deviation from WT. In other words, the role of glycosylation in receptor signalling still hangs in the balance but from all indications, it cannot be discarded.

The protease deficient strain tested due to the recurrent issue of degradation did not offer any noticeable remedy especially for CLR and RAMP3 (Fig 5.14). This is not unexpected as there have been mixed reports about the successes recorded in the use of these strains in alleviating the problem

of proteolytic degradation (reviewed in Macauley-Patrick *et al.*, 2005). The protease-deficient strain SMD1163 (*his4 pep4 prb1*), used in this study, has disruption in the genes encoding proteinase A (*PEP4*) and B (*PRB1*) which results in the inactivation of vacuolar aspartyl protease and some vacuolar carboxypeptidase (see Chapter 1 for details). It is possible that these proteases are not mainly responsible for the degradation of CLR and RAMP proteins expressed herewith. Interestingly, the protease-deficient strain offered a very noticeable difference for RAMP1 in terms of protein quality (Fig 5.14). While faint and multiple bands were seen for RAMP1 expressed in the WT (X33) strain, a very distinct band was observed for RAMP1 expressed in SMD1163 (Fig 5.14). It is possible at this instance that this strain has contributed to RAMP1's resistance to proteolytic degradation. This implies that the benefits of this strain (and possibly other protease deficient strains) could be protein-specific, meaning that a strain could be successfully used to express a particular protein but expresses another poorly.

RAMP2 was resistant to proteolytic degradation, at least under the same circumstance as other proteins (Fig 5.14), irrespective of the expression host strain. This makes this a little more complicated but interesting. Moreover, going by the data obtained from AUC analysis of CLR and RAMP2 and -3, RAMP2 showed no peaks that could be likened to aggregation unlike the case of CLR and RAMP3. Instead, it showed a very distinct peak corresponding to the glycosylated monomer at ~ 20,000 Da and a much smaller peak corresponding to a trimer at ~ 60,000 Da. The reasons for these varying observations are difficult to predict. RAMP2 has only one glycosylation site at N130 (Kusano *et al.* 2011). It is uncertain whether the aggregation of CLR and RAMP3, from the several broad peaks observed for them in Fig 5.13, are influenced by their glycosylation status. Possible reasons and underlining mechanisms for these varying observations are still open to experimental findings.

Another interesting feature in the results obtained in this study is that the proteins were found to be folded. This was interesting because being the N-terminal ECD and individually expressed for the first time, it was not certain what effect this would pose on the overall secondary structures of these receptor proteins. The circular dichroism (CD) plots of RAMP1, -2 and -3 (Fig 5.16) revealed spectra

typical of rich  $\alpha$ -helical content with troughs at around 208 and 222 nm and a peak at near 190nm (Kelly *et al.*, 2005). Crystal structures exist for RAMP1 (Kusano *et al.*, 2008) and RAMP2 (Kusano *et al.*, 2011) which showed these proteins to be predominantly (up to 90%)  $\alpha$ -helical. Although no similar structure currently exists for RAMP3, it is assumed that they show similar structural architecture as RAMP1 and -2. The slightly altered CD spectra of RAMP3 could be as a result of its heavy heterogeneous glycosylation. Glycosylation of proteins is capable of altering their CD spectra albeit to a reasonably low degree (Liu *et al.*, 2007). This could also be a reflection of the protein quality as an aftermath of protein glycosylation. CLR on the other hand showed that it contains anti-parallel  $\beta$ -strands as well as  $\alpha$ -helix and random coil (Fig 5.16). Like RAMP1 and -2, crystal structures exist for CLR ECD (ter Haar *et al.*, 2010) and the overall constituent of this protein is reflected in the far UV spectra with the antiparallel  $\beta$ -strand been predominant over  $\alpha$ -helix with a converged minimum between 210 and 220 nm and a peak at about 192 nm (Kelly *et al.*, 2005). Moreover, the spectra obtained here for CLR ECD agrees to an appreciable extent with that reported for the mouse CRFR2 $\beta$  ECD (Perrin *et al.*, 2003).

In conclusion, the successful production of the CLR and RAMPs ECD in *P. pastoris* eukaryotic expression system has been established. The heavy heterogeneous glycosylation of the proteins, especially for CLR and RAMP3 which was a bottleneck, was overcome by deglycosylating the proteins and this has also provided more information about homo-dimerization of these receptor proteins independent of one another. Also, the successful immobilization of the CLR ECD on SPR makes this system promising as a tool for investigating novel interacting partners of RAMPs. Obtaining these receptor proteins as individually folded moiety was another highlight of the successes recorded in this work, although it is not yet known how this may affect CLR/RAMPs interaction. Above all, this strategy has now produced the soluble N-terminal extracellular domains of CLR and RAMP1, -2 and -3 which could now be used for further biophysical characterization. This looks promising and could be extended to determining novel receptor-receptor heterodimerization and non-receptor partners.

## Chapter 6: General discussion and Conclusion

The CGRP receptor has a more complex pharmacology than other GPCRs. This is largely due to CLR's associations with the receptor activity modifying protein 1 (RAMP1) and receptor component protein (RCP), which are required to form a functional receptor. A clear mechanism of receptor signalling is yet to be established for this receptor. This has not been helped by the lack of a full receptor structure for this receptor or any other family B GPCRs. Using (mainly alanine scan) mutagenesis, this thesis has dug further into this mechanism by investigating residues within the N- and C-termini with the aim of better understanding how this receptor performs its signalling. Results obtained have been interpreted based on the general paradigm of the family B GPCRs and available information from the family A. The soluble N-terminus of the CLR and RAMP1, -2 and -3 (RAMPs) have also been produced and characterized to help develop a system that studies RAMPs interaction with other receptors.

Chapter 3 investigated the N-terminal extracellular domain (ECD) of the CLR as a component of the CGRP receptor. Mutation of all residues selected for investigation significantly affected, to a varying degree,  $\alpha$ CGRP potency,  $\alpha$ CGRP-mediated receptor internalization and/or cell surface expression. The exception to this was D90A that showed significant, but relatively negligible, effect on cAMP stimulation and behaved like wild type in other parameters employed for analysing the mutants. These residues are mainly localised within the  $\beta$  hairpin structure and loop of CLR ECD. The localization of these residues, which have now been shown to be crucial for receptor signalling, implies that these strictly structurally conserved  $\beta$  hairpin and loop structures are highly significant for the distinct ligand binding and activation prototype of this family of receptors. This is a possible explanation for the works of Stroop *et al.* (1995) and Bergwitz *et al.* (1996), for instance, where hybrid receptors (from within the B family) of varying N-terminus and Juxta domain with C-terminus displayed similar pattern of ligand-receptor interaction with corresponding composite ligands (reported in Chapter 1). Despite the positions occupied by these residues within the N-terminal ECD of CLR as well as the significant effects their mutations showed on receptor signalling when compared to wild type, it could not be ascertained if they directly interact with CGRP. This is, as earlier stated, as a result of the

unavailability of CGRP binding data which would indicate an effect of mutation on  $K_D$ . Thus their role in ligand binding remains, to some extent, speculative.

Although the N-terminus has been widely recognised as mainly required for ligand binding, investigation on the ECD also suggests a possible role in regulating the pathways for receptor coupling Gs or  $\beta$ -arrestin. This is suggestive from the result obtained for the K103A mutant receptor (Chapter 3), although there could also be some other residues within the ECD with similar role acting individually or synergistically. Possible mechanisms for this might be that the residue(s) influences the orientation of CGRP receptor in such a manner that it is committed to Gs coupling and not  $\beta$ -arrestin and vice versa. It is also possible that such residue affects the conformation of the ECD and consequently stabilizes the EC loops, and possibly the TM, in a conformation that influences Gs coupling or  $\beta$ -arrestin binding. This thus shows the need for more information on the ECD/TM bundle interface.

As mentioned earlier, while crystal/NMR structures exist for the ECD of receptors within the B family with great deal of similarity observed in structural signature and ligand-binding pattern, there is still inconsistency observed especially in receptor-ligand interaction (Miller *et al.*, 2012). An opposite orientation of ligand in ligand-receptor interaction observed for the VPAC1 receptor when compared to other family B members (Couvineau *et al.*, 2011) makes this more controversial but interesting. Some of these differences are believed to define individual specific ligand binding and it is therefore in itself a valuable reason why the study of each receptor is of high importance. Moreover, since the peptides of the secretin family appear to characteristically fall into 3 or more subgroups (e.g. the glucagon and calcitonin subfamily; see Chapter 1 for further details), it is very likely that these ligands exhibit some unique variations in specifically associating with their receptors.

Even though much of the information available on the structural mechanism of family B GPCRs activation remain hypothetical, what remains an undisputable fact is that the N-terminal ECD of this family of receptors is very important for ligand binding at least for those with solved structures of ligand-bound ECD. The work carried out in Chapter 3, together with the work of Barwell *et al.*

(2010), has further reiterated this for the CGRP receptor. More specifically and importantly, the work has mapped out, for the first time, residues constituting a putative ligand binding core, which most likely represents the cleft that binds the C-terminal and helical mid-region of CGRP.

Chapter 4 dwelt on helix 8 and associated C-terminal region. As far as the B family of GPCRs is concerned, the presence of H8 still remains hypothetical, implying that little is known about it (see Chapter 1). Despite the minimal information available for this part of the receptor, especially in the family B and C, the H8 has not been widely studied across all families of GPCRs. This further makes this work of high importance.

The work identified several residues at this region required for receptor signalling. Most notable are A393 and W399 within the 8<sup>th</sup> helix and I404 in the region immediately flanking H8 at its C-terminal end. The mutation of these residues showed significant effects on 3 or more of the parameters employed in the investigations. Although A393 appeared particularly important especially considering the effects of its mutation (i.e. A393D) on basal activity and  $\alpha$ CGRP-mediated internalization of the receptor, these effects might be a reflection of the charge on the Asp residue with which this residue was substituted rather than the actual absence of A393. Among other inferences that could be drawn from this, an interesting one to point out is that the amphipathic nature of this helix is plausibly a required feature for receptor signalling. Moreover, it has been observed that the charged residues component of H8 across various GPCR families are richly basic (Paker and Parker, 2010). This might not be unconnected with the orientation of H8 relative to the phospholipid bilayer, where it lies at an interface between the membrane bilayer and the cytoplasm. So, it is possible that A393D mutation has disrupted this charge balance and, as earlier suggested, it would be interesting to see what effect a Leu mutation will have.

Also peculiar to GPCRs H8, in the A family, is the phenomenon of palmitoylation. Although (as described in Chapter 1) this is not present in all family A members with solved structures and while one might be tempted to think that this feature might not be required for helix 8 anchoring to the membrane, it is possible that there are some other elements in other family A GPCRs (without the

palmitoylation site) that play similar role. In Chapter 4, a similar anchoring role was suggested for W399, which may also involve Y402. The former is invariantly conserved in receptors of family B. This may therefore represent a common theme/model by which these receptors stabilize or anchor their H8, when present, to the phospholipid membrane bilayer just like the common signatures observed in the N-termini among receptors of this family.

Results obtained in Chapter 4 showed that the H8 of CLR influences  $\beta$ -arrestin and G-protein coupling and has a subtle effect on receptor cell surface expression. This is in agreement with what have been reported for family A GPCRs where their H8 participate in conformational changes accompanying receptor activation (Hoersch *et al.*, 2008) and  $\beta$ -arrestin-dependent receptor internalization (Kirchberg *et al.*, 2011). Conversely, as discussed in Chapter 4, there have been disagreements in the roles reported for H8. Taking a close look at their amino acid composition particularly at this region, it is tempting to think that the differences in reported roles might relate to the distribution of Ser/Thr residues (which are potential phosphorylation sites) within H8 and associated C-terminal region. For instance, no Ser/Thr residues are present among the amino acid residues spanning helix 8 and even among the first 8 residues beyond H8 (see Fig 4.1) in human CLR while in its ‘sister receptor, the human CTR, 2 Thr residues are found. Also, while there are no Ser/Thr in H8 and associated C-terminal region of human CLR, VIPR1 and SCTR, other family members possess Ser and/or Thr within this region.

Apart from the aforementioned probable mechanisms by which these residues could affect receptor signalling, it is important to add that some might act by influencing ligand binding to the receptor. Although the region within which these residues are localized rules out any possibility of their direct involvement in ligand interaction, it is possible that they affect receptor conformational changes between the resting and active states, thus influencing ligand binding and hence, receptor signalling. Some residues within this same region (H8) of CLR have been found to significantly affect  $K_D$  in AM receptor (Kuwasako *et al.*, 2011). It is however not known if they show similar effect in CGRP receptor due to the lack of ligand binding data.

Also investigated in Chapter 4 was F163 at the bottom of TM1. A common stabilizing interaction for H8 among family A GPCRs is with ICL1 and/or the bottom of TM1 (see Chapter 1). CLR H8 may make similar interaction as postulated by Vohra *et al.* (2013) and F63 may play a key role in this. More effort is needed to be put into the molecular structure determination of CLR, as well as other family B receptors, to give clearer picture of the conformation of and interactions by this helical segment and associated C-terminal region.

In Chapter 5, the ECDs of CLR and RAMP1, -2 and -3 were recombinantly produced and were subjected to biophysical characterizations especially analytical ultracentrifugation (AUC) and circular dichroism (CD). A series of troubleshooting experiments were performed to alleviate some technical problems most especially the heavy glycosylation of the soluble protein fragments, which affected their resolution by SDS PAGE. CLR/RAMP interaction was also tested using surface plasmon resonance (SPR).

The various successes recorded in structural studies of the ECDs of family B receptors and RAMPs, coupled with the size of and role played by this domain, has given a clear impression that the domain could be used as a means to learn more about these receptors and probably even other non-family B receptors, albeit to a limited extent. RAMPs have already been found to associate with some other GPCRs like VPAC1 and -2, PTH1 and CRF1 receptors in family B (Christopoulos *et al.*, 2003; Wootten *et al.*, 2013) and even the calcium sensing receptor in family C (Bouschet *et al.*, 2005). Although it is not known whether the ECD of these receptors also interact with RAMPs ECD, the realization of this in the case of CLR/RAMP1 (ter Haar *et al.*, 2010; Koth *et al.*, 2010) and CLR/RAMP2 (Kusano *et al.*, 2011) indicates that this might also work with these other receptors. Moreover, Chapter 5 showed that the ECD of these receptor proteins may form dimers and trimers which suggests they could make similar interactions with other receptor proteins at the N-terminus. In other words, the system reported in Chapter 5 could be extended to producing and screening novel partners of RAMPs considering the complexity currently involved in producing stable full length of these receptors. Moreover, the *Pichia pastoris* expression system used here produced reasonably large (up to an average of 4 mg) amounts of protein from as little as 2 g of wet cells and the SPR on the

other hand requires as little as around 0.3 mg of protein sample. However it is important to state that this SPR system still requires some amount of testing and optimization.

In conclusion, gaining full understanding of the mechanism of ligand binding and activation of the CGRP receptor is very demanding. The complete comprehension of this molecular mechanism entails understanding the various conformational changes through which the receptor passes in the course of ligand recognition and binding to G-protein coupling and beyond. This currently remains elusive especially in the absence of the molecular structure of the TM domain of this receptor. Aside from the on-going quest to understand how this receptor is activated, new interests are emerging, with more focus on the roles RAMPs (and possibly other non-GPCR molecules) may play in the signalling of GPCRs. Overall, while there has been laudable research on this receptor and its modifying proteins, the task of unfolding their mechanism of signalling is also far from done. Thus, the hunt of this information remains fascinating and novel.

## References

- Aldecoa A, Gujer R, Fischer JA and Born W. (2000). Mammalian calcitonin receptor-like receptor/receptor activity modifying protein complexes define calcitonin gene-related peptide and adrenomedullin receptors in *Drosophila Schneider 2* cells. *FEBS Lett.* 471(2-3):156-60
- Ali, H. and Haribabu, B. (2006) Transmembrane Signaling protocols. 2nd Edition. *Humana Press*. Pp 54-55.
- Alibés A, Nadra AD, De Masi F, Bulyk ML, Serrano L and Stricher F. (2010). Using protein design algorithms to understand the molecular basis of disease caused by protein-DNA interactions: the Pax6 example. *Nucleic Acids Res.* 38(21):7422-31
- Amara, S. G., Jonas, V., Rosenfeld, M. G., Ong, E. S., & Evans, R. M. (1982). Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. *Nature*, 298, 240-244.
- Andre, N., Cherouati, N., Prual, C., Steffan, T., Zeder-Lutz, G. et al (2006) Enhancing functional production of G protein-coupled receptors in *Pichia pastoris* to levels required for structural studies via a single expression screen. *Protein Science* 15: 1115-1126
- Aoki, H., Ahsan, N.M. and Watabe, S. (2003) Heterologous expression in *Pichia pastoris* and single-step purification of a cysteine proteinase from northern shrimp. *Protein Expression and Purification* 31: 213-221
- Armbruster, B.N. and Roth, B.L. (2004) Mining the Receptorome. JBC Papers in Press. Manuscript 1-23
- Bailey, R. J., Bradley, J. W. I., Poyner, D. R., Rathbone, D. L., & Hay, D. L. (2010). Functional characterization of two human receptor activity-modifying protein 3 variants. *Peptides*, 31(4), 579-584.
- Bailey, R. J., & Hay, D. L. (2006). Pharmacology of the human CGRP1 receptor in Cos 7 cells. *Peptides* 27(6), 1367-1375.
- Ballesteros JA, Jensen AD, Liapakis G, Rasmussen SG, Shi L, Gether U, Javitch JA. (2001) Activation of the beta 2-adrenergic receptor involves disruption of an ionic lock between the cytoplasmic ends of transmembrane segments 3 and 6. *J. Biol. Chem.* 276, 29 171– 29 177
- Banerjee, S., Evanson, J., Harris, E., Lowe, S., Speth, R., Thomasson, K., & Porter, J. (2006). Correction: Identification of specific calcitonin-like receptor residues important for calcitonin gene-related peptide high affinity binding. *BMC Pharmacol*, 6(1), 14.
- Barwell J, Wheatley M, Conner AC, Taddese B, Vohra S, Reynolds CA, Poyner DR. (2013). The activation of the CGRP receptor. *Biochem Soc Trans.* 1;41(1):180-4
- Barwell J, Gingell JJ, Watkins HA, Archbold JK, Poyner DR, Hay DL. (2012). Calcitonin and calcitonin receptor-like receptors: common themes with family B GPCRs? *Br J Pharmacol.* 166(1):51-65.
- Barwell J, Conner A, Poyner DR. (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(10):1906-16
- Barwell, J., Miller, P.S., Donnelly, D. and Poyner, D.R. (2010) Mapping interaction sites within the N-terminus of the calcitonin gene-related peptide receptor; the role of residues 23–60 of the calcitonin receptor-like receptor. *Peptides.* 31(1-3): 170–176.

- Benovic JL, Kuhn H, Weyland I, Codina J, Caron MG and Lefkowitz RJ (1987). Functional desensitization of the isolated  $\beta$ -adrenergic receptor by the  $\beta$ -adrenergic receptor kinase: Potential role of an analog of the retinal protein arrestin (48-kDa protein). *Proc Natl Acad Sci USA* 84:8879–8882
- Bergwitz, C., Gardella, T. J., Flannery, M. R., Potts, J. T., Kronenberg, H. M., Goldring, S. R., & Juppner, H. (1996). Full Activation of chimeric receptors by hybrids between parathyroid hormone and calcitonin. Evidence for a common pattern of ligand-receptor interaction. *J Biol Chem*, 271(43), 6469-26472.
- Bill RM, Henderson PJF, Iwata S, Kunji ERS, Michel H, Neutze R, Newstead S, Poolman B, Tate CG and Vogel H (2011). Overcoming barriers to membrane protein structure determination. *Nat Biotechnol*, 29: 335-340.
- Bissantz C, Bernard P, Hibert M, Rognan D. (2003). Protein-based virtual screening of chemical databases. II. Are homology models of G-protein coupled receptors suitable targets? *Proteins Struct Funct Genet* 50:5–25.
- Bockaert J, Fagni L, Dumuis A and Marin P. (2004). GPCR interacting proteins (GIP). *Pharmacol Ther.* 103(3):203-21.
- Bühlmann N, Aldecoa A, Leuthäuser K, Gujer R, Muff R, Fischer JA and Born W. (2000). Glycosylation of the calcitonin receptor-like receptor at Asn(60) or Asn(112) is important for cell surface expression. *FEBS Lett.* 486(3):320-4.
- Campbell, M.K. and Farrell, S.O. (2006) *Biochemistry*. 5th Edition. Cengage Learning. Pp 355-6.
- Chan KY, Pang RT, Chow BK (2001). Functional segregation of the highly conserved basic motifs within the third endloop of the human secretin receptor. *Endocrinology* 142: 3926–3934
- Chen Q, Pinon DI, Miller LJ, Dong M (2010). Spatial approximations between residues 6 and 12 in the amino-terminal region of glucagon-like peptide 1 and its receptor: a region critical for biological activity. *J Biol Chem* 285: 24508–24518.
- Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, et al. (2007) High-resolution crystal structure of an engineered human  $\beta_2$ -adrenergic G protein-coupled receptor. *Science* 318:1258–65.
- Childs, G.V. (2003) *Membrane Structure and Function*. Cell Biology Graduate Programme, University of Texas medical Branch (UTMB)  
[http://cellbio.utmb.edu/cellbio/membrane\\_intro.htm](http://cellbio.utmb.edu/cellbio/membrane_intro.htm)
- Christopoulos, A., Christopoulos, G., Morfis, M., Udawela, M., Laburthe, M., Couvineau, A., Kuwasako, K., Tilakaratne, N., & Sexton, P. M. (2003). Novel receptor partners and function of receptor activity-modifying proteins. *J Biol Chem*, 278(5), 3293-3297.
- Chugunov AO, Simms J, Poyner DR, Dehouck Y, Rooman M, Gilis D, Langer I. (2010). Evidence that interaction between conserved residues in transmembrane helices 2, 3, and 7 are crucial for human PAC1 receptor activation. *Mol. Pharmacol.* 78, 394–401.
- Conklin BR and Bourne HR. (1993). Structural elements of G  $\alpha$  subunits that interact with G  $\beta$   $\gamma$ , receptors, and effectors. *Cell* 73(4):631-41
- Conner, M., Hicks, M.R., Dafforn, T. et al (2008) Functional and Biophysical Analysis of the C-Terminus of the CGRP-Receptor; a Family B GPCR. *Biochemistry* 47: 8434-8444
- Conner, A.C., Simms, J., Barwell, J., Wheatley, M. and Poyner, D. (2007) Ligand Binding and Activation of the CGRP receptor. *Biochemical Society Transaction* 35 (4): 729-732

- Conner AC, Simms J, Conner MT, Wootten DL, Wheatley M, Poyner DR. (2006). Diverse functional motifs within the three intracellular loops of the CGRP(1) receptor. *Biochemistry* 45, 12 976–12 985
- Conner, A. C., Hay, D. L., Simms, J., Howitt, S. G., Schindler, M., Smith, D. M., Wheatley, M., & Poyner, D. R. (2005). A key role for transmembrane prolines in calcitonin receptor-like receptor agonist binding and signalling: implications for family B G- protein-coupled receptors. *Mol Pharmacol*, 67(1), 20-31.
- Conner AC, Hay DL, Howitt SG, Kilk K, Langel U, Wheatley M, Smith DM, Poyner DR. (2002). Interaction of calcitonin-gene-related peptide with its receptors. *Biochem Soc Trans* 30(4):451-5
- Cooper, G. J., Willis, A. C., Clark, A., Turner, R. C., Sim, R. B., & Reid, K. B. (1987). Purification and characterization of a peptide from amyloid-rich pancreases of type 2 diabetic patients. *Proc Natl Acad Sci U S A*, 84(23), 8628-8632.
- Coopman K, Wallis R, Robb G, Brown AJ, Wilkinson GF, Timms D, Willars GB. (2011). Residues within the transmembrane domain of the glucagon-like peptide-1 receptor involved in ligand binding and receptor activation: modelling the ligand-bound receptor. *Mol. Endocrinol.* 25, 1804– 1818
- Couvineau A, Laburthe M (2011). VPAC receptors: structure, molecular pharmacology and interaction with accessory proteins. *Br J Pharmacol* 166:42–50
- Couvineau A, Lacapere JJ, Tan YV, Rouyer-Fessard C, Nicole P, Laburthe M. (2003) Identification of cytoplasmic domains of hVPAC1 receptor required for activation of adenylyl cyclase. Crucial role of two charged amino acids strictly conserved in class II G protein-coupled receptors. *J. Biol. Chem.* 278, 24 759– 24 766.
- Daulat, A.M., Maurice, P and Jockers, R (2008) Recent methodological advances in the discovery of GPCR-associated proteincomplexes. *Trends in Pharmacological Sciences* 30(2): 72=78
- Durham, P.L. (2004) CGRP-Receptor Antagonists – A Fresh Approach to Migraine Therapy? *New England Journal of Medicine.* 350: 11
- Devi, A.L. (2005) The G protein coupled receptors handbook. Illustrated Edition. Springer. Pp 4-10
- Donnelly D. 1997 The arrangement of the transmembrane helices in the secretin receptor family of G-protein-coupled receptors. *Febs Lett.* 409, 431–436
- Dong M, Lam PC, Pinon DI, Hosohata K, Orry A, Sexton PM, Abagyan R, Miller LJ. (2011). Molecular basis of secretin docking to its intact receptor using multiple photolabile probes distributed throughout the pharmacophore. *J. Biol. Chem.* 286, 23 888 – 23 899
- Dong M, Lam PC, Gao F, Hosohata K, Pinon DI, Sexton PM, Abagyan R, Miller LJ. (2007) Molecular approximations between residues 21 and 23 of secretin and its receptor: development of a model for peptide docking with the amino terminus of the secretin receptor. *Mol. Pharmacol.* 72, 280 – 290.
- 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(2), 1167-1175.
- Doods, H., Hallermayer, G., Wu, D., Entzeroth, M., Rudolf, K., Engel, W., & Eberlein, W. (2000). Pharmacological profile of BIBN4096BS, the first selective small molecule CGRP antagonist. *Br J Pharmacol*, 129(3), 420-423.

- Egea SC, Dickerson IM. (2012). Direct interactions between calcitonin-like receptor (CLR) and CGRP-receptor component protein (RCP) regulate CGRP receptor signaling. *Endocrinology* 153(4):1850-60.
- Estall JL, Koehler JA, Yusta B and Drucker DJ. (2005). The glucagon-like peptide-2 receptor C terminus modulates beta-arrestin-2 association but is dispensable for ligand-induced desensitization, endocytosis, and G-protein-dependent effector activation. *J Biol Chem.* 280(23):22124-34
- Fergusson, S.S. (2001). Evolving concepts in G protein-coupled receptor endocytosis: the role in receptor desensitization and signalling. *Pharmacol Rev.* 53(1):1-24.
- Findlay DM, Houssami S, Lin HY, Myers DE, Brady CL, Darcy PK et al. (1994). Truncation of the porcine calcitonin receptor cytoplasmic tail inhibits internalization and signal transduction but increases receptor affinity. *Mol Endocrinol* 8: 1691–1700.
- Fitzsimmons, T. J., Zhao, X., & Wank, S. A. (2003). The extracellular domain of receptor activity-modifying protein 1 is sufficient for calcitonin receptor-like receptor function. *J Biol Chem*, 278(16), 14313-14320.
- Flahaut, M., Pfister, C., Rossier, B., & Firsov, D. (2003). N-Glycosylation and conserved cysteine residues in RAMP3 play a critical role for the functional expression of CRLR/RAMP3 adrenomedullin receptor. *Biochemistry* 42(34), 10333-10341.
- Flanagan, C.A. (2005) A GPCR That Is Not “DRY”. *Molecular Pharmacology* 63: 1-3.
- Fraser, N.J. (2006) Expression and functional purification of a glycosylation deficient version of the human adenosine 2a receptor for structural studies. *Protein Expression and Purification* 49: 129-137
- Fredriksson R, Schiöth HB. (2005) The repertoire of G-protein-coupled receptors in fully sequenced genomes. *Mol. Pharmacol.* 67:1414–25
- Fredriksson R, Lagerström MC, Lundin LG, Schiöth HB. 2003 The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* 63, 1256 – 1272.
- Frimurer TM, Bywater RP. 1999 Structure of the integral membrane domain of the GLP1 receptor. *Proteins* 35, 375–386
- Garcia GL, Dong M, Miller LJ. (2012). Differential determinants for coupling of distinct G proteins with the Class B secretin receptor. *Am. J. Physiol. Cell Physiol.* 302, C1202 – C1212.
- Gaudin P, Maoret JJ, Couvineau A, Rouyer-Fessard C, Laburthe M. (1998). Constitutive activation of the human vasoactive intestinal peptide 1 receptor, a member of the new class II family of G protein-coupled receptors. *J Biol Chem.* 273:4990–4996
- GenWay Biotech, Inc. (2013). Protein Expression <http://www.genwaybio.com/technologies/protein-expression>
- Gether U, Asmar F, Meinild AK, Rasmussen SG. (2002). Structural basis for activation of G-protein-coupled receptors. *Pharmacol Toxicol* 91(6):304-12.
- Gether, U. (2000) Uncovering Molecular Mechanism Involved in Activation of G Protein-Coupled Receptors. *Endocrine Reviews* 21(1): 90-113

- Grace, C. R. R., Perrin, M. H., DiGrucchio, M. R., Miller, C. L., Rivier, J. E., Vale, W. W., & Riek, R. (2004). NMR structure and peptide hormone binding site of the first extracellular domain of a type B1 G protein-coupled receptor. *Proc Natl Acad Sci U S A*, *101*(35): 12836-12841.
- Grace, C. R. R., Perrin, M. H., Gulyas, J., DiGrucchio, M. R., Cattle, J. P., Rivier, J. E., Vale, W. W., & Riek, R. (2007). Structure of the N-terminal domain of a type B1 G protein-coupled receptor in complex with a peptide ligand. *Proc Natl Acad Sci U S A*, *104*(12), 4858-4863.
- Gruenberg J, Maxfield FR (1995). Membrane transport in the endocytic pathway. *Curr Opin Cell Biol*. *7*(4):552-63.
- Guerois R, Nielsen JE and Serrano L. (2002). Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J Mol Biol*.*320*(2):369-87
- Gujer R, Aldecoa A, Bühlmann N, Leuthäuser K, Muff R, Fischer JA and Born W. (2001). Mutations of the asparagine117 residue of a receptor activity-modifying protein 1-dependent human calcitonin gene-related peptide receptor result in selective loss of function. *Biochemistry*. *40*(18):5392-8.
- Hanson, B.J. (2006) Multiplexing Fluo-4 NW and a GeneBLAzer® Transcriptional Assay for High-Throughput Screening of G-Protein-Coupled Receptors. *Journal of Biomolecular Screening* *11*(6): 644 – 651
- Harikumar, K. G., Simms, J., Christopoulos, G., Sexton, P. M., & Miller, L. J. (2009). Molecular basis of association of receptor activity-modifying protein 3 with the family B G protein-coupled secretin receptor. *Biochemistry* *48*(49), 11773-11785.
- Harmar, A.J. (2001) Family-B G-protein-coupled receptors. *Genome Biol*. *2*(12): reviews 3013. 1-10
- Hausdorff WP, Bouvier M, O'Dowd BF, Irons GP, Caron MG and Lefkowitz RJ (1989) Phosphorylation sites on two domains of the b2 -adrenergic receptor are involved in distinct pathways of receptor desensitization. *J Biol Chem* *264*:12657–12665.
- Hawtin SR, Simms J, Conner M, Lawson Z, Parslow RA, Trim J, Sheppard A and Wheatley M. (2006). Charged extracellular residues, conserved throughout a G-protein-coupled receptor family, are required for ligand binding, receptor activation, and cell-surface expression. *J Biol Chem*. *281*(50):38478-88
- Hay DL, Walker CS, Poyner DR. (2011). Adrenomedullin and calcitonin gene-related peptide receptors in endocrine-related cancers: opportunities and challenges. *Endocr Relat Cancer*. *13*:18(1):C1-14
- Hay, D.L., Poyner, D.R. and Sexton, P.M. (2006) GPCR modulation by RAMPs. *Pharmacology and Therapeutics* *109*: 173-197
- Hay DL, Christopoulos G, Christopoulos A, Sexton PM (November 2004). "Amylin receptors: molecular composition and pharmacology". *Biochem. Soc. Trans*. *32*
- He, R., Browning, D.D. and Ye, R.D. (2001) Differential Roles of the NPXXY Motif in Formyl Peptide Receptor Signalling. *The Journal of Immunology* *166*: 4099 - 4105
- Heroux M, Hogue M, Lemieux S, Bouvier M. (2007) Functional calcitonin gene-related peptide receptors are formed by the asymmetric assembly of a calcitonin receptor-like receptor homo-oligomer and a monomer of receptor activity-modifying protein-1. *J Biol Chem* *282*:31610–31620.
- Higgins, D.R. and Cregg, J.M. (1998) Methods in Molecular Biology. In: Pichia Protocols. Humana Press, New Jersey. Pp 1-16

- Hilairt, S., Foord, S. M., Marshall, F. H., & Bouvier, M. (2001b). Protein-protein interaction and not glycosylation determines the binding selectivity of heterodimers between the calcitonin receptor-like receptor and the receptor activity-modifying proteins. *J Biol Chem*, 276(31): 29575-29581.
- Hill, S.J., Ganellin, C.R., Timmerman, H., Schwartz, J.C., Shankley, N.P., Young, J.M., Schunack, W., Leri, R. and Hass, H.L. (1997) International Union of Pharmacology. XIII. Classification of Histamine Receptors. *Pharmacological Reviews* 49(3): 253-270
- Hinson, J. P., Kapas, S., & Smith, D. M. (2000). Adrenomedullin, a multifunctional regulatory peptide. *Endocr Rev*, 21(2), 138-167.
- Hjorth, S. A., Orskov, C., & Schwartz, T. W. (1998). Constitutive activity of glucagon receptor mutants. *Mol Endocrinol*, 12(1), 78-86.
- Hoare, S.R. (2005) Mechanisms of peptide and nonpeptide ligand binding to Class B G-protein-coupled receptors. *Drug Discov. Today* 10, 417–427
- Hollinger, M.A. (2003) Introduction to Pharmacology. 2nd edition. CRC Press. Pp 3
- Holmes, W.J., Darby, R.A.J., Wilks, M.D.B., Smith, R. and Bill, R.M. (2009). Developing a scalable model of recombinant protein yield from *Pichia pastoris*: the influence of culture conditions, biomass and induction regime. *Microbial Cell Factories*, 8, 35.
- Holtmann MH, Hadac EM, Miller LJ (1995). Critical contributions of amino-terminal extracellular domains in agonist binding and activation of secretin and vasoactive intestinal polypeptide receptors. Studies of chimeric receptors. *J Biol Chem* 270: 14394–14398.
- Höppener, J. W. M., Ahrén, B., & Lips, C. J. M. (2000). Islet amyloid and type 2 diabetes mellitus. *N Engl J Med*, 343(6), 411-419.
- Howitt SG, Kilk K, Wang Y, Smith DM, Langel U, Poyner DR (2003). The role of the 8-18 helix of CGRP8-37 in mediating high affinity binding to CGRP receptors; coulombic and steric interactions. *Br J Pharmacol* 138(2): 325-332.
- Huber, T., Menon, S. and Sakmar, T.P. (2008) Structural Basis for Ligand Binding and Specificity in Adrenergic Receptors: Implications for GPCR-Targeted Drug Discovery. *American Chemical Society* 47 (42): 11013-11023
- Hulme, E.C., (2013). GPCR activation: a mutagenic spotlight on crystal structures. *Trends Pharmacol Sci*. 34(1):67-84
- Hulme EC, Lu ZL, Ward SD, Allman K. and Curtis CA (1999). The conformational switch in 7-transmembrane receptors: the muscarinic receptor paradigm. *Eur J Pharmacol*.30;375(1-3):247-60.
- Insel, P.A., Tang, C., Hahntow, I. and Michel, M.C. (2007) Impact of GPCRs in clinical medicine: Monogenic diseases, genetic variants and drug targets. *Biochimica et Biophysica Acta* 1768: 994 – 1005.
- Ittner LM, Luessi F, Koller D, Born W, Fischer JA, Muff R (2004). Aspartate (69) of the calcitonin-like receptor is required for its functional expression together with receptor-activity-modifying proteins 1 and -2. *Biochem Biophys Res Commun* 319: 1203–1209
- Jaakola V-P, Griffith MT, Hanson MA, Cherezov V, Chien EYT, Lane JR, et al. (2008). The 2.6 Angstrom crystal structure of a Human A2A adenosine receptor bound to an antagonist. *Science* 322:1211–7

- Jackson, D.A., Symons, R.H. and Berg, P. (1972) Biochemical Method for Inserting New Genetic Information into DNA of Simian Virus 40: Circular SV40 DNA Molecules Containing Lambda Phage Genes and the Galactose Operon of *Escherichia coli*. *Proc. Nat. Acad. Sci.* 69(10): 2904-2909.
- Johnston CA, Kimple AJ, Giguere PM, Siderovski DP (2008). Structure of the parathyroid hormone receptor C terminus bound to the G-protein dimer G $\beta$ 1 $\gamma$ 2. *Structure* 16: 1086–1094.
- Jin, L., Briggs, S.L., Chandrasekhar, S., Chirgadze, N.Y., Clawson, D.K., Schevitz, R.W., Smiley, D.L., (...), Zhang, F. (2000) Crystal structure of human parathyroid hormone 1-34 at 0.9-Å resolution. *Journal of Biological Chemistry*, 275 (35), pp. 27238-27244.
- Johnston, C.A., Kimple, A.J., Giguere, P.M. and Siderovski, D.P. (2008) Structure of the parathyroid hormone receptor C-terminus bound to the G-protein dimer G $\beta$ 1 $\gamma$ 2. *Structure* 16(7): 1086-1094
- Kelly SM, Jess TJ, Price NC (2005). How to study proteins by circular dichroism. *Biochim Biophys Acta.* 2005 Aug 10;1751(2):119-39
- Kirkpatrick A, Heo J, Abrol R, Goddard WA (2012). Predicted structure of agonist-bound glucagon-like peptide 1 receptor, a class B G protein-coupled receptor. *Proc Natl Acad Sci U S A.* 109(49):19988-93
- Kitamura, K., Kangawa, K., Kawamoto, M., Ichiki, Y., Nakamura, S., Matsuo, H., & Eto, T. (1993). Adrenomedullin: a novel hypotensive peptide isolated from human pheochromocytoma. *Biochem Biophys Res Commun*, 192(2), 553-560.
- Knowles TJ, Finka R, Smith C, Lin YP, Dafforn T and Overduin M. (2009). Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *J Am Chem Soc.* 131(22):7484-5
- Knudsen SM, Tams JW, Fahrenkrug J (2001). Functional roles of conserved transmembrane pralines in the human VPAC (1) receptor. *FEBS Lett* 503: 126–130.
- Kobilka, B.K. (2007) G protein coupled receptor structure and activation. *Biochimica et Biophysica Acta* 1768: 794 – 807
- Koller, D., Born, W., Leuthauser, K., Fluhmann, B., McKinney, R. A., Fischer, J. A., & Muff, R. (2002). The extreme N-terminus of the calcitonin-like receptor contributes to the selective interaction with adrenomedullin or calcitonin gene-related peptide. *FEBS Letters*, 531(3), 464-468.
- Koole C, Wootten D, Simms J, Miller LJ, Christopoulos A, Sexton PM. (2012). Second extracellular loop of human glucagon-like peptide-1 receptor (GLP-1R) has a critical role in GLP-1 peptide binding and receptor activation. *J Biol Chem.* 287(6):3642-58
- Koth CM, Abdul-Manan N, Lepre CA, Connolly PJ, Yoo S, Mohanty AK et al. (2010). Refolding and characterization of a soluble ectodomain complex of the calcitonin gene-related peptide receptor. *Biochemistry* 49: 1862–1872
- Krauss, G. (2008) *Biochemistry of signal transduction and regulation*. 4th Edition. Wiley-VCH. Pp 222-5
- Krieger E, Koraimann G. and Vriend G. (2002). Increasing the precision of comparative models with YASARA NOVA--a self-parameterizing force field. *Proteins.* 47(3):393-402

- Kristiansen, K. (2004) Molecular mechanisms of ligand binding, signaling, and regulation within the superfamily of G-protein-coupled receptors: molecular modeling and mutagenesis approaches to receptor structure and function. *Pharmacol Ther.* 103(1):21-80
- Kroeze, W.K. and Roth, B.L. (2006) Screening receptorome. *Journal of Psychopharmacology* 20(4): 41-48
- Kumar S, Pioszak A, Zhang C, Swaminathan K, Xu HE (2011). Crystal structure of the PAC1R extracellular domain unifies a consensus fold for hormone recognition by Class B G-protein coupled receptors. *PLoS One* 6: e19682
- Kurihara, H., Shindo, T., Oh-Hashi, Y., Kurihara, Y., & Kuwaki, T. (2003). Targeted Disruption of Adrenomedullin and alphaCGRP Genes Reveals Their Distinct Biological Roles. *Hypertens Res*, 26(Suppl), S105-S108.
- Kusano, S., Kukimoto-Niino, M., Hino, N., Ohsawa, N., Okuda, K., Sakamoto, K., Shirouzu, M., Shindo, T. and Yokoyama, S. (2011). Structural basis for extracellular interactions between calcitonin receptor-like receptor and receptor activity-modifying protein 2 for adrenomedullin-specific binding. *Protein Sci.* 21(2):199-210
- Kusano, S., Kukimoto-Niino, M., Akasaka, R. et al (2008) Crystal Structure of the Human Receptor Activity-Modifying Protein 1 Extracellular Domain. *Protein Sci.* 17: 1907-1914
- Kuwasako K, Kitamura K, Nagata S, Hikosaka T, Kato J (2011). Structure-function analysis of helix 8 of human calcitonin receptor-like receptor within the adrenomedullin 1 receptor. *Peptides* 32: 144–149
- Kuwasako K, Kitamura K, Nagata S, Hikosaka T, Kato J (2010). Function of the cytoplasmic tail of human calcitonin receptor-like receptor in complex with receptor activity-modifying protein 2. *Biochem Biophys Res Commun* 392: 380–385
- Kuwasako, K., Kitamura, K., Ito, K., Uemura, T., Yanagita, Y., Kato, J., Sakata, T., & Eto, T. (2001). The seven amino acids of human RAMP2 (86) and RAMP3 (59) are critical for agonist binding to human adrenomedullin receptors. *J Biol Chem*, 276(52), 49459.
- Laan, E.B, Chupin, V., Killian, J.A. and Kruijff, B. (2004) Stability of KcsA Tetramer Depends on Membrane Lateral Pressure. *Biochemistry* 43: 4240-4250
- Lambert, D.G. (2004). Drugs and receptors. *Contin Educ Anaesth Crit Care Pain* 4 (6): 181-184.
- Langer I, Langlet C, Robberecht P. (2005) Effect of inactivating mutations on phosphorylation and internalization of the human VPAC2 receptor. *J. Mol. Endocrinol.* 34, 405–414.
- Langer I, Robberecht P. (2005) Mutations in the carboxy-terminus of the third intracellular loop of the human recombinant VPAC1 receptor impair VIPstimulated [Ca<sup>2+</sup>]<sub>i</sub> increase but not adenylate cyclase stimulation. *Cell Signal* 17, 17 – 24.
- 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. *J Headache Pain* 9:151–157
- Leff, P. (1995) The two-state model of receptor activation. *Trends Pharmacol Sci.* 16(3):89-97.
- Li, J., Zhao, H., Supowit, S. C., DiPette, D. J., & Wang, D. H. (2004). Activation of the renin-angiotensin system in alpha-calcitonin gene-related peptide/calcitonin gene knockout mice. *J Hypertens*, 22(7), 1345-1349.

- Limbird, L.E. (1996) Cell surface receptors: a short course on theory and methods. 2nd edition. Springer. Pp 1-20
- Cereghino JL, Cregg JM. (2000). Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol Rev* 24(1):45-66.
- Lodish, H., Berk, A., Zipursky, L., Matsudaira, P., Baltimore, B and Darnell, J. (2000). Protein structure and function. In: Molecular Cell Biology. Chapter 3. Fourth Edition. W.H. Freeman and Company.
- Lopez de Maturana R, Treece-Birch J, Abidi F, Findlay JB, Donnelly D (2004). Met-204 and Tyr-205 are together important for binding GLP-1 receptor agonists but not their N-terminally truncated analogues. *Protein Pept Lett* 11: 15–22.
- Lou, H., & Gagel, R. F. (1998). Alternative RNA processing--its role in regulating expression of calcitonin/calcitonin gene-related peptide. *J Endocrinol*, 156(3), 401-405.
- Lundstrom, K. (2005) Structural genomics of GPCRs. *TRENDS in Biotechnology* 23(2): 103-8
- Lundstrom, K.H. and Chiu, M.L. (2006) G-protein coupled receptor in drug discovery. Illustrated Edition. CRC Press. Pp 19-63
- Macauley-Patrick S, Fazenda ML, McNeil B and Harvey LM. (2005). Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*. 22(4):249-70
- Mallee, J.J., Salvatore, C.A., LeBourdelle, B., Oliver, K.R., Longmore, J., Koblan, K.S. *et al.* (2002) Receptor activity-modifying protein 1 determines the species selectivity of non-peptide CGRP receptor antagonists. *J. Biol Chem* 277(16): 14294-14298
- Mann RJ, Al-Sabah S, de Maturana RL, Sinfield JK, Donnelly D. (2010). Functional coupling of Cys-226 and Cys-296 in the glucagon-like peptide-1 (GLP-1) receptor indicates a disulfide bond that is close to the activation pocket. *Peptides* 31(12):2289-93.
- Marin EP, Krishna AG, Zvyaga TA, Isele J, Siebert F, Sakmar TP (2000) The amino terminus of the fourth cytoplasmic loop of rhodopsin modulates rhodopsin-transducin interaction. *J Biol Chem* 275:1930–1936
- Mathi SK, Chan Y, Li X, Wheeler MB. (1997) Scanning of the glucagon-like peptide-1 receptor localizes G protein-activating determinants primarily to the N terminus of the third intracellular loop. *Mol. Endocrinol.* 11, 424– 432
- May, L.T., Leach, K., Sexton, P.M. and Christopoulos, A. (2007) Allosteric Modulation of G Protein–Coupled Receptors. *Annu. Rev. Pharmacol. Toxicol.* 47:1-51
- McCusker EC, Bane SE, O'Malley MA, Robinson AS. (2007). *Biotechnol Prog.* 23:540–547
- McLatchie, L. M., Fraser, N. J., Main, M. J., Wise, A., Brown, J., Thompson, N., Solari, R., Lee, M. G., and Foord, S. M. (1998) *Nature.* 393, 333-339.
- Mellman I (1996). "Endocytosis and molecular sorting". *Annual Review of Cell and Developmental Biology* 12: 575–625
- Mierke, D. F., and Pellegrini, M. (1999). Parathyroid hormone and parathyroid hormone-related protein model systems for the development of an osteoporosis therapy. *Curr Pharm Des*, 5(1), 21-36.
- Miller LJ, Dong M, Harikumar KG. (2012). Ligand binding and activation of the secretin receptor, a prototypic family B G protein-coupled receptor. *Br J Pharmacol.* 166(1):18-26.

- Miller LJ, Chen Q, Lam PC, Pinon DI, Sexton PM, Abagyan R, Dong M. (2011). Refinement of glucagon-like peptide 1 docking to its intact receptor using mid-region photolabile probes and molecular modeling. *J Biol Chem.* 286(18):15895-907
- Miret JJ, Rakhilina L, Silverman L and Oehlen B. (2002). Functional expression of heteromeric calcitonin gene-related peptide and adrenomedullin receptors in yeast. *J Biol Chem.* 277(9):6881-7
- Monaghan P, Thomas BE, Woznica I, Wittelsberger A, Mierke DF, Rosenblatt M (2008). Mapping peptide hormone-receptor interactions using a disulfide-trapping approach. *Biochemistry* 47: 5889–5895.
- Moore EL, Gingell JJ, Kane SA, Hay DL, Salvatore CA (2010). Mapping the CGRP receptor ligand binding domain: Tryptophan-84 of RAMP1 is critical for agonist and antagonist binding. *Biochem Biophys Res Commun* 394: 141–145.
- Moore, C.A., Milano, S.K., and Benovic, J.L. (2007). Regulation of receptor trafficking by GRKs and arrestins. *Annu Rev Physiol.* 69:451-82.
- Muff R, Born W and Fischer JA. (2001). Adrenomedullin and related peptides: receptors and accessory proteins. *Peptides.* 22(11):1765-72.
- Muller, D.J., Wu, N. and Palczewski, K. (2008) Vertebrate membrane Proteins: Structure, Function, and Insights from Biophysical Approaches. *Pharmacological Reviews* 60: 43-78
- Murakami, M. and Kouyama, T. (2008). Crystal structure of squid rhodopsin. *Nature* 453:363–8.
- Nabhan C, Xiong Y, Xie LY, Abou-Samra A.B. (1995) The alternatively spliced type II corticotropin-releasing factor receptor, stably expressed in LLCPK-1 cells, is not well coupled to the G protein(s). *Biochem Biophys Res Commun.* 212:1015–1021
- Nag K, Sultana N. and Hirose S. (2012). Calcitonin receptor-like receptor (CLR) influences posttranslational events of receptor activity-modifying proteins (RAMPs). *Biochem Biophys Res Commun.* 24;418(4):824-9.
- Nanga RP, Brender JR, Vivekanandan S and Ramamoorthy A. (2011). Structure and membrane orientation of IAPP in its natively amidated form at physiological pH in a membrane environment. *Biochim Biophys Acta.* 1808(10):2337-42
- Nelson, D.L. and Cox, M.M. (2005). *Lehninger Principle of Biochemistry.* 4th Edition. W.H. Freeman and Co., New York. Pp
- Neumann JM, Couvineau A, Murail S, Lacapere JJ, Jamin N, Laburthe M (2008). Class-B GPCR activation: is ligand helix-capping the key? *Trends Biochem Sci* 33: 314–319.
- Neves, S.R., Ram,P.T. and Iyengar, R. (2002) G Protein Pathways. *Science* 296( 5573): 1636 – 1639
- Nichols, P.L., Brand, J., Briggs, M. et al (2010) Potent oxadiazole CGRP receptor antagonists for the potential treatment of migraine. *Bioorganic & Medicinal Chemistry Letters* 20: 1368-1372
- Nussenzveig DR, Thaw CN, Gershengorn MC. (1994) Inhibition of inositol phosphate second messenger formation by intracellular loop one of a human calcitonin receptor. Expression and mutational analysis of synthetic receptor genes. *J Biol Chem.* 269:28123–28129
- Offermanns, S. and Rosenthal, W. (2008) *Encyclopedia of Molecular Pharmacology.* 2nd Edition. Springer. Pp 917

- Olesen, J., Diener, H. C., Husstedt, I. W., Goadsby, P. J., Hall, D., Meier, U., Pollentier, S., Lesko, L. M., & BIBN 4096 Clinical Proof of Concept Study Group (2004). Calcitonin gene-related peptide receptor antagonist BIBN 4096 BS for the acute treatment of migraine. *N Engl J Med*, *350*(11), 1104-1110.
- Ogoshi M, Inoue K, Naruse K, Takei Y (2006). Evolutionary history of the calcitonin gene-related peptide family in vertebrates revealed by comparative genomic analyses. *Peptides* *27*(12): 3154-3164.
- Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., Le Trong, I., Teller, D.C., Okada, T., Stenkamp, R.E., Yamamoto, M. and Miyano, M. (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* *289*(5480): 733-4.
- Parrill AL. (2008). Crystal structures of a second G protein-coupled receptor: triumphs and implications. *ChemMedChem* *3*:1021-3.
- Park, S. H., Das, B. Casagrande, F. et al. (2012). "Structure of the chemokine receptor CXCR1 in phospholipid bilayers." *Nature* *491*(7426): 779-783.
- Park, P.S., Lodowski, D.T. and Palczewski, K. (2007) Activation of G Protein-Coupled Receptors: Beyond Two-State Models and Tertiary Conformational Changes. *Annu Rev Pharmacol Toxicol.* *48*: 107-141
- Parker, M.S., and Parker, S.L. (2010). The fourth intracellular domain of G-protein coupling receptors: helicity, basicity and similarity to opsins. *Amino Acids.* *38*(1):1-13.
- Parthier C, Reedtz-Runge S, Rudolph R, Stubbs MT. (2009) Passing the baton in class B GPCRs: peptide hormone activation via helix induction? *Trends Biochem. Sci.* *34*, 303- 310.
- Parthier, C., Kleinschmidt, M., Neumann, P., Rudolph, R., Manhart, S., Schlenzig, D., Fangha, J., Rahfeld, J., Demuth, H. and Stubbs, M.T. (2007) Crystal structure of the incretin-bound extracellular domain of a G protein-coupled receptor. *PNAS* *104*(35): 13942-13947
- Parton, R.G. and Simons, K. (2007). "The multiple faces of caveolae". *Nature Reviews Molecular Cell Biology* *8* (3): 185-94
- Patil SM, Xu S, Sheftic SR, Alexandrescu AT (2009). Dynamic alpha-helix structure of micelle-bound human amylin. *J Biol Chem* *284*(18): 11982-11991.
- Perrin MH, DiGruccio MR, Koerber SC, Rivier JE, Kunitake KS, Bain DL, Fischer WH, Vale WW. (2003). A soluble form of the first extracellular domain of mouse type 2beta corticotropin-releasing factor receptor reveals differential ligand specificity. *J Biol Chem* *278*(18):15595-600.
- Petty, H.R. (1993) Molecular biology of membranes: structure and function. Illustrated edition. Springer. Pp 26-28
- Pisegna JR, Moody TW, Wank SA. (1996). Differential signaling and immediate-early gene activation by four splice variants of the human pituitary adenylate cyclase-activating polypeptide receptor (hPACAP-R). *Ann N Y Acad Sci.* *805*:54-64
- Pioszak, A.A., Naomi, R.P., Suino-Powell, K. and Xu, H.E. (2008) Molecular Recognition of Corticotropin-releasing Factor by Its G-protein-coupled Receptor CRFR1. *Journal of Biological Chemistry* *283* (47): 32900-32912
- Pioszak, A.A. and Xu, H.E. (2008) Molecular recognition of parathyroid hormone by its G protein-coupled receptor. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 5034-5039

- Pochet, R. and Donato, R. (2000) Calcium: the molecular basis of Calcium action in biology and medicine. Illustrated Edition. Springer. Pp 548
- Poyner, D. R., Sexton, P. M., Marshall, I., Smith, D. M., Quirion, R., Born, W., Muff, R., Fischer, J. A., and Foord, S. M. (2002). *Pharmacol. Rev.* 54, 233–246
- Poyner, D. R., Andrew, D.P., Brown D., Bose C. and Hanley M.R.. (1992). *Br J Pharmacol.* 105 (2), 441-7.
- Punn A, Chen J, Delidaki M, Tang J, Liapakis G, Lehnert H, Levine MA, Grammatopoulos DK. (2012) Mapping structural determinants within third intracellular loop that direct signalling specificity of type 1 corticotropin-releasing hormone receptor. *J. Biol. Chem.* 287, 8974–8985.
- Rasmussen SG., Rasmussen SG, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Pardon E, Calinski D, Mathiesen JM, Shah ST, Lyons JA, Caffrey M, Gellman SH, Steyaert J, Skiniotis G, Weis WI, Sunahara RK, Kobilka BK. (2011). Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* 477, 549– 555
- Rasmussen, S. G. F., Choi, H.-J., Rosenbaum, D. M., Kobilka, T. S., Thian, F. S., Edwards, P.C., Burghammer, M., Ratnala, V. R. P., Sanishvili, R., Fischetti, R. F., Schertler, G. F. X., Weis, W. I., & Kobilka, B. K. (2007). Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature* 450(7168), 383-387.
- Reddi, O.S. (2000) Recombinant DNA Technology. A laboratory Manual. Allied Publishers. Pp 1-20
- Ren, H., Yu, D., Ge, B., Cook, B., Xu, Z. and Zhang, S. (2009) High-Level production, Solubilization and Purification of Synthetic Human GPCR Chemokine Receptors CCR5, CCR3, CXCR4 and CX3CR1. *PLoS ONE* 4(2): 1-15
- Roh, J., Chang, C. L., Bhalla, A., Klein, C., & Hsu, S. Y. T. (2004). Intermedin is a calcitonin/calcitonin gene-related peptide family peptide acting through the calcitonin receptor-like receptor/receptor activity-modifying protein receptor complexes. *J Biol Chem* 279(8), 7264-7274.
- Rosenbaum, D. M., Cherezov, V., Hanson, M. A., Rasmussen, S. G. F., Thian, F. S., Kobilka, T. S., Choi, H.-J., Yao, X.-J., Weis, W. I., Stevens, R. C., & Kobilka, B. K. (2007). GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science*, 318(5854), 1266-1273.
- Scheindlin, S. (2001) A brief history of pharmacology. *Modern Drug Discovery* 4(5): 87-88, 91
- Rudolf, K., Eberlein, W., Engel, W., Pieper, H., Entzeroth, M., Hallermayer, G., & Doods, H. (2005). Development of human calcitonin gene-related peptide (CGRP) receptor antagonists. 1. Potent and selective small molecule CGRP antagonists. 1-[N2-[3,5-dibromo-N-[[4-(3,4-dihydro-2(1H)-oxoquinazolin-3-yl)-1-piperidiny]carbonyl]-Dtyrosyl]-1-lysyl]-4-(4-pyridinyl)piperazine: the first CGRP antagonist for clinical trials in acute migraine. *J Med Chem*, 48(19), 5921-5931.
- Runge S, Thogersen H, Madsen K, Lau J, Rudolph R (2008). Crystal structure of the ligandbound glucagon-like peptide-1 receptor extracellular domain. *J Biol Chem* 283(17): 11340-11347.
- Runge S, Gram C, Brauner-Osborne H, Madsen K, Knudsen LB, Wulff BS (2003). Three distinct epitopes on the extracellular face of the glucagon receptor determine specificity for the glucagon amino terminus. *J Biol Chem* 278: 28005–28010.
- Salvatore, C. A., Hershey, J. C., Corcoran, H. A., Fay, J. F., Johnston, V. K., Moore, E. L., Mosser, S. D., Burgey, C. S., Paone, D. V., Shaw, A. W., Graham, S. L., Vacca, J. P., Williams, T. M.,

- Koblan, K. S., & Kane, S. A. (2008). Pharmacological characterization of MK-0974 [N-[(3R,6S)-6-(2,3-difluorophenyl)-2-oxo-1-(2,2,2-trifluoroethyl)azepan-3-yl]-4-(2-oxo-2,3-dihydro-1H-imidazo[4,5-b]pyridin-1-yl)piperidine-1-carboxamide], a potent and orally active calcitonin gene-related peptide receptor antagonist for the treatment of migraine. *J Pharmacol Exp Ther*, 324(2), 416-421.
- Sasaki, K., Dockerill, S., Adamiak, D.A., Tickle, I.J., Blundell, T. (1975). X ray analysis of glucagon and its relationship to receptor binding. *Nature* 257 (5529): 751-757
- Sheikh SP, Vilardarga JP, Baranski TJ, Lichtarge O, Iiri T, Meng EC, Nissenson RA, Bourne HR. (1999). Similar structures and shared switch mechanisms of the beta2-adrenoceptor and the parathyroid hormone receptor. Zn(II) bridges between helices III and VI block activation. *J. Biol. Chem.* 274, 17 033 – 17 041.
- Schioth, H.B. and Fredriksson, R. (2005) The GRAFS classification system of G-protein coupled receptors in comparative perspective. *General and Comparative Endocrinology* 142 (1-2): 94 – 101
- Schipani E, Langman CB, Parfitt AM, Jensen GS, Kikuchi S, Kooh SW, Cole WG, Juppner H. (1996) Constitutively activated receptors for parathyroid hormone and parathyroid hormone-related peptide in Jansen's metaphyseal chondrodysplasia. *N. Engl. J. Med.* 335, 708 – 714
- Schulte, G. and Bryja, V. (2007). The Frizzled family of unconventional G-protein-coupled receptors. *Trends Pharmacol. Sci.* 28(10):518-25
- Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F and Serrano L (2005). The FoldX web server: an online force field. *Nucleic Acids Res.*33:W382-W388.
- Seifert, R. and Weiland, T. (2005) G protein-coupled receptor as drug targets: analysis of activation and constitutive activity. Illustrated Edition. John Wiley & Sons. Pp 31
- Serrano-Vega MJ, Magnani F, Shibata Y, Tate CG. (2008). Conformational thermostabilization of the b1-adrenergic receptor in a detergent-resistant form. *Proc Natl Acad Sci USA* 105:877–82.
- Seth, S.D. and Seth, V. (2009) Textbook of Pharmacology. 3rd edition. Elsevier. Pp 45-70
- Sexton PM, Albiston A, Morfis M, Tilakaratne N. (2001). Receptor activity modifying proteins. *Cell Signal.* 13(2):73-83.
- Sexton, P. M., Findlay, D. M., & Martin, T. J. (1999). Calcitonin. *Curr Med Chem*, 6(11),1067-1093.
- Shimizu, N., Guo, J., & Gardella, T. J. (2001). Parathyroid hormone (PTH)-(1-14) and -(1-11) analogs conformationally constrained by alpha-aminoisobutyric acid mediate full agonist responses via the juxtamembrane region of the PTH-1 receptor. *J Biol Chem.* 276(52): 49003-49012.
- Simms, J., Hay, D. L., Wheatley, M., & Poyner, D. R. (2006). Characterization of the structure of RAMP1 by mutagenesis and molecular modeling. *Biophys J*, 91(2), 662-669.
- Smith, R.D. and Veenstra (2003) Proteome Characterization and Proteomics. Illustrated Edition. Academic Press. Pp 57-70
- Smith, S.O. (2010). Structure and Activation of the Visual Pigment Rhodopsin. *Annu. Rev. Biophys.* 39:309–28
- Standfuss, J., Edwards, P.C., D'Antona, A., Fransen, M., Xie, G. Oprian, D.D. and Schertler, G.F. (2011). The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature* 471,656–660

- Steiner, S., Muff, R., Gujer, R., Fischer, J. A., & Born, W. (2002). The transmembrane domain of receptor-activity-modifying protein 1 is essential for the functional expression of a calcitonin gene-related peptide receptor. *Biochemistry* 41(38), 11398-11404.
- Stoorvogel W, Strous GJ, Geuze HJ, Oorschot V, Schwartz AL (1991). "Late endosomes derive from early endosomes by maturation". *Cell* 65 (3): 417–27
- Stroop, S. D., Kuestner, R. E., Serwold, T. F., Chen, L., & Moore, E. E. (1995). Chimeric human calcitonin and glucagon receptors reveal two dissociable calcitonin interaction sites. *Biochemistry* 34(3), 1050-1057.
- Sun, C., Song, D., Davis-Taber, R. A., Barrett, L. W., Scott, V. E., Richardson, P. L., Pereda-Lopez, A., Uchic, M. E., Solomon, L. R., Lake, M. R., Walter, K. A., Hajduk, P. J., & Olejniczak, E. T. (2007). Solution structure and mutational analysis of pituitary adenylate cyclase-activating polypeptide binding to the extracellular domain of PAC1-RS. *Proc Natl Acad Sci U S A*, 104(19), 7875-7880.
- Swift S, Leger AJ, Talavera J, Zhang L, Bohm A, Kuliopulos A (2006) Role of the PAR1 receptor 8th helix in signaling: the 7-8-1 receptor activation mechanism. *J Biol Chem* 281:4109–4116
- Taylor, C.K., Smith, D.D., Hulce, M. and Abel, P.W. (2006) Pharmacological Characterization of Novel  $\alpha$ -Calcitonin Gene-Related Peptide (CGRP) Receptor Peptide Antagonists That Are Selective for Human CGRP Receptors. *Journal of Pharmacology and Experimental Therapeutics* 319: 749-757
- ter Haar E, Koth CM, Abdul-Manan N, Swenson L, Coll JT, Lippke JA, Lepre CA, Garcia-Guzman M, Moore JM. (2010) Crystal structure of the ectodomains complex of the CGRP receptor, a class-B GPCR, reveals the site of drug antagonism. *Structure* 18, 1083– 1093
- Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol*. 72(2):211-22
- Tobin, A.B. and Wheatley, M. (2004). G-protein-coupled receptor phosphorylation and palmitoylation. In: Receptor Signal Transduction Protocols. *Methods Mol Biol*. 259:275-81.
- Tolun, A. A., Dickerson. I. M., Malhatra, A. (2006). Overexpression and purification of human calcitonin gene-related peptide-receptor component protein in Escherichia coli. *Protein Expr Purif*. 52(1):167-74
- Topiol S, Sabio M. (2009). X-ray structure breakthroughs in the GPCR transmembrane region. *Biochem Pharmacol*. 78(1):11-20
- Triggle, D.J. (2000) Pharmacological receptors: a century of discovery – and more. *Pharmaceutica Acta helvetiae* 74: 79-84.
- Tseng CC, Lin L. (1997) A point mutation in the glucosedependent insulinotropic peptide receptor confers constitutive activity. *Biochem. Biophys. Res. Commun*. 232, 96–100.
- Udawela, M., Christopoulos, G., Morfis, M., Christopoulos, A., Ye, S., Tilakaratne, N., & Sexton, P. M. (2006a). A critical role for the short intracellular C-terminus in receptor activity modifying protein function. *Mol Pharmacol*, 70(5), 1750-1760.
- Udawela, M., Christopoulos, G., Tilakaratne, N., Christopoulos, A., Albiston, A., & Sexton, P. M. (2006b). Distinct receptor activity-modifying protein domains differentially modulate interaction with calcitonin receptors. *Mol Pharmacol*, 69(6), 1984-1989.

- Underwood CR, Garibay P, Knudsen LB, Hastrup S, Peters GH, Rudolph R, et al. (2010). Crystal structure of glucagon-like peptide-1 in complex with the extracellular domain of the glucagon-like peptide-1 receptor. *J Biol Chem* 285(1): 723-730.
- Uniprot. <http://www.uniprot.org/uniprot/B3BT11>
- Unson CG, Wu CR, Jiang Y, Yoo B, Cheung C, Sakmar TP, Merrifield RB. (2002). Roles of specific extracellular domains of the glucagon receptor in ligand binding and signaling. *Biochemistry* 41(39):11795-803.
- Van Durme J, Delgado J, Stricher F, Serrano L, Schymkowitz J. and Rousseau F. (2011). A graphical interface for the FoldX forcefield. *Bioinformatics*. 27(12):1711-2
- van Koppen, C.J. and Jakobs, K.H. (2004). Arrestin-independent internalization of G protein-coupled receptors. *Mol Pharmacol*. 66(3):365-7
- Venkatakrishnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF, Babu MM. (2013). Molecular signatures of G-protein-coupled receptors. *Nature* 494(7436):185-94.
- Vohra S, Taddese B, Conner AC, Poyner DR, Hay DL, Barwell J, Reeves PJ, Upton GJ, Reynolds CA. (2013). Similarity between class A and class B G-protein-coupled receptors exemplified through calcitonin gene-related peptide receptor modelling and mutagenesis studies. *J R Soc Interface*. 12;10(79):20120846
- Vohra, S., Chintapalli, S.V., Illingworth, C.J.R., Reeves, P.J., Mullineaux, P.M., Clark, H.S.X., Dean, M.K., Upton†, G.J.G. and Reynolds, C.A. (2007) Computational studies of Family A and Family B GPCRs. *Biochemical Society Transactions* 35 (4): 749-754.
- Wang C, Wu H, Katritch V, Han GW, Huang XP, Liu W, Siu FY, Roth BL, Cherezov V, Stevens RC. (2013). Structure of the human smoothed receptor bound to an antitumour agent. *Nature* 497(7449):338-43
- Warne, T., Serrano-Vega, M. J., Baker, J. G., Moukhametzianov, R., Edwards, P. C., Henderson, R., Leslie, A. G. W., Tate, C. G., & Schertler, G. F. X. (2008). Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* 454(7203), 486-491.
- Watkins HA, Rathbone DL, Barwell J, Hay DL, Poyner DR. (2012). Structure-activity relationships for alpha calcitonin gene-related peptide. *Br J Pharmacol*. doi: 10.1111/bph.12072
- Wedemeyer, C., Neuerburg, C., Pfeiffer, A., Heckeley, A., Bylski, D., Knoch, F. v., Schinke, T., Hilken, G., Gosheger, G., Knoch, M. v., Löer, F., & Saxler, G. (2007). Polyethylene particle-induced bone resorption in alpha-calcitonin gene-related peptide-deficient mice. *J Bone Miner Res*, 22(7), 1011-1019.
- Wheatley M, Wootten D, Conner MT, Simms J, Kendrick R, Logan RT, Poyner DR, Barwell J. (2011). Lifting the lid on GPCRs: the role of extracellular loops. *Br J Pharmacol*. 165(6):1688-703
- White JF., Noinaj, N., Shibata, Y., Love, J., Kloss, B., Xu, F., Gvozdenovic\_Jeremic, J., Shah, P., et al. (2012) Structure of the agonist-bound neurotensin structure. *Nature* 490, 508– 513
- Wootten D, Simms J, Miller LJ, Christopoulos A, Sexton PM. (2013). Polar transmembrane interactions drive formation of ligand-specific and signal pathway-biased family B G protein-coupled receptor conformations. *Proc Natl Acad Sci U S A*. 110(13):5211-6.
- Wu B, Chien EY, Mol CD, Fenalti G, Liu W, Katritch V et al. (2010). Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* 330: 1066–1071.

- Yasuda D, Okuno T, Yokomizo T, Hori T, Hirota N, Hashidate T, Miyano M, Shimizu T, Nakamura M (2009) Helix 8 of leukotriene B4 type-2 receptor is required for the folding to pass the quality control in the endoplasmic reticulum. *FASEB J* 23:1470–1481
- Yeagle, P.L. and Lee, A.G. (2002) Membrane protein structure. *Biochimica et Biophysica Acta (BBA) – Biomembranes* 1565 (2): 143.
- Zeder-lutz, G., Cherouati, N., Rheinart, C., Pattus, F. and Wagner, R. (2006) Dot-blot immunodetection as a versatile and high-throughput assay to evaluate recombinant GPCRs produced in the yeast *Pichia pastoris*. *Protein Expr Purif.* 50(1):118-27
- Zhang, Z., Winborn, C.S., Marquez de Prado, B. and Russo, A.F. (2007) Sensitization of Calcitonin Gene-Related Peptide Receptors by Receptor Activity-Modifying Protein-1 in the Trigeminal Ganglion. *Journal of Neuroscience* 27(10): 2693-2703.
- Zhang J, Ferguson SSG, Barak LS, Menard L and Caron MG (1996) Dynamin and b-arrestin reveal distinct mechanisms for G protein-coupled receptor internalization. *J Biol Chem* 271:18302–18305.

## **Appendices**

### **Sequencing primers for HA CLR pcDNA3.1(-)**

T7 primer

Forward: 5' TAATACGACTCACTATAGGGAAACCC 3'

TM2 primer

Forward: 5' ATCTGTTCTTCTCATTTGTTTGTA ACT 3'

TM4 primer

Forward: 5' CCTTCAGGTCGCCATGGAATCAGCAC 3'

BGH primer

Reverse: 5' ATCTGTTCTTCTCATTTGTTTGTA ACT 3'

### **Sequencing primers for 10xHis pPIC9K\_MepNet**

$\alpha$ -Factor primer

Forward: 5' ACTACTATTGCCAGCATTGC 3'

3' AOX1 primer

Reverse: 5' GCAAATGGCATTCTGACATCC 3'

### **Oligonucleotide primers used to generate CLR mutants**

W72A

Forward: 5' CAGAACCTGGGATGGAGCCCTCTGCTGGAACGATG 3'

Reverse: 5' CATCGTTCCAGCAGAGGGCTCCATCCCAGGTTCTG 3'

D90A

Forward: 5' GCAGCTCTGCCCTGCCTACTTTCAGGACTTTGATC 3'

Reverse: 5' GATCAAAGTCCTGAAAGTAGGCAGGGCAGAGCTGC 3'

Y91A

Forward: 5' GCAGCTCTGCCCTGATGCCTTTCAGGACTTTGATC 3'

Reverse: 5' GATCAAAGTCCTGAAAGGCATCAGGGCAGAGCTGC 3'

F92A

Forward: 5' CTCTGCCCTGATTACGCCAGGACTTTGATCCATC 3'

Reverse: 5' GATGGATCAAAGTCCTGGGCGTAATCAGGGCAGAGC 3'

D94A

Forward: 5' CTGCCCTGATTACTTTCAGGCCTTTGATCCATCAGAAAAAG 3'

Reverse: 5' CTTTTTCTGATGGATCAAAGGCCTGAAAGTAATCAGGGCAG 3'

F95A

Forward: 5' CCCTGATTACTTTCAGGACGCCGATCCATCAGAAAAAGTTAC 3'

Reverse: 5' GTAACTTTTTCTGATGGATCGGCGTCCTGAAAGTAATCAGGG 3'

K103A

Forward: 5' GATCCATCAGAAAAAGTTACAGCCATCTGTGACCAAGATGGAAAC 3'

Reverse: 5' GTTCCATCTTGGTCACAGATGGCTGTAACTTTTTCTGATGGATC 3'

R119A

Forward: 5' GACATCCAGCAACGCCACATGGACAAATTATACC 3'

Reverse: 5' GGTATAATTTGTCCATGTGGCGTTGCTTGCTGGATGTC 3'

Y124A

Forward: 5' CAACAGAACATGGACAAATGCCACCCAGTGTAATGTTAACAC 3'

Reverse: 5' GTGTTAACATTACTGGGTGGCATTGTCCATGTTCTGTTG 3'

F163A

Forward: 5' CTTATCTCGCTTGGCATAGCCTTTTATTTCAAGAGCC 3'

Reverse: 5' GGCTCTTGAAATAAAAGGCTATGCCAAGCGAGATAAG 3'

F163M

Forward: 5' CTGCTTATCTCGCTTGGCATAATGTTTTATTTCAAGAGCCTAAG 3'

Reverse: 5' CTTAGGCTCTTGAAATAAACATTATGCCAAGCGAGATAAGCAG 3'

A393D

Forward: 5' CTTTAATGGAGAGGTTCAAGACATTCTGAGAAGAACTGG 3'

Reverse: 5' CCAGTTTCTTCTCAGAATGTCTTGAACCTCTCCATTAAG 3'

I394A

Forward: 5' GAGAGGTTCAAGCAGCCCTGAGAAGAACTGGAATC 3'

Reverse: 3' GATTCCAGTTTCTTCTCAGGGCTGCTTGAACCTCTC 3'

L395A

Forward: 5' GGAGAGGTTCAAGCAATTGCCAGAAGAACTGGAATCAATAC 3'

Reverse: 5' GTATTGATTCCAGTTTCTTCTGGCAATTGCTTGAACCTCTCC 3'

R396A

Forward: 5' GAGAGGTTCAAGCAATTCTGGCCAGAACTGGAATCAATAC 3'

Reverse: 5' GTATTGATTCCAGTTTCTGGCCAGAATTGCTTGAACCTCTC 3'

R397A

Forward: 5' GAGGTTCAAGCAATTCTGAGAGCCAACCTGGAATCAATACAAAATC 3'

Reverse: 5' GATTTTGTATTGATTCCAGTTGGCTCTCAGAATTGCTTGAACCTC 3'

W399A

Forward: 5' GTTCAAGCAATTCTGAGAAGAAACGCCAATCAATACAAAATCCAATTTGG 3'

Reverse: 5' CCAAATTGGATTTTGTATTGATTGGCGTTTCTTCTCAGAATTGCTTGAAC 3'

N400A

Forward: 5' CAATTCTGAGAAGAACTGGGCCCAATACAAAATCCAATTTGG 3'

Reverse: 5' CCAAATTGGATTTTGTATTGGGCCAGTTTCTTCTCAGAATTG 3'

Y402A

Forward: 5' GAGAAGAACTGGAATCAAGCCAAAATCCAATTTGGAAAC 3'

Reverse: 5' GTTCCAAATTGGATTTTGGCTTGATTCCAGTTTCTTCTC 3'

K403A

Forward: 5' GAGAAGAAACTGGAATCAATACGCCATCCAATTTGGAAACAGCTTTTC 3'

Reverse: 5' GAAAAGCTGTTTCCAAATTGGATGGCGTATTGATTCCAGTTTCTTCTC 3'

I404A

Forward: 5' GAAACTGGAATCAATACAAAGCCCAATTTGGAAACAGCTTTTC 3'

Reverse: 5' GAAAAGCTGTTTCCAAATTGGGCTTTGTATTGATTCCAGTTTC 3'

F406A

Forward: 5' GGAATCAATACAAAATCCAAGCCGGAACAGCTTTTCCAACCTC 3'

Reverse: 5' GAGTTGGAAAAGCTGTTTCCGGCTTGGATTTTGTATTGATTCC 3'

### **T8-HA CLR**

RPRRRNCATM**ALPVTALLPLALLHAARPDYAS**YPYDVPDYASLGGPSLEGS**AE**LEESPEDS  
IQLGVTRNKIMTAQYECYQKIMQDPIQQAEGVYCNRTWDGWLCWWDVAAGTESMQLCPDY  
FQDFDPSEKVTKICDQDGNWFRHPASNRTWTNYTQCNVNTHEKVKTALNLFYLTIIHGGLSI  
ASLLISLGIFFYFKSLSCQRITLHKNLFFSFVCNSVVTIIHLTAVANNQALVATNPVSCKVSQFI  
HLYLMGCNYFWMLCEGIYLHTLIVVAVFAEKQHLMWYYFLGWGFPLIPACIHAIARSLYYN  
DNCWISSDTHLLYIIHGPICAALLVNLFFLLNIVRVLITKLKVTHQAESNLYMKAVRATLILVP  
LLGIEFVLPWRPEGKIAEEVYDYIMHILMHFQGLLVSTIFCFNGEVQAILRRNWNQYKIQFG  
NSFSNSEALRSASYTVSTISDGPYSHDCPSEHLNGKSIHDIENVLLKPENLYN-

### **CD33-*myc* RAMP1**

M**PLLLLLPLLWAGALAME**EQKLISEEDL**L**HGSCQEANYGALLRELCLTQFQVDMEAVGETLW  
CDWGRTIRSYRELADCTWHMAEKLGCFWPNAEVDREFFLAVHGRYFRSCPISGRAVRDPPGSI  
LYPFIVVPITVTLVLTALVVWQSKRTEGIV

Colour code

Red - Signal peptide

Blue - Epitope tag

Magenta - First residue of mature protein