An insight into the activation mechanism of the calcitonin-gene-related peptide receptor

James Barwell

Doctor of Philosophy

Aston University

February 2010

Aston University

An insight into the activation mechanism of the calcitonin-gene-related peptide receptor

James Barwell

2010

Summary

The calcitonin-gene- related peptide (CGRP) receptor is unique among G-protein coupled receptors (GPCRs) as it consists of at least three proteins: calcitonin receptor like receptor (CLR), receptor activity modifying protein (RAMP)1 and receptor component protein (RCP). An endogenous agonist for this curious receptor is αCGRP, which is a sensory nerve-derived peptide made up of 37 amino acids. αCGRP acts as a potent vasodilator having pronounced effects on arterioles and capillaries. Understanding the pharmacodynamics of the CGRP receptor would not only be of academic interest but may have pharmaceutical benefit as the receptor has been associated with the onset of migraines and implicated in Raynaud's syndrome.

However, the architecture of the CGRP receptor remains elusive and consequently identifying the orthosteric binding site of αCGRP remains a challenge. The primary aim of this thesis was to identify functionally important residues in the extracellular face of the CGRP receptor. Three areas of interest were selected including the extreme N-terminus of the CLR, extracellular loop 1 (ECL1) of the CLR and its associated transmembrane (TM) regions, and finally extracellular loop 3 (ECL3) of the CLR and its juxtamembrane regions. A site-directed mutagenesis (SDM) strategy was used to investigate these regions, primarily substituting the innate residues of CLR with alanine and assessing the mutation on multiple criteria including a functional cAMP assay, cell-surface expression, total expression, agonist-mediated internalisation and αCGRP binding. The results are interpreted and discussed taking into consideration contemporary concepts surrounding Secretin-like GPCRs. Moreover, the thesis also contains details of RAMP purification and discusses the advantages and pitfalls of this approach.

Overall the thesis provides novel data that furthers insight into the complex phenomenon of CGRP receptor activation. Site-directed mutants have been identified that affect αCGRP binding, receptor signal transduction, the CLR/RAMP1 interface and the integrity of the protein complex structure.

Keywords: calcitonin receptor like receptor, extracellular loop, G-protein coupled receptors, receptor activity modifying protein, site-directed mutagenesis.

I would like to dedicate this to my loving parents
Maureen Barwell and Fred Barwell

Acknowledgements

I would like to thank my supervisor Dr D. Poyner for all his time and guidance throughout the course of this degree. I would also like to thank Dr R. Bill and her research group. Specifically I would like to thank Dr M. Jamshad who demonstrated many laboratory techniques to me with regards to working with *Pichia pastoris* and together we continued work on human RAMP family expression and purification.

I would also like to thank Dr J. Simms for many helpful discussions, it was really appreciated. Furthermore, I would like to thank Dr A. Conner who shared his protocols, data and mutant constructs to the benefit of this project. I would also like to thank Dr D. Donnelly and Dr P. Miller who conducted binding experiments on the extreme CLR N-terminal mutants.

I would also like to thank the British Heart Foundation for funding and supporting this project.

Contents **Page**

Abbreviations

AM	Adrenomedullin
AMBER	Assisted model building with energy refinement
AUC	Analytical ultracentrifugation
β_2AR	β2 adrenoceptor
β-OG	n-Octyl-β-D-glucopyranoside
BIBN4096BS	N-[2-[[5-amino-l-[[4-(4-pyridinyl)-l-piperazinyl]carbonyl]pentyl]amino]-1-[(3,5-dibromo-4-hydroxyphenyl)methyl]-2-oxoethyl]-4-(1,4-dihydro-2-oxo-3(2H)-quinazolinyl
BiFC	Bimolecular fluorescence complementation
BRET	Bioluminescence resonance energy transfer
BMGY	Buffered complex glycerol media
BMMY	Buffered complex methanol media
cAMP	Cyclic adenosine 3', 5' monophosphate
CaSR	Calcium sensing receptor
cDNA	Complimentary deoxyribonucleic acid
CGRP	Calcitonin gene-related peptide
CHARMM	Chemistry at Harvard macromolecular mechanics
CLR	Calcitonin receptor like receptor
COSY	Correlation spectroscopy
cpm	counts per minute
CRF	Corticotrophin-releasing factor
CRFR	Corticotropin-releasing hormone receptor
CTR	Calcitonin receptor
dd	Double distilled
DDM	Dodecyl-β-D-maltopyranoside
DEER	Double electron-electron resonance
DFIRE	Distance-scaled, finite, ideal-gas reference state
dNTP	Deoxyribonucleotide triphosphate
DOPE	Discrete optimized protein energy
EC50	Half maximal effective concentration of agonist
ECD	Extracellular domain
ECL	Extracellular loop

ELISA	Enzyme-linked immunosorbent assay
Emax	Maximal response
EMeP	European membrane protein consortium
EPR	Electron paramagnetic resonance
FRET	Fluorescence resonance energy transfer
g	Gravity
GABA	γ aminobutyric acid
GαCT	C-terminal peptide derived from the Gα subunit of transducin
GB/SA	Generalized Born/surface area
GCR1	G-protein coupled receptor 1
GIPR	Gastric inhibitory polypeptide receptor
GHRHR	Growth hormone releasing hormone receptor
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
GROMACS	Groningen machine for chemical simulations
Gs	G-protein stimulatory
GTP	Guanosine triphosphate
HA	Haemagglutinin
hr	hour/s
IC50	Half maximal inhibition concentration
ICL	Intracellular loop
L	Litre
LB	Luria-Bertani
L-BFGS	Limited-memory Broyden-Fletcher-Goldfarb-Shanno
LDTRs	Ligand-directed trafficking of receptor stimulus
LIP	Loops In Protein
M	Molar concentration
mg	Milligram
min	Minute/s
ml	Millilitre

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]
mM	Millimolar
MTS	Methanethiosulfonate
Ni-NTA	Nickel nitrilotriacetic acid
nM	Nanomolar
NMR	Nuclear magnetic resonance
OD	Optical density
OPLS	Optimized potentials for liquid simulations
PACAP	Pituitary adenylyl cyclase-activating protein
PAC1R	Pituitary adenylyl cyclase-activating protein receptor
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	polymerase chain reaction
PDB	Protein data bank
PDZ	Post synaptic density protein, Drosophila disc large tumor suppressor, and zonula occludens-1 protein
pEC50	Negative logarithm of EC50
pH	Negative logarithm of the hydrogen ion concentration
pIC50	Negative logarithm of IC50
PLOP	Protein local optimization program
PTH	Parathyroid hormone
PTHR	Parathyroid hormone receptor
RAPDF	Residue-specific all atom conditional probability discriminatory function
RAMP	Receptor activity modifying protein
RCP	Receptor component protein
RCSB	Research collaboratory for structural bioinformatics
rpm	revolutions per minute
RMSD	Root mean square deviation
SB-273779	N-methyl-*N*-(2-methylphenyl)-3-nitro-4-(2-thiazolylsulfinyl)-nitrobenzanilide
SCAM	Substituted cysteine accessibility method
SCR	Short consensus repeat

SCTR	Secretin receptor
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulphate
SEM	Standard error of the mean
ssNMR	Solid state nuclear magnetic resonance
TEMED	N,N,N',N'-Tetramethylethylenediamine
TEV	Tobacco etch virus
TM	Transmembrane domain
Tris	Tris(hydroxymethyl)aminoethane
TROSY	Transverse relaxation optimised spectroscopy
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 and sorbital

Table 1. Standard amino acid abbreviations. Standard amino acid abbreviations used interchangeably throughout text.

Amino acid	3-letter	1-letter	Amino acid	3-letter	1-letter
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

Equation contents **Page**

Figure contents **Page**

Table contents **Page**

Chapter 1: General Introduction

1.1 Pharmacology and the role of membrane proteins

1.1.1 The 'receptive substance'

In 1905 the Cambridge physiologist J.N. Langley explored the biological role of nicotine and curare on skeletal muscle and proposed the activity of such compounds was mediated *via* a 'receptive substance' (discussed by Rang, 2006). The 'receptor concept' as it came to be known suggested that the plasma membrane was not merely a simple barrier but actually facilitated communication between the extracellular and intracellular environments of cells. This biomembrane barrier defines the boundaries between the cytosol and the extracellular fluid, the existence of which is universally recognised as a fundamental precondition for life. Further research has identified that plasma cell membranes are made from proteins and amphipathic lipids that together spontaneously form bilayers in aqueous solution (Tanford, 1978).

Cuthbert (2006) described how the emerging scientific discipline of 'Pharmacology' took on the challenge of finding an explanation of the physiological processes underpinning both the beneficial and harmful effects of endogenous and medicinal compounds. This challenging mission can be dated from 1847 with the establishment of the first pharmacology department founded by Rudolph Bucheim in the University of Dorpat. At that time, the knowledge of organic chemical structures was rudimentary, since that date pharmacology has grown to become an extensive multidisciplinary subject drawing upon its roots in chemistry, biochemistry, pharmacy, physiology and bioinformatics (see Hill, 2006). Among many significant milestones, perhaps one of the most fruitful was the quantitative validation of the 'receptor theory'. Hill (1909) and Langmuir (1918) can be credited with this triumph. These pioneers applied the Law of Mass Action (which states that the rate of a chemical reaction is proportional to the product of the concentrations of reactants) to drug-receptor interactions. The Hill-Langmuir equation elegantly summarises the crucial relationship between receptor occupancy and drug concentration- a cornerstone of modern pharmacological thought (see Figure 1).

$$[AR] = \frac{[R_T] \times [A]}{[A] + K_a}$$

Figure 1. The Hill-Langmuir binding isotherm. [AR] represents the proportion of occupied receptors. [A] represents the concentration of ligand. K_a represents the dissociation equilibrium constant that is derived from the association rate constant and dissociation rate constant. $[R_T]$ represents the total receptor density. The Hill-Langmuir binding isotherm was taken from Leach *et al.,* (2010).

Receptor theory has steadily evolved from its inception in the early years of the twentieth century. It became clear that some drugs acted as agonists whereas others had antagonistic effects. Adopting a statistical approach, Gaddum (1937) demonstrated that ligand binding was not only reversible in certain cases, but also competitive in the sense that two drugs can actively 'compete' for the same binding site (see Hill, 2006). The now historic 'Gaddum equation' utilised this concept making it possible to predict the proportion of binding sites occupied by a certain ligand when two ligands competing for a common binding site were in equilibrium. Unfortunately, radioligand-binding experiments only began in 1965 (see Paton and Rang, 1965) and as this provided the most empirical measure for receptor occupancy, Gaddum's 1937 equation had no immediate experimental use (see Colquhoun, 2006).

Later, in an innovative leap Schild (1947) provided the first reliable measure for antagonist affinity, based solely on physiological response measurements. The relationship between receptor occupancy and receptor activation was ambiguous at this time and Schild decided to bypass this thorny issue by assuming that the agonist response remained constant irrespective of whether other receptors were occupied by an antagonist (see Colquhoun, 2006). This approach emphasised the importance of the 'dose-ratio' (i.e. the increase in agonist concentration required to overcome antagonist competition) and this also had the advantage of being elegantly presented as rightward parallel shifts of equilibrium log (agonist concentration)–response curves. Schild Plots (graphs with log (antagonist) on the X-axis and log (dose ratio -1) on the Y-axis) can be subsequently generated to produce pA_2 values. Numerically, pA_2 can be defined as

"the negative logarithm to base 10 of the molar concentration of antagonist that makes it necessary to double concentrations of agonist needed to elicit the original submaximal response obtained in the absence of an antagonist." (Schild, 1947).

Consequently, under the right experimental conditions pA_2 values can be used to estimate the antagonist affinity constant (K_b), giving an indication of the antagonist affinity (see Neubig *et al.*, 2003).

Of course, affinity alone is an insufficient explanation of receptor activity- a drug's ability to activate a receptor (or its *efficacy*) is another essential component in understanding this pharmacological process. R.P. Stephenson (1956) was the first pharmacologist to operationally distinguish between receptor occupancy and activation (see Rang, 2006). At present efficacy is measured by functional assays to determine parameters such as maximal response (Emax) and half maximal effective concentration (EC50) (see Strange, 2007).

However, the concept of efficacy still perplexes the modern-day pharmacologist. Hypothetical activation models, such as the two-state model and ternary complex model, are based around the concept of a 'conformational equilibrium'. This proposes that receptors exist in either an 'active' or 'inactive' state and the conformational equilibrium can shift depending on the presence of a functional ligand. In other words, an agonist has the capacity to change the balance of the equilibrium by promoting receptor activation. The conformational explanation has the theoretical benefit of encompassing both ion-channel and guanine nucleotide –binding protein coupled receptor (GPCR) activity, as well as offering an explanation for modern notions such as constitutive activation. Unfortunately, 'equilibrium – conformational' models struggle to explain other complicated phenomenon such as collateral efficacy where a ligand causes the receptor to express some, but not all, of its repertoire of activities (see Kenakin, 2007). Consequently, receptors may not be simple on/off switches.

Modern day pharmacologists aim to discover the underpinning mechanisms responsible for receptor activation and inactivation by studying the molecular events responsible for such states. Most receptors are a macromolecule or an assembly of macromolecules embedded within the membrane. There are at least two basic types of integral membrane proteins: α-helical bundle proteins and β-barrel proteins (see

Cowan and Rosenbusch, 1994). Membrane proteins are essential for energy conversion, transport, signal recognition, and transduction (see Kandt *et al.*, 2007) and make up approximately 30% of the total coding sequences in the human genome (Wallin and Heijne, 1998 and Liu *et al.*, 2002). Yet, despite their abundance and physiological importance it has been particularly difficult to obtain high- resolution structural data for this broad class of proteins. This is evident in an examination of the research collaboratory for structural bioinformatics (RCSB) Protein Data Bank (PDB) where a little over 180 unique structures have been deposited (White, 2009). Determining the structure of such proteins is essential for understanding their functions, interactions and architectures and this often necessitates both two- and three-dimensional modelling.

1.1.2 Common strategies used to determine protein structure

Classical biophysical techniques have thrived in successfully elucidating the three-dimensional structures of soluble proteins. For soluble proteins, X-ray crystallography is the leading method for discerning the atomic resolution (see Pusey *et al.* 2005). Once a protein has been successfully cloned, over-expressed and purified to homogeneity, a three-dimensional crystal is grown. A fine-grained resolution between 2.5-3Å is necessary to interpret the structural arrangements of the amino acid side chains (Fujiyoshi, 1998).

In spite of advances in protein purification (from bacterial, yeast and mammalian cell expression systems), X- ray crystallography remains both intrinsically difficult and time consuming when applied to membrane proteins (see Kobilka, 2006). One methodological problem is the need to introduce a detergent to isolate the protein from its lipid surroundings but the detergent interacts with the membrane protein and that can inhibit crystal formation (Yeagle and Albert, 2006). Moreover, identifying a detergent that is compatible with the crystallisation process and does not denature the purified protein can be problematic. For example, GPCRs have a tendency to be more stable in non-ionic detergents with relatively long alkyl chains (Kobilka, 2006). Yet, such detergents can produce relatively large micelles that interfere with crystal formation (Ostermeier and Michel, 1997). Another complexity in crystal nucleation and growth, especially in eukaryotic proteins, is post-translational modifications that can lead to aggregation and misfolding (McPhearson, 1998). Other technical concerns associated with protein crystallography are large flexible regions of a protein (see

Cherezov *et al.,* 2007 and Rasmussen *et al.,* 2007) and the thermostability of the protein (see Serrano-Vega *et al.,* 2008). Despite the limitations surrounding this procedure, pivotal findings have been yielded by this technique, which have changed the face of pharmacology. For example, the elucidation of the crystal structure of 'ground state' bovine rhodopsin (Palczewski *et al.,* 2000) influenced understanding of the entire superfamily of GPCRs.

Nuclear magnetic resonance (NMR) spectroscopy has also been applied to the chaotic 'Wild West' of structural membrane biology (see Torres *et al.*, 2003). Traditionally, solution NMR methods have been limited to small proteins (≤35 kDa) (see Wand *et al.,* 1998). Larger macromolecules 'tumble' too slowly in solution, which results in resonance-band broadening and intensity loss. In attempt to compensate for this, isotope labelling and multidimensional experiments were introduced to allow investigators to differentiate between spectroscopy peaks. Moreover, the recent application of transverse relaxation optimised spectroscopy (TROSY) and deuteration to solution NMR, prolongs the magnetisation of the protein to extend the window of opportunity in which a signal can be detected. Yet, only a handful of membrane structures have been identified *via* this technique e.g. the dimeric transmembrane domain of glycophorin A (MacKenzie *et al.,* 1997).

This small protein region was solubilised in aqueous dodecylphosphocholine detergent micelles (MacKenzie *et al.,* 1997). Sadly, this approach cannot be universally applied because protein-detergent micelles have a tendency to become too large to be analysed especially for α- helical conformations (Torres *et al.,* 2003). Despite obvious drawbacks, solution NMR has successfully elucidated high-resolution structures of extracellular and intracellular domains of membrane proteins e.g. the N-terminal extracellular domain of the corticotrophin-releasing factor receptor (CRFR; Grace *et al.,* 2004). Moreover, two-dimensional NMR techniques such as correlation spectroscopy (COSY) and nuclear overhauser effect spectroscopy (NOSEY) can be used to determine distance restraints within proteins that can aid molecular dynamic investigations.

Solid state NMR (ssNMR) is a relatively new approach for studying membrane proteins, and has the specific advantage that proteins over 100kDa can be studied (Tycko, 2001). However, with ssNMR a new set of limitations has to be overcome.

The presence of anisotropic chemical shifts and coupling within an ssNMR sample masks the observable NMR spectra of the experimental macromolecule. Andrew *et al.,* (1958) discovered that the anisotropic effects could be suppressed when the sample is spun rapidly at a specific angle (54.74°). The method is now referred to as 'magic angle spinning'. Unfortunately, designing a rotor capable of maintaining the required speed is both expensive and technically demanding. Furthermore, electro-magnetic pulse sequences are often needed to obtain a better indication of the nuclear dynamics. In addition, being able to correlate nuclear rotational position with pulse frequency is another technical problem. Yet, two-dimensional ssNMR has successfully provided detail information on conformational changes of ligands upon receptor docking. For example, Luca *et al.,* (2003) used ssNMR to provide a detailed account of neurotensin binding.

It is clear that biophysical techniques have a key role in protein structure elucidation. Although some techniques remain in their methodological infancy, in time many of the currently formidable conceptual and technical difficulties will be overcome. The European membrane protein consortium (EMeP; www.e-mep.org.) shared this optimism, as their scientific objective between 2004 - 2009 was to deliver 50 purified proteins suitable for crystallisation trials. The EMeP consortium successfully published 11 structures but given that 18 leading European laboratories were involved this highlights the complexity of membrane protein elucidation.

Understandably, many pharmacologists have adopted other empirical strategies to comprehend receptor functioning. Systematic identification of natural ligands and production of synthetic drugs has provided a 'chemical tool-kit' (see Hill, 2006) to explore receptor activity, along with establishing a 'working' receptor taxonomy. Modern high-throughput screening protocols frequently take advantage of focused ligand libraries directed against receptor targets to increase the probability of discovering lead compounds (see Liu, Li and Hu, 2004). This 'brute-force' procedure has discovered a plethora of ligands, which are often functional mimetics. However, identifying novel and selective compounds remain both laborious and expensive and can be unrewarding. Consequently, understanding the mechanisms that govern receptor functioning may not only be cost-effective but may be the most pragmatic strategy in discovering novel drug leads (Bundries *et al.* 2001).

Since the advent of cloning many research groups have taken greater note of the successes in molecular biology, particularly the mapping of the human genome and the ever-increasing sophistication in biochemical reagents. Consequently, manipulating complimentary deoxyribonucleic (cDNA) to synthesise engineered proteins has become a viable option for most laboratories. Site-directed mutagenesis (SDM) and chimeric proteins have been used extensively to probe ligand structure-activity relationships in membrane receptors. SDM is often hypothesis-driven based on several criteria such as amino acid conservation, characteristics of native amino acids (e.g. hydrophobic patches, charged residues, prolines in transmembrane domains, etc), molecular models or naturally occurring polymorphisms. However, SDM can also be non-hypothesis driven whereby native residues are consecutively mutated.

Alanine scans have traditionally been implemented in the study of membrane proteins particularly in alpha helical bundles as it is assumed the small, non-reactive methyl side-chain will not distort the native secondary structure, yet shed light on the function of the native residue. Consequently, alanine scanning is a powerful tool to probe ligand-protein and protein-protein interfaces. Hulme *et al.,* (1999) outlined four stereotypical roles of amino acids within a receptor that can be inferred with SDM, using the muscarinic receptor as a model GPCR system (see Figure 2). The first category of residues is referred to as 'filler'; a mutation would be tolerated because the side chain is generally unimportant. The second category is a 'stabiliser' residue; mutating this site reduces the overall stability of the receptor, albeit whether in the active or inactive state. A typical characteristic of such mutations is a decrease in cell surface expression brought on by impaired folding and cell surface trafficking. Yet, the mutant would not be expected to alter transduction from inactive to active conformations. Hulme *et al.,* (1999) extended the 'stabiliser' category to also contain 'ligand anchor' residues. This subset was defined as residues that were capable of making ligand interactions, whilst the receptor was in its inactive conformation. Consequently, mutating a 'ligand anchor' residue would decrease ligand affinity but not necessarily affect receptor efficacy. The third category of residues was referred to as 'constraint' residues. A 'constraint' residue is expected to participate in intra-molecular interactions that hold the receptor in its inactive conformation. Consequently, mutating this residue could lead to an increase in receptor basal activity or/and agonist affinity and efficacy. The final category was referred to as 'activator'

residues as such residues are expected to participate in receptor signal transduction. The category could be further differentiated into 'ligand-transducer' or 'guanine nucleotide-binding protein (G-protein) transducer' residues.

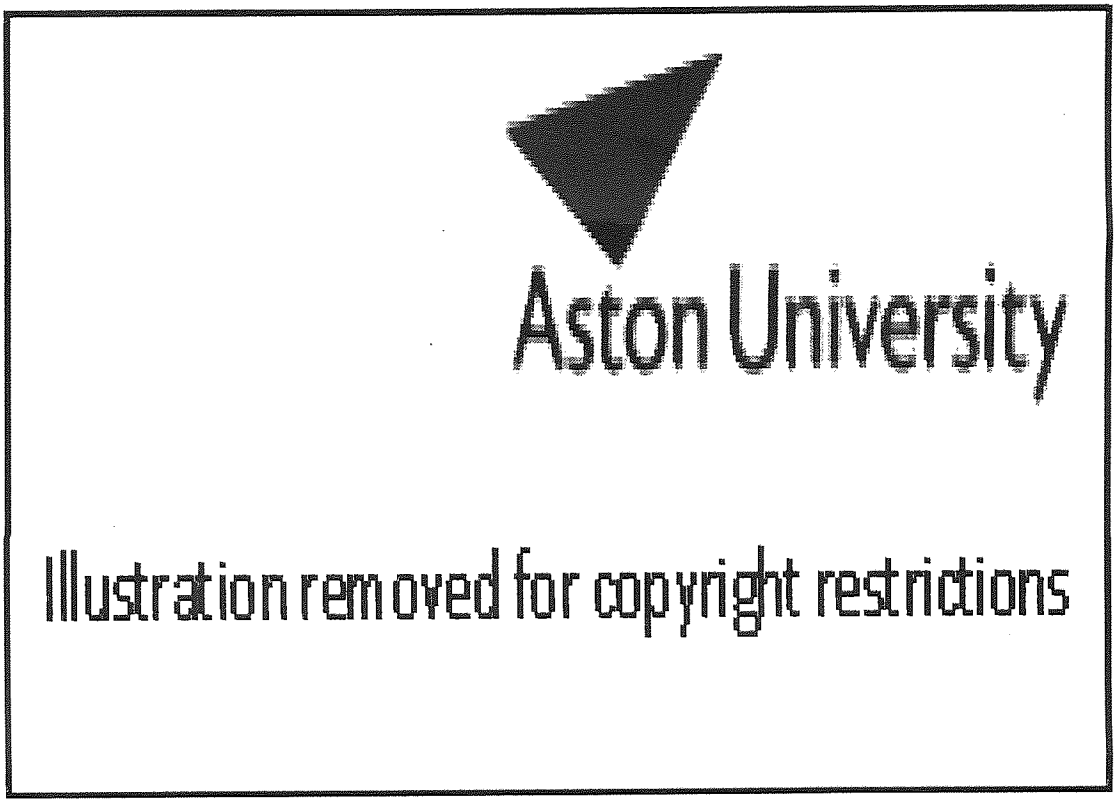


Figure 2. Taxonomy of residues and the effect of the mutation. Taken from Hulme *et al.*, (1999).

However, even with a working SDM taxonomy interpreting alanine mutants can be problematic. The categories outlined by Hulme *et al.*, (1999) are not necessarily mutually exclusive as certain residue may have dual functions, referred to by Hulme as 'double agents'. Furthermore, Clackson and Wells's (1995) pioneering work on the human growth hormone receptor suggested only a small fraction of interface residues account for the majority of the binding energy between ligand/protein and protein/protein interfaces. Therefore, the relative effect of an alanine substitution is dependent on its impact on the overall total binding energy of the interface. If the alanine mutant substantially changes the total binding energy it is referred to as a 'hot

spot'. However, the alanine substitution may be in the binding interface but not make a significant difference on the total binding energy. This could give rise to false-negative results if not interpreted correctly. Holst *et al.,* (1998) suggested that steric hindrance mutagenesis could be used to minimise this problem. The addition of larger side chains in the presumed binding region could directly impair ligand interaction and disrupt neighboring amino acid side chain conformations. Other mutant scanning approaches include lysine scanning. This has been used to investigate the nicotinic acetylcholine receptor to distinguish between core hydrophobic from surface hydrophilic orientations of native side chain residues (Sine *et al,* 2002).

The substituted cysteine accessibility method (SCAM) is a technique where native residues get consecutively substituted to cysteines and analysing the rate at which sulfhydryl specific reagents such as biocytin maleimide or derivatives of methanethiosulfonate (MTS) interact, can determine whether the cysteine side chain is exposed to an aqueous environment. Consequently, inferences on the physico-chemical environment of the native residue can be made, which can lead to the identification of residues involved in binding cavities and water channels. For example, Javitch *et al.,* (2003) made 10 consecutive cysteine mutants within extracellular loop (ECL) 2 of the dopamine 2 receptor, which led to the conclusion that the loop was likely to be folded into the ligand-binding crevice of the receptor.

Cysteine mutations can also be used to discover distance restraints within and between proteins. Engineering disulphide bonds within a protein, before and after activation can give detailed information on the conformational changes that are essential for signal transduction in receptors. For example, disulphide bonds were successfully engineered in the parathyroid hormone receptor (PTHR) 1 between the top of transmembrane domain (TM) 2 and TM7 suggesting that these regions are in close spatial proximity (Thomas *et al.,* 2008). Interestingly, in the presence of the reducing agent iodine, the basal receptor was able to accommodate a disulphide bond between F238C/F447C but in the presence of both iodine and the endogenous peptide agonist (parathyroid hormone [PTH]) the disulphide bond was unable to form. The authors suggest that this observation provides novel insight into the initial stages of agonist induced signal transduction (Thomas *et al.,* 2008).

Cysteine mutants have also been extensively used to incorporate biophysical probes. Site-directed spin labelling is an approach that takes advantage of the capability of electron paramagnetic resonance (EPR) spectroscopy, which can detect structural changes within a millisecond time frame. Hubbell and co-workers introduced pairs of sulfhydryl-reactive spin labels on the cytoplasmic face of the GPCR rhodopsin. The authors concluded that there was a significant reduction in side chain packing at the cytoplasmic surface during activation (Hubbell *et al.,* 2003). This 'flowering' effect upon activation was previously hypothesised when a spin label was introduced at intracellular loop (ICL) 2 and at different locations in TM6 (Farrens *et al.,* 1996). This made it possible to measure the relative changes in distances between TM3 and TM6, respectively. This elegant study concluded that the foot of TM3 remained relatively static upon photoactivation while TM6 under went a rigid body movement, shifting in a counter- clockwise direction (as viewed from the extracellular side) to end up 8Å away from ICL2. Recently, Altenbach *et al.,* (2008) has used double electron-electron resonance (DEER) spectroscopy to determine conformational rearrangements in rhodopsin. DEER involves insertion of pairs nitroxide spin labels and is regarded as advantageous over other spin-directed spin labelling approaches because it can make long distance measurements (20-60Å) and measurements are made at high resolution (within 1Å). Altenbach *et al.,* (2008) inserted 16 pairs of nitroxide spin labels in the cytoplasmic outer surfaces of rhodopsin's helices and found a 5Å outward movement of TM6, and smaller movements in TM1 and TM7.

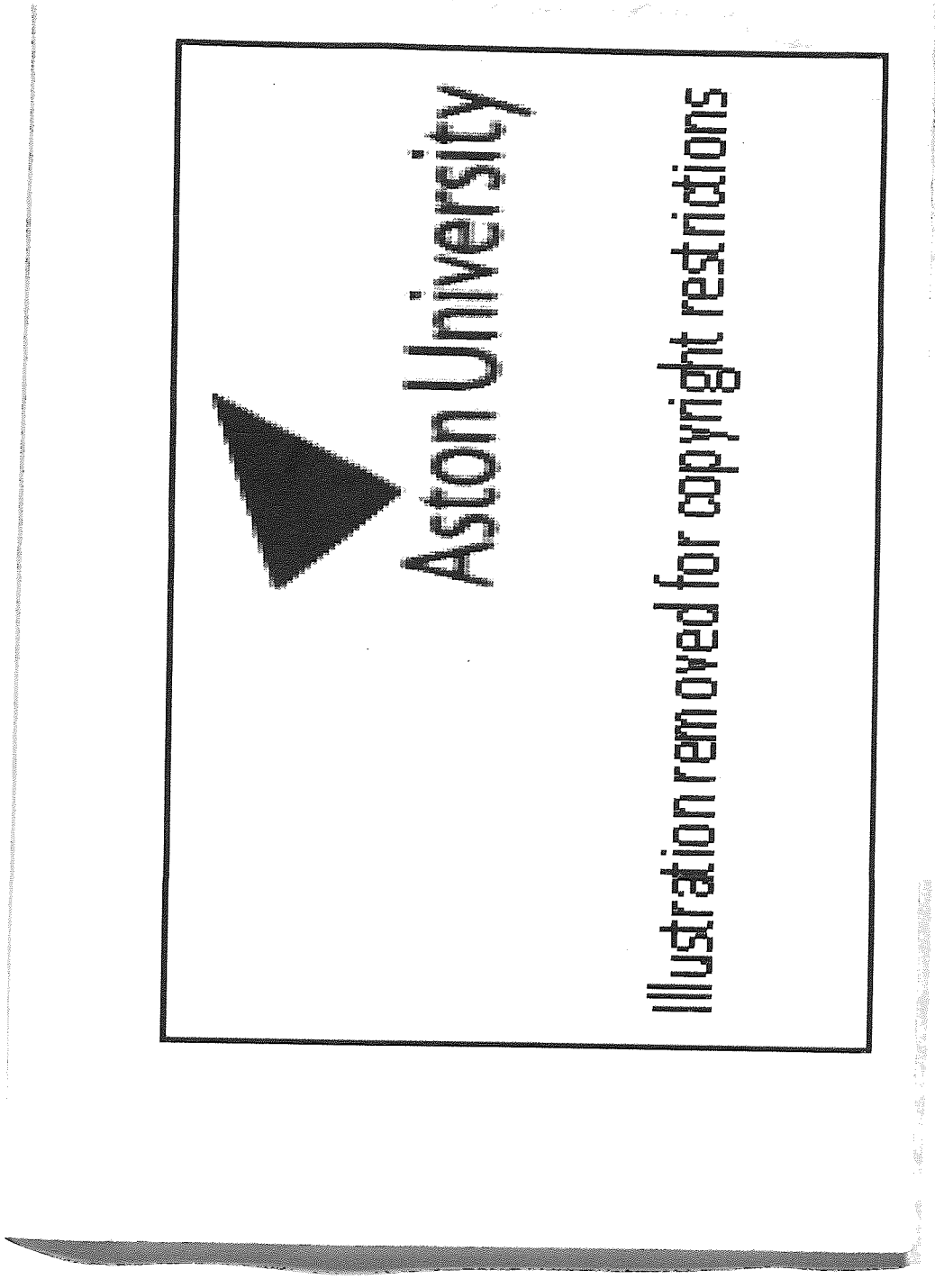


Figure 3. TM6 movement upon rhodopsin activation. Taken from Altenbach *et al.*, (2008) showing the difference in projection contours of the spin locations on the cytoplasmic face of bovine rhodopsin between the inactive and active state.

Engineering metal ion binding sites within proteins is another strategy to probe the tertiary structure of a protein along with its conformational changes required for function. Sheikh *et al.*, (1996, 1999) constructed a Zn^{2+} ion bridge between TM3 and TM6 in retinal rhodopsin and in the β-adrenergic receptor by generating bis-histidine metal ion-binding sites. It was concluded that the metal ion bridge constrained TM6 mobilisation and as a result prevented the activation of the intracellular G-protein (i.e. in these cases transducin and G-protein stimulatory [Gs]). Moreover, using an

evolutionary trace method, Sheikh *et al.*, (1999) were able to engineer a Zn^{2+} ion bridge in the cognate position in the Secretin-like PTHR1. Interestingly, the metal-ion bridge in the parathyroid receptor also impaired the ability of the receptor to induce an intracellular cascade. In spite of the obvious lack of sequence homology between the two families the authors postulated that a shared activation mechanism involving TM6 locomotion is present in both Rhodopsin-like and Secretin-like GPCRs.

Although only a select number of experimental strategies have only been briefly highlighted here, it is evident that there is currently a plethora of experimental techniques that can be successfully employed to investigate membrane proteins. Furthermore, the rapid advance in computer technology has allowed the realm of bioinformatics to have an immediate but sustaining impact on the field of pharmacology.

1.2 GPCRs

1.2.1 GPCR taxonomy

GPCRs or seven TM segment receptors form a large and functionally diverse 'super-family' of cell surface proteins (see Karchin, 2002). The receptors comprise seven stretches of about 20–35 consecutive amino acid residues that show a relatively high degree of hydrophobicity and form α-helices that span the plasma membrane in an anti-clockwise manner when viewing from the extracellular surface (see Schwartz *et al.*, 2006). Alternating ICLs and ECLs connect the α-helical segments. The 'seven TM receptor' label has been considered more apt as the receptors can bind to other proteins, such as ion channels, arrestins and tyrosine kinases (see Hill, 2006). In contrast, the GPCR label was based on the ability of this type of receptor to mediate intracellular pathways *via* 'large' heterotrimeric G-proteins. For consistency, the term 'GPCR' will be used herein.

In 1994, Martin Rodbell and Alfred Gilman were awarded the Nobel Prize for Physiology/Medicine for their discovery and research into G-proteins. Rodbell and his team studied the β-adrenergic receptor and discovered that both adrenaline and guanosine triphosphate (GTP) were needed to stimulate the adenylate/cyclic adenosine 3', 5' monophosphate (cAMP) system. This work led onto the elucidation

of the GTP 'switch' located on the α-subunit of the G-protein, which is mediated by receptor activation and initiates the intracellular cascade of messengers. Gilman and his collaborators validated Rodbell's work after they identified this elusive 'transducer' molecule as a G-protein by purifying the protein from mutated lymphoma cells.

The GPCR super-family comprises over 800 members, where approximately 350 are non-sensory GPCRs (Jacoby *et al.*, 2006). These receptors have a crucial role in many physiological systems based on their ability to detect a wide array of endogenous ligands including hormones, neurotransmitters, odorants, pheromones, tastants, protons and even photons (Jacoby *et al.*, 2006). Malfunctions in GPCR functioning causes a wide range of pathologies including cognitive disorders, cardiovascular diseases and obesity (see Horn *et al.* 2000). Consequently, they are the most successful therapeutic targets for drug development; reports suggest 25-50% of all marketed prescription drugs act on GPCRs (Topiol and Sabio, 2009).

In an attempt to develop a 'working' GPCR taxonomy, the receptors have been classified on various criteria, including structural analysis of motifs and globular domains, phylogenetic analysis, analysis of endogenous/exogenous ligand(s) and analysis of transduction pathway. Despite these efforts, no consensus on the most useful taxonomy has yet been reached (see Schioth and Fredricksson, 2005). However, the most frequently referenced taxonomy is probably the A-F system. Here both vertebrates and invertebrate GPCRs are grouped into six families based on >20% amino acid homology within the transmembrane helix domains (Kolakowski, 1994). Metazoan GPCRs are represented in the first three families (called A, B, C) of the A-F system. However, the taxonomy omits to include newly identified additional mammalian GPCR families such as adhesion receptors, frizzled/smoothened receptors, the vomeronasal 1 receptor and the taste 2 receptor.

The shortcomings of the A-F system have led to the increasing popularity of the GRAFS classification system (Fredriksson *et al.*, 2003). The GRAFS taxonomy is based on strict phylogenetic criteria, which divides human GPCRs into five distinct families: Glutamate (G), Rhodopsin (R), Adhesion (A), Frizzled/Taste2 (F), and Secretin (S). An important characteristic of the GRAFS taxonomy is the distinction

between adhesion and secretin receptors, a decision based around major differences within the extracellular N-terminal region.

The Rhodopsin-like family comprises of more than 90% of the GPCR superfamily. The GRAFS taxonomy divides this family into four predominant groups: α, β, γ, and δ. In general, the α-group contains the amine binding GPCRs and prostaglandin receptors. The β-group includes small molecule/peptide GPCRs. The γ-group contains chemokine receptors, neuropeptide receptors such as somatostatins, galanin, and opioids. Finally, the δ-group includes the large group of olfactory receptors, purine receptors and glycoprotein receptors. Defining characteristics of this extensive family include the conserved residues GN in TM1, (N/S)LXXXD (where X can be any amino acid) in TM2 , the DRY motif or D(E)-R-Y(F) located between TM3 and intra cellular loop 2, CWXP motif in TM6, the NSXXNPXXY motif in TM7 and finally HX located in the cytoplasmic helix. Other defining features include a short N-terminus domain and palmitoylated cysteine located in the carboxyl-tail (see Figure 4).

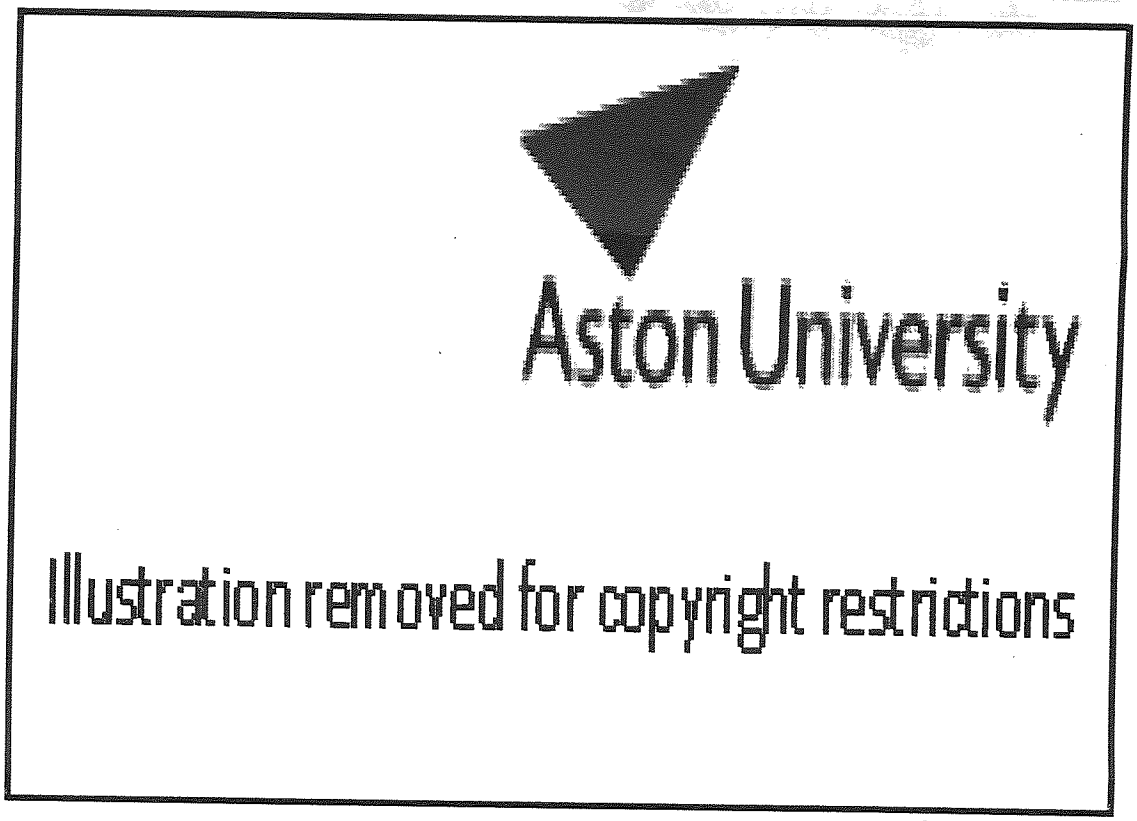


Figure 4. Two dimensional schematic of bovine rhodopsin the prototypical Rhodopsin-like GPCR. Taken from Palczewski *et al.*, (2000). Filled black circles highlight some important residues. Grey residues highlight residues that were unable to be determined in the original crystal structure.

The Secretin-like family includes receptors for medium sized peptide hormones typically 25-50 residues in length such as secretin, calcitonin, glucagon, corticotrophin-releasing factor (CRF), PTH, amylin and the calcitonin gene related peptide (CGRP). Pham and Sexton (2003) suggested that this family of receptors shared between 30%-50% sequence homology with five distinctive features:

1) A large N-terminal domain (150-180 amino acids) that typically consists of six cysteine residues, two conserved tryptophan residues and an aspartate reside, which may have an essential role in ligand binding (see Sun *et al.*, 2007)

2) Putative N-glycosylation sites on the extracellular domain that could have a pivotal role in receptor expression or/and high affinity binding.

3) A conserved cysteine residue in extracellular loop 1 and 2, which is assumed to produce a disulphide bond that is conserved across all GPCRs .

4) Absence of the DRY motif and the palmitoylated cysteine in the carboxyl tail.

5) A unique distribution of prolines within the transmembrane domains.

The Glutamate-like family comprises of eight metabotropic glutamate receptors, two γ aminobutyric acid (GABA) receptors, a single calcium sensing receptor (CaSR) and taste receptor type 1. The family is characterised by an extremely large N- terminal region (approx. 600 amino acids in length), which in glutamate receptors is believed to form a binding pocket consisting of two lobes which acts in essence like a'venus fly trap' (Schioth and Fredricksson, 2005). Apart from a putative disulphide bridge between TM3 and ECL2, the Glutamate-like family does not have any resemblance to the Secretin and Rhodopsin family (Gether, 2000).

For the sake of brevity, discussions of the two remaining families has not been included, see Schioth and Fredricksson (2005) for a more extensive discussion.

1.2.2 The rhodosin/opsin influence

In June 2000, a three-dimensional crystal structure for a GPCR was published (Palczewski *et al.*, 2000). This landmark was a crystal structure of bovine rhodopsin in its ground state. The crystal provided atomic detail of the apo-protein opsin, a prototypical family A GPCR, which was covalently bound to the chromophore, 11-*cis*-retinal.

To date, fifteen bovine rhodopsin (or opsin) crystal structures have been determined at different inactive, intermediate and activate conformations (see Lodowski *et al.*, 2009 for comparative analysis). Moreover, two crystal structures of squid rhodopsin have been deposited in the PDB (Murakami *et al.*, 2008; Shimamura *et al.*, 2008). Consequently, rhodopsin as a GPCR system is the best understood.

Phototransduction involves transition steps that lead up to activation of the G-protein, transducin. Photon absorption causes isomerisation around the C_{11}= C_{12} double bond of retinal (Nakamichi *et al.,* 2007). Here on in, light absorption spectroscopy can detect receptor photointermediates. The first intermediate which can be identified is bathorhodopsin (absorption maximum 500nm) where the receptor initially accommodates the strained all-*trans* retinal conformer. The receptor begins to thermally relax through the intermediates lumirhodopsin (497nm) and Meta I rhodopsin (480nm). The transition between Meta I rhodopsin to the active receptor conformation (Meta II- 380nm) has been defined by two characteristics (see Lodowski *et al.,* 2009).

The first is the deprotonation of the Schiff base linking all-*trans*-retinal to Lys-296. This event can be clearly observed as a spectral shift of the absorbance maximum from 500 nm to 378 nm. It has been suggested that Glu-113 acts as a counterion to allow the deprotonation event to occur (Sakmar *et al.,* 1989 and Arnis *et al.,* 1994). However, the recent crystal structure of opsin bound to a C-terminal peptide derived from the Gα subunit of transducin (GαCT) suggests that Glu-113 may not be able to fulfill this role as it is spatially too far away from Lys-296 (Scheerer *et al.,* 2009).

The second characteristic that defines Meta II is the disruption of the 'ionic lock'. Although not fully understood, it is assumed that a diffuse network of intra-molecular bonds stabilise the inactive conformation of GPCRs. The crystal structure of ground state rhodopsin revealed that the arginine of the highly conserved D(E)-R-Y(F) motif located at the boundary between TM3 and intracellular loop 2 interacts with an acidic glutamate located at the cytoplasmic end of TM6 (Palczewski *et al.,* 2000). The crystal structure of opsin bound to GαCT supports this notion as the ionic interaction is disrupted. However, the authors suggest that the arginine of the D(E)-R-Y(F) plays a dual role. After, the arginine is liberated from the ionic lock it is free to reside in the GαCT binding cavity, where it directly interacts with the backbone carbonyl group of GαCT at Cys-347.

The plethora of crystal structures that have helped to characterise the structure of rhodopsin/opsin, coupled with recent NMR studies (Ahuja *et al.,* 2009a and Ahuja *et al.,* 2009b) has provided novel details on how the highly conserved CWXP motif in

TM6 plays an important role in regulating receptor activation. The tryptophan (Trp-265) within the CWXP motif has been identified as a key structural element that facilitates rhodopsin activation (Schwartz *et al.,* 2006). It has been proposed that retinal acts as a 'clamp' which forces the side-chain of Trp-265 to be located between the retinal polyene chain and Lys-296 (found on TM7- see Crocker *et al.,* 2006 and Ahuja and Smith, 2009). Upon activation ECL2 gets displaced, which in turn allows the outward rotation of the extracellular end of TM5 (Ahuja *et al.,* 2009a). Consequently, the β-ionone ring moves toward TM5 and therefore Trp-265 is not under the same steric restraints (see Ahuja and Smith, 2009). The proline in the CWXP motif is able to act as a flexible hinge, which straightens TM6. This movement pushes the intracellular region of TM6 away from TM3 towards the lipid bilayer (Schwartz *et al.,* 2006).

The change in the orientation of the side chain of Trp-265 that induces receptor activation has been coined the 'rotamer toggle switch' (Schwartz *et al.,* 2006). This hypothesis could adequately explain the 'rigid body' movements observed in TM6 and TM7 during receptor activation, which could be conceptually compared to a vertical seesaw where highly conserved prolines within the core of the receptor act as the pivot. Therefore upon activation, the extracellular regions of TM6 and TM7 bend inwards towards TM3 but the cytoplasmic face of the receptor 'flowers' to expose epitopes to allow G-protein and arrestin coupling (see Figure 5).

The NPXXY$(x)_{5,6}$F motif, which resides between TM7 and cytoplasmic helix 8, has also been identified as critical for stabilising the rhodopsin ground state. The highly conserved tyrosine (Tyr-306) interacts with Phe-313 located on helix 8 (Palczewski *et al.,* 2000). Moreover, a high resolution crystal revealed that a water molecule could associate with the NPXXY$(x)_{5,6}$F motif to facilitate a stabilising backbone interaction with TM6 (Okada *et al.,* 2002). The opsin/GαCT crystal structure reveals that these interactions are disrupted upon activation, where Tyr-306 extends into the TM bundle to prevent an inward tilt of TM6.

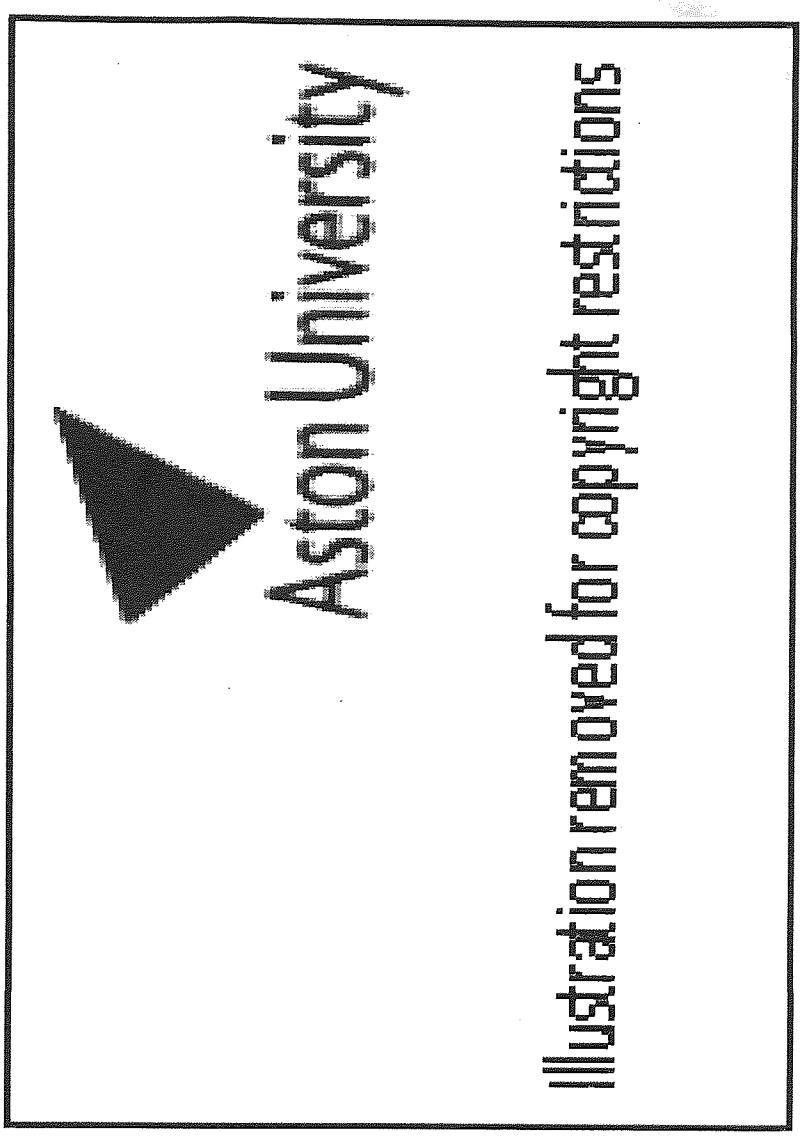


Figure 5. Arrangement of helices in the inactive and active conformation of rhodopsin/opsin. Taken from Scheerer *et al.,* (2008). a) is a view from the cytoplasm of Ops*–Gαt CT. Ops* shown in ribbon (orange) and surface (grey) representation, the Gαt peptide (ILENLKDCGLF) is shown as a ribbon and stick model (blue). b) is a view from the cytoplasm of dark-state rhodopsin (PDB accession 1U19; green ribbon and grey surface) and superposed GαCT (blue) is shown.

Despite these successes, it is sensible to note the pitfalls in homology model design when using bovine rhodopsin as a template. Bovine rhodopsin has potentially unique features that may not be shared globally with other GPCRs. TM5 has the most variable primary sequence throughout the Rhodopsin-like family (Bywater, 2005). In bovine rhodopsin, a short region within TM5 adopts an under-wound π- helix conformation potentially caused by Pro-215. Since prolines have a cyclic structure

where the aliphatic side chain binds to both the backbone nitrogen and α-carbon within the amino acid, the residues unique structure can disrupt the integrity of the helix by introducing kinks and changing the properties of the α-helix. However, the impact of a proline is difficult to predict since the amino acid does not always disrupt an α-helix (Yohannan *et al.,* 2004). The proline in TM5 is highly conserved throughout class A GPCRs but its role is ambiguous and this should be remembered when constructing a homology model. Furthermore, bovine rhodopsin contains a lysine (Lys-296) in TM7 that is covalently bound to retinal.

The pair of glycine residues in TM2 in rhodopsin could also produce problems in homology modelling. Glycine is the simplest amino acid as its side chain contains a solitary hydrogen atom. Consequently, a glycine residue is particularly flexible as it is not under the same steric pressures as the other 19 natural amino acids. A single glycine residue within the TM region would not be sufficient enough to distort a TM domain but suggests two consecutive glycines could explain the twisted phenotype seen within TM2 (Bywater *et al.,* 2005). However, Bywater *et al.* (2005) toy with the idea that the highly conserved proline found in TM2 in other Rhodopsin-like GPCRs, (located 2-3 residues away from the double glycine residues in bovine rhodopsin), could produce a similar curved topology seen in TM2.

1.2.3 A plethora of new GPCR crystals

In 2007 to 2008 consecutive breakthroughs led to the elucidation of several GPCR structures determined by X-ray crystallography. The human β2 adrenoceptor (β_2AR) was the first ligand operated GPCR to be determined (Rasmussen *et al.*, 2007). The initial work produced two similar structures of carazolol-bound β_2AR-Fab5 complexes determined at a 3.4Å/3.7Å resolution. The 3.4Å structure contained a tobacco etch virus (TEV) cleavage site after amino acid 24 of the N-terminus to facilitate crystallisation. Carazolol is an inverse agonist that would reduce constitutive activity, which in turn reduced the inherent flexibility of the protein. Fab5 is a fragment of a monoclonal antibody (Mab5) that is specific to the ICL3 of β_2AR. The addition of Fab5 provided additional conformational stability, while increasing the polar surface of the receptor to encourage crystal contacts. Unfortunately, the extracellular regions of the receptor were not visible in these structures. Although, the TM bundle had a somewhat more open conformation compared with bovine

rhodopsin, the TM Cα atom backbone shared a 1.56Å root mean square deviation (RMSD), suggesting a similar architecture.

Within the same month, the structure of the β_2AR/T4 lysozyme fusion protein bound to carazolol was elucidated with a 2.4Å resolution (Cherezov *et al.,* 2007). The T4 lysozyme replaced ICL3 of β_2AR to facilitate crystallisation. This high-resolution structure was far more informative than its predecessor as atomic detail of the extracellular domains and the pharmacophore of carazolol are present. The TM arrangement has clear differences to bovine rhodopsin. Most notably, TM1 does not possess a proline; consequently the TM is not kinked. The authors hypothesise this feature produces a more open platform that in turn allows an accessible ligand to bind.

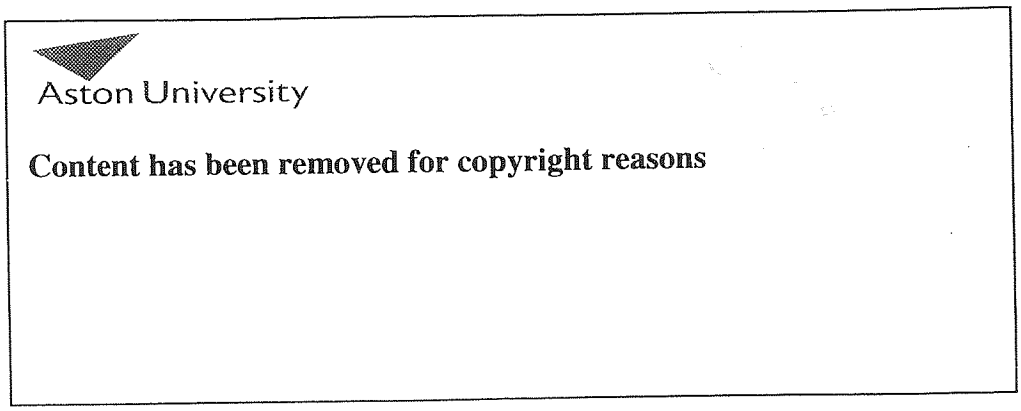
Another, key structural difference between β_2AR and bovine rhodopsin is ECL2. In bovine rhodopsin ECL2 folds to produce a tight antiparallel β-hairpin structure that forms a lid over the 11-*cis*- retinal by plunging into the TM-bundle (Palczewski *et al.,* 2000). In contrast, ECL2 of the β_2AR contains an α-helical segment and is more exposed to the aqueous environment, presumably to facilitate ligand entry. ECL2 of β_2AR is stabilised by two disulphide bonds. Interestingly, the location of the carazolol binding site in β_2AR is comparable to the retinal site in bovine rhodopsin. As noted previously, carazolol is an inverse agonist that should reduce basal activity of β_2AR. To this end, it is interesting to note that the side-chain orientation Trp-286 on TM6, residue predicted to act as a 'toggle switch' in rhodopsin, is in the inactive conformation. However, unlike rhodopsin where this rotamer state is directly governed by retinal interactions, carazolol seems to indirectly impose this conformational state by interacting with Phe-289 and Phe-290.

Assigning functionality to other key motifs is a considerable challenge. The proposed 'ionic lock' is not present in the initial β_2AR structures. It has been proposed that the absence of the 'ionic lock' could be due to the experimental manipulations of the proteins (i.e. the addition of an antibody and insertion of T4 lysozyme). Contrastingly, the effects of an inverse agonist (e.g. carazolol) may propagate different receptor conformations compared to a classical antagonist, possible evidence for graded activation (Topiol and Sabio, 2009).

In 2008 the T4 lysozyme strategy was employed to elucidate the structure of the β_2AR bound to timolol, a first generation beta-blocker that has been characterised as an inverse partial agonist (Hanson *et al.,* 2008). The defining feature of this structure was it revealed two cholesterol-binding sites formed by TM1 and 2 and by TM3 and 4. The presence of two cholesterol molecules has been suggested to increase the stability of the protein and influence the mobility of TM domains in the receptor complex (Hanson *et al.,* 2008).

In July 2008, the structure of the turkey β1 adrenergic receptor bound to the antagonist cyanopindolol was elucidated to 2.7Å (Serrano-Vega *et al.,* 2008). To achieve crystallisation certain modifications including truncation and the incorporation of six mutations that increased the thermostability of the protein had to be made. The TM bundle was similar to the previous high-resolution β_2AR structure, as the Cα atom backbone share of 0.25Å RMSD. Moreover, the ECL regions are also very similar apart from the presence of a sodium ion in ECL2 that is believed to stabilise the short α helical conformation of the loop. However, analysis of the ICL regions, particularly ICL2, was of considerable interest. The proposed 'ionic lock' that was thought to be essential in maintaining the inactive state is not present in this structure. However, ICL2 contains a short alpha helix, where a tyrosine (Tyr-149) in this structure is in direct contact with the aspartate (Asp-138) of the DRY motif. Consequently, the authors speculate that it is ICL2 that acts as switch to enable G-protein activation.

In October 2008 the crystal structure of the human A_{2A} adenosine receptor bound to an antagonist was elucidated (Jaakola *et al.,* 2008). The T4L strategy was again applied to determine the 2.6Å structure. The structure has a similar TM bundle when compared to the adrenergic receptors but slight variations are noted that allow the receptor to accommodate a different set of ligands. Furthermore, ECL2 lacks the secondary structural elements seen in both bovine rhodopsin and the adrenergic receptors as three disulphide bonds support its random architecture. However, the A_{2A} adenosine receptor crystal structure provides further support for the 'rotamer toggle switch' as the antagonist has a 14Å contact area with Trp-246, restricting the side chains mobility encouraging an inactive state. ICL2 of this structure also contains a short alpha helix where the same hydrogen interactions are exhibited as the turkey β1 adrenergic receptor.



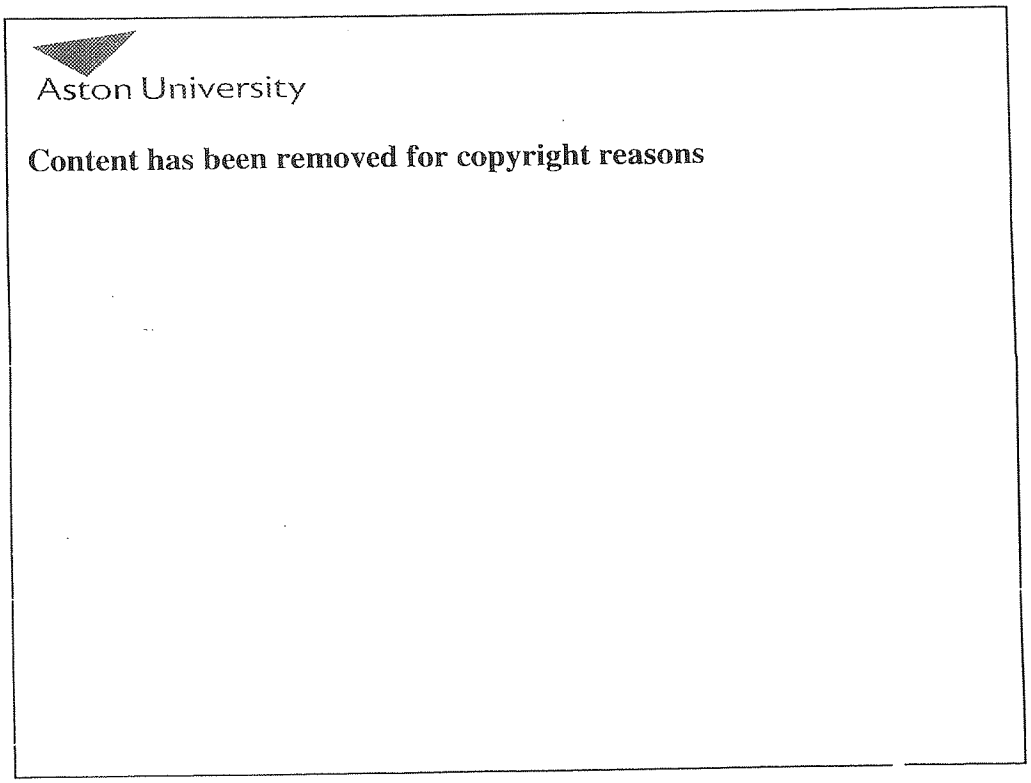


Figure 6. Comparison of novel GPCR crystal structures. Taken from Rosenbaum *et al.*, (2009). A shows inactive bovine rhodopsin (purple), turkey β_1AR (orange) and human A_{2A} adenosine receptor (green) are each superimposed on the human β_2AR structure (blue). B shows extracellular views of rhodopsin, the β_2AR and the A_{2A} adenosine receptor. The ligands are shown as spheres.

It is clear that 2008 was an unprecedented success in understanding ligand mediated GPCRs. Yet, the precise molecular determinants that govern Rhodopsin-like GPCR activation remain unclear. This can partly be ascribed to the naïve way receptor activation has been traditionally perceived as an all or nothing on/off switch. Concepts such as graded activation and 'ligand-directed trafficking of receptor stimulus (LDTRs)', which can be described as

"the ability of different ligands to stabilize specific conformations preferentially, each associated with its own repertoire of stimuli and signalling behaviours, to the relative exclusion of other possible receptor states" (Christopoulous *et al.*, 2007)

have to be appreciated when trying to identify molecular switches found inherently in GPCRs. The diverse range of agonists coupled with their functional idiosyncrasy suggests that there are multiple ways in which a GPCR could be stabilised to trigger an intracellular cascade (Schwartz and Rosenklide, 1996). Yet, there is strong empirical evidence that supports global activation models such as the 'rotamer toggle switch' suggesting GPCRs may share a global mode of activation despite the lack of a universal 'lock' (Schwartz *et al.*, 2006).

1.3 Binding and activation models associated with the Secretin-like GPCRs

1.3.1 The Secretin family recognition fold

Larger ligands such as endogenous peptides have a tendency to bind to the receptors extracellular domains, whereas smaller ligands, such as the bioamines, bind to the core of the receptor in and around the helical bundles. Secretin-like GPCRs respond to medium sized peptides, which implies a diffuse pharmacophore within the extracellular surface, making the large N-terminal region a site of interest.

In 2004, the NMR structure of the N-terminal extracellular domain (ECD) of the mouse CRFR-2β (a prototypical Secretin –like receptor) was elucidated (Grace *et al.*, 2004), and even more recently the mouse CRFR-2β ECD in association with an antagonist, astressin, was published (Grace *et al.*, 2007). This breakthrough revealed that between residues 39-133 the N-terminus contained a common protein fold referred to as a short consensus repeat (SCR) or sushi domain. The SCR fold is comprised of two antiparallel β-sheet regions that presumably end because of the presence of two highly conserved prolines (Pro-72 and Pro-83). The architecture of the polypeptide fold is stabilised by three disulphide bonds and by an ionic interaction between Asp-65 and Arg-101, which is located within the core of the construct sandwiched between the aromatic rings of Trp-71 and Trp-109. The ECD also has two disordered segments comprising residues 45–58 (loop 1) and residues 84–98 (loop 4-

nomenclature from Parthier *et al.,* 2009). Interestingly, upon antagonist binding the largest difference observed in the ECD is the conformation of loop 4. Loop 4 contains highly conserved residues throughout the CRFR family and adopts a structured architecture when the antagonist binds. The authors imply that this change would have low entropy costs and could provide an induced fit mechanism for peptide docking.

To date, a slice variant of the pituitary adenylate cyclase-activating polypeptide type 1 receptor (PAC1R) ECD (Sun *et al.,* 2007), the gastric inhibitory polypeptide receptor (GIPR) ECD (Parthier *et al.,* 2007), the glucagon-like peptide 1 receptor (GLP1R) ECD (Runge *et al.,* 2008) and the PTHR1 ECD (Pioszak and Xu 2008) and human CRFR1 ECD (Pioszak *et al.,* 2008) have been elucidated *via* NMR or X-ray crystallography, with the majority in complex with their associated ligand or respective analog. It has become apparent that a common SCR fold is shared amongst this protein family, referred to as the *'secretin family recognition fold'* (Parthier *et al.,* 2009). However, whether all Secretin-like ECD contain the ionic interaction between the highly conserved aspartic acid and arginine has been questioned (Runge *et al.,* 2008). Moreover, a key feature that has become clearly evident in the latter structures compared to the initial mouse CRFR-2β ECD is an additional α helix located at the extreme N-terminus (Sun *et al.,* 2007, Parthier *et al.,* 2007, Runge *et al.,* 2008, Pioszak and Xu 2008, Pioszak *et al.,* 2008).

The main structural discrepancies found between the Secretin-like ECD structures are located in the loop regions. The PTH1R ECD has an elongated loop1 that is unable to be structurally deciphered by X-ray crystallography. Furthermore, the topology of the other disordered loop, referred to as loop 4 in the mouse CRFR-2β ECD, has been debated (see Parthier *et al.,* 2009). The initial mouse CRFR-2β ECD and PAC1R ECD suggested that this loop lay 'above' the final disulphide bond as the loop approached β-strand 4. Yet, a later revised mouse CRFR-2β ECD along with the remaining Secretin- like ECDs suggests that the loop resides 'below' the terminal disulphide bond. Consequently, questions regarding the reliability of this PAC1R ECD structure have been posed.

Figure 7. The Secretin family recognition fold. Taken from Parthier *et al.* (2009). a) Common structural elements of Secretin-like GPCR ECDs as observed in the GIPR ECD. The domains are stabilized by three conserved disulphide bridges (yellow sticks). b) Superposition of the polypeptide backbones of CRFR-2β ECD (red; PDB code: 2JND), CRFR1 ECD (green; PDB code: 3EHU), GIPR ECD (grey, in cartoon representation; PDB code: 2QKH), GLP1R ECD (yellow; PDB code: 3C5T) and PTH1R ECD (blue, PDB code: 3C4 M). c) The aberrant topology of loop 4 in PAC1R ECD (pink and lilac; PDB code: 2JOD) superimposed on the GIPR ECD (light green and grey). The direction of the main chain of loop 4 (indicated by arrows) in $PAC1R_s$ ECD is opposite to that in GIPR ECD and the other ECDs.

1.3.2 Two-step model of ligand binding

A two-step model of ligand binding has been proposed for the natural ligands of Secretin-like GPCRs (Grace *et al.*, 2004). The first step involves the C-terminal region of the ligand to interact with the solvent exposed N-terminal ECD. Apart for the exception of $PACAP_{6-38}$ binding derived from the controversial PAC1R ECD structure (Sun *et al.*, 2007) it has become apparent that Secretin-like ligands share a common ligand/receptor interface. The endogenous ligands all appear to have the propensity to adopt an amphipathic α-helix conformation upon binding and fit into the same hydrophobic groove in the cognate position in each ECD. Amusingly, high affinity binding in Secretin-like ligands has been referred to as the 'hot dog in a bun' model (Pioszak and Xu, 2008) –see Figure 8.

Figure 8. The 'hot dog in a bun' model. Taken from Parthier *et al.* (2009). a) Superposition of the ECD-ligand complexes after structural alignment of the ECDs (GIPR ECD shown as grey cartoon with surface representation). The bound ligands are coloured as follows: astressin (red; PDB code: 2JND), $PACAP_{6-38}$ (pink; PDB code: 2JOD), GIP_{1-42} (orange; PDB code: 2QKH), exendin-4_{9-39} (yellow; PDB code: 3C5T), PTH_{15-34} (blue; PDB code: 3C4 M), CRF_{22-41} (green; PDB code: 3EHU). Note the binding mode of ECD-bound $PACAP_{6-38}$, which is substantially different to those of the other ligands (N and C termini are labelled). b) View rotated about a horizontal axis by 90°.

The second event of the two-step model of binding involves the N-terminal region of the agonist penetrating into the transmembrane region of the receptor. This tethering may exert enough tension to cause a conformational change and induce activation (see Dong *et al.* 2006). Interestingly, it has been suggested that Secretin-like peptide ligands contain a helix N-capping motif (Neumann *et al.*, 2008). Helix N-capping not only protects and stabilises a peptide but also can introduce a specific local fold, which in this case may help explain how Secretin-like GPCR activation is achieved. Interestingly, the only Secretin-like GPCR peptide ligand family that did not possess a helix N-capping motif was the calcitonin ligand family. However, the calcitonin family contains a disulphide bond between Cys-1 and Cys-7. This modification is predicted to give rise to a conformation similar to that of a helix N-capped peptide (Neumann *et al.*, 2008).

1.3.3 An endogenous agonist?

Lawrence Miller and his research group have suggested that activation in the secretin receptor is not only achieved by the agonist but by a conformational change in the N-terminal ECD, which in turn reveals an endogenous motif that induces receptor activation by interacting with ECL3 (Dong *et al.*, 2006).

The group produced an array of synthetic peptides including a tripeptide corresponding to Trp-48, Asp-49, Asn-50 within the receptor. Interestingly, the WDN peptide was capable of producing a cAMP response. This motif corresponds to Trp-64, Asp-65, Asn-66 of the mouse CRFR-2β, which forms a tightly bound loop region between β-strand 1 and 2. Dong *et al.*, (2006) mimicked this tightly bound loop within the WDN ligand by applying diaminoproprionic acid linkers across the ends of the peptide. This subtle change actually increased the efficacy of the peptide. Another beneficial adaptation to the peptide was fatty acid acylation, which enhanced the peptides hydrophobicity to allow it to target the core of the receptor.

Although Miller's provocative hypothesis has to be carefully considered it does struggle to adequately explain pre-existing data within the Secretin-like family of receptors. It fails to explain how the (1–14) fragment of PTH activates intact and ECD truncated PTH1R (Luck *et al.*, 1999). Moreover, Laburthe *et al.*, (2007) has highlighted the difficulty in applying this hypothesis to the Vasoactive intestinal peptide receptor (VPAC) 1. Asn-69 in VPAC1 would belong to the W-67 D-68 N- 69 motif. However, Asn-69 is glycosylated by a 9kDa carbohydrate (Couvineau *et al.*, 1996). Therefore the carbohydrate moiety would be predicted to face the aqueous environment of the receptor and taking into account steric hindrances it would be difficult to foresee how the WDN motif in this receptor could facilitate activation.

1.3.4 Structure is better conserved than sequence

The tertiary structure of the Secretin-like transmembrane domains remains elusive. The Secretin-like family are remote homologues of the Rhodopsin-like family. Therefore, using standard conservation alignment procedures to construct a homology model is unachievable. When the similarity between primary sequences decreases below 30% (Rost, 1999), the alignment enters the so-called 'twilight zone' where a

large number of gaps and alignment errors are inevitable (Johnson and Overington, 1993).

Consequently, the application of *de novo* or *ab initio* GPCR modelling based on first principles maybe a fruitful endeavour in studying Secretin-like GPCRs. For example, Predix Pharmaceuticals Ltd published details of their PREDICT method for generating *in silico* GPCRs, which only relies on the primary sequence of the protein and the physioco-chemical properties of the membrane environment. This algorithmic approach that no longer relies on a template structure has successfully produced structures of ground state bovine rhodopsin and models that adhere to known experimental data for the dopamine 2 receptor, neurokinin 1 receptor and the neuropeptide 1 receptor (Shacham *et al.,* 2004).

Yet, all GPCRs bind to similar effectors e.g. G-proteins, β-arrestins, etc. Therefore, a tentative hypothesis that GPCRs share a global geometric conformation has been assumed (Frimurer and Bywater, 1999). Ingenious alignment strategies have been employed to allow the Secretin-like GPCR researcher to take advantage of the high-resolution Rhodopsin-like crystal structures. An early attempt by Donnelly (1997) analysed multiple sequence alignments of the Secretin-like family. Hydropathic analysis identified the putative helical domains and a fourier-transform method was used to define the direction of the internal face of each helix. Taking into account further considerations, such as highly conserved residues are not likely to face the phospholipid bilayer, a two dimensional model of the TM domain of the rat GLP1 receptor was constructed.

Noteworthy, is the 'cold-spot' alignment method proposed by Frimurer and Bywater (1999). A cold-spot is defined as a pair of conserved residues, which are the same distance apart in the primary sequence in both Secretin-like and Rhodopsin-like GPCRs. The alignment pivots on such sites as it is assumed structure would be most highly conserved at this position.

More recently, Chugunov *et al.,* (awaiting publication) constructed a Secretin-like and Rhodopsin-like GPCR alignment based on the analysis of four alignments to produce a VPAC1 TM model. The first alignment used in the analysis was an alignment proposed by Bissantz *et al.,* (2004), in her automated GPCR modelling procedure.

The second alignment used was based on the Frimurer and Bywater (1999) cold-spot method. The third alignment was produced by mGenThreader (McGuffin, 2000) *via* the BioInfoBank Meta Server (Ginalski, 2003). Alignment four was a manually built alignment based on a vague implementation of the cold-spot method. The authors noted that the alignments varied in predicting the corresponding positions of TM1, 2, 4 and 5. When constructing a global alignment this could potentially produce 48 alignments. Here on in, an elaborate model building procedure was implemented. Initially, the TM models were assessed on their α helical packing ability using the 'membrane score' approach (Chugunov *et al.,* 2007). Furthermore, to ensure that the variable (less conserved) residues faced the lipid environment, a variability moment vector was used to analyse the models (Du *et al.,* 1994). An alignment that satisfied this rigorous selection process was chosen. The authors note that the final alignment had a better 'membrane score' compared to the four parent alignments and conclude that their alignment supersedes its predecessors.

Christopher Reynolds and his research group have noted that G-protein coupled receptor 1 (GCR1), a plant GPCR, shares homology with Rhodopsin-like, Secretin-like and cAMP GPCRs (Vohra *et al.,* in preparation). Reynolds has exploited this observation to address the Rhodopsin-like/Secretin-like GPCR alignment problem. However, there are certain regions, particularly in TM5, where the GCR1 alignment is not clear-cut. TM5 has long been thought to be a critical domain for ligand recognition as the sequence is variable in Rhodopsin-like GPCRs (see Schwartz *et al.,* 2006). Moreover, as Secretin-like GPCRs have two highly conserved asparagines that are 6 residues apart it is very difficult to apply polarity violation principles, as both of these residues cannot both reside in the internal environment of the TM bundle.

Although Vohra *et al.,* primarily used GCR1 to guide the alignment; other bioinformatical tools were also used. Maximum lagged correlations between Rhodopsin-like and Secretin-like GPCRs based on hydrophobicity, amino acid conservation and variability were made. Importantly, this approach allows the orientation of TM helices not to be determined by the dependent variables (e.g. hydrophobicity, amino acid conservation and variability) but by the pattern of these variables, hence each TM domain should be similar in both Rhodopsin-like and Secretin-like GPCRs. Careful consideration was made to take into account GPCR dimerisation, since patterns of conservation on the external lipid-facing helices are

similar in both Rhodopsin-like and Secretin-like GPCRs. Moreover, the final alignment proposed by Vohra *et al.*, took into account the extensive SDM of the TM domain of the calcitonin-like receptor produced by Conner and Poyner (unpublished data).

Interestingly, five out of seven helical alignments are identical when the Chugunov *et al.*, alignment is compared to the Vohra *et al.* alignment. Moreover, the TM1 alignment is only out by one residue. However, a significant difference is seen in the TM5 alignment. At present, it remains debatable which alignment is correct.

The speculative alignments of Secretin-like receptors all assume that *'structure is better conserved than sequence'* (Frimurer and Bywater, 1999). Consequently, Secretin-like GPCRs may have functionally equivalent motifs as the Rhodopsin-like family. Interestingly, the highly conserved Secretin-like YLH motif located on TM3 is in the cognate position of the E/DRY motif. However, the functional significance of this has not yet been determined. Intriguingly, a highly conserved glutamate located one turn above the L in the YLH motif has been predicted to be involved in an ionic interaction with a highly conserved arginine and histidine located on TM2. Mutagenesis data derived from the CLR supports this concept as R173 at the base of TM2, H177 further up TM2 and E233 in TM3 appear to act as a functional triplet (Conner and Poyner, unpublished data).

The counterpart of the NPXXY motif is the VAVLY motif in Secretin-like GPCRs. It is plausible that the equivalent motif may play a similar role in stabilising the proteins conformation. This line of thought is further supported by the analysis of a synthetic peptide that corresponded to the C-terminal tail of CLR (Conner *et al.*, 2008). This provided evidence that both families contained a helix that lies parallel to the cytoplasmic face of the membrane, referred to as helix 8 (Palczewski *et al.*, 2000).

The CWXP Rhodopsin-like motif located on TM6 does not appear to have a highly conserved counterpart in Secretin-like GPCRs. The tryptophan involved in the 'rotamer toggle switch' is often a tyrosine, tryptophan or phenylalanine in Secretin-like GPCRs. However, in Secretin-like GPCRs a proline residue is not found two residues below this position. However, TM prolines do contribute to signal transduction in Secretin-like GPCRs. It has been proposed Pro-343 on TM6 of CLR

acts like a 'hinge' to facilitate G-protein coupling (Conner *et al.,* 2006). This suggests a similar activation mechanism as seen in Rhodopsin-like GPCRs. Yet, detailed descriptions on the molecular determinants that control the 'hinge –like' proline remain elusive.

1.3.5 Loop domains are important in Secretin-like GPCRs

Loop regions connect the secondary structural elements of proteins and are regarded as important functional domains (see Fiser *et al.* 2000). This is evident in GPCR research as the ICLs facilitate G-protein recognition (see Scheerer *et al.,* 2009) and the extracellular loops ECLs may contain key determinants for ligand binding (see Grace *et al.,* 2004).

Despite having valuable implications for GPCR-based drug discovery, only a handful of studies have focused on ECL domains (see Hauser *et al.*, 2007). Although experimental evidence is sparse, the ECLs in Secretin-like GPCRs have been regarded as important determinants for receptor activation for over a decade. Earlier work by Holtmann *et al.,* (1996) used secretin/vasoactive intestinal polypeptide receptor chimeras to investigate the role of ECL1 and ECL2 in secretin receptor activation. H189 and K190, which were predicted to reside in the C-terminal half of ECL1, were found to critical in both secretin binding and activation. Meanwhile, F257, L258, N260 and T261 that were predicted to be located on the N terminal half of ECL2 were also required for secretin activation and binding.

Bergwitz *et al.,* (1997) investigated ligand selectivity in the PTHR2, which responds to PTH but not to PTH-related peptide. In contrast, the PTHR1 responds to both PTH and PTH-related peptide with equal potency. The difference in selectivity had previously been attributed to two divergent residues between the ligands PTH and PTH-related peptide. Residue 23, which is phenylalanine in PTH-related peptide and a tryptophan in PTH, was shown to govern high affinity binding (Gardella *et al.,* 1996). In contrast, residue 5 modulates receptor activation, which is a histidine in PTH-related peptide and isoleucine in PTH (Behar *et al.,* 1996). Bergwitz *et al.,* (1997) took advantage of the previous observations and designed PTH-related peptide analogues that contained a tryptophan at position 23, which allowed high affinity binding to the PTHR2 and then tested signaling on the PTHR2 with analogues that

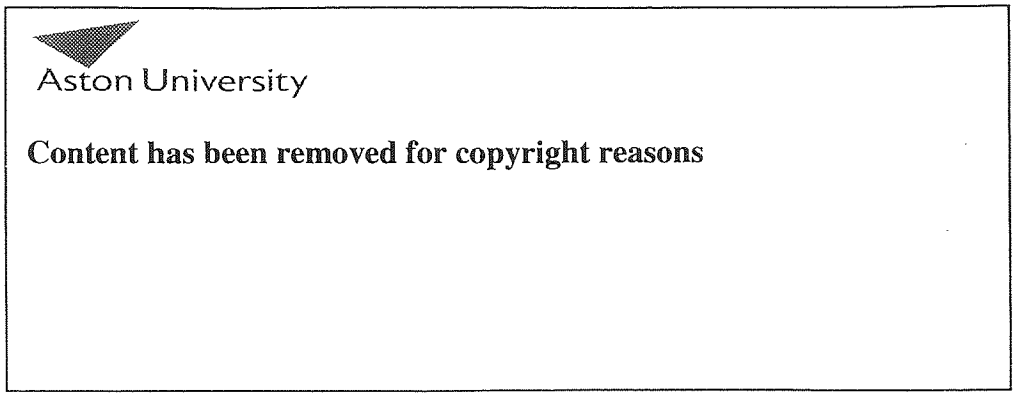
either contained a histidine or isoleucine at position 5. Moreover, Bergwitz *et al.*, (1997) tested the analogues on PTHR1/PTHR2 chimeras followed up by SDM. The chimera results suggested that ligand selectivity involved ECL1, ECL2 and TM3. SDM of ECL1 did not reveal a single site that was important for signalling. However, SDM on ECL2 and TM3 revealed that I244 predicted to be located near the extracellular end of TM3 and Y318 near the C-terminal end of ECL2 were responsible for ligand selectivity of PTH-related peptide at position 5.

Moreover, photoaffinity labelling studies and mutagenesis strategies on rat and human PTH receptors have identified putative epitopes located in and around ECL3 (Lee *et al.*, 1995 and Bisello *et al.*, 1998). More recently, a novel disulphide trapping approach has been used to map out the interaction sites between PTH and its cognate PTHR1 (Monaghan *et al.*, 2008). The results suggested that the first residue of PTH (Ser-1) was found to be close to four residues, three of which reside in the top of TM6 (see Figure 9).

The ECL domains have been identified as important in other Secretin-like GPCRs. Assil-Kishawi *et al.*, (2002) found that K16 of the ligand sauvagine was capable of cross-linking with K257, located in ECL2 of CRFR1. More recently, Gkountelias *et al.*, (2009) conducted an alanine scan of ECL2 of CRFR1 (Leu-251 to Val-266). The results suggested that only W259A and F260A reduced sauvagine potency and binding affinity. Moreover, using various truncated ligands the authors were able to determine that W259 and F260 were most likely to interact with positions 8 to 10 of sauvagine.

Runge *et al.*, (2003) suggested that three distinct epitopes in the glucagon receptor (GLR) extracellular face governed ligand selectivity. Runge and co-workers were able to deduce that the top of TM2, ECL2 and ECL3 were essential in facilitating glucagon-like- peptide 1 activation. The elegant study involved GLR and GLP1R chimeras that were probed with four glucagon analogues, which incorporated divergent GLP1R mutations. Specifically, the top of TM2 chimera rescued Glu3-glucagon, ECL2 and associated juxtamembrane chimera rescued Ser12-glucagon and ECL3 and the proximal segments of TM6 and TM7 rescued Ala2-glucagon. Runge *et al.*, (2003) proposed a potential binding model depicted in Figure 9.

However, triangulating current information to postulate a general model for Secretin-like ligand docking on the TM bundle is not feasible. For example, comparing the Runge *et al.,* (2003) model against the Monaghan *et al.,* (2008) model shows clear discrepancies (see Figure 9). Although, both models suggest that the ligand is spread across the extracellular face of the TM bundle, the position of the N-terminus of the ligand is in opposing positions. The first residue of the ligand in Runge *et al.,* (2003) is predicted to be buried in the crevice between TM1, 2 and 7, whereas, the Monaghan *et al.,* (2008) model suggests that the first residue of PTH is located near TM5 and 6, which is located on the opposite side of the TM bundle. It could be hypothesised that both binding sites are critical for receptor activation and it depends on the entrance pathway of the ligand that determines which activation site is preferred. Further investigation into the orientation of the N-terminus ECD relative to the TM bundle in Secretin-like GPCRs may shed light on this controversial area.



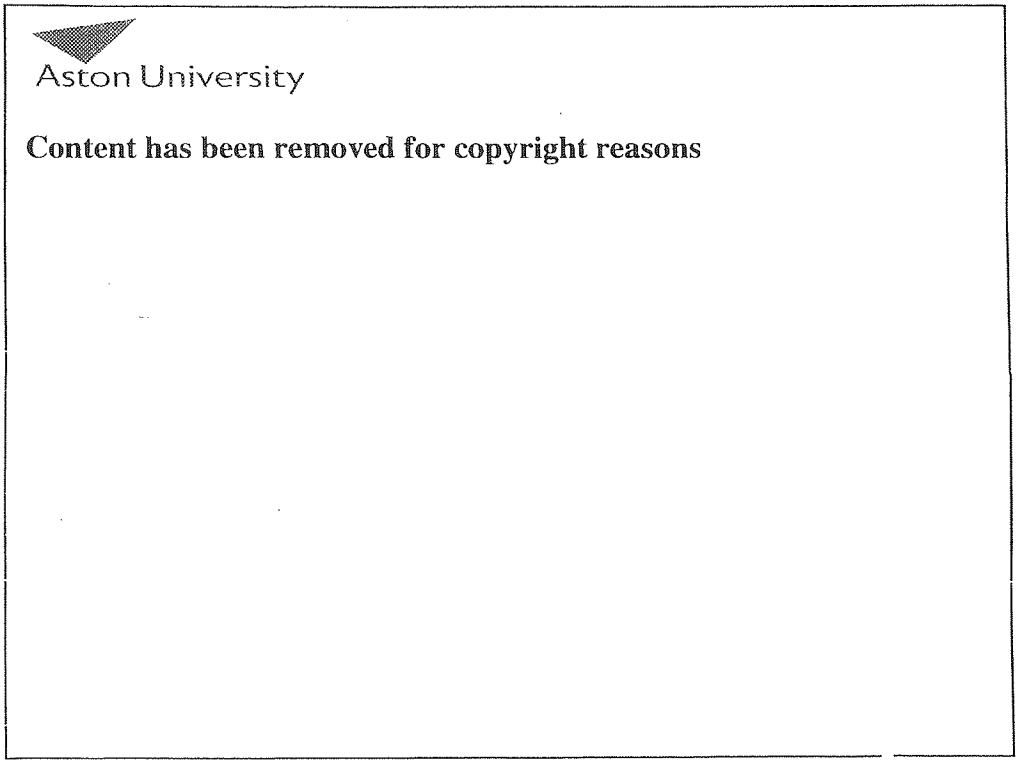


Figure 9. The difference between the Runge *et al.*, (2003) binding model compared to the Monaghan *et al.*, (2008) binding model. a) Taken from Runge *et al.*, (2003) showing Ser-2 of glucagon is close to the extracellular end of TM7 of glucagon receptor, Gln-3 is close to the extracellular end of TM2, and Lys12 is close to ECL2 and/or the proximal helices TM4 and/or TM5. b) Taken from Monaghan *et al.*, (2008) showing a model of the interaction of position 1 of PTH with PTHR1 obtained by molecular dynamic simulations. The four sites in PTHR1 that formed disulphide bonds with Cys1-PTH are shown in green; the backbone of the ligand is shown in yellow (nitrogen = blue; oxygen = red).

1.3.6 Determining loop structure

Loop regions are highly divergent both in size and amino acid composition. Given the difficulties in elucidating entire GPCRs by X-ray crystallography or *via* solution NMR, numerous groups have adopted the 'divide and conquer' approach to gain detailed information on GPCR loop domains using loop mimetic peptides. For example, Bellot *et al.,* (2009) synthesised a peptide of ICL3 of the rat vasopressin receptor 2 and elucidated the solution NMR structure in dodecylphosphocholine micelles. More specifically to Secretin-like GPCRs Plati *et al.,* (2008) using an NMR approach determined how the Gαs subunit is able to interact with a seven-residue peptide mimetic of the C-terminal region of ICL3 in the PTH1 receptor.

However, when there is no or little experimental evidence available, molecular modelling techniques are often implemented to gain insight into the structure of the loop domain. Loop prediction is difficult due to the loop's inherent flexibility and has been described as a 'mini protein folding problem' (Fiser and Sali, 2000). A plethora of loop prediction methods have been published but broadly loop prediction methodology can be divided into two approaches: database searches and *ab initio* methods. Database search is a knowledge-based approach where a segment of a protein with a known three-dimensional structure is obtained from a database, which can fit in between the stem regions of a loop. The stems are defined as the main-chain atoms that precede and follow the loop, but are not part of it. A clear advancement in this area was the Loops In Protein (LIP) database developed by Michalsky *et al.,* (2003), which reports accurate results for loops between nine and fourteen residues in length.

The pioneers of applying an *ab initio* approach to loop prediction were Moult and James (1986) and Bruccoleri and Karplus (1987). Generally, the *ab initio* method is based on a conformational search (usually randomly) in a given environment, and selection of near native conformers is guided by a scoring or energy function (see Fiser and Sali, 2000). Consequently, loop prediction accuracy depends on the effectiveness of the conformational search and on the quality of the scoring function, which usually depends on the quality of the force field used to evaluate the conformational energy (Xiang *et al.,* 2002). Soto *et al.,* (2008) suggested that most *ab initio* loop-modeling procedures can accurately predict up to seven residues and

therefore the field is striving to increase loop accuracy in conjunction with loop length to fourteen residues.

At present the five most popular packages for predicting loop conformations are Modeller (Fiser and Sali, 2000), RAPPER (de Bakker *et al.*, 2003), Loopy (Xiang *et al.,* 2002) Rosetta (Rohl *et al.,* 2004) and Protein local optimization program (PLOP; Jacobsen *et al.*, 2004 and Zhu and Pincus *et al.,* 2006). To begin with all methods produce a large ensemble of loop conformers. Modeller produces the initial ensemble based on a sum of many spatial restraints that include the bond length, bond angle, and improper dihedral angle terms from the Chemistry at Harvard macromolecular mechanics (CHARMM)-22 force field and uses a relaxation method that uses both conjugate gradient minimisation and molecular dynamics with simulated annealing to produce the final set of loop conformers (Fiser and Sali, 2000). RAPPER uses fine-grained residue-specific phi/psi propensity tables for conformational sampling (de Bakker *et al.*, 2003). Loopy uses a modified version of the random tweak algorithm, which generates loop conformations that are open at one end and closed by using phi/psi residue constraints between the loop and stem region. Rosetta uses a dual approach as the conformational search begins with database-derived fragments of protein structure and assembles them with a Monte Carlo procedure and simulated annealing tethers the loop to the stem. PLOP uses an elaborate dihedral angle buildup procedure.

Moreover, many scoring functions have been proposed to evaluate the large ensemble of loop conformers, in an attempt to identify native or near native structures. Fiser and Sali (2000) suggested that a good loop prediction was equal to or below ~1.5Å global RMSD, whereas a bad prediction was above ~3.0Å global RMSD. Broadly, scoring functions can be divided into two categories: physical based potentials and knowledge (or statistical) based potentials (Zhang *et al.,* 2004). As the name suggests physical based potentials are derived from the law of physics, typically well-characterised force-fields are used to evaluate the energetics of the system. However, for large-scale application to loop prediction certain approximations have to be used, such as the implementation of implicit solvent models to mimic the native environment (de Bakker *et al.*, 2003 and Jacobsen *et al.*, 2004).

In contrast, a statistical based potential is obtained directly from statistical analysis of known protein structures (see Yang *et al.*, 2008). Common statistical-based potentials include residue-specific all atom conditional probability discriminatory function (RAPDF) developed by Samudrala and Moult (1998), distance-scaled, finite, ideal-gas reference state (DFIRE) developed by Zhou and Zhou (2002) and discrete optimized protein energy (DOPE) developed by Shen and Sali (2006). The updated DFIRE based potential referred to as DFIRE 2.0, appears to yield the most accurate results (Yang *et al.*, 2008). However, some procedures use both physical and statistical-based potentials to guide the loop selection procedure. For example, de Bakker *et al.*, (2003) scores the loop conformers with RAPDF and selects the top 50 conformers for further analysis using the Assisted model building with energy refinement (AMBER) forcefield with a Generalised Born/surface area (GB/SA) solvation model. The accuracy of Modeller, RAPPER, Loopy, Rosetta and PLOP loop procedures are summarised in figure 10.

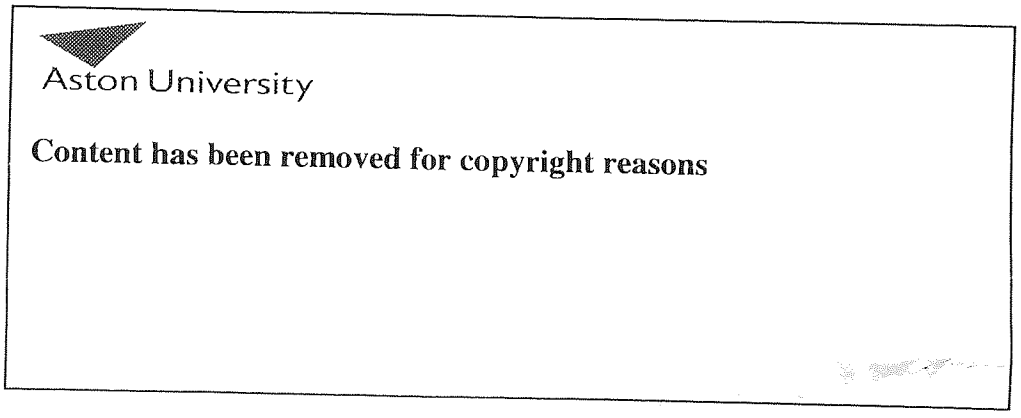


Figure 10. Summary of reported accuracy of loop packages. Mean average summary of popular loop packages taken from Soto *et al.*, (2008)

It is clear that PLOP is the most accurate program. However, the main drawback of PLOP is that it is not computationally efficient as it can take days to weeks to generate loop conformers. Soto *et al.*, (2008) developed a loop procedure called LoopBuilder that aimed to be as accurate as PLOP and be efficient. The procedure involved a modified version of loopy where the conformers were scored by DFIRE and then using a mechanics-based energy minimisation step to refine the top 50

conformers, using the optimized potentials for liquid simulations (OPLS)/SBG-NP force field. The procedure is very similar to de Bakker *et al.*, (2003) but Soto *et al.*, (2008) suggest the marked improvement in accuracy is primarily due to the superiority of DFIRE over RAPDF. Although, LoopBuilder is capable of building 8-12 loops with an accuracy of 1.31-2.65Å RMSD, it still remains slightly less accurate than PLOP by approximately 0.3-0.9Å RMSD.

Taking into consideration Chris Reynold's Family A to Family B GPCR alignment it suggests that certain Secretin-like GPCR loop domains are amenable to an *ab initio* loop procedure. For example, the CLR TM bundle may be approximated on the basis that ICL1 is predicted to be nine residues in length, ECL1 is eight residues in length, ICL2 is nine residues, ICL3 is eleven residues in length, and ECL3 has 10 residues. The most problematic loop to model would be ECL2, which is 26 residues in length. However, there is a cysteine residue predicted to be 14 residues into the loop that is suspected to participate in a highly conserved disulphide bond in TM3. Consequently, given the speculative nature of Secretin-like GPCR models it could feasible to use this cysteine as an anchor region and split this loop into two segments. The first of which would still be 13 residues long and great caution would have to be taken when interpreting results, while the second segment would be 12 residues in length.

Another concern when generating loop regions for a Secretin-like GPCR is the impact of the N-terminal ECD on the ECLs. The current drawback in the field is deciphering the relative position of the N-terminal ECD in relation to the TM bundle, since the two domains are linked by a flexible loop. Although, attempts using loose restraints between the WDN motif and ECL3 have been used to estimate the relative position of the ECD (Dong *et al.*, 2008) there is only one study, to my knowledge, that has identified a specific contact site between the two domains. Vilardaga *et al.*, (1997) investigated the role of cysteine residues in the rat secretin receptor using serine mutations. In addition, to the cysteines that are believed to be involved in stabilising the ECD in all Secretin-like ECDs, a putative disulphide bond was predicted to be located between Cys-11 (located in the extreme N-terminus of the receptor) and Cys-186 (located at the top of TM2).

Cysteine residues in the cognate positions are also found in the human Secretin receptor (SCTR), VPAC1, VPAC2 and in the PAC1R. Unfortunately, the only ECD

structure that has been elucidated out of these four domains is the human splice variant of PAC1R (Sun *et al.*, 2007). This structure has already been heavily criticised (Parthier *et al.*, 2008) and the first 9 residues of the NMR structure do not correlate with the native PAC1R sequence and consequently the cysteine of interest in the extreme N-terminus has not been elucidated. The lack of structural information inspired Taylor *et al.*, (2003) to produce an *ab initio* model of the rat secretin ECD. Then, the putative disulphide bond (C11-C186) was used as a distance restraint to dock the rat secretin ECD on to its predicted TM bundle. Based on the assumption that disulphide bonds typically stabilise the pre-existing fold of the protein it could be hypothesised that the extreme N-terminus of Secretin-like ECDs are all in close proximity to the top of TM2 and ECL1.

However, the cognate cysteines (C11-C186) in the rat secretin receptor are not found in all Secretin-like receptors e.g. the CLR. Furthermore, sequence homology in the extreme N-terminus of Secretin-like receptors is extremely low before the putative N-terminal helix. This makes it difficult to determine the equivalent C11 in the rat secretin receptor.

1.4 The CGRP system

1.4.1 Calcitonin family of peptides

The calcitonin family of peptides comprises of six members including calcitonin, amylin, adrenomedullin (AM), two distinct forms of calcitonin gene-related peptide (αCGRP and βCGRP), and a recently discovered member called intermedin or AM 2 (see Roh *et al.*, 2004). The family as a whole does not share primary sequence homology but its members are grouped together on their predicted secondary structure (Poyner *et al.*, 2002). In spite of structural similarities and a certain degree of pharmacological overlap, unique pharmacological profiles have been reported for these peptides both *in vitro* and *in vivo* systems (Poyner *et al.*, 2002).

Calcitonin is a hormone derived from the C-cells of the thyroid gland and inhibits oestoclast mediated bone reabsortion. Amylin is produced in the β-cells of the pancreas and causes a reduction in nutrient intake. AM is 52 amino acids in length and is predominantly found in vascular tissue and is believed to increase pulmonary

blood flow (Wimalawansa, 1996). Intermedin exists as both a 40 and 47 amino acid neuropeptide and has been found to reduce blood pressure and gastric emptying by acting promiscuously on CGRP and adrenomedullin receptors (Rho *et al.,* 2003).

αCGRP is a sensory nerve-derived peptide made up of 37 amino acids (Amara *et al.,* 1982). Alternate splicing of the calcitonin gene, which is located on the short arm of chromosome 11, regulates αCGRP synthesis over calcitonin, which is determined in a tissue-specific manner. Later, it was found that βCGRP is transcribed from its own distinct gene, which is also located on chromosome 11 (Steenbergh *et al.,* 1985). It is assumed that the second gene arose through gene duplication but it remains ambiguous which gene first appeared in evolution (Wimalawansa, 1996).

αCGRP and βCGRP only differ by three amino acids in humans (see Poyner *et al.,* 2002). Although, αCGRP and βCGRP are both widely distributed throughout the central and peripheral nervous system, differences in expression have been reported in both primary sensory neurons and enteric autonomic neurons in the rat (Mulderry *et al.,* 1988) and in the human hypothalamus (Henke *et al.,* 1987).

1.4.2 The role of αCGRP and related ligands

αCGRP acts as a potent vasodilator having pronounced effects on arterioles and capillaries (Brain and Grant, 2004). αCGRP is widely distributed throughout the nervous system as well as in the cardiovascular system, implying important physiological significance (see Arulmani *et al.,* 2004). At very high concentrations αCGRP can reduce blood pressure (Brain and Grant, 2004). However, it seems that αCGRP does not play a fundamental role in regulating blood pressure. The neuropeptide acts near its site of release causing an increase in local blood flow, which in turn can cause facial flushing, oedema and inflammation (see Brain and Cox, 2006). CGRP-like peptides have also been associated with nociception, glucose uptake and the stimulation of glycolysis in skeletal muscles (see Van Rossum *et al.,* 1998).

Interestingly, αCGRP has also been associated with a number of vascular diseases. Raynaud's disease is relatively rare characterised by vasospasms in both fingers and toes and is associated with changes in plasma levels of αCGRP (Brain and Grant,

2004). Furthermore, the aetiology of vascular headaches including migraines may partly be ascribed to increased levels of CGRP (Lassen, 2002. and Juhasz, 2003). The exact neuronal mechanism underpinning migraine onset remains elusive but there are several ways in which αCGRP may be implicated. Firstly, αCGRP can act on cerebrovasculature smooth muscle to cause the cranial blood vessels to dilate, which leads to neurogenic inflammation. Neurogenic inflammation can be further enhanced by αCGRP activity since CGRP receptors are located on dural mast cells, which also release pro-inflammatory cytokines upon activation. Moreover, CGRP receptors are found in abundance on the trigeminal ganglion, where the trigeminovascular system within the brain has been directly associated with intense migrainous pain (see de Prado and Russo, 2006). CGRP receptors are also located at the caudal brainstem, which has also been associated with nociception (Edvinsson, 2004).

A family of tryptamine-based drugs called triptans are the current favoured therapeutic agent for relieving acute migrainous pain. Triptans are 5-HT_{1B} and 5-HT_{1D} agonists. Inexplicably one- third of patients do not respond to triptans (Geraud *et al.*, 2003). Furthermore, patient tolerance of this medication is a potential problem. The drugs can increase the risk of cardiovascular disease in susceptible hypertensive patients (Martin and Goldstein, 2005).

A new approach using CGRP receptor antagonists to treat migraine has been suggested (see de Prado and Russo 2006). Chiba *et al.* (1989) found that removing the first seven amino acids from human αCGRP to produce $CGRP_{8-37}$ dose dependently displaces ^{125}I-[Tyr]rat CGRP binding in rat liver plasma membranes, without activating the receptor. This was the first CGRP receptor antagonist to be identified but unfortunately this CGRP fragment has no real clinical benefits due to its short-half life *in vivo*. Other truncated CGRP fragments have also been developed that bind with higher affinity to CGRP receptors, such as human $CGRP_{27-37}$ (Rist *et al.* 1999), but still these remain limited by their inherent short-half lives (see de Prado and Russo, 2006). Consequently the identification of alternative types of selective antagonist is a research priority.

The identification of N-[2-[[5-amino-l-[[4-(4-pyridinyl)-l-piperazinyl]carbonyl]pentyl]amino]-1-[(3,5-dibromo-4-hydroxyphenyl)methyl]-2-oxoethyl]-4-(1,4-dihydro-2-oxo-3(2H)-quinazolinyl (BIBN4096BS), a dipeptide-like

compound, marked a breakthrough in this quest to identify a low molecular weight CGRP receptor antagonist (Doods *et al.*, 2000). BIBN4096BS measurements of affinity to membranes isolated from human neuroblastoma cell lines (SK-N-MC) have revealed that the compound binds to human CGRP receptors with a markedly higher affinity (150 times) than the $CGRP_{8-37}$ fragment (Doods *et al.*, 2000). Moreover, this relatively small compound is highly selective since Doods *et al.* (2000) demonstrated non-interactivity with 75 different sets of receptor and enzyme systems.

BIBN4096BS is also capable of differentiating between species as it preferentially binds to primate CGRP receptors (most notably human and marmoset) over rodent CGRP receptors. This may be because BIBN4096BS is able to discriminate between these receptors based on a single amino acid, (identified as Trp-74) which is located on the extracellular surface of the primate receptor activity modifying protein (RAMP)-1.

Alternative CGRP receptor antagonists have since been developed such as N-methyl *N*-(2-methylphenyl)-3-nitro-4-(2-thiazolylsulfinyl)-nitrobenzanilide (SB-273779) that acts as a reversible and competitive CGRP antagonist, albeit less potently than BIBN4096BS but is considered functionally distinct since both rodent and porcine CGRP receptors are inactivated by the compound (Aiyar *et al.*, 2001). Moreover, Merck Research Laboratories have developed a potent and orally active antagonist referred to as 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] (MK-0974; Salvatore *et al.*, 2007).

Figure 11. Chemical structures of three CGRP receptor antagonists: BIBN4096BS, MK-0974 and SB-273779. Taken from Miller *et al.*, (2010).Perspective drawings of BIBN4096BS, MK-0974 and SB-273779.

1.4.3 The structure of αCGRP

The structure of αCGRP has been investigated by an array of conformational and modelling techniques (Lynch and Kaiser, 1988; Manning, 1989; Hubbard *et al.*, 1991; Breeze *et al.*, 1991; Hakala & Vihinen, 1994). Conner *et al.*, (2002) suggested that the neuropeptide can be broadly differentiated into four domains. The N-terminus domain comprises of the first seven amino acids and adopts a ring-like structure held together by a disulphide bridge between Cys-2 and Cys-7. In conjunction with the 'two-step' model of binding this ring-like structure is able to directly interact with juxta-membrane regions of the CGRP receptor to induce activation. Although at very low potency, fragment $\alpha CGRP_{1\text{-}15}$ is the shortest active fragment to be identified (see Maggi *et al.* 1990).

The second domain specified by Conner *et al.* (2002) lies between residues 8-18. This forms an amphipathic α-helix (Lynch and Kaiser, 1988) when exposed to a hydrophobic environment (Hubbard *et al.*, 1991). The domain has been shown to mediate high affinity binding to CGRP receptors since the deletion of this secondary structure causes a 50-100 fold decrease in affinity (Rovero *et al.*, 1992). Furthermore,

introducing a proline at position 16, which causes a kink in the α-helix, is also detrimental to affinity (Wisskirchen *et al.*, 1999).

A SDM study determined the orientation of the ligand α-helix upon binding to the receptor (Howitt *et al.*, 2003) Arg-11 and Arg-18 (located on the hydrophilic face of the α-helix) had previously been shown to play a vital role in promoting high affinity binding (Poyner *et al.*, 1998). Arginine side chains are positively charged making them hydrophilic in nature. When αCGRP binds, hydrophilic Arg-11 and Arg-18 could potentially interact with the extracellular fluid, which would orientate the hydrophobic face of the helix towards the receptor. Alternatively, since Arg-11 and Arg-18 are positively charged it is also possible that the residues play a more direct role in ligand binding by forming non-covalent interactions with either the receptor or with other regions within the peptide. Howitt *et al.*, (2003) substituted the pair of arginines with serine (uncharged but hydrophilic) and then with glutamic acid (negatively charged) on the understanding that if the two arginines were interacting with water molecules then these amino acid substitutions would not significantly impact on high-affinity binding. However, the substitutions did show a decrease in αCGRP affinity, supporting the latter model of binding. Moreover, the spatial constraints of the hydrophobic side of the helix were also examined by substituting leucine (at positions 12, 15 and 16) with phenylbenzoic acid residues. A large decrease in affinity was noted in Leu-16, implying that the size and geometry of this residue is essential for normal functioning.

The third domain, between residues 19-27, appears to have no stringent constraints on its composition (Conner *et al.*, 2002). However, two-dimensional NMR spectroscopy coupled with molecular modeling has suggested the presence of a β-turn between Ser-19 and Gly-21 (Boulanger *et al.*, 1995).

The fourth domain that occurs after this hinge-like region incorporates residues 28-37. It has been suggested that this region is essential for receptor binding but not essential for activation for the reason that C-terminal analogues (the shortest being $[D^{31}, P^{34}, F^{35}]$-$CGRP_{30-37}$) always act as antagonists (see Carpenter *et al.*, 2001). Although, the C-terminal region is predominantly disordered, two turn structures within this region (centered on Pro-28 and Gly-33) expose the side chain of Thr-30

putting it in close proximity to Val-32 and Phe-37 and hence form a putative binding domain (see Carpenter *et al.*, 2001 and Conner *et al.*, 2002).

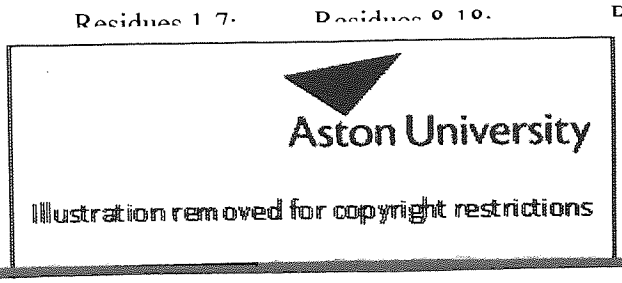


Figure 12. The structure of human αCGRP. The amino acid sequence of αCGRP is shown and the four putative domains outlined by Conner *et al.*, (2002) are shown.

An observation by Robinson *et al.*, (2009) was that αCGRP has an 80% α-helical content in 50% trifluoroethanol, an environment that is expected to mimic binding conformations.

1.5 The CGRP receptor

In humans, only one CGRP receptor has been successfully cloned and molecularly deciphered. The CGRP receptor is unique among G-protein coupled receptors consisting of at least three proteins: calcitonin receptor like receptor (CLR), RAMP1 and receptor component protein (RCP) (see Poyner *et al.*, 2002).

The CGRP receptor is able to activate G_s proteins to increase cellular levels of cAMP, and to a lesser extent other G-proteins such as G_q have been suggested to be activated by the CGRP receptor (see Muff *et al.*, 2001).

Controversially, Kapas and Clark (1995) proposed that two related Rhodopsin-like receptors, referred to as RDC1 and L1/G10D, act as CGRP and adrenomedullin receptors. This hypothesis has been heavily criticised since the results have not been successfully replicated (see Kennedy *et al.*, 1998). Consequently, RDC1 and L1/G10D are no longer considered to be CGRP receptors (see Poyner *et al.*, 2002).

1.5.1 CLR

The CLR was discovered in 1993 (Njuki *et al.,* 1993, Fluhmann *et al.,* 1995). CLR is a typical Secretin-like GPCR. It has a relatively large N-terminus domain that contains a twenty-two amino acid signal peptide, six cysteines and three potential N-glycosylation sites. At present, little is known about CLR's tertiary structure. In spite of the lack of atomic detail speculative models of activation based on SDM and molecular modeling have been put forward that mirror Rhodopsin-like modes of activation. For example, Conner *et al.,* (2006) suggested that agonist binding would allow Pro-343 to act as a 'hinge' to allow the outward rigid body movement of TM6. In turn, this would allow ICL2 to swing away from the core of the receptor allowing G-protein coupling.

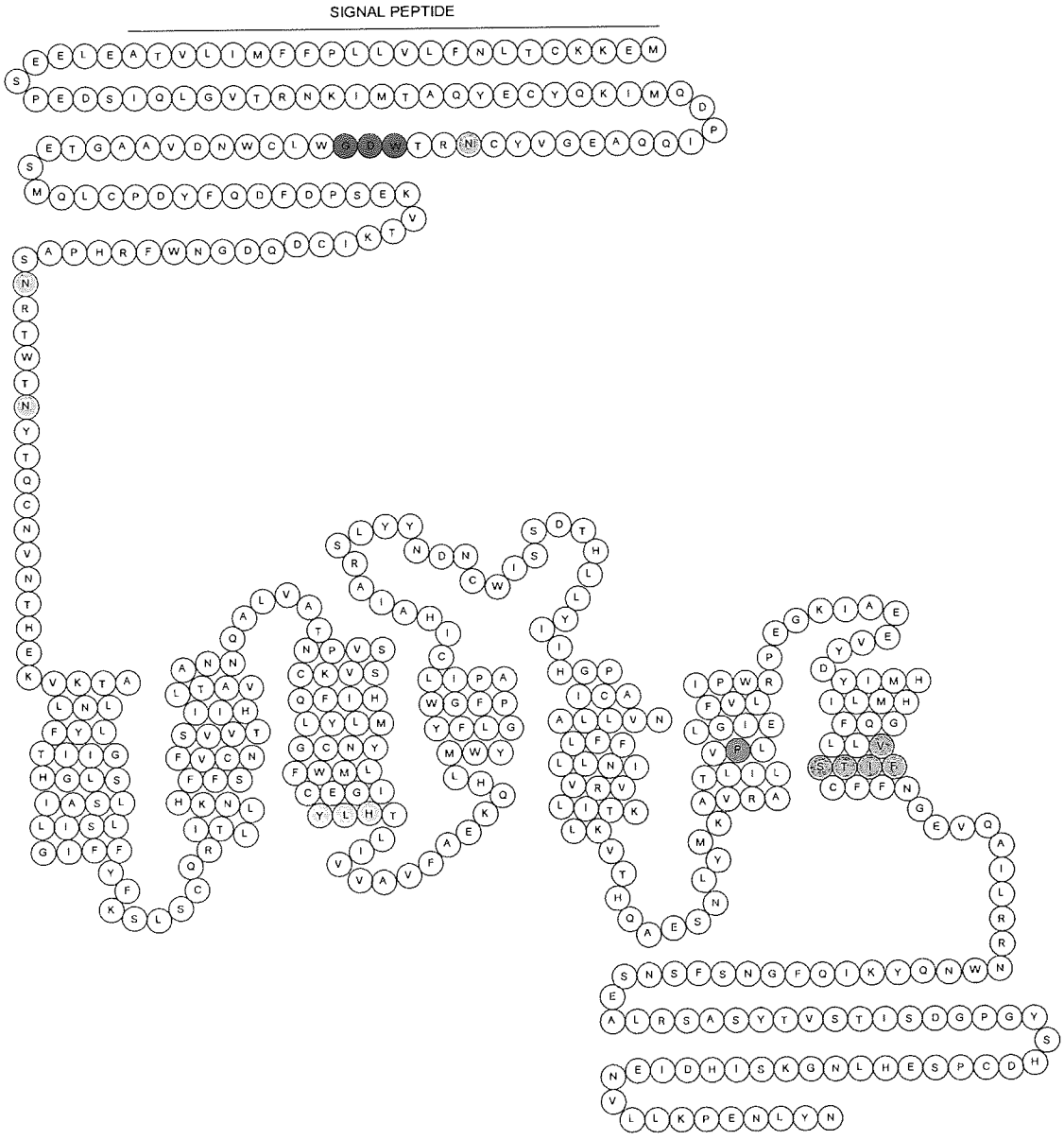


Figure 13. Snake diagram of the CLR representing the putative TM domains. The relative positions of the TM domains have been approximated from the Vohra *et al.*, alignment. The three putative N-glycosylation sites are highlighted yellow. The consenus Secretin –like WDN/G epitope within the ECD is highlighted red. The consenus Secretin-like YLH epitope located in TM3 is highlighted pale blue. Pro-343 in TM6 is highlighted green. The VAVLY consenus Secretin-like epitope located in TM7 is actually VSTIF in the CLR and is highlighted pink.

1.5.2 The RAMP family

1.5.2.1 The pharmacological role of the RAMPs

The International Union of Pharmacology (IUPHAR) guidelines refer to the CLR as a receptor but this is inaccurate as CLR is unable to bind to any known ligand independently (see Poyner *et al.*, 2002). McLatchie (1998) discovered a membrane protein family called the receptor activity modifying proteins (RAMPs), which consists of RAMP1, 2 and 3 in humans. The identification of the RAMP family was part of a revolution in GPCR research as it became evident that GPCRs were not merely monomeric units but could be part of a multifaceted signaling complex (see Hay *et al.*, 2006).

Initially, the RAMPs were known to be calcitonin receptor family accessory proteins. RAMP association with the calcitonin receptor (CTR) or with CLR generates six distinct receptor phenotypes with different specificities for CGRP, AM, intermedin, amylin and calcitonin (Bühlmann *et al.*, 1999, Leuthäuser *et al.*, 2000, Aldecoa *et al.*, 2000, Muff *et al.*, 1999).

The CTR does not require RAMP association to be functional but RAMP1, RAMP2 and RAMP3/CTR complexes result in amylin sensitive receptors (see Hay, *et al.*, 2006). These are the amylin 1, 2 and 3 receptors. RAMP1 association with CLR yields a CGRP receptor, while RAMP2/CLR complexes produce an AM receptor, the AM1 receptor (Poyner, *et al.*, 2002). In contrast, CLR/RAMP3 makes a promiscuous receptor, which responds to both AM and CGRP (see Conner *et al.*, 2004), although preference is for AM; this is referred to as the AM2 receptor.

RAMP association also occurs in other GPCR systems. Immunofluoresence confocal microscopy discovered that the RAMPs interact with at least four other Secretin-like receptors (Christopoulos *et al.*, 2003). The VPAC1 receptor interacts with all three RAMPs, while the glucagon receptor (GLR) and PTHR1 associate only with RAMP2 and the PTH2 receptor only interacts with RAMP3. Recently, it has also been found that RAMP3 interacts with the secretin receptor, an interaction mediated by TM6 (Harikumar *et al*, 2009). RAMP1 and RAMP3 have also been found to associate with a Glutamate-like GPCR, the CaSR (Bouschet *et al.*, 2005).

The functional importance of RAMP association appears to be different depending on the GPCR complex (Sexton *et al.*, 2009). McLatchie *et al.*, (1998) suggested that RAMP association governs cell surface expression of CLR. This chaperone role of RAMP1 and RAMP2 seems to be shared in the CaSR (Bouschet *et al.*, 2005). CLR/RAMP association occurs in the endoplasmic reticulum and continues into the Golgi apparatus to allow terminal glycosylation; the oligomeric receptor is then transported to the cell surface (Hilairet *et al.*, 2001b).

RAMP association clearly alters the pharmacological profile of CLR and CTR. Theoretically; the RAMPs may allosterically modify the conformation of CLR or CTR to reveal the binding site/s (Foord *et al.*, 1999). Equally, the RAMPs themselves may possess the necessary epitopes to allow ligand binding and may even mask certain binding sites to determine ligand selectivity (see Hilairet *et al.*, 2001a).

There is evidence that the RAMPs also influence receptor signaling. For example, Christopoulos *et al.*, (2003) noted that the VPAC1/RAMP2 complex increased agonist mediated phosphoinositide hydrolysis, implying that RAMP2 may improve G_q accessibility. Moreover, the ability of the Amylin receptors to stimulate a cAMP response was severely impaired when the C-terminal tail of the RAMPs was truncated (Udawela *et al.*, 2006). However, overexpression of G_s partially recovered the receptors ability to produce a cAMP response, suggesting that the C-terminal region of the RAMPs may play a role in efficient G-protein coupling. Importantly, RAMP C-terminal truncations did not impair the cAMP response in CLR based receptors in this study (Udawela *et al.*, 2006).

1.5.2.2 The structure of RAMP proteins

The human RAMP family share approximately 30% sequence homology (see Fitzsimmons *et al.*, 2003). The structure of all three RAMPs can be broken down into three domains; the N-terminus ECD, a single TM domain and a short intracellular cytoplasmic domain.

The ECDs of all three RAMPs have a signal peptide region and share four conserved cysteines, with RAMP1 and RAMP3 having an additional pair of cysteines. Yet, each

RAMP has its own pattern of N-glycosylation. RAMP1 does not contain any sites of glycosylation, whereas, RAMP2 has one putative N-terminal glycosylation site and, in contrast, RAMP3 contains potentially four sites of N- glycosylation.

In 2008, the crystal structure of the RAMP1 ECD at a 2.4Å resolution was published (Kusano *et al.*, 2008 –see figure 14). The crystal shows that the domain contains three conventional α helices and a 3_{10} helix located between α helix 1 and α helix 2. Importantly, Leu-39 located in α helix 1 is kinked, causing an irregular hydrogen bond network around this residue. The RAMP1 ECD is stabilised predominantly by core hydrophobic interactions and three conserved disulphide bonds.

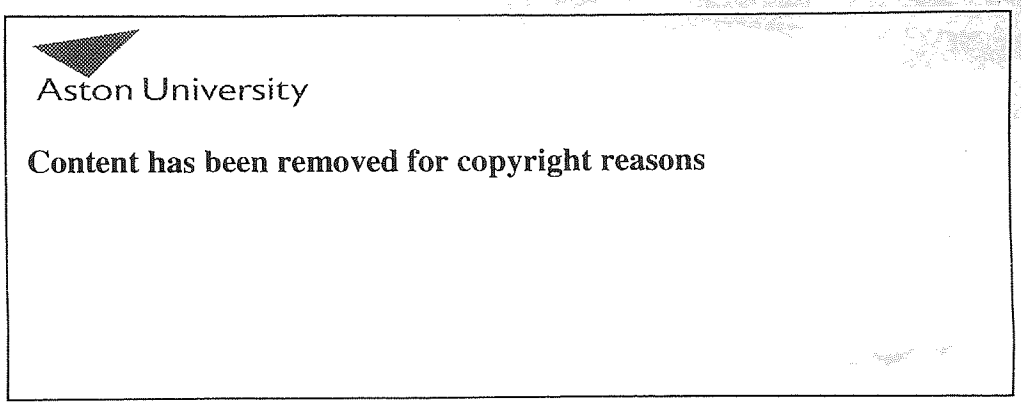


Figure 14. RAMP1 ECD crystal structure and a sequence alignment of the three human RAMPs. a) Taken from Kusano *et al.*, (2008) showing the tri-helical ECD of RAMP1 (red ribbon) and the three stabilising disulphide bonds (yellow sticks). PDB accession 2YX8. b) Taken from Conner *et al.*, (2004) identical residues located within each RAMP are highlighted in grey. Potential glycosylation sites are highlighted by **N**.

Solvent accessible surface analysis of the RAMP1 crystal highlighted that a hydrophobic concave area is formed between α-helix 2 and α-helix 3. Kuwasako *et al.*, 2003) mutated residues within this hydrophobic patch and found cell surface expression and αCGRP binding was impaired (Kuwasako *et al.*, 2003). Taking this into consideration, Kusano *et al.*, (2008) suggest that Phe-93, Leu-94, His-97 and Phe-101 could form the CLR interface.

Furthermore, Kusano *et al.*, (2008) suggested the identity of residues that may form a ligand-binding pocket in the CGRP receptor. High affinity binding of BIBN4096BS is facilitated by Trp-74 located on RAMP1 (Mallee *et al.*, 2002). It is predicted that the side chains of Arg-67, Asp-71, Glu-78, and Trp-84 face the solvent and as they are in close proximity to Trp-74 it has been assumed they could form a ligand-binding site.

Mutations at position 74 in RAMP1 and RAMP3 do alter adrenomedullin binding (Qi *et al.*, 2008 and Robinson *et al.*, 2009). However, the ligand-binding pocket identified may not necessarily be the orthosteric-binding site. For example, in the Amylin 1 receptor the antagonist effects of BIBN4096BS may be mediated through an allosteric binding site (Hay *et al.*, 2006).

The functional importance of the single TM domain of RAMP1 has been debated. Fitzsimmons *et al.,* (2003) found that the TM domain was required for full receptor sensitivity to CGRP but suggested that the ECD of RAMP1 alone could still produce similar Emax values, albeit a 4000 fold decrease in potency. Steiner *et al.,* (2002) have described the TM domain as essential for CLR association and found that sequential truncations of this domain abolished receptor trafficking. A detailed phylogenetic analysis that focused on residue functional divergence suggested that certain amino acids located in the TM domain of mammalian RAMPs are potentially important (Benítez-Paéz and Cárdenas-Brito, 2008). Furthermore, the residues highlighted appear to be located on the same face of the helix suggesting a potential CLR interface. Yet, the functional significance of the identified residues has not been experimentally tested.

Although the RAMP family only contain a C-termini tail ~9 amino acids in length, the short domain may play a significant role in receptor functioning. Steiner *et al.,* (2002) suggested that RAMP1 contained a novel endoplasmic reticulum retentive motif (QSKRT) that was overridden in the presence of CLR. Kuwasako *et al.,* (2006) found that the C-tail of RAMP2 was imperative for AM receptor cell surface expression, an observation that was not seen in RAMP1/CLR and RAMP3/CLR complexes. Taking into account the results of sequential truncations of RAMP2 and swapping the C-tail domain in RAMP2 and RAMP3, the authors suggest that the highly conserved SK motif is required for AM receptor cellular trafficking.

Interestingly, Kuwasako *et al.,* (2006) also suggest that the deletion of the C-tail in RAMP3 enhanced receptor internalisation. Bomberger *et al.,* (2005a) noted that agonist induced desensitisation was not observed in CLR/RAMP-3 when in the presence of the Na^+/H^+ exchanger regulatory factor-1 adaptor protein (NHERF). The authors suggested that through a chain of interactions mediated by PDZ (see abbreviations), the receptor was able to tether to the actin cytoskeleton thus preventing endocytosis. Moreover, the C-tail of RAMP3 is the most divergent as it contains a classical PDZ I motif (DTLL). The PDZ I motif influences the post-endocytic process of the receptor complex since the motif can associate with the hexameric ATPase, *N*-ethylmaleimide-sensitive factor, to promote rapid resensitation of the receptor (Bomberger *et al.,* 2005b).

1.5.3 RCP

RCP is an intracellular peripheral membrane protein, which is 148 amino acids in length. Using stable cell lines that express RCP antisense RNA, to diminish the presence of RCP, revealed that RCP did not affect expression or high affinity binding of CGRP and AM receptors but reduced signal transduction (Evans *et al.*, 2000). RCP has also been found to co-immunoprecipitate with CLR implying a direct interaction (Evans *et al.*, 2000). Consequently, it has been proposed that RCP acts as a dynamic regulator of G-protein coupling, adding another level of sophistication to the CGRP system (see Tolun *et al.*, 2006).

An attempt has been made to overexpress RCP in an *Escherichia coli* (*E.coli*) based expression system (Tolun *et al.*, 2006). However, the tertiary structure of this protein has not yet been elucidated. This could be due to the proteins tendency to aggregate (Tolun *et al.*, 2006).

1.6 The stoichiometry of the CGRP receptor

The stoichiometry of CGRP receptor is unclear. A cross-linking study suggested that the CGRP receptor was a 1:1 complex between RAMP1 and CLR (Hilairet *et al.*, 2001b). However, a novel bimolecular fluorescence complementation (BiFC) with bioluminescence resonance energy transfer (BRET) approach suggested the CGRP receptor contained two CLRs with only one RAMP1 (Heroux *et al.*, 2007). However, Sexton *et al.*, (2009) suggested that Heroux *et al.*, (2007) data did not rule out the possibility that individual RAMP1 monomers may individually bind to a CLR dimer.

The possibility that the CGRP receptor contains a CLR dimer is thought-provoking and in tune with the paradigm shift that Secretin-like GPCRs can form dimers. For example, co-immunprecipitation and fluorescence resonace energy transfer (FRET) analysis revealed that the rabbit CTR (C1a isoform) and the Δ exon 13 variant of the rabbit CTR could form both homo- and heterodimers (Seck *et al.*, 2003). Interestingly, cell surface expression of the rabbit CTR C1a isoform was inhibited by Δ exon 13 variant dimerisation.

Moreover, Secretin-like GPCR dimerisation does not appear to be limited to the CLR and CTR based receptors. Ding *et al.,* (2002) using a BRET approach revealed that wild-type secretin receptors can form homodimers but the heterodimerisation of wild-type secretin receptor with a mis-spliced Δexon 3 secretin receptor variant can inhibit normal receptor functioning in pancreatic cancer cell lines. Recent efforts have focused on discovering the molecular mechanisms that govern Secretin-like GPCR dimerisation. Lisenbee and Miller (2006) made partial and full N-terminal and C-terminal truncations of the secretin receptor and using BRET and FRET analysis concluded that it was the receptor core that was responsible for dimerisation. A TM segment peptide-competition assay, where synthetic peptides that resemble the native TM domains are mixed with the intact receptor in an attempt to compete for interacting domains and hinder dimerisation (Hebert *et al.,* 1996), was neatly applied to secretin receptor homodimerisation (Harikumar *et al.,* 2007). The study identified that the lipid exposed face of TM4 was essential for dimerisation and SDM revealed that Gly-243 and Ile-247 within TM4 were important molecular determinants (Harikumar *et al.,* 2007).

A follow up study carefully mapped out the dimer interface by systematically introducing fourteen cysteines in to TM4, in an attempt to promote covalent interactions between the secretin receptors when in the presence of cuprous phenanthroline (Gao *et al.,* 2009). Five cysteine mutations resulted in disulphide attachment as assessed by the peptide-competition assay. Moreover, the binding profile of the wild type dimeric secretin receptor was compared to the monomeric secretin receptor, caused by the double mutation G243A and I247A. Interestingly, the dimeric secretin receptor demonstrated a G-protein coupled high affinity binding state and exhibited negative binding cooperativity, unlike the monomeric receptor. This observation allowed the authors to further scrutinise the five cysteine mutations found to form disulphide bonds. It was found that secretin receptors that contained a disulphide bond at positions 243, 247 and 250 retained the high affinity state unlike the disulphide bonds at position 240 and 246. Consequently, this allowed the authors to reduce the rotational position of the corresponding TM4. This enabled the construction of a putative molecular model of the dimeric secretin receptor.

1.7 Aims and objectives

Currently, the architecture of the CGRP receptor remains elusive. Furthermore, current understanding of the orthosteric binding site of αCGRP is limited. The primary aim of this thesis is to identify functionally important residues within the extracellular face of the CGRP receptor. The extreme N-terminus of the CLR, ECL1 of the CLR and its associated TM regions, and finally ECL3 of the CLR and its juxtamembrane regions have been selected as regions of interest. A systematic SDM strategy is the method of choice to probe each region within the CLR. The extreme N-terminus of the CLR has been chosen because it may provide novel insight into high affinity C-terminal binding of αCGRP and may help outline the RAMP1 interface. The ECLs have been selected on the basis that these regions may be involved in receptor activation either directly by interacting with the N-terminal domain of αCGRP or by facilitating signal transduction. Another key objective is to generate speculative molecular models of the domains of the CLR. Triangulating the mutagenesis data with the hypothetical molecular models may provide informative details of important epitopes within the CGRP receptor. The final objective of this thesis is to determine whether the human RAMP family is amenable to *Pichia pastoris* (*P. pastoris*) expression and purification.

Chapter 2: General methods

2.1 Production and analysis of CLR site-directed mutations

2.1.1 Materials

Human αCGRP was from Calbiochem (Beeston, Nottingham, U.K). Peptides were dissolved in 1mM of acetic acid and stored as aliquots at –20°C in non-stick microcentrifuge tubes (Thermo Life Sciences, Basingstoke, U.K). Unless otherwise specified, chemicals were from Sigma-Aldrich U.K. [8-^{3}H] cAMP, NH_4 salt was purchased from Amersham Biosciences (Chalfont, U.K) and the specific radioactivity was 42Ci/mmol. The radioligand ^{125}I-hCGRP was from PerkinElmer Life and Analytical Sciences (Waltham, M.A.). The specific radioactivity of ^{125}I-hCGRP was 2200Ci/mmol and the site of iodination was His-10.

2.1.2 Expression constructs

Dr. S. M. Foord (Glaxo-Smith Kline, Stevenage, U.K) kindly provided both the human CLR cDNA that contained a T8 signal peptide and an N-terminal heamagglutinin (HA) epitope tag in the mammalian vector pcDNA3.1- (Invitrogen, Renfrew, U.K) and the human RAMP1 cDNA with a CD33 signal peptide and N-terminal *myc* epitope tag also incorporated in the pcDNA3.1- vector. Introduction of the novel signal peptides, epitope tags and linker regions did not affect the pharmacology of the receptor (McLatchie *et al.*, 1998 and Fraser *et al.*, 1999). These vectors were used for all subsequent mutagenesis experiments (see Figure 15 and 16).

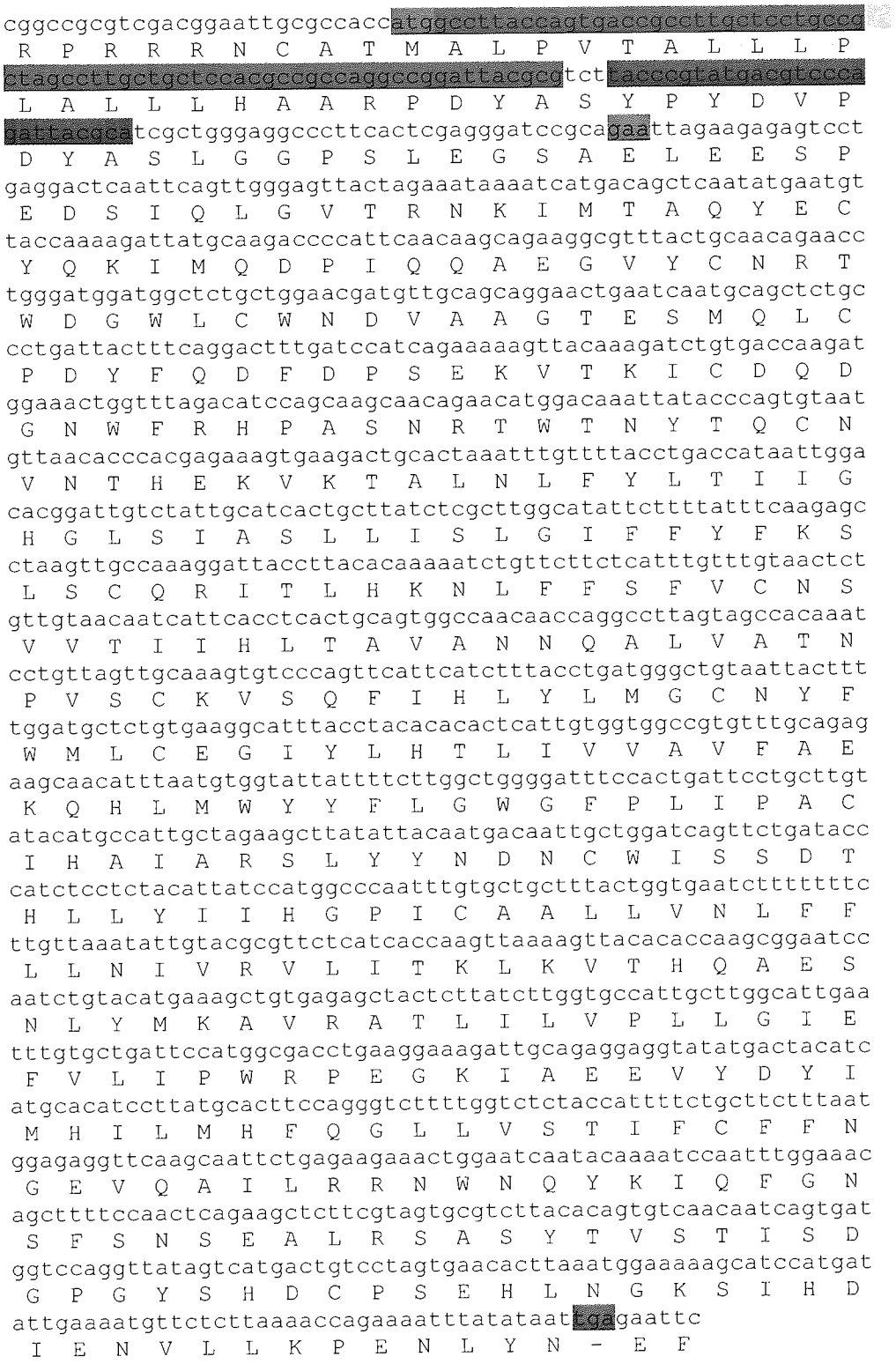


Figure 15. The cDNA sequence and translated sequence of T8-HA CLR.

The start codon (ATG) is highlighted in pink. The SignalP 3.0 neural network results (Bendtsen *et al.*, 2004) suggest that the cleavable signal peptide is 24 residues in length (highlighted in red). The signal peptide incorporates the T8 signal peptide sequence, which was taken from the T-cell surface glycoprotein CD8 alpha chain (Swiss-prot accession number: P01732). The HA epitope tag is highlighted in blue. The codon for the beginning of CLR is highlighted in cyan. The stop codon (TGA) is highlighted in green.

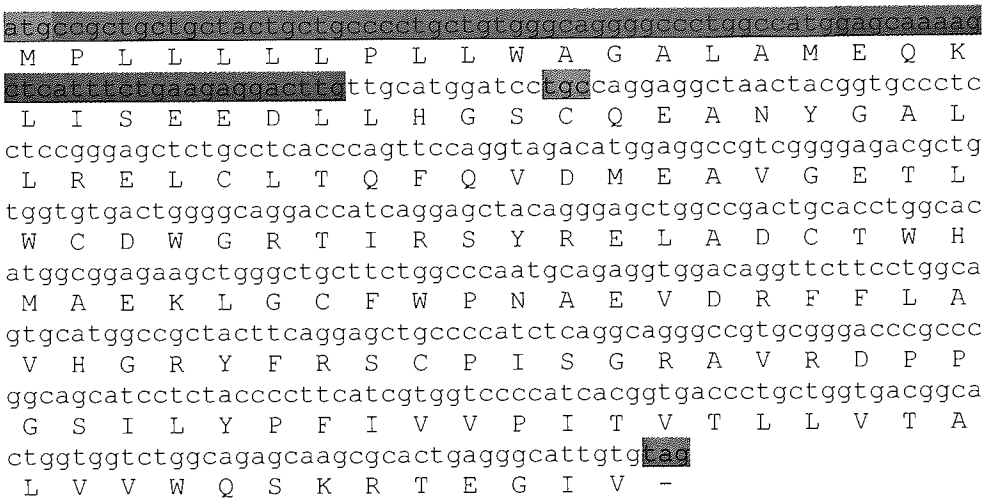


Figure 16. The cDNA sequence and translated sequence of CD33-*myc* RAMP1. The start codon (ATG) is highlighted in pink. The amino acid CD33 signal peptide sequence was taken from myeloid cell surface antigen CD33 (Swiss-prot accession number: P20138) and is highlighted in red. The *myc* epitope tag is highlighted in blue. The codon for the beginning of RAMP1 is highlighted in cyan. The stop codon (TAG) is highlighted in green.

2.1.3 Defining the regions of interest within the CLR

Three regions of the human CLR were investigated by means of site directed mutagenesis: the extreme N-terminus, ECL1 and ECL3. The extreme N-terminus of CLR was defined as residues E23-A60 (Ittner *et al.*, 2005). In an attempt to define the ECL regions a consensus prediction strategy was used to estimate the location of the transmembrane helices of CLR (see Cuthbertson *et al.*, 2005). Eleven different topology programs were examined. Including, SPLIT-4 (Juretic *et al.*, 2002), TMHMM2 (Krogh *et al.*, 2001), TMAP (Persson and Argos, 1997), HMMTOP-2 (Tusnády and Simon, 1998, 2001), TMPred (Hoffmann and Stoffel, 1993), ALOM2 (Nakai and Kanehisa, 1992), PHD (Rost *et al.*, 1996), TOPPRED2 (Claros and Von Heijne, 1994), DAS (Cserzo *et al.*, 1997), MEMSAT3 (Jones *et al.*, 2007) and SOSUI (Hirokawa *et al.*, 1998).

A majority vote procedure was used to determine the TM domains. However, taking into account that even the best topology predictors can have, on average, an error rate of two turns of a helix (Cuthbertson *et al.*, 2005), it was decided that residues H194 – M223 would undergo mutation to ensure the whole of ECL1 was incorporated into the analysis along with residues F349-M373 for the ECL3 region. This overestimation of the length of these regions ensured that the corresponding juxtamembrane regions

were also investigated. The amino acid numbering system is based on Swiss-Prot accession code Q16602.

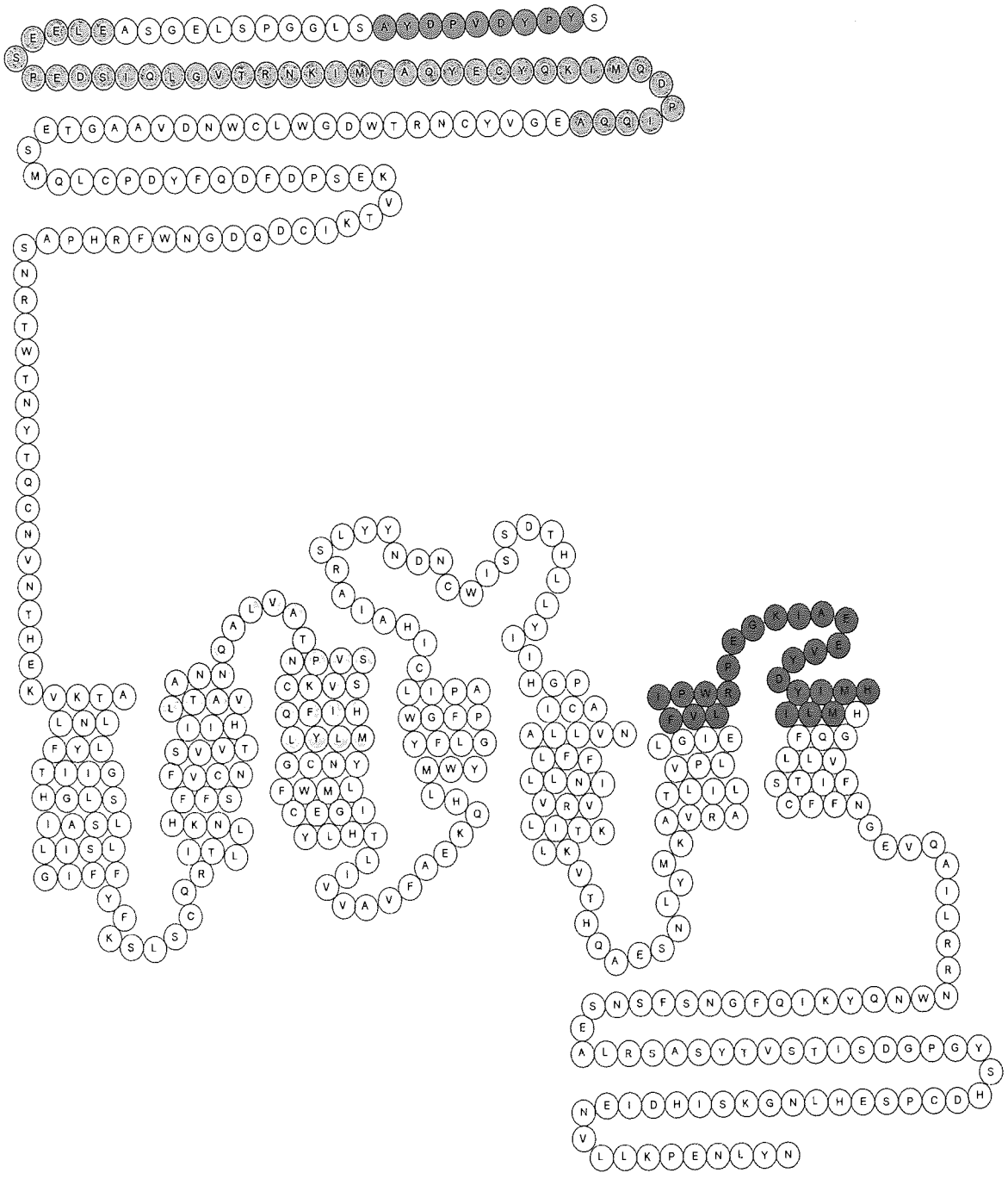


Figure 17. Residues selected for SDM on the mature HA CLR transcript. The relative positions of the TM domains have been approximated from the Vohra *et al.*, alignment. The residues of the HA epitope tag located on the the N-terminus of the mature CLR transcript have been highlighted green. The residues that make up the extreme N-terminus of CLR are highlighted pink. The ECL1 and its corresponding TM residues are highlighted pale blue whereas residues thar encompass ECL3 and its associated TM regions are highlighted red.

2.1.4 Site-directed mutagenesis

Mutagenesis was carried out using the QuikChange II site-directed mutagenesis kit ™ (Stratagene, Cambridge, U.K), following the manufacturers instructions. Forward and reverse oligonucleotide primers were designed with base changes to incorporate amino acid point mutations alanine or leucine to the final CLR constructs (see appendix for primers) and synthesised by Invitrogen. The plasmid of interest was extracted using the GenElute™ HP Plasmid Miniprep Kit, following the Spin Method. The plasmid DNA was eluted in 50µl sterile water and stored at -20°C. The entire gene of interest coding region was confirmed by the Functional Genomics Laboratory (Birmingham University, U.K). The four oligonucleotide primers used to sequence the HA CLR gene are found in the appendix (referred to as T7 primer, TM2 primer, TM4 primer and BGH primer). The plasmid DNA was then amplified to concentrations between 0.5 to 1mg/ml using the High Purity Plasmid Maxiprep System designed by Marligen Biosciences, Inc. The concentration and purity of maxiprep plasmid DNA was assessed by the NanoDrop 1000 (Thermo Scientific, U.K). Note, Dr A.Conner provided mutant constructs for C212A, H194A, H219A, Y221A, P353A, W354A, I360A and H370A.

a)

```
                                    ***
Forward: 5' GAGGTATATGACTACATCGCCCACATCCTTATGCACTTC 3'
Reverse: 5' GAAGTGCATAAGGATGTGGGCGATGTAGTCATATACCTC 3'
                                    ***
 GC content: 46.15%                 Location: 1156-1194
 Melting temp: 75.1°C               Mismatched bases: 3
 Length: 39 bp                      Mutation: Substitution
 5' flanking region: 18 bp          Forward primer MW: 11861.86 Da
 3' flanking region: 18 bp          Reverse primer MW: 12110.99 Da
```

b)

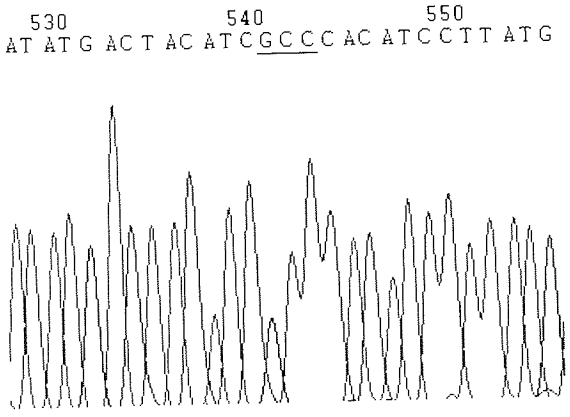


Figure 18. Confirmation of M369A mutant cDNA. a) The oligonucleotide primers used to generate M369A are shown with thier physico-chemical properties. The mutation within the primer is highlighted by *. b) Validation of the mutant confirmed by the Functional Genomics Laboratory (Birmingham University, U.K). The mutation within the cDNA gene is underlined.

2.1.5 Cell culture and transfection

COS-7 cells were cultured in Dulbecco's Modified Eagles Medium supplemented with L-glutamine, 10% (v/v) foetal bovine serum and 5% (v/v) penicillin/streptomycin in a humidified 95% air/5% CO_2 atmosphere. For transfection, the cells were plated onto either 24 or 48 well plates or 100mm dishes and grown to ~80% confluency.

Equal amounts of the HA CLR (either wild type [WT] or mutant) vector and *myc* RAMP1 vector were simultaneously transfected. Each well of a 48 well plate was treated with 1μg total DNA (e.g. 0.5μg HA CLR and 0.5μg *myc* RAMP1). 24 well plates were treated with 2μg DNA per well. 100mm dishes were transfected with 10μg of total DNA.

The transfection solution used to transfect 1μg of total DNA included 6μl of 10mM polyethyleneimine and 45μl of 5% glucose, which was allowed to incubate for 30 mins at room temperature. The transfection can be scaled up proportionally, that is the amount of total DNA can increase relative to transfection solution. Next, an appropriate amount of full media was added. A well of a 48 well plate contained a total volume (i.e.total amount of DNA, transfection solution and full media) of 500μl. A well of a 24 well plate also had a final total volume of 500μl and a 10mm dish had a final total volume of 9ml.Characterisation of expressed receptors was always performed 48hr after transfection.

2.1.6 Assessment of cAMP production

48 well plates were transiently transfected with WT condition (HA CLR/ *myc* RAMP1) alongside a mutant condition (HA mutant CLR/ *myc* RAMP1) in every experiment. Each assay point was assayed in duplicate and at least three independent experiments with independent transfections were conducted for each WT and mutant condition comparison. Growth medium was removed from the cells and replaced with 100μl of DMEM containing 500μM isobutyl methyl xanthine for 30 min. αCGRP in the range 1pM to 100nM was added for a further 10 min. The concentration of ligand varied over six orders of magnitude. The medium was then removed and replaced with ice-cold ethanol (95-100 % v/v). The ethanol was evaporated off using a fan. Next, 100μl of assay buffer (20mM HEPES and 5mM EDTA, pH =7.5) was added to each well. The 48 well plate/s were then allowed to shake for 5 min. 50μl of cell extract solution was then transferred to 1.5ml tubes. Next, 2μl (74kBq) of [8-^{3}H] cAMP, NH_4 salt was diluted in 4ml of assay buffer. Then, 50μl of the diluted tritiated cAMP was added to each 1.5ml tube. Next, 100μl of 0.02% w/v cAMP-dependent protein kinase (protein kinase A from bovine heart purchased from Sigma-Aldrich product code: P5511) in 1mM sodium citrate with 2mM dithiothreitol (pH 6.5) was added. The mixture was left to incubate at 4°C for 2 hr. 100μl of 5% w/v activated charcoal containing 0.2% w/v bovine serum albumin was added to each tube and left for 5 mins. The tubes underwent centrifugation at 10,000 rpm for 5 min using a bench top centrifuge. Next, 200μl of supernatant was transferred into a scintillation vial and 4ml of Optiphase 'Hisafe' 2 scintillant was added. The Packard 1600TR liquid scintillation analyzer was used to count each sample.

2.1.7 Enzyme-linked immunosorbant assay (ELISA) to determine cell-surface expression of CGRP receptor by probing for HA CLR

24 well plates were transiently transfected with WT condition (HA CLR/ *myc* RAMP1) alongside a mutant condition (HA mutant CLR/ *myc* RAMP1) in every experiment. A negative control of *myc* RAMP1/empty pcDNA3.1(-) was transiently transfected. The *myc* RAMP1/ empty pcDNA3.1(-) condition was used as the basal value for this system as it adequately controls for non-specific antibody binding as the cells in this negative control were exposed to a similar transfection procedure. At least three independent experiments with three replicates per experiment were conducted for each WT and mutant condition comparison. The transfected cells were treated with 3.7% formaldehyde for 15 min after aspiration of growth medium. The cells were then washed three times with 0.5ml of PBS. Non-specific binding of the antibody was blocked with 2% BSA in PBS for 45 min. The cells were treated with 250µl of primary antibody (mouse, anti-HA antibody H9658 [Sigma-Aldrich] diluted 1:2000 in PBS with 1% BSA) for 1hr and the cells were washed again three times with 0.5ml PBS. A further block step was performed for 15 min before the cells were incubated with 250µl secondary antibody (anti-mouse, horseradish peroxidase conjugated #7076 [Cell Signaling Technology] diluted 1:2000 in PBS) for 1hr. The cells were washed a further three times before development with *SIGMAFAST* ™ o-phenylenediamine tablets according to the manufacturer's instructions. Reactions were terminated with 100µl/well of 1M H_2SO_4. The Biotek EL800 Universial Microplate reader using the 490nm filter was used to quantitfy the peroxidise product.

2.1.8 CGRP receptor surface expression after agonist dependent internalisation

The ELISA procedure outlined above was conducted but wells were treated with 100nM human αCGRP at 37°C for 1hr in full medium.

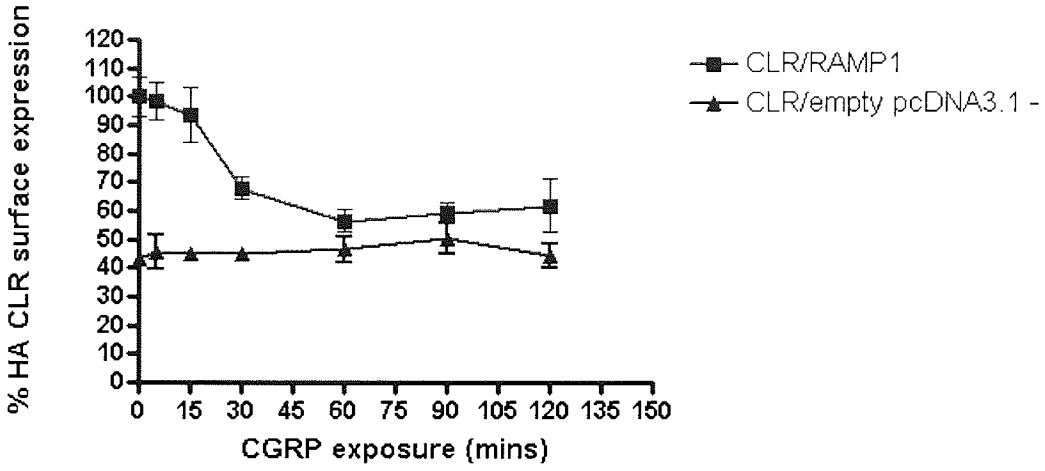


Figure 19. Time course of agonist dependent internalisation. COS-7 cells were co-transfected with HA CLR/myc RAMP1 alongside a negative control (*myc* RAMP1/empty pcDNA3.1-). Two cell surface ELISAs probing for the CLR N-terminal HA epitope tag were performed. Each assay point was assessed in triplicate with each point in the graph representing the mean ± SEM of n=6. HA CLR cell surface expression was normalised so 0% equalled the mean *myc* RAMP1/empty pcDNA3.1- and 100% equalled the HA CLR/*myc* RAMP1 cell surface expression in the absence of αCGRP stimulation. Time points of αCGRP exposure include 5, 15, 30, 60, 90 and 120 mins.

2.1.9 Total CLR expression

Total CLR receptor expression was assessed on mutant receptors that were found to alter cell surface expression. The ELISA procedure outlined above was conducted but after the transfected cells were fixed with 3.7% formaldehyde for 15 min the cells were permeabilised with 0.1% Triton-X 100 in PBS for 1hr. Cell surface expression of the selected mutants was again assessed in these experiments. The cell surface expression of mutant receptors in relation to WT in these experiments was comparable to initial findings.

2.1.10 Crude membrane preparation

Transfected cells were grown in 100mm tissue culture dishes (Orange Scientific, Belgium). Each tissue culture dish was washed with 2ml ice cold PBS. Then 1ml of ice cold homogenization buffer (20mM HEPES, 1mM EGTA, 10mM $MgCl_2$ and SIGMAFAST^{TM} Protease Inhibitor cocktail tablet EDTA-free, pH 7.5) was added to

the tissue culture dishes. The cells were scraped and put into a 50ml falcon tube on ice. The cells were then exposed to three 15 sec bursts of an Ultra-Turrax T25 tissue homogenizer at the maximum setting, placing the suspension on ice for 45 sec between bursts. The homogenate was centrifuged at 20,000 × *g* for 20 min, 4°C. The pellet was suspended in 1ml of binding buffer (20mM HEPES, 2mM $MgCl_2$, pH 7.5). Protein concentration was determined by the NanoDrop 1000 (Thermo Scientific).

2.1.11 αCGRP inhibition binding assay

Firstly, 10μl of unlabelled αCGRP was added to 1.5ml non-stick microcentrifuge tubes. The concentration of unlabelled αCGRP ranged from 1μM to 1pM in a final 100μl volume. Next, 10μl containing ~20pM of radioactive-labelled CGRP (~50,000 cpm) is added to each tube. Total radioactive binding was assessed by not adding unlabelled αCGRP. Finally, 80μl of membrane preparation is added to each tube containing 200μg total protein. The reaction was incubated for 1 hour at 24°C. The incubation was terminated by microcentrifuging the tubes for 5 min at 14,000 x *g*, 4 °C. The supernatant was discarded from each tube. Next, each tube was washed twice with distilled water, taking care not to dislodge the pellet. The tubes were then placed into the Packard Combra Auto-Gamma counter and analysed for 5 min per tube.

2.1.12 Data analysis

GraphPad Prism 4 (Graphpad Software Inc., San Diego, CA) was used to generate non-linear regression sigmoidal concentration-response curves to fit raw cAMP production data (see Equation 1).

$$Y = \text{Bottom} + \frac{(\text{Top} - \text{Bottom})}{(1 + 10\,\hat{}\,((\text{LogEC50} - X)))}$$

Equation 1. Sigmoidal dose-response equation used to fit cAMP data. Bottom refers to the lowest plateau and Top to the highest value plateau. X is the logarithm of concentration. Y is the response which starts at the bottom and goes to the top following a sigmoidal shape. Hill slope assumed to be of unity.

The cAMP dose-response curves were then normalised from 0% to 100% based on WT Top and Bottom values generated by GraphPad Prism 4. As a WT dose-response curve was always performed alongside a mutant dose-response curve a two-tailed unpaired *t* test was used to compare WT and mutant pEC50 values from each independent experiment, which was replicated a minimum of three times. The Emax and basal activity of each mutant was assessed by the mean of the top and bottom of each dose-response curve. The mutant receptors mean Emax and basal activity was expressed as a percentage that corressponded to WT normalisation. A meaningful difference was noted if the mean size of effect differed by 20% or more.

Cell-surface expression of the CGRP receptor was approximated by probing for HA CLR (in both WT and mutant receptor conditions) when co-expressed with *myc* RAMP1. The raw data for each independent experiment was normalised to WT cell-surface expression where the mean negative control (*myc* RAMP1/ empty pcDNA3.1-) was equal to 0% and the mean WT expression was equal to 100%. Next, each normalised cell-surface percentage of WT receptor expression from each independent experiment was compared to the mutant receptors cell-surface expression percentages, which were relative to the WT normalisation, using a Mann Whitney U test.

Agonist mediated internalisation of the CGRP receptor (both WT and mutant receptors) was approximated by a cell surface ELISA taking into account the difference in cell surface expression levels between CGRP receptors that have or have not been exposed to 100nM of human αCGRP for an 1hr. WT receptor cell-surface expression values (in the absence and presence of αCGRP) were normalised to the mean WT cell-surface expression in the absence of αCGRP (to equal 100%) and

negative control (*myc* RAMP1/empty pcDNA3.1-) that equalled 0%. WT receptor agonist mediated internalisation was determined by subtracting the percent mean WT cell-surface expression (i.e.100%) from all the normalised percent values for WT receptor cell-surface expression after αCGRP treatment. Next, mutant receptor cell-surface expression values (in the absence and presence of αCGRP) were normalised to the mean mutant receptor cell-surface expression (equal to 100%) and negative control (equal to 0%). Next, the mutant receptor agonist mediated internalisation was determined by subtracting the percent mean mutant cell-surface expression (i.e.100%) from the values for mutant receptor cell surface expression after αCGRP treatment. Finally, WT agonist mediated internalisation values were compared to the mutant agonist mediated internalisation values using a Mann Whitney U test.

The raw data of total expression of WT and mutant CGRP receptors was normalised in each independent experiment so that the mean WT total expression equalled a 100% and the the mean negative control (*myc* RAMP1/ empty pcDNA3.1- treated with Triton-X 100) was equal to 0%. A Mann Whitney U test was then used to compare WT and mutant percent total expression.

The GraphPad Prism 4 equation referred to as the one-site competition non-linear regression curve was used to fit inhibition αCGRP binding data. The one-site competition equation is the same as Equation 1 stated above. Next, a two-tailed unpaired *t* test was used to compare WT and mutant pIC_{50} values.

2.2 Method for RAMP purification using *P. pastoris*

2.2.1. Equipment and reagents

Uniplate 24-deep well plates (Whatman, U.K)
Bugstopper siliconized rubber cap with Erlenmeyer vent (Whatman, U.K)
20ml sterile universal (Sarstedt, U.K)
Protran Nitrocellulose transfer membrane (Whatman, U.K)
Tween-20 (Sigma-Aldrich, U.K)
Marvel (Premier Foods Ltd, U.K)
6xHis Monoclonal Antibody; Albumin Free (Clontech, U.K)
Goat Anti-mouse IgG (Fab Specific) peroxidase conjugate (Sigma-Aldrich, U.K)

EZ-ECL chemiluminescence solution (Geneflow, U.K)
EmulsiFlex -C3 (Avestin, U.K)
Homogenizer PTFE Pestle/S.S Rod 15ml (Scientific Laboratory Supplies, U.K)

2.2.2. Media and stock solutions

2.2.2.1 10x Yeast nitrogen base (YNB)

134.0g of yeast nitrogen base was dissolved in double distilled (dd) H_2O to a total volume of 1L and filter sterilised. The medium was stored at 4°C.

2.2.2.2 500x Biotin (0.02%)

20mg biotin was dissolved in ddH_2O to a total of 100mL and filter sterilised. It was then stored at 4°C.

2.2.2.3 10x Glycerol (10%)

100ml of glycerol was mixed with 900ml of ddH_2O. It was filter sterilised and stored at room temperature.

2.2.2.4 10x Methanol (5%)

5ml methanol was mixed with 95ml ddH_2O and filter sterilised. The media was then stored at 4°C.

2.2.2.5 10x Glucose (20%)

200.0g glucose was dissolved in water to a total of 1L ddH_2O and autoclaved at 121°C for 20 min then cooled to room temperature. The solution was stored at 4°C.

2.2.2.6 1M Potassium phosphate buffer pH 6.0

A 1M solution of K_2HPO_4 was made by dissolving 174.2g in ddH_2O to a total volume of 1L. A 1M solution of KH_2PO_4 was made by dissolving 136.1g in water to a total

volume of 1L. Next 132ml of 1M K_2HPO_4 was mixed with 868ml KH_2PO_4 and the pH set to 6.0 using a pH meter and phosphoric acid. The solution was autoclaved and stored at room temperature.

2.2.2.7 Buffered complex glycerol/methanol media (BMGY/BMMY)

Dissolve 7.0g yeast extract, 14.0g peptone in 490ml ddH_2O in 1000ml bottle. Autoclave and allow to cool to room temperature. Add 70ml 1M potassium phosphate buffer pH 6.0, 70ml 10x YNB stock, 1.4ml 500x Biotin. For BMGY media, add 70ml 10x glycerol or for BMMY media, add 70ml 10x methanol to give a final volume of 700ml media.

2.2.2.8 Yeast peptone dextrose (YPD)

20.0g peptone, and 10g yeast extract were dissolved in ddH_2O to a total volume of 900ml. For agar plates 20.0g agar was added. The solution was autoclaved at 121°C for 20 min then cooled to room temperature before adding 100ml 10x glucose and stored at 4°C.

2.2.2.9 Extract peptone dextrose medium with sorbitol (YPDS + Zeocin)

10.0g yeast extract, 182.2g sorbitol and 20.0g of peptone were dissolved in 900ml of ddH_2O. 20.0g of agar was then added. The solution was autoclaved at 121°C for 20 min. Next, 100ml of 20% dextrose (filter-sterilise dextrose before use) was added. Cool solution to ~60°C and add the appropriate amount of Zeocin from 100mg/ml stock solution. The plates containing Zeocin were stored at + 4°C.

2.2.2.10 Laemmli sample buffer

The SDS sample buffer used was taken from Laemmli *et al.*, (1970). 4x Laemmli Buffer consists of 2.4ml 1M Tris pH 6.8, 0.8 g SDS stock, 4ml 100% glycerol, 0.01% bromophenol blue, 1ml β-mercaptoethanol and 2.8ml ddH_2O.

2.2.2.11 Fermentation basal salts medium

0.93g calcium sulphate, 18.2g potassium sulphate, 14.9g Magnesium sulfate-7H_2O, 4.13g potassium hydroxide, 40.0g Glycerol and 26.7ml of 85% phosphoric acid were dissolved in ddH_2O to make 1L.

2.2.2.12 PTM_1 trace salts

6.0g cupric sulphate-5H_2O, 0.08g sodium iodide, 3.0g manganese sulfate-H_2O, 0.2g sodium molybdate-2H_2O, 0.02g boric acid, 0.5g cobalt chloride, 20.0g zinc chloride, 65.0g ferrous sulfate-7H_2O, 0.2 g Biotin, 5.0 ml sulphuric acid were dissolved in a final volume of 1L ddH_2O.

2.2.3 Protein identification

2.2.3.1. Sodium dodecyl sulphate polyacrylamide gel electrophoresis gels

Sodium dodecyl sulphate (SDS) polyacrylamide gel electrophoresis (PAGE) gels contained 12% polyacrylamide in the separating gel and 4% polyacrylamide in stacking gel (see Table 2 and 3)

Table 2. Method for making separating gels. APS was made fresh each time. TEMED was always added last. Isopropanol was used to flatten gel and then washed off with ddH_2O . The gel was then dried with filter paper.

12 % Separating Gel	Per gel
Polyacrylamide 30% Water Tris-HCl 1.5M, pH 8.8 SDS 10% Ammonium persulphate (APS) 20% TEMED	2.3ml 1.8ml 1.5ml 60µl 20µl 4.5µl

Table 3. Method for making stacking gels. APS was made fresh each time. TEMED was always added last. The comb was then added and removed when gel had set.

4% Stacking Gel	Per gel
Polyacrylamide 30%	0.3ml
Water	1.5ml
Tris-HCl 0.5M, pH 6.8	0.6ml
SDS 10%	25µl
Ammonium persulphate (APS) 20%	10µl
TEMED	2.5µl

2.2.3.2 SDS PAGE

5µl 4x Laemmli sample buffer was added to 15µl of protein sample/s. The samples were heated to 70°C for 10 min and then spun briefly. The samples were loaded on to the SDS PAGE gels in the presence of 1x running buffer (3.03g Tris-base, 18.8g glycine and 1g SDS, made up to 1L with ddH_2O). Then using the Bio-Rad PROTEAN 3 cell the SDS PAGE was run at 200 volts until the dye front reached the bottom of the gel.

2.2.3.3 Western blotting

After SDS PAGE the stacking gel was removed from the separating gel. The stacking gel was then discarded. The separating gel was then placed into transfer buffer (3.03g Tris-base, 14.4g glycine and 200ml of methanol, made up to 1L with ddH_2O). Next the Whatman PROTRAN nitrocellulose transfer membrane was cut to approximately 6 x 9 cm and the Whatman 3mm chromatography (filter) paper was cut into 6 rectangles approximately 8 x 10 cm. The gel, the membrane, the filter paper and the fibre pads were left to equilibrate in the transfer buffer for 30 mins. Next, using the Bio-Rad colour coded easy lock cassette a transfer sandwich was made (see Figure 20). Briefly, the Bio-Rad colour coded cassette was placed with the black-side down. Then a fiber pad was placed on the black side followed by 3 filter paper sheets. Then the gel was placed on the filter paper followed by the nitrocellulose membrane. Next, an additional 3 sheets of filter paper was added and a glass tube was gently rolled over the sandwich to remove air bubbles. The final fiber pad was added and the cassette was locked with the provided white latch. The cassette was inserted into the colour-

coded electrophoretic blotting cell. The electrophoretic blotting cell along with a Bio-ice cooling unit was added to the Bio-Rad PROTEAN 3 cell. The cell was then filled with transfer buffer. The transfer was conducted at 100V for 1hr.

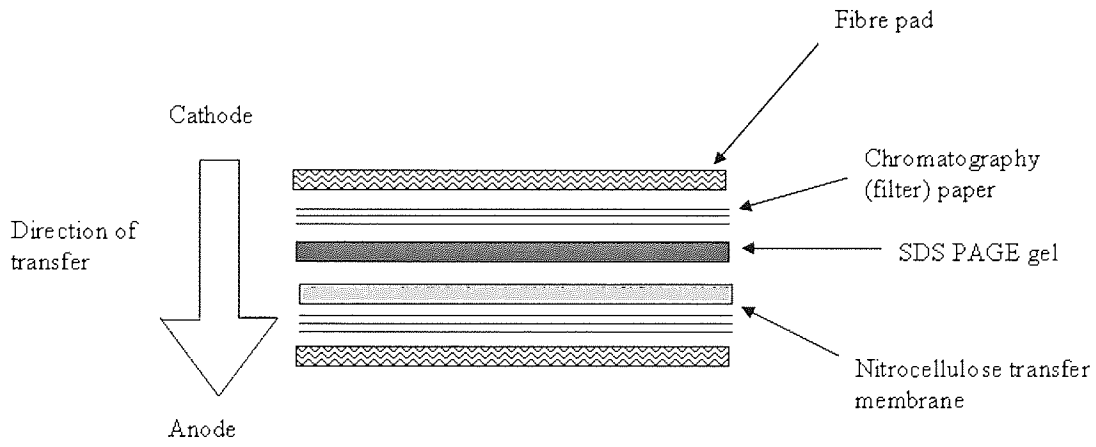


Figure 20. A schematic representation of the Western blot transfer sandwich. Each layer of the Western blot transfer sandwich has been depicted in the order they were assembled. The direction of the electrophoretic transfer is also stated. The SDS PAGE gel is highlighted light blue and the nitrocellulose transfer membrane is highlighted yellow.

After transfer, the nitrocellulose membrane underwent a blocking step where the membrane was exposed to PBS containing 5% Marvel for 1 hr at room temperature. Next, the primary antibody was added to the blocking buffer at a dilution of 1:5000 for 1 hr at room temperature while undergoing moderate rocking. Then the membrane was washed in PBS + 0.2% Tween-20, twice for 5 mins. The secondary antibody was then added with the blocking buffer at a dilution of 1:5000 for 1 hr at room temperature, which was rocked moderately. Again, the membrane was washed in PBS + 0.2% Tween-20, twice for 5 mins. EZ-ECL chemiluminescence solution was used in accordance with the manufacturerss instructions to detect protein. Molecular mass markers used were the Protometric ladder (National diagnostic Ltd), the MBI fermentas plus pageRuler and the Geneflow wide range protomarker.

2.2.3.4 Coomassie Brilliant Blue R-250

The stain solution is 0.25g of Coomassie Brilliant Blue R-250 was dissolved in 100ml of methanol:acetic acid solution (500ml methanol, 400ml of ddH_2O and 100ml acetic acid). The SDS-polyacrylamide gel was immersed in stain solution for a minimum of 4 hr. The stain solution was then discarded. The methanol:acetic acid solution was then used to de-stain gels.

2.2.3.5 Silver stain

The Silver Stain Plus kit (BioRad, U.K) was used for SDS-polyacrylamide gel staining.

2.2.4 Molecular biology

Expression clones of human RAMP1, RAMP2 and RAMP3 in pcDNA3.1(-) were subcloned into pPICZB vectors. The putative signal peptide was removed (residues 1-26 for RAMP1, residues 1-35 for RAMP2, and residues 1-27 for RAMP3, respectively). Sense primers contain *EcoRI* restriction site and yeast initiation consensus sequence (Nyblom *et al.*, 2007). Antisense primers contain *XhoI* restriction site. Restriction site selection was based on NEBcutter v2.0 results (Vincze *et al.*, 2003) and the multiple cloning site of pPICZB. For identification and purification purposes a C-terminal hexa histidine tag was also incorporated.

Table 4. RAMP oligonucleotide primers. The restriction sites have been underlined. The yeast consensus sequence has been highlighted bold.

RAMP1	Sense: 5'-GGGGGGAATTC**AAAATG**TCTTGCCAGGAGGCTAACTAC-'3 Antisense: 5'-GGGGGCTCGAGTTAGTGATGGTGATGGTGATGCACAATGCCCTCAGTGCG-'3
RAMP2	Sense: 5'-GGGGGGAATTC**AAAATG**TCTAATCCCCACGAGGCCCTG-'3 Antisense: 5'-GGGGGCTCGAGTTAGTGATGGTGATGGTGATGGGCCTGGGCCTCACTGTC-'3
RAMP3	Sense: 5'-GGGGGGAATTC**AAAATG**TCTTGCAACGAGACAGGCCTG-'3 Antisense: 5'-GGGGGCTCGAGTTAGTGATGGTGATGGTGATGCAGCAGCGTGTCGGTGCG-'3

A temperature gradient polymerase chain reaction (PCR) was used to generate RAMP PCR products (see Table 5 for PCR reagents and PCR program). The QIAquick Gel Extraction Kit (Qiagen, U.K) was used to extract PCR products. The PCR products for each construct were then pooled together and extracted again with the QIAquick Gel Extraction Kit.

Table 5. RAMP PCR protocol. This procedure was found to be successful at all gradient annealing temperatures in RAMP1, RAMP2 and RAMP3 PCRs.

PCR reagents	PCR program	Annealing gradient temperatures
15µl *pfu* buffer 100pmol sense primer 100pmol antisense primer 3µl dNTP (10 mM) 4ng of template plasmid 2µl *pfu* DNA polymerase 125.2µl ddH_2O Total volume = 150µl Add 10µl to 0.2ml sterile PCR tubes	1 = 95°C for 2 min 2 = 95°C for 1 min 3 = 45°C to 65°C for 1 min 4 = 72°C for 3 min 5 = Repeat step 2, 3 and 4 (35 times) 6 = 72°C for 4 min 7 = 4°C until collection	1 = 45.0°C 2 = 45.5°C 3 = 46.5°C 4 = 48.2°C 5 = 50.5°C 6 = 53.4°C 7 = 56.7°C 8 = 59.6°C 9 = 61.8°C 10 = 63.4°C 11 = 64.6°C 12 = 65.0°C

The pooled PCR product and vector (pPICZB) underwent a double digestion with *EcoRI* and *XhoI*, which was incubated at 37°C for 2 hr. The reaction was purified using the Qiagen MinElute PCR purification kit, following the manufacturer's guidelines. To prevent unnecessary complications in ligation the digested vector was

treated with Antarctic Phosphatase (following manufacturer's guidelines), which catalyzes the removal of 5′ phosphate groups from DNA preventing self-ligation (Sambrook, *et al,* 1989). Next, the insert was ligated to the vector using T4 ligase (following manufacturer's guidelines) and was incubated at 12°C for 20 hr followed by 65°C for 20 min to inactivate the ligase. The ligated plasmid was then transfomed into XL-1 sub-cloning grade competent cells (Stratagene, U.K) using heat shock method (42°C for 45 sec) and allowing cells to recover in Luria-Bertini (LB) medium for 1 hr. Then the cells were plated onto Zeocin/LB agar plates overnight. The plasmid of interest was extracted using the GenElute^{TM} HP Plasmid Miniprep Kit (Sigma-Aldrich), following the Spin Method. The plasmid DNA was eluted in 50µl sterile water and stored at -20°C. Plasmid DNA sequences were confirmed by the Functional Genomics Lab (Birmingham University, UK). The plasmid DNA was then amplified to concentrations between 0.5 to 1mg/ml using the Marligen High Purity Plasmid Maxiprep System.

2.2.5 *P. pastoris* transformation

2.2.5.1 Competent *P. pastoris* cells

5ml of *P. pastoris* cells in YPD were incubated at 30°C and agitated at 250 rpm overnight. The overnight culture was diluted to an optical density $(OD)_{600}$ of 0.15 – 0.20 in a volume of 50ml YPD. The yeast culture was then grown at 30°C with 250 rpm agitation until an OD_{600} of 0.8 – 1.0 was achieved. The culture was then centrifuged (500 x *g* for 5 min at room temperature) and the supernatant was discarded. The pellet was then resuspended in 9ml of ice-cold BEDS solution (BEDS solution was composed of 10mM bicine-NaOH, pH 8.3, 3% (v/v) ethylene glycol, 5% (v/v) dimethyl sulfoxide (DMSO), and 1M sorbitol) supplemented with 1ml 1.0M dithiothreitol. Cells were incubated for 5 min at 100 rpm in a 30°C shaking incubator. The cells were then centrifuged at 500 x *g* for 5 min at room temperature and then the cells were resuspended in 250µl (0.005 volumes) of BEDS solution without DTT and aliquoted into 40µl volumes. Aliquots were stored at -80°C for up to 6 months.

2.2.5.2 *P. pastoris* electroporation

4μg of *PmeI* linearized plasmid was added to a 40μl aliquot of competent cells in an electroporation cuvette and incubated for 2 min on ice. The samples were electroporated using the Eppendorf multiporator™ (Hamburg, Germany) 1700 V, 15 mS pulse length. Immediately after electroporation, the samples were resuspended in 0.5ml 1.0M sorbitol and 0.5ml YPD and incubated for 1 hr, 30°C. Next, the electroporated cells were plated on YPDS agar containing concentrations of Zeocin 250μg/ml for the selection of integrants. Following yeast recovery, the transformed cells were plated onto YPDS agar containing 250μg Zeocin. Plates were incubated at 30°C for 3-4 days or until colonies were distinguishable. Colonies were selected and re-plated on YPDS agar with 250μg/ml of Zeocin and incubated for a following 3-4 days at 30 °C.

2.2.6 Screening for RAMP expression colonies

10 *Pichia pastoris* clones were selected for expression for each RAMP construct. A single colony of *Pichia pastoris* was incubated overnight in 3ml BMGY media at 30°C at 220 rpm. The OD_{600} was determined in the morning. Next, using the 24-deep-well plate 3ml BMMY was added with the appropriate volume of the BMGY culture to achieve an OD_{600} of 1. Cells were left to incubate for approximately 53 hours at 30°C while shaking at 220 rpm (methanol was added to make the total concentration equal to 1% at 24 and 48 hours post induction). Then, 100μl of each sample was harvested after 53 hours and centrifuged at 10,000 *x g* for 5 min. The cell pellets were resuspended in 60μl of SDS-PAGE sample buffer. The samples were then heated for 10 min at 98°C, and spun briefly for 1 min at 13,000 rpm. Samples were then used for Western blot analysis.

2.2.7 Fermentation

The below procedure adheres to the Invitrogen guidelines for *Pichia pastoris* fermentation although taking into consideration that only 2L Applikon bioreactors were used.

2.2.7.1 Inoculum seed flask preparation

Baffled flasks containing 50ml of BMGY were inoculated with a 10ml overnight culture of a colony from the YPDS/Zeocin plates or from a frozen glycerol stock. Cultures were grown at 30°C (250-300 rpm) for 24 hr. 2L bioreactors were inoculated to ~1 OD_{600}.

2.2.7.2. Glycerol batch phase

The 2L bioreactor was sterilised with 1L fermentation basal salts medium containing 4% glycerol. Next, the temperature was set to 30°C, agitation and aeration was applied to the system. The pH of the fermentation basal salts medium was adjusted to 5.0 with 28% ammonium hydroxide. 4.35ml filter sterilised PTM_1 trace salts was added to the basal salts medium. The bioreactor was inoculated from the culture generated in the inoculum shake flasks. Dissolved oxygen was maintained above 20% by adding oxygen when needed. The batch culture was grown until the glycerol was completely consumed, which was approximated by the dissolved oxygen level.

2.2.7.3. Glycerol fed-batch phase

50% w/v glycerol feed containing 12ml PTM_1 trace salts was added at a feed rate of 18.5 ml/hr for ~4 hours.

2.2.7.4 Methanol fed-batch phase

A 100% methanol feed containing 12ml PTM_1 trace salts was added with an initial feed rate to 1.92 ml/hr. The feed rate was then increased to 3.96ml/hr/liter and after 2 hours the methanol feed rate was increased to 7.62ml/hr. This feed rate was maintained throughout the remainder of the fermentation, which was ~25 hr.

2.2.8 Preparation of yeast membrane

The cell pellet was resuspended in ice-cold breaking buffer (50mM $NaPO_4$, pH 7.4, 5% glycerol, 2mM EDTA, 100mM NaCl) at a ratio of 2:1 to pellet (v/w) and yeast protease inhibitors were added. The cells were passed through an Emulsiflex-C3 cell

disrupter four times. The unbroken cells and cell debris was removed by centrifugation (10,000 x *g*, for 30 min, at 4°C). The supernatant was collected and ultracentrifuged (100,000 x *g*, for 90 min, at 4°C). The supernatant was discarded and the membrane pellet was re-suspended in ice-cold 20mM HEPES buffer containing 50mM NaCl, 10% glycerol (pH 5.5) using a glass homogenizer at a ratio of 10ml buffer per gram of pellet.

2.2.9 Solubilisation of RAMPs

The ProFoldin Membrane Protein Extraction Kit was used. In brief, 0.3mg of membrane was added to 100µl of each of the 12 detergent solutions. The samples were incubated overnight at 4°C. Non-solubilized material was removed by centrifugation at 18,000 x *g* for 2 hr at 4°C. Large scale solubilisation involved increasing the volumes proportionally.

2.2.10 Purification of RAMPs

2.2.10.1 RAMP2 purification

RAMP2 purification used the 1.5ml nickel nitrilotriacetic acid (Ni-NTA) Qiagen superflow column. Protein purification under native conditions using gravity flow was the method of choice. First the storage buffer from the column was allowed to drain through. Then 10ml of buffer NPI-10 (50mM $NaH_2PO_4.2H_2O$, 300mM NaCl and 10mM imidazole pH 8.0) was allowed to drain through the column. Then solubilised RAMP2 fraction was added to the column and allowed to drain through. Then two wash steps were performed with 10ml of buffer NPI-20 (50mM $NaH_2PO_4.2H_2O$, 300mM NaCl and 20mM imidazole pH 8.0). However, instead of a one-step elution a manual gradient elution was used with increasing amounts of imidazole (100mM, 150mM, 200mM, 250mM and 300mM, respectively) in 50mM $NaH_2PO_4.2H_2O$, 300mM NaCl, pH 8.0). Protein elution samples were concentrated using vivaspin 20ml concentrators (Vivascience, U.K) centrifuged at 4000 x *g* in a swing-out centrifuge rotor. Overnight dialysis removed imidazole and NaCl.

2.2.10.2 RAMP3 purification

1ml of Ni-NTA agarose resin (Qiagen, U.K) was added to 10ml of equilibriation buffer (50mM $NaH_2PO_4.2H_2O$, 300mM NaCl and 10mM imidazole pH 7.2) and then centrifuged at 5000 *x g*. The membrane sample was solubilised and resuspended in resuspension buffer (50mM $NaH_2PO_4.2H_2O$, 300mM NaCl and 10mM imidazole pH to 7.2) and was added to the resin. The mix was left overnight on a shaker at room temp. The sample was centrifuged at 5000 x *g*. All the supernatant was removed apart from the last 10ml, which was used to resuspend the resin. The resuspended resin was poured into a 20ml column and allowed to settle for 30 minutes. Next, the resin underwent two wash steps with 10ml wash buffer (40mM imidizole, 50mM $NaH_2PO_4.2H_2O$, 300mM NaCl, pH 7.2). Then a manual gradient elution was implemented using an increasing amount of imidazole in 10 ml of elutions containing 50mM $NaH_2PO_4.2H_2O$, 300mM NaCl, pH 7.2. Protein elution samples were concentrated with vivaspin 20ml concentrators at 4000 x *g* in a swing-out centrifuge rotor. Overnight dialysis removed imidazole and NaCl.

2.2.11 Circular dichroism

50µl of dialysed RAMP sample (diluted to 0.4 mg/ml for RAMP2 and 2.6mg/ml for RAMP3 as assessed by the NanoDrop 1000) was used for circular dichroism (CD) analysis. CD spectra of samples were collected using cell pathlengths of 0.1mm using the Jasco J-715 spectropolarimeter. Standard settings were used (sensitivity = 100mdeg, data pitch = 0.5nm, scanning mode = continuous, scanning speed = 100 nm/min, response = 1 sec and Band width = 2.0nm). The CD spectrum was then buffer corrected against 50mM $NaH_2PO_4.2H_2O$ containing 1% Soultion 8 pH 7.2.

2.3 Bioinformatics

2.3.1 Construction of CLR-ECD

The protein sequences of the human Secretin receptor family as defined by Fredriksson *et al.*, (2003) were downloaded from the EMBL GenBank (www.ncbi.nlm.nih.gov/Genbank/). The signal peptide for each construct was predicted by SignalP 3.0 (Bendtsen *et al.*, 2004) and removed. Note, SignalP 3.0 was

unable to identify a signal peptide for the GLP2R (GenBank accession NP_004237) and therefore removed from analysis. A multiple sequence alignment of the Secretin receptor family was made using the Tcoffee server (Poirot et *al.*, 2003).

A comparative protein structure of ECD of CLR, from residues 23-134 was generated using Modeller9v3 (Sali and Blundell, 1993 and Fiser *et al.*, 2003). The comparative model was based on two templates GIPR ECD (Pathier *et al.*, 2007) and PTHR1 ECD (Pioszak and Xu, 2008). The GIPR ECD was selected based on the structures high-resolution (1.90Å). The CLR ECD and GIPR ECD alignment contained a large seven-residue gap region (located between H114- T119 of the CLR primary sequence). The PTHR1 ECD was selected because it provided a sufficient template for this region. Moreover, multiple templates have been found to improve the accuracy of homology models (Chakravarty, *et al.*, 2008).

Initially, Modeller9v3 generated 500 models. The models were ranked by the Modeller9v3 energy objective function. The top 20 structures were retained and the stereochemical quality was assessed by PROCHECKv3.5.4 (Laskowski *et al.*, 1993 and Laskowski *et al.*, 2001). Based on overall and residue-by-residue geometry a structure was selected.

CLR ECD contained three long loop regions, loop 1(D55-R67), loop 3 (N76-G81) and loop 4 (P89-S98). The program Loopy (Xiang *et al.*, 2002 and Soto *et al.*, 2008) was used to generate initial loop conformations. 2,000 initial loop conformations were generated for loops that contained 10 residues or less. Whereas, 4,000 initial loop conformations were generated for loops that contained more than 10 residues (Soto *et al.*, 2008). Note, for loop1 only residues between D55-Y64 were used to generate initial loop conformations due to the presence of a disulphide bond at position C65. Next, Loopy attached the generated loops to the protein by a random tweak method. Loopy then performed a fast energy minimisation on torsional space and then allowed the program Scap to predict and build the protein side-chains (Xiang *et al.*, 2002).

The initial loop conformations were ranked by DFIRE 2.0 statistical energy function (Yang *et al.*, 2008). The top 50 loop conformers were retained. The filtered ensemble then underwent a physical based scoring method first described by de Bakker *et al.*, (2003) using AMBER99 in the presence of GB/SA solvation model implemented in

the TINKER program package (http://dasher.wustl.edu/tinker/). In summary, the program MINIMIZE performed a limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) minimisation on the 50 loop conformers until the convergence criterion (RMSD 0.1 kcal/mol/Å) was reached or when the number of minimization steps exceeded 100. Only atoms belonging to the loop region of interest were allowed to move during the minimization while the remaining protein remained locked.

Minimized fragments were subsequently ranked by total energy potential by the program ANALYZE in the TINKER package. The top 10 conformers were retained and visually inspected. A conformation from this ensemble was selected taking into consideration both the architecture of the loops in other Secretin-like GPCR ECDs and the total energy potential score of the loop conformer.

The two long loop regions of the GIPR ECD were constructed in the presence and absence of the ligand GIP and compared to the crystal structure to gauge the accuracy of this loop procedure when dealing with Secretin-like GPCR ECDs. Taking into consideration that the loop process outlined here is aimed to predict loop conformers in solvent and loop 1 in Secretin-like ECD is particularly flexible (Grace *et al.*, 2007) it is clear that this procedure has the capability of making reasonable predictions according to Fiser *et al.*, (2000) scoring system (see Table 6).

Table 6. Summary of long loop construction using the GIPR-ECD. The GIPR-ECD (PDB accession code: 2QKH) was used to assess the accuracy of the loop construction protocol. Loop 1 and 4 of the GIPR-ECD was constructed in the absence and presence of incretin GIP (1-42). The top 10 loop conformers that were retained for visual inspection were assessed against the crystal structure loop conformation using global RMSD (N, Cα, C, O), where all loop residues and stem residues were incoporated into analysis. The table shows the loop with the best global RMSD compared to the crystal structure loop. Also the mean and median global RMSD of the top 10 loop conformers compared to the crystal structure loop has been reported.

Loop domain	Length	Best loop conformer (global RMSD)	Mean of top 10 conformers (global RMSD)	Median of top 10 conformers (global RMSD)
Loop 1 without ligand	8 residues	2.11Å	3.24Å	3.36Å
Loop 1 with ligand	8 residues	1.79Å	3.49Å	3.40Å
Loop 4 without ligand	12 residues	1.84Å	3.99Å	3.29Å
Loop 4 with ligand	12 residues	2.55Å	3.53Å	3.34Å

The H++ web server (http://biophysics.cs.vt.edu/H++/) was used to calculate the protonation states of titratable sites in CLR ECD (external dielectric constant = 80, internal dielectric constant = 6, salinity = 0.15 M, pH = 7.2). Groningen machine for chemical simulations (GROMACS) v4.0 (Hess *et al.*, 2008), utilising the OPLS-AA/L force field parameters was used to perform the steepest descent energy minimisation on the CLR-ECD in the presence of an explicit SPC/E water molecules and neutralising Na^+ ions.

2.3.2 Low resolution docking to produce CGRP-ECD

The CLR ECD was docked against itself with GRAMMv1.03 (Vakser *et al.*, 1995) using the generic low resolution settings to produce 1000 complexes. These were ranked based on residue level pair potentials scoring, where each potential corresponds to the likelihood of a trans-interface between a pair of residue types, using the 3D-Dock suite (Moont *et al.*, 1999). The top 20 constructs were retained.

Taking into consideration that the C-terminal region of the ECDs is assumed to face the lipid bilayer *in vivo*, constructs where the C-terminal regions of the ECDs that were not in the orientation to accommodate this for both ECDs were excluded. The remaining constructs shared a similar overall architecture; a selected dimer interface was refined using the MULTIDOCK program from the 3D-Dock suite. The RAMP1 ECD crystal structure (Kusano *et al.*, 2008) was downloaded from the RCSB Protein Data Bank (www.rcsb.org/pdb/). The selenomethionine labels were changed to conventional methionine residues with the Swiss PDB viewer DeepView v4.0. RAMP1 ECD was docked onto the CLR-ECD complex using the strategy described above. The resulting trimer was refined with Gromacs v4.0, utilising the OPLS-AA/L force field parameters to perform a steepest descent energy minimization in the presence of an explicit SPC/E water model with neutralising Na^+ ions.

2.3.3 Construction of the CLR TM domain

Two CLR TM domain models were constructed. The Vohra *et al.*, (awaiting publication) alignment was used to align the putative TM segments of CLR against the TM domains in bovine rhodopsin (Okada *et al.*, 2004- PDB accession code 1U19) to resemble the inactive conformation of CLR and opsin (Scheerer *et al.*, 2008- PDB accession code 3DQB) to resemble the active conformation of CLR. Modeller9v3 generated 500 models. The models were ranked by Modeller9v3 energy objective function. The top 20 structures were retained and stereochemical quality was assessed by PROCHECKv3.5.4 (Laskowski *et al.*, 1993 and Laskowski *et al.*, 2001). Based on overall and residue-by-residue geometry a structure was selected.

The loop domains were constructed using the same procedure described above. However, ECL2 was divided into two segments. Cys-212 is assumed to participate in a highly conserved disulphide bond with Cys-282. The position of Cys-212 was approximated based on the initial model selected using Modeller9v3. Dr A. Conner generously provided mutagenesis data of an alanine scan that incorporated most of the residues predicted within ECL2 (see Table 7). This experimental work guided conformer selection when the loops were being visually assessed.

Table 7. Summary of the fold effects of CLR ECL2 mutations compared to WT CLR when co-expressed with RAMP1. Personal communication from Dr. A Conner. WT HA CLR/ *myc* RAMP1 (WT receptor) was transiently transfected into COS-7 cells alongside the mutant HA CLR/ *myc* RAMP1. Fold effects for cAMP production was determined from a minimum of three independent αCGRP stimulated dose-response curves (each assay point was assessed in duplicate). The mean EC50 of the WT receptor was compared to the mean EC50 of the mutant receptor to deduce fold differences. A two-tailed independent t-test was used to assess statistical differences between pEC50 values of WT and mutant receptors, where *, **, *** represent significantly different from WT ($p > 0.05$, $p > 0.01$ and $p > 0.001$ respectively). The IC50 fold effects were deteremined from at least three independent αCGRP inhibition binding assays, where the mean IC50 for the WT receptor was compared to the mutant receptor mean IC50. A one-way ANOVA followed by a Dunnett's test was used to determine statistical differences between WT pIC50 values and mutant pIC50 values.

Mutant	cAMP production EC50 fold effects after αCGRP stimulation (compared to WT receptors)	αCGRP inhibition IC50 fold effects (compared to WT receptors)
R274A	43.65 decrease*	25.70 decrease**
S275A/L276A	1.86 decrease	Experiments not conducted
Y277A	9.33 decrease*	1.55 decrease
Y278A	6.76 decrease*	26.92 decrease**
N279A	2.10 decrease	Experiments not conducted
D280A	19.05 decrease**	6.17 decrease*
N281A	2.19 decrease	Experiments not conducted
C282A	10.96 decrease**	5.12 decrease*
W283A	26.91 decrease**	11.22 decrease**
I284A	0.91 increase	Experiments not conducted
S285A	5.62 decrease*	Abolished***
S286A/D287A	1.78 decrease	Experiments not conducted
T288A	28.18 decrease*	7.24 decrease*
H289A	2.19 increase	Experiments not conducted

Reggio *et al.*, (2006) stated that LOOPY was applicable to GPCR loop prediction. Table 8 summarises the results of predicting ECL1 and ECL3 of inactive bovine rhodopsin using the method described above. The results suggest that the loop prediction strategy is very good at predicting these two examples

Table 8. Summary of loop construction of ECL1 and ECL3 of ground-state bovine rhodopsin. Only chain A from the 1U19.pdb file was used in analysis. ECL1 and 3 of 1U19 chain A was constructed. The top 10 loop conformers that were retained for visual inspection were assessed against the crystal structure loop conformation using global RMSD (N, Cα, C, O). All loop residues and stem residues were incoporated into analysis. The table shows the loop with the best global RMSD in relation to the crystal structure loop aswell as the mean and median global RMSD of the 10 loop conformers compared to the crystal structure loop.

Loop Domain	Length	Best loop conformer (global RMSD)	Mean of top 10 conformers (global RMSD)	Median of top 10 conformers (global RMSD)
Bovine Rhodopsin ECL1 (PDB accession: 1U19)	8 residues	1.40Å	1.93Å	1.88Å
Bovine Rhodopsin ECL3 (PDB accession: 1U19)	8 residues	0.74Å	1.47Å	1.54Å

The ProPka program (Li *et al.*, 2005) *via* the PDBQPR server (see Dolinsky *et al.*, 2007) was used to assign the protonation states of the titratable groups in each CLR TM domain model, using the CHARMM parameters set at pH 7.4. The PDBQPR server has been used previously with membrane proteins. For example, meaningful homology models of the human histamine H4 receptor have used the PDBQPR server with success (Jojart *et al.*, 2007).

The CLR TM models were then orientated on their Z-axis based on the relative position of Tyr-165 and Tyr-367. The CHARMM (c35b3) GBSW module, which contains Im *et al.*, (2003) membrane GB/SA application was used. The all-atom param22/cmap force-field in the presence of a 32Å implicit membrane was set up. Then 1500 steps of a steepest descent energy minimisation followed by 5000 steps of adopted basis Newton-Raphson minimisation (or until the RMSD was less than 0.001 kcal/mol Å) was conducted. Note, TM models contained an acetylated N-terminus and a N-Methylamide C-terminus. This is an approximation for the fact that there are missing residues in the model since only a fragment of the protein has been attempted

to be modeled. This step prevents unnecessary large electrostatic attractive forces between the helical ends during energy minimisation.

Chapter 3: The functional role of ECL1 and its associated TM regions in the CLR

3.1 Introduction

Loops connect secondary structural elements of proteins and may contribute to receptor affinity and efficacy. The architecture of ECLs may help orientate the TM bundle and provide key molecular determinants for ligand binding (see Runge *et al.*, 2003). The functional role of ECL1 in Secretin-like GPCRs is unclear. Pham and Sexton (2003) implied that the loop was associated with agonist recognition. A conclusion based on a series of chimera studies that noted that ECL1 participates in the ligand-receptor complex interface in the VPAC1 (Olde *et al.*, 1998), secretin receptor (SCTR; Holtmann *et al.*, 1995) and PAC1R (Hashimoto *et al.*, 1997).

The idea that ECL1 may participate in the orthosteric binding site adheres with the two-step activation model (Grace *et al.*, 2004). However, the mechanistic role of ECL1 in the CLR remains elusive. Identifying important residues in this region of the CLR could enhance understanding of the orthosteric binding site and may provide further information on signal transduction mechanism within the CGRP receptor.

Consequently, an alanine-scan has been used to investigate ECL1 and its corresponding TM regions. Endogenous alanine residues within this region were substituted to leucine residues to probe the significance of the size and/or geometry of the innate methyl side chain. Mutant receptors were assessed on multiple criteria including cAMP production, agonist-mediated desensitisation, cell-surface and total expression and αCGRP binding in an attempt to determine the functional role of the inherent residues within the protein.

3.2 Results

3.2.1 Stimulation of cAMP production

Each mutant was challenged with human αCGRP and cAMP production was measured (see Table 9 for pEC50 comparisons). C212A dose-response curves were conducted by Dr A.Conner and pEC50 values were provided, which suggested a ~33 fold decrease. A reduction in αCGRP potency (as assessed by significant differences in pEC50 values compared to WT receptors) was also observed in the mutant receptors (in the order of the magnitude of EC50 fold decrease); L195A (~30 fold), A199L (~20 fold), V198A (~11 fold) and H219A (~11 fold- see Figure 21).

In contrast, an increase in αCGRP potency was found in L220A (~25 fold), A203L (~11 fold), A206L (~9 fold), L222A (~6 fold), (see Table 9 and Figure 22). Furthermore, the mean Emax of L222A was found to increase by 40.1% ± 11.2% and mean basal activity was 27.0% ± 11.3% higher when compared to WT. The mean Emax of L220A also increased by 34.4% ± 19.5%, relative to WT. A modest increase in mean Emax was found in A203L (18.7% ± 8.5%) and a slight increase in A206L (13.9% ± 14.3%).

The mean pEC50 values of L204A, P209A, V210A, Q216A and M223A were not found to be significantly different to WT. However, the mean basal activity and the mean Emax in each of these mutants were higher when compared to WT. Furthermore, the mean Emax of T207A was found to be higher while the mean basal activity resembled WT (see Table 10 and Figure 23).

Stimulation of cAMP production on the remaining mutant receptors not discussed above resembled WT. The mean pEC50 values were found not to be significantly different. Moreover, the size of effect of the percent mean Emax and mean basal activity did not exceed 20% in either direction relative to WT.

Table 9. The ability of ECL1 mutant receptors to stimulate cAMP compared to the WT receptor. Values are pEC50 means ± S.E.M. $p < 0.05$ is represented by *. $p < 0.01$ is represented by ** and $p < 0.001$ is represented by ***. pEC50 mutant values were compared to WT using an independent two-tailed t-test. +C212A was provided by Dr A. Conner.

Mutant	N	pEC50 WT	pEC50 Mutant	Mutant	N	pEC50 WT	pEC50 Mutant
H194A	5	9.97 ± 0.48	9.41 ± 0.53	P209A	4	9.91 ± 0.33	9.85 ± 0.14
L195A	4	10.36 ± 0.25	8.87 ± 0.09**	V210A	4	9.95 ± 0.29	10.44 ± 0.23
T196A	3	9.73 ± 0.21	10.18 ± 0.10	S211A	4	9.49 ± 0.33	9.60 ± 0.21
A197L	4	10.09 ± 0.29	10.05 ± 0.39	C212A+	3	9.62 ± 0.76	8.10 ± 0.43**
V198A	5	9.67 ± 0.17	8.87 ± 0.26*	K213A	5	9.49 ± 0.40	8.78 ± 0.39
A199L	6	10.20 ± 0.31	8.87 ± 0.32*	V214A	4	9.83 ± 0.39	9.97 ± 0.37
N200A	3	9.80 ± 0.17	9.62 ± 0.28	S215A	6	9.39 ± 0.33	9.41 ± 0.22
N201A	4	9.38 ± 0.16	9.88 ± 0.29	Q216A	3	9.63 ± 0.13	9.94 ± 0.13
Q202A	4	9.49 ± 0.43	9.17 ± 0.32	F217A	4	9.94 ± 0.43	9.37 ± 0.22
A203L	5	9.75 ± 0.22	10.77 ± 0.35*	I218A	4	9.27 ± 0.22	9.60 ± 0.25
L204A	3	9.73 ± 0.13	9.68 ± 0.17	H219A	3	9.41 ± 0.17	8.35 ± 0.13**
V205A	4	9.62 ± 0.21	10.04 ± 0.30	L220A	4	9.53 ± 0.17	10.93 ± 0.13***
A206L	5	9.43 ± 0.29	10.37 ± 0.19*	Y221A	4	9.73 ± 0.42	9.01 ± 0.51
T207A	4	9.98 ± 0.30	10.50 ± 0.20	L222A	5	9.95 ± 0.21	10.75 ± 0.10*
N208A	3	9.85 ± 0.11	9.97 ± 0.18	M223A	4	9.71 ± 0.18	9.65 ± 0.15

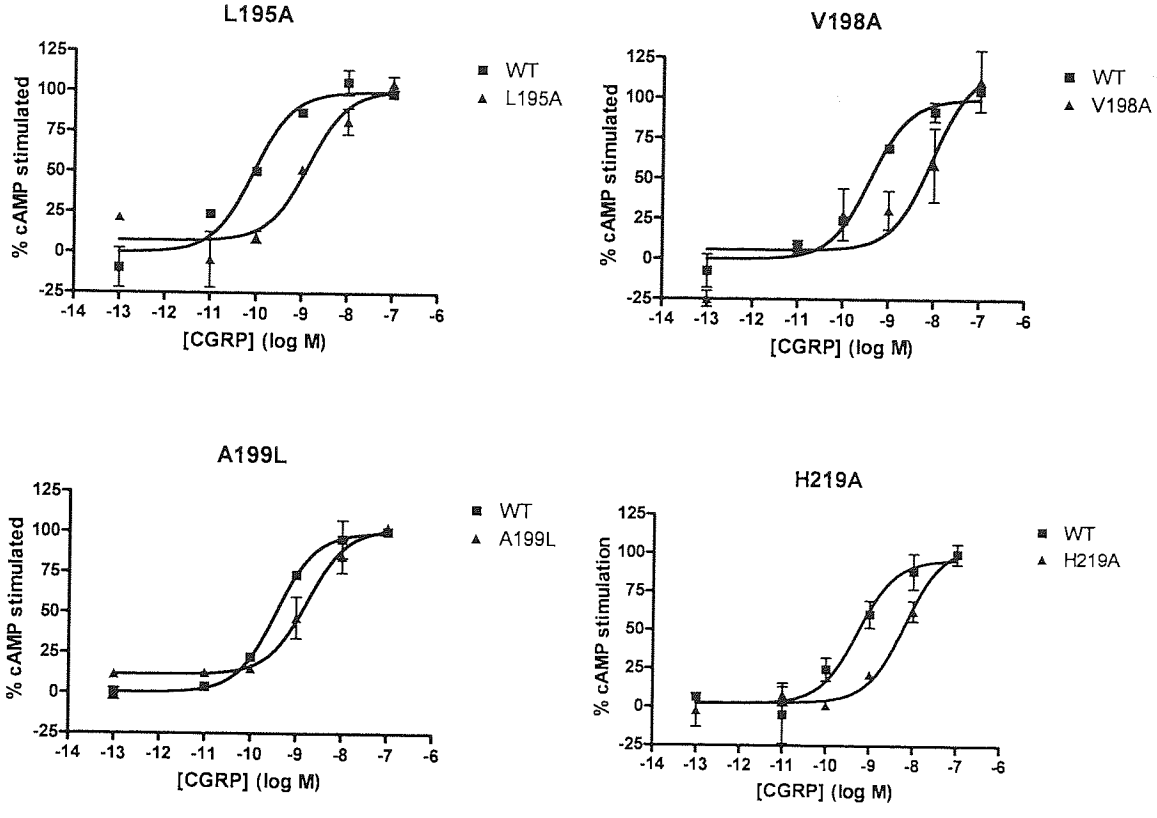


Figure 21. Representative dose-response curves of ECL1 mutants that showed a significant decrease in αCGRP potency. Sigmoidal concentration-response curves comparing the WT receptor and mutant receptors (L195A, V198A, A199L, and H219A) capability of stimulating cAMP after αCGRP activation are shown. Each WT and mutant receptor concentration-response comparison curve is a representative example from an ensemble of at least three independent experiments. Each assay point was performed in duplicate where each point on the graph represents the mean ± S.E.M. The line of best fit for each concentration-response curve was fitted with GraphPad Prism 4 see Equation 1.

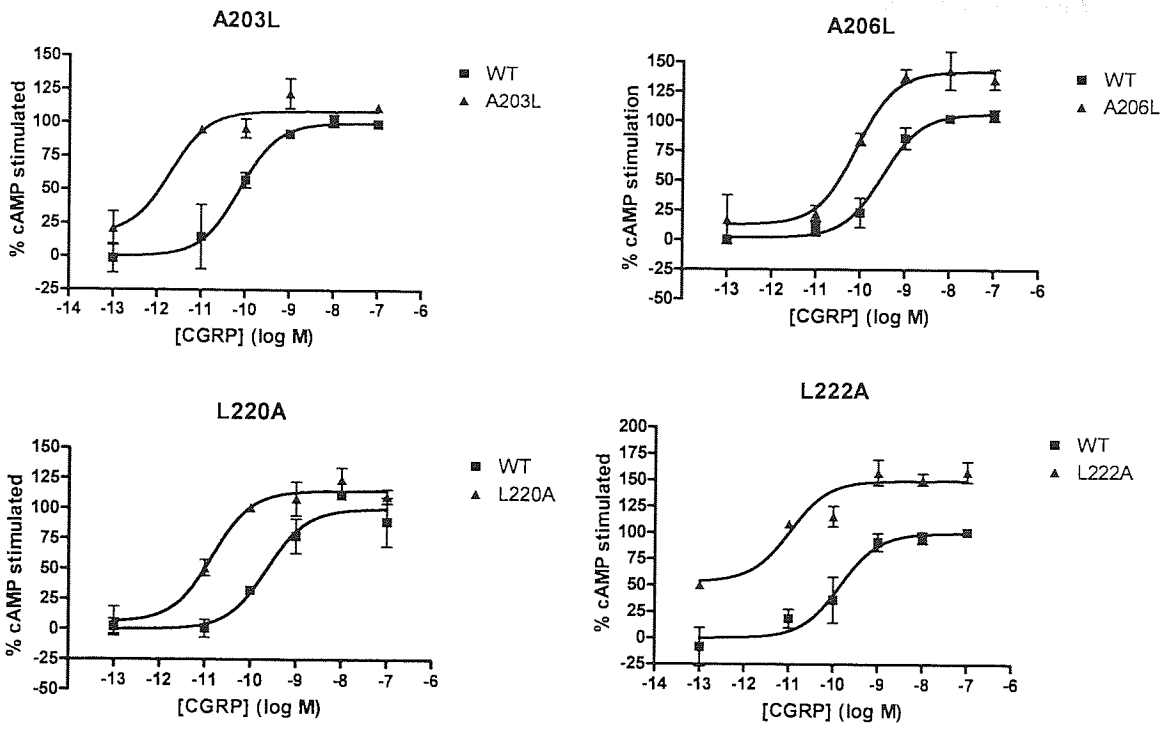


Figure 22. Representative dose-response curves of ECL1 mutants that showed a significant increase in αCGRP potency. Sigmoidal concentration-response curves comparing the WT receptor and mutant receptors (A203L, A206L, L220A and L222A) capability of stimulating cAMP after αCGRP activation are shown. Each WT and mutant receptor concentration-response comparison curve is a representative example from an ensemble of at least three independent experiments. Each assay point was performed in duplicate where each point on the graph represents the mean ± S.E.M. The line of best fit for each concentration-response curve was fitted with GraphPad Prism 4 see Equation 1.

Table 10. ECL1 mutant receptors that showed altered mean basal activity and Emax. The WT and mutant cAMP dose-response comparison curves were normalised from 0% to 100% based on WT Top and Bottom values generated by GraphPad Prism 4. The Emax and basal activity of each mutant was assessed by the mean of the top and bottom of each dose-response curve. The mutant receptors mean Emax and basal activity was expressed as a percentage that corresponded to WT normalisation. A meaningful difference was noted if the mean size of effect differed by 20% or more. Values reported are mutant percent means ± S.E.M (% WT).

Mutant	N	Mean basal activity (% WT)	Mean Emax (% WT)
L204A	3	37.3 ± 18.5	142.3 ± 13.4
P209A	4	41.5 ± 5.2	122.5 ± 5.0
V210A	4	33.7 ± 8.9	117.9 ± 10.7
Q216A	3	22.7 ± 8.1	144.0 ± 13.7
M223A	4	13.2 ± 11.1	128.8 ± 16.2
T207A	4	0.4 ± 7.7	122.5 ± 9.3

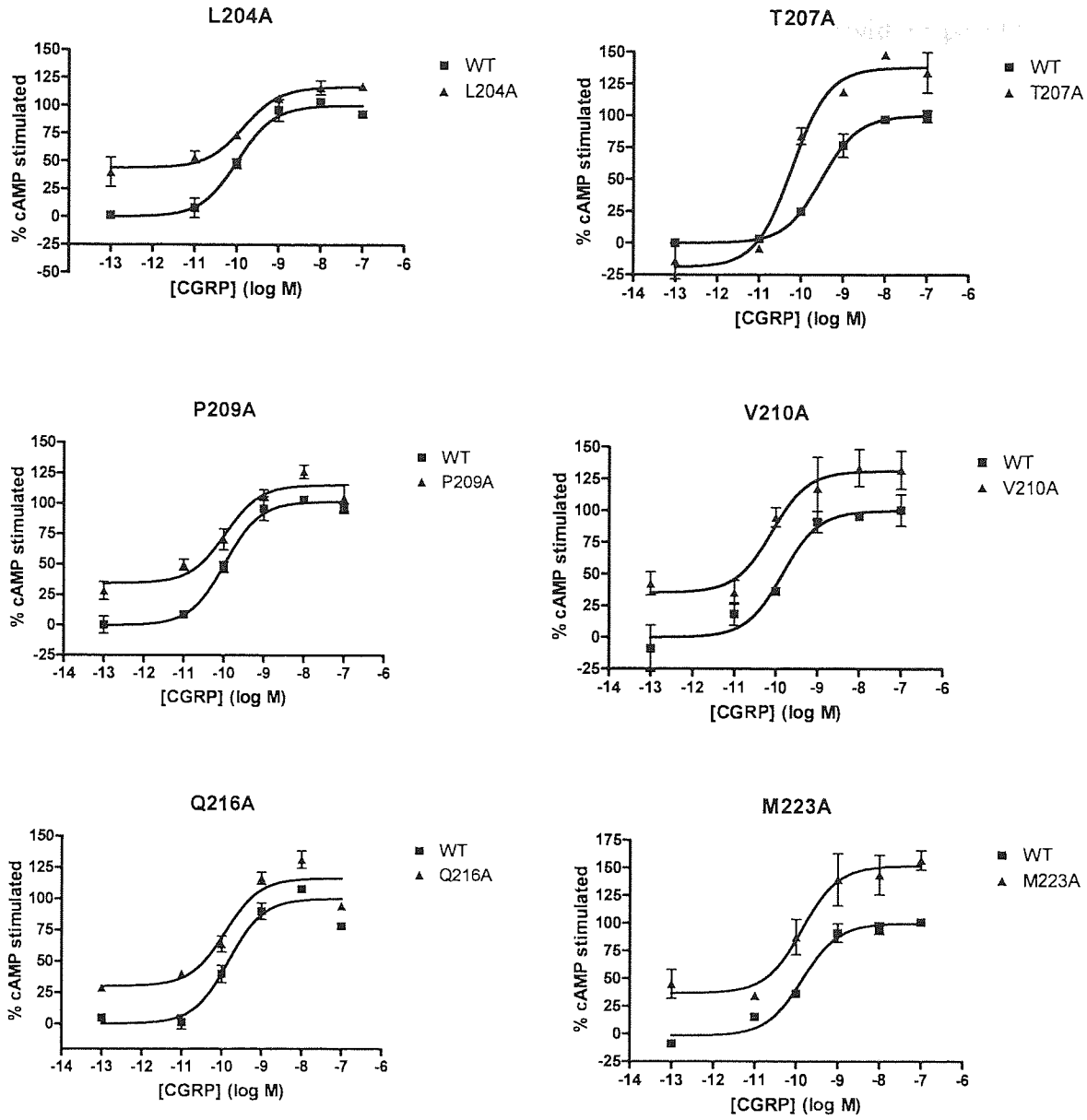


Figure 23. Representative dose-response curves of ECL1 mutants that showed altered basal activity and/or Emax. Sigmoidal concentration-response curves comparing the WT receptor and mutant receptors (L204A, T207A, P209A, V210A, Q216A and M223A) capability of stimulating cAMP after αCGRP activation are shown. Each WT and mutant receptor concentration-response comparison curve is a representative example of an ensemble of at least three independent experiments. Each assay point was performed in duplicate where each point on the graph represents the mean ± S.E.M. The line of best fit for each concentration-response curve was fitted with GraphPad Prism 4 see Equation 1.

3.2.2 Cell surface receptor expression

Expression of receptors was measured using a cell-surface ELISA (Table 11). Statistically significant reductions in cell-surface expression were seen in nine

mutants (N200A, V205A, N208A, C212A, F217A, I218A, H219A, L220A and M223A). The largest reduction in cell surface expression was 41.03% (with respect to WT) in M223A but overall the cell surface reductions seen in these mutants were fairly modest. Although, it is interesting that four consecutive mutants (F217A, I218A, H219A, L220A) reduce cell surface expression.

Four mutants were found to increase cell surface expression significantly, V198A, A199L, A206L and P209A, respectively.

Table 11. Cell surface expression of ECL1 mutant receptors. Cell surface expression ELISA was used to probe for the presence of the HA epitope. Mutant HA CLR/*myc* RAMP1 receptors were compared with WT HA CLR /*myc* RAMP1 receptors. 3-6 independent experiments that contained triplicate data points were used in analysis. The raw data for each independent experiment was normalised where the mean WT receptor cell-surface expression equalled a 100% and the mean negative control (*myc* RAMP1/ empty pcDNA3.1-) was equal to 0%. Values reported are mutant means ± S.E.M (% WT). Mutant cell surface expression was compared to WT receptor using a Mann-Whitney U test ($p < 0.05$ is represented by *, $p < 0.01$ represented by **, and $p < 0.001$ represented by ***).

Mutant	Cell surface expression (% WT)	Mutant	Cell surface expression (% WT)
H194A	106.5 ± 9.9	P209A	133.0 ± 9.0*
L195A	92.7 ± 5.4	V210A	111.9 ± 12.4
T196A	96.0 ± 9.2	S211A	92.62 ± 11.6
A197L	88.2 ± 5.3	C212A	69.71 ± 10.3*
V198A	133.4 ± 11.6***	K213A	94.54 ± 9.1
A199L	130.2 ± 8.2**	V214A	88.17 ± 8.9
N200A	74.8 ± 6.0*	S215A	101.6 ± 6.8
N201A	91.0 ± 8.3	Q216A	107.3 ± 8.4
Q202A	81.4 ± 11.9	F217A	72.78 ± 6.7*
A203L	112.4 ± 6.9	I218A	73.05 ± 8.4**
L204A	105.9 ± 5.6	H219A	68.81 ± 7.2***
V205A	81.3 ± 3.3	L220A	79.21 ± 6.1*
A206L	114.5 ± 4.9**	Y221A	107.3 ± 9.9
T207A	111.6 ± 11.5	L222A	94.33 ± 9.6
N208A	68.88 ± 7.3**	M223A	58.97 ± 8.3***

3.2.3 Total receptor expression

Total receptor expression probing for the HA epitope was assessed on mutant receptors that were either found to have a significantly different mean pEC50 and/or cell surface expression. Overall the results show that the mutant cDNA was successfully transfected, transcribed and translated. A significant, yet modest, increase in total expression was observed in V205A and A206L. Furthermore, a significant difference in total expression was seen in C212A and H219A but again the size of the effect was only modest (see Table 12).

Table 12. Total expression of ECL1 mutant receptors. Total expression of HA-tagged receptors both mutant and WT were analysed when co-transfected with *myc* RAMP1. At least 3 independent experiments containing triplicate data points were used in analysis. Total expression in the mutant condition was normalised to the WT condition (equal to 100%) and negative control (myc RAMP1/empty pcDNA3.1-, after 0.1% Triton X 100, which was equal to 0%). A Mann-Whitney U test was used to assess statistical differences between WT and mutant receptors were *, **, *** represent $p < 0.05$, 0.01 and 0.001 significance levels.

Mutant	Total expression (% WT mean ± S.E.M)
L195A	111.5 ± 8.5
V198A	87.6 ± 13.5
A199L	115.5 ± 5.6
N200A	104.8 ± 7.3
A203L	110.2 ± 8.5
V205A	115.0 ± 4.8*
A206L	116.5 ± 3.1**
N208A	110.4 ± 3.8
P209A	106.8 ± 5.3
C212A	80.0 ± 7.9*
F217A	98.9 ± 6.0
I218A	118.2 ± 12.7
H219A	75.7 ± 16.5***
L220A	101.6 ± 4.7
L222A	106.4 ± 6.0
M223A	101.7 ± 6.2

3.2.4 αCGRP mediated internalisation

αCGRP mediated internalisation was severely impaired or abolished in L195A and A199L and significantly reduced in C212A and V198A. In contrast, N208A and H219A were found to internalise moderately more readily than WT (see Table 13).

Table 13. ECL1 mutant receptors capability of agonist (αCGRP) mediated internalisation compared to the WT receptor. Agonist mediated internalisation of the CGRP receptor (both WT and mutant receptors) was approximated by a HA epitope probing cell surface ELISA taking into account the difference in cell surface expression levels between CGRP receptors that have or have not been exposed to 100nM of human αCGRP for an 1hr. Percent mean ± S.E.M agonist mediated internalisation was determined by 3-6 independent experiments that contained triplicate data points. A Mann Whitney U test was used to compare mutant and WT percent agonist internalisation where $p < 0.05$ is respresented by *, $p < 0.01$ is represented by **, and $p < 0.001$ is represented by ***.

Mutant	WT receptor internalisation (% mean ± S.E.M)	Mutant receptor internalisation (% mean ± S.E.M)
H194A	47.70 ± 4.9	58.14 ± 4.7
L195A	53.99 ± 2.8	1.67 ± 3.2***
T196A	51.98 ± 5.5	67.44 ± 3.2
A197L	54.66 ± 2.9	52.04 ± 3.6
V198A	55.52 ± 3.0	30.32 ± 3.7***
A199L	64.04 ± 4.1	8.39 ± 7.7***
N200A	58.81 ± 3.0	67.90 ± 5.6
N201A	56.81 ± 3.9	63.03 ± 2.8
Q202A	60.38 ± 4.9	66.48 ± 8.2
A203L	56.37 ± 4.0	62.21 ± 3.3
L204A	47.72 ± 5.2	49.54 ± 4.3
V205A	51.32 ± 7.4	64.42 ± 4.0
A206L	56.97 ± 2.0	61.03 ± 2.2
T207A	53.11 ± 2.1	58.75 ± 10.3
N208A	47.87 ± 5.3	65.65 ± 3.1*
P209A	55.38 ± 5.8	55.90 ± 6.3
V210A	62.01 ± 4.1	61.18 ± 6.0
S211A	54.01 ± 3.1	58.33 ± 2.8
C212A	58.04 ± 4.7	37.29 ± 9.7*
K213A	52.00 ± 4.5	51.53 ± 6.1
V214A	46.50 ± 2.9	59.87 ± 5.4
S215A	66.03 ± 7.2	64.07 ± 5.1
Q216A	50.47 ± 2.5	55.27 ± 3.3
F217A	45.10 ± 8.0	44.78 ± 7.4
I218A	48.68 ± 8.5	66.06 ± 9.8
H219A	55.42 ± 3.1	75.07 ± 6.9*
L220A	55.44 ± 3.3	58.27 ± 4.5
Y221A	51.50 ± 4.6	63.57 ± 4.1
L222A	48.54 ± 7.0	56.60 ± 7.6
M223A	48.32 ± 3.2	60.59 ± 4.5

3.2.5 Inhibition of ^{125}I-hCGRP radioligand binding

Inhibition of ^{125}I-hCGRP radioligand assays were performed on mutant receptors that were either found to have a significantly different mean pEC50 and/or agonist mediated internalisation when compared to WT (see Table 14). The pIC50 of four mutants were significantly reduced when compared to WT; L195A, A199L, C212A and H219A (Figure 24). In contrast, the pIC50 of L200A showed an increase compared to WT (Figure 25).

Table 14. Apparent affinities of αCGRP for ECL1 mutant receptors, estimated by inhibition of radioligand binding. Mean ± S.E.M pIC50 WT and mutant values shown. An independent two-tailed t-test was used to assess statistical differences where $p < 0.05$ is represented by * and $p < 0.01$ is represented by **. N.M.B. stands for no measurable binding.

Mutant	N	pIC50 for WT receptor (mean ± S.E.M)	pIC50 for mutant receptors (mean ± S.E.M)
L195A	5	9.01 ± 0.40	N.M.B
V198A	4	8.86 ± 0.40	8.07 ± 0.47
A199L	4	8.88 ± 0.25	7.26 ± 0.13**
A203L	4	9.33 ± 0.31	9.77 ± 0.64
A206L	5	9.00 ± 0.13	9.19 ± 0.19
N208A	5	9.03 ± 0.09	8.90 ± 0.21
C212A	4	9.04 ± 0.09	6.94 ± 0.49**
H219A	4	9.01 ± 0.07	7.91 ± 0.44
L220A	5	9.36 ± 0.20	10.21 ± 0.36
L222A	3	9.84 ± 0.03	10.76 ± 0.17**

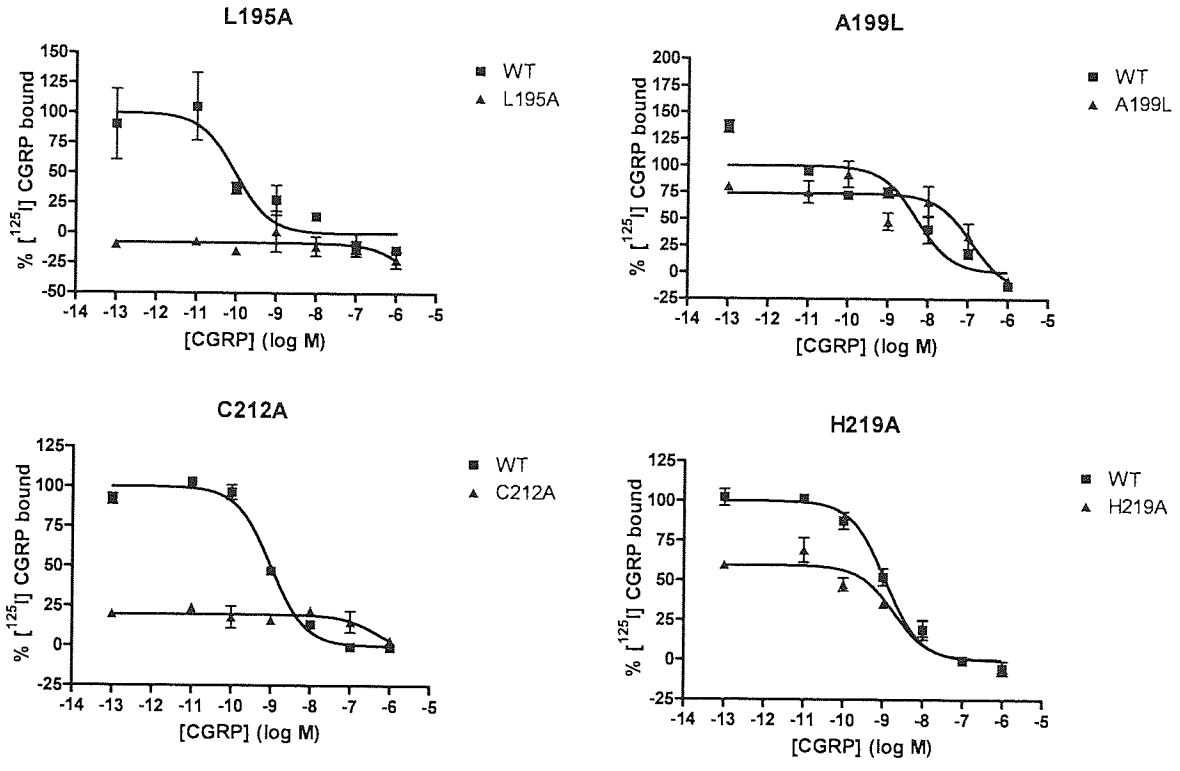


Figure 24. Representative inhibition curves of ECL1 mutant receptors that significantly impaired CGRP binding. Sigmoidal αCGRP inhibition curves comparing the WT receptor and mutant (L195A, A199L, C212A and H219A) receptors are shown. The receptors capability of binding the radioligand ^{125}I-hCGRP (~20pM) was assessed aswell as the ability of the non-radioactive αCGRP to displace the radioactive ligand over six log concentrations. Each WT and mutant receptor curve is a representative example of an ensemble of at least three independent experiments. Each assay point was performed in duplicate where each point on the graph represents the mean ± S.E.M. The line of best fit for each comparisin curve was fitted with GraphPad Prism 4 see Equation 1. The concentration of COS-7 cell membranes was 2mg/ml.

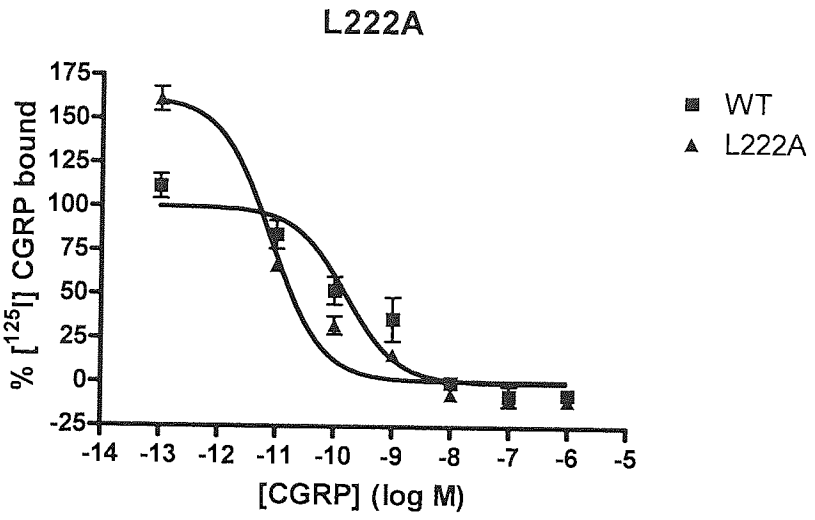


Figure 25. Representative inhibition curve of L222A compared to WT, which was showed to significantly enhance CGRP binding. Sigmoidal αCGRP inhibition curves comparing the WT receptor and L222A receptors were conducted. The receptors capability of binding the radioligand ^{125}I-hCGRP (~20pM) was assessed, aswell as the ability of the non-radioactive αCGRP to displace the radioactive ligand over six log concentrations. Each WT and mutant receptor curve is a representative example of an ensemble of at least three independent experiments. Each assay point was performed in duplicate where each point on the graph represents the mean ± S.E.M. The line of best fit for each comparisin curve was fitted with GraphPad Prism 4 see Equation 1. The concentration of COS-7 cell membranes was 2mg/ml.

3.2.6 Summary of important ECL1 mutations

ECL1 mutations that were found to significantly alter the potency of αCGRP to stimulate cAMP (pEC50), cell-surface expression, total expression, agonist-mediated internalisation and/or CGRP binding (pIC50) are summarised in table 15.

Table 15. The ECL1 mutations that significantly altered the pharmacology of the CGRP receptor. EC50 fold effects and IC50 fold effects were determined by sigmoidal concentration-response curves comparing the WT receptor and mutant receptors capability of either stimulating cAMP after αCGRP exposure or the inhibition of ^{125}I-hCGRP binding, respectively. Both the EC50 and IC50 fold effects represent the mean fold difference from at least three independent experiments. A two-tailed independent t-test was used to assess statistical differences. Cell surface expression, total expression and αCGRP mediated internalisation values represent the mean difference in % WT. A Mann Whitney U test was used to test for statistical differences between WT and mutant receptors on these three parameters, which were determined by at least three independent experiments. Mutants found not to be significantly different from WT condition are represented by N.S. Experiments that have not been conducted are represented by -. N.M.B stands for no measurable binding.

Mutant	EC50 fold effects after αCGRP stimulation (compared to WT receptors)	Cell surface expression (% WT mean difference)	Total expression (% WT mean difference)	αCGRP mediated internalisation (% WT mean difference)	IC50 fold effects (compared to WT receptors)
L195A	~30 fold decrease	N.S.	N.S.	96.9% decrease	N.M.B
V198A	~11 fold decrease	33.4% increase	N.S.	45.4% decrease	N.S.
A199L	~20 fold decrease	30.2% increase	N.S.	86.9% Decrease	~42 fold decrease
N200A	N.S.	25.2% decrease	N.S.	N.S.	-
A203L	~11 fold increase	N.S.	N.S.	N.S.	N.S.
V205A	N.S.	N.S.	15.0% increase	N.S.	-
A206L	~9 fold increase	14.5% increase	16.5% increase	N.S.	N.S.
N208A	N.S.	31.1% decrease	N.S.	37.1% Increase	N.S.
P209A	N.S.	33.0% increase	N.S.	N.S.	-
C212A	~33 fold decrease	30.3% decrease	20.0% decrease	N.S.	~126 fold decrease
F217A	N.S.	27.2% decrease	N.S.	N.S.	-
I218A	N.S.	27.0% decrease	N.S.	N.S.	-
H219A	~11 fold decrease	31.2% decrease	24.3% decrease	35.5% Increase	N.S.
L220A	~25 fold increase	20.8% increase	N.S.	N.S.	N.S.
L222A	~6 fold increase	N.S.	N.S.	N.S.	~8 fold increase
M223A	N.S.	41.0% decrease	N.S.	N.S.	N.S.

3.3 Discussion

Currently, it is unclear how extracellular loops in Secretin-like GPCRs contribute to receptor functioning. A systematic alanine scan was conducted on ECL1 of CLR and its corresponding juxtamembrane regions. Mutant CGRP receptors were assessed on multiple criteria including; cAMP accumulation, cell surface expression, total expression, agonist mediated internalisation and CGRP radioligand binding. This pragmatic approach found that this region of the receptor is required for normal receptor pharmacology and has revealed key molecular determinants that facilitate receptor activation.

A cluster of three residues; L195, V198 and A199 that are located at the top of TM2 and predicted to face into the TM bundle are required for CGRP receptor functioning (see Figure 26). Agonist mediated internalisation and cAMP accumulation was severely decreased in all three residues. Competitive radioligand binding assays show that αCGRP binding was severely impaired in L195A and A199L and reduced, although not significantly, in V198A. The pharmacological phenotype of these mutant receptors can not be attributed to the inability of the receptor to reach the cell surface. L195A cell surface expression is comparable to WT whilst cell surface expression in A199L and V198A was significantly increased. The precise reason for this is speculative but it could partially be attributed to the observation that the receptors can not undergo agonist mediated internalisation, and therefore presumably can not undergo constitutive activation (basal level) internalisation resulting in more receptors at the surface. However, further experiments are required to clarify this hypothesis.

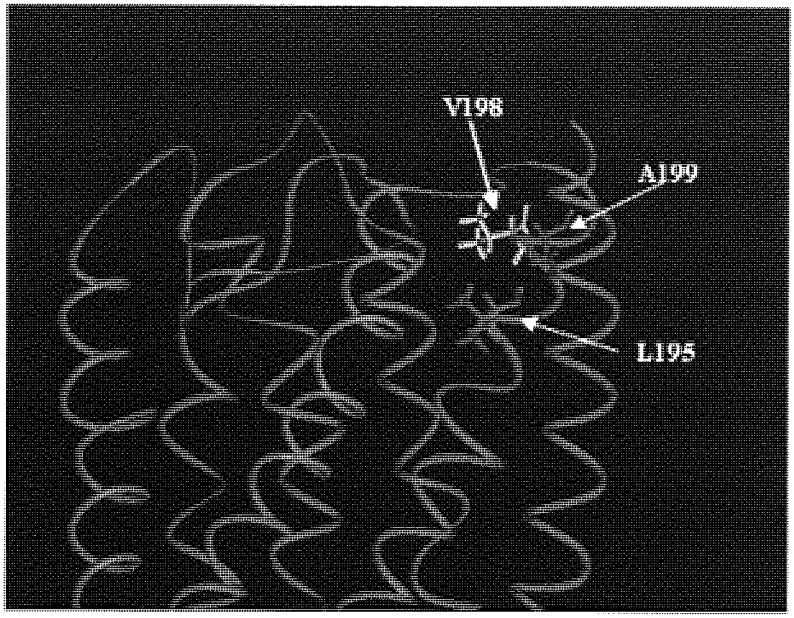


Figure 26. L195, V198 and A199 are predicted to reside at the top of the exofacial end of TM2. The assumed inactive CLR monomeric TM bundle model generated from the ground-state rhodopsin (PDB accession: 1U19) template. a) Side view of CLR TM bundle represented as a solid blue ribbon. The side chains are represented as sticks. L195 is highlighted green, V198 is highlighted yellow and A199 is highlighted red.

Intriguingly, the three residues are not highly conserved amongst the Secretin-like GPCRs. The consensus residue at position L195 is actually an aspartic acid, which is part of the highly conserved KD motif. The consensus residue at position V198 is a leucine and A199 is a tyrosine (see Figure 27). Previously, Langer *et al.*, (2003) found that mutating the KD motif reduced the ability of the VPAC1 receptor to stimulate cAMP production. This observation is in line with previous studies on the VPAC1 receptor (see Du *et al.*, 1997) and on other Secretin-like GPCRs, such as the secretin receptor (Di Paolo *et al.*, 1999). Although, primary sequence is not conserved these observations support the importance of this region in Secretin-like GPCRs.

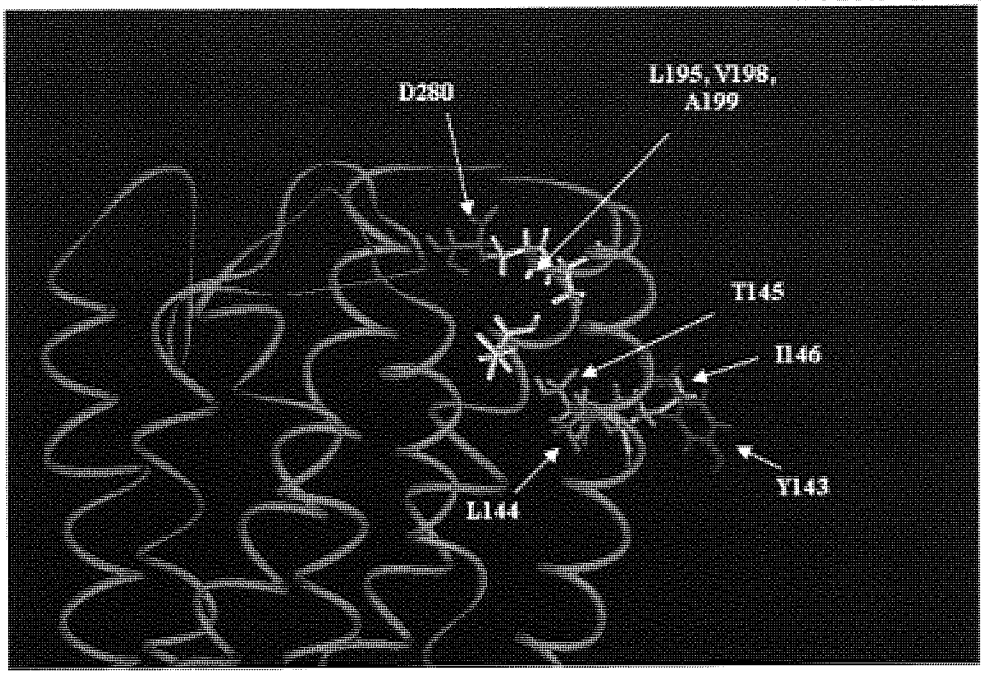
```
                 *  **
CLR         VTIIHLTAVANNQA
CRFR1       TWFVVQLTMSP--E
CRFR2       MWFLLQLVDH---E
CTR         IIIIHLVEVVPNGE
GIPR        AILSRDRLLPRPGP
GLP1R       SVFIKDAALKWMYS
GLP2R       AVLVKDVVFYNSYS
GLR         SVLVIDGLLRTRYS
GHRHR       AVFLKDAALFHSDD
SCTR        SNFIKDAVLFSSDD
VIPR2       SVLVKDDVLYSSSG
VIPR1       AVFIKDLALFDSGE
PACAPR      SVFIKDWILYAEQD
PTHR1       SIFVKDAVLYSGAT
PTHR2       SIFVKDRVVHAHIG
```

Figure 27. Human Secretin-like GPCR alignment of the C-terminal amino acids of the exofacial end of the TM2 domain. * represents residues of interest L195, V198 and A199. Taken from alignment of all human Secretin-like GPCRs using the Tcoffee server.

However, it remains difficult to specify the actual molecular mechanism in which the L195, V198 and A199 cluster contributes to receptor functioning. The triplet cluster may be directly involved in αCGRP binding or the mutations in this area may indirectly disrupt the orthosteric binding site or prevent signal transduction. It is predicted that these residues are located near the top of TM1 and in close proximity to ECL2. Specifically, the triplet cluster is in close proximity to D280 of ECL2 (see Figure 28), which has been found to significantly reduce cAMP production and CGRP binding when mutated to alanine (Conner *et al.*, see Chapter 2).

Photolabelling studies have been conducted on the CTR with the [Bpa^{19}] salmon calcitonin analogue (8-32), which was found to cross-link between residues C134-K141 located just above TM1 (Pham *et al.*, 2005). This evidence suggests that calcitonin-like ligands reside near TM1. More recently, Dong *et al.*, (2009) found that the potency of small-molecule agonists, which are specific to the CTR, is severely impaired when Y150, L151, A152 and I153 were deleted or mutated. These residues are predicted to residue in the exofacial end of TM1.The equivalent residues in the CLR are Y143, L144, T145 and I146 (see Figure 28). It is plausible that this space between ECL2, TM1 and TM2 may contribute to the orthosteric binding site and/or provide a molecular switch necessary for activation of calcitonin-like ligands including αCGRP.

a)



b)

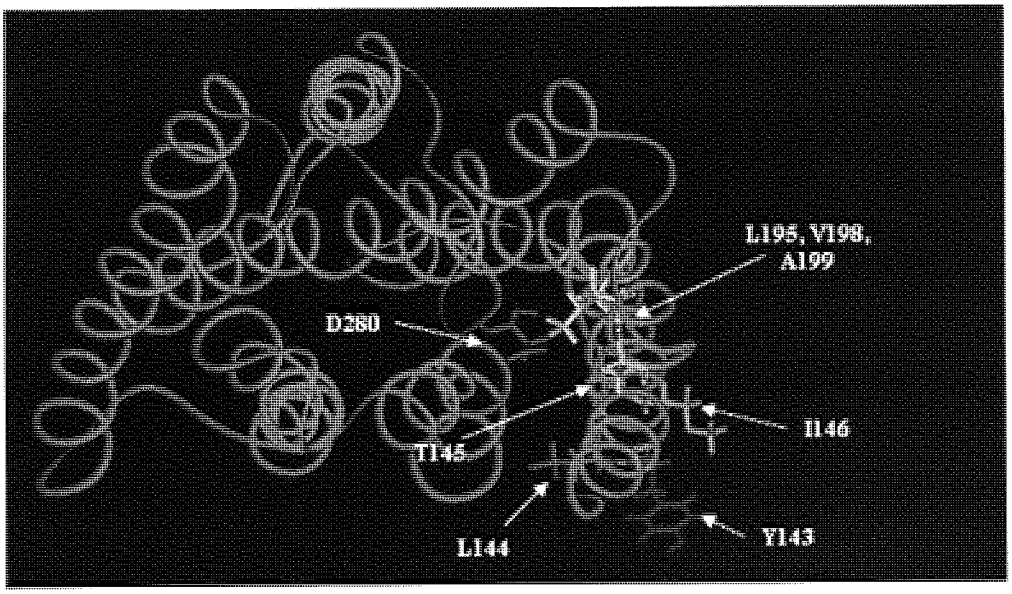


Figure 28. The predicted position of D280 relative to the triplet cluster L195, V198 and A199 within the CLR TM domain. Inactive CLR monomeric TM bundle model generated from the ground-state rhodopsin template (PDB accession: 1U19). a) Side view of CLR TM bundle represented as a solid cyan blue ribbon. Side chains are represented as sticks. L195, V198 and L198 triplet cluster is highlighted yellow, D280 is highlighted red and Y143, L144, T145 and I146 highlighted purple to pink. b) Extracellular view of CLR TM bundle. Colour scheme same as above.

The pharmacology of the CGRP receptor was severely impaired by C212A. C212 is predicted to be involved in the highly conserved disulphide bond with C282 located in ECL2. Consequently, disruption of this covalent bond would be expected to cause a loss of receptor functioning.

Interestingly, H219A was found to significantly impair cAMP production, αCGRP binding and cell surface expression. Yet, H219A readily underwent agonist mediated internalisation. H219 is predicted to reside on the third turn of TM3 facing into the bundle between TM2 and ECL2. H219 may directly be involved in αCGRP binding. However, the contrasting efficacy effects yielded by H219A, which is suggestive that β-arrestin binding is preferred over G-protein coupling, implies that H219 plays a role in signal transduction. L220A and L222A increased the potency of αCGRP in stimulating cAMP production and L222A was found to significantly enhance αCGRP binding.

H219A, L220A and L222A were found to effect αCGRP efficacy. Interestingly, all three residues are predicted to be located in the middle of TM3. The residues are predicted to be within the centre of the TM domain, an optimal position to contribute to signal transduction within the receptor (Figure 29). Visually comparing the inactive and active monomeric CLR TM bundles suggests that TM3 does not undergo a large rigid body movement but rather a subtle rotational movement upon activation. The rotational movement changes the relative positions of H219, L220 and L222. Consequently, H219, L220 and L222 may have a direct role in propagating receptor activation by contributing to the complex signal transduction process within the CGRP receptor. However, this intriguing hypothesis needs further investigation to be validated.

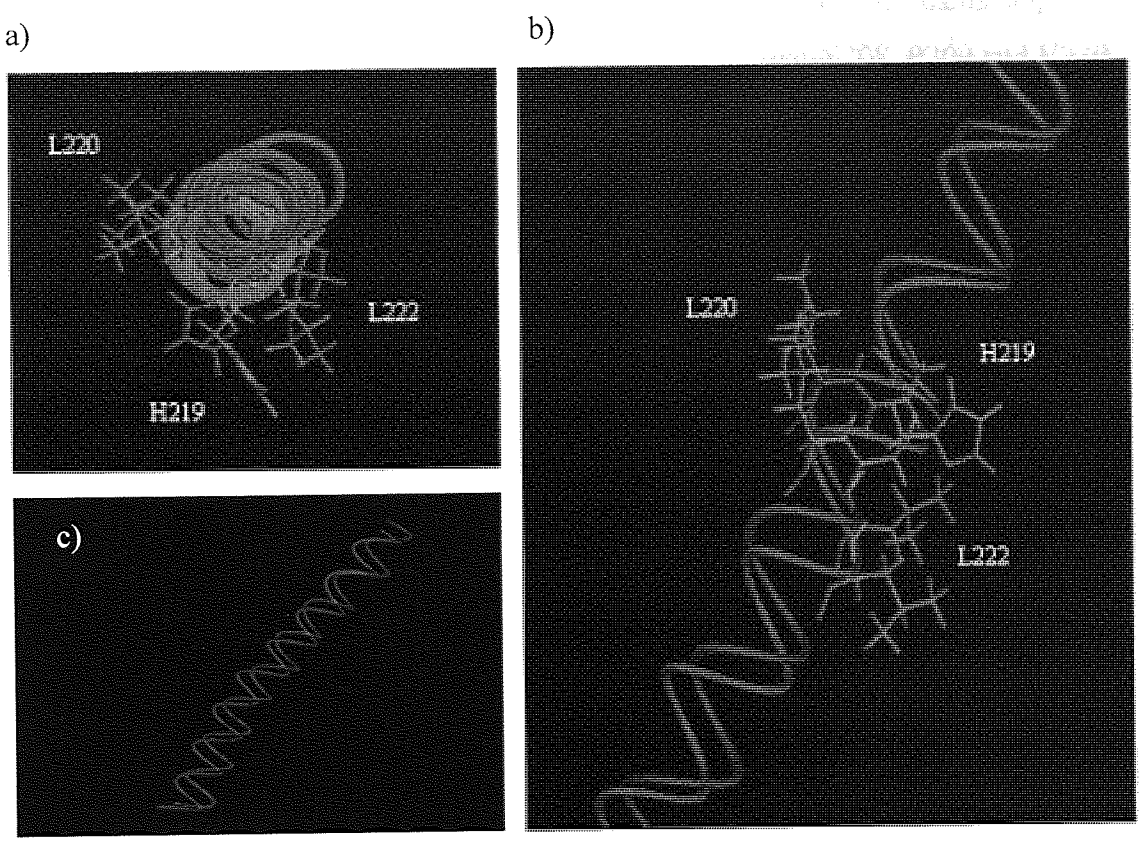


Figure 29. Comparing the predicted positions of H219, L220 and L222 within TM3 from the inactive and active CLR TM domain models. The assumed inactive CLR TM domain was generated using ground-state rhodopsin as a template (PDB accession: 1U19) and the assumed active CLR TM domain was generated from the acid opsin structure (PDB accession: 3DQB). a) Shows the extracellular view of TM3. The inactive TM3 helix is highlighted blue and the associated side chains of H219, L220 and L222 are highlighted pink. The active TM3 helix is highlighted red and the associated side chains of H219, L220 and L222 are highlighted green. b) Side view of TM3 showing the predicted orientations of H219, L220 and L222. Colour scheme same as above. c) Shows the side view of inactive and active TM3. Colour scheme same as above.

A203L and A206L were found to increase the potency of αCGRP at stimulating cAMP production. Both residues are predicted to reside in the loop region itself. The effects of these mutants are only subtle as binding assays suggest a minor increase in αCGRP affinity, albeit not significant. Yet, it is difficult to comprehend from the TM bundle model alone how these mutants could contribute to signal transduction. Consequently, it is reasonable to conclude that A203L and A206L may contribute directly or indirectly to the orthosteric binding site.

N208A was found to moderately increase agonist mediated internalisation. Moreover, L204A, P209A, V210A, Q216A and M223A were found to increase the mean basal

activity and the mean Emax implying that the native amino acid side chains help maintain the receptor in its inactive state. It is noteworthy that N208, P209 and V210 are located at the very top of TM3 and Q216 and M223 are located mid-way down TM3. Although the rigid body movements of TM3 are relatively small (as assessed by visually observing the inactive and active CLR TM bundles) it would be appealing to speculate that TM3 holds the CLR TM bundle in its inactive state and the subtle changes of the helix upon ligand docking helps to govern receptor signal transduction.

In conclusion, it is clear that this region of the CLR is essential in the normal functioning of the CGRP receptor. Novel molecular determinants have been found that enhance and impair both the affinity and efficacy of the receptor. This work could be used in the future to aid molecular model building and rational drug design.

Chapter 4: ECL3 and the associated regions of TM6 and TM7 in CLR are important in CGRP receptor pharamacology

4.1 Introduction

Lawson and Wheatley (2004) described the third extracellular loop (ECL3) of Rhodopsin-like GPCRs *'as more than just a protein bridge linking TM6 and TM7'*. However, due to the little or no size and sequence conservation across the GPCR superfamily the only defining feature of this loop is that it connects TM6 and TM7. Yet, it is thought that GPCR activation involves rigid body movements of TM6 and TM7 (Schwartz *et al.,* 2006). Consequently ECL3 may determine the range of flexibility in this region of the receptor and may assist or even govern receptor activation.

There is growing interest in the architecture and function of ECL3 in human Secretin-like GPCRs after Dong *et al.* (2006) hypothesised that a 'hidden agonist' sequence within the N-terminal ECD, referred to as the WDN motif (although the equivalent is a WDG motif in the CLR), could potentially interact with ECL3 and/or associated transmembrane domains to induce activation in the secretin receptor. Yet, the implications for other Secretin-like GPCRs, including CLR, remain questionable (see Labuthe *et al.,* 2007). In contrast, the more widely quoted 'two-domain' model of activation for Secretin-like ligands suggests that the N-terminal region of the agonist penetrates into the transmembrane region of the receptor to induce activation (see Grace *et al.,* 2004). In line with this latter model, ECL3 may contain the molecular determinants, which facilitate agonist selectivity and receptor activation.

Direct experimental evidence that ECL3 contributes to Secretin-like GPCR activation is currently sparse. Photoaffinity labelling and mutagenesis studies on rat and human parathyroid receptors have identified important epitopes located in and around ECL3 (Lee *et al.,* 1995 and Bisello *et al.,* 1998). More recently, a novel disulphide trapping approach, which has considerable advantages over conventional photoaffinity labelling strategies, has been used to map out the interaction site/s of parathyroid hormone (PTH) and its cognate receptor (see Monaghan *et al.,* 2008 and Chapter 1 for more details). The results suggested that the first residue of PTH (Ser-1) was found to be close to the juxtamembrane region of TM6.

In an attempt to discern whether ECL3 of CLR is important in CGRP receptor pharmacology a systematic alanine/leucine scan has been conducted on ECL3 and its corresponding juxtamembrane regions.

4.2 Results

4.2.1 Stimulation of cAMP production

Each mutant was challenged with human αCGRP and cAMP production was measured (see table 16). E357A was found to significantly decrease EC50 by ~33 fold compared to WT. The mean basal activity of E357A was found to slightly increase by 18.3% ± 12.0% but the mean Emax was found to decrease by 41.73% ± 17.9% compared to WT. I360A was also found to significantly decrease EC50 by ~7 fold compared to WT but basal activity and Emax resembled WT. L351A was not found to significantly impair EC50 but the mean Emax was found to have decreased by 28.9% ± 5.9%, while the basal activity resembled WT (see Figure 30)

The mean pEC50 values of V350A, I352A and R355A, were not found to be significantly different to WT. However, an increase of 20% or higher was found in the mean basal activity and the mean Emax in each of these mutants (see Table 17 and Figure 31). The mean Emax of F349A, K359A, A361L, Y367A and M369A was also found to have increased by 20% or more and although an increase in basal activity at varying magnitudes was noted in each mutant, the basal activity of the mutants was not found to surpass the 20% threshold. Furthermore, the mean basal activity of Y365A was found to have increased above 20% and the Emax had increased approaching 20% (see Table 17 and Figure 31).

Table 16. Summary of pEC50 values for ECL3 mutant receptors compared to WT. Values are pEC50 means ± S.E.M. $p < 0.05$ is represented by *. $p < 0.01$ is represented by ** and $p < 0.001$ is represented by ***. pEC50 mutant values were compared to WT using an independent two-tailed t-test.

Mutant	N	pEC50 WT (mean ± S.E.M)	pEC50 Mutant (mean ± S.E.M)
F349A	4	10.59 ± 0.34	10.72 ± 0.28
V350A	3	9.89 ± 0.02	9.74 ± 0.15
L351A	3	9.36 ± 0.16	9.54 ± 0.04
I352A	6	10.00 ± 0.36	9.33 ± 0.15
P353A	5	9.79 ± 0.27	9.06 ± 0.21
W354A	4	9.08 ± 0.36	8.91 ± 0.28
R355A	4	9.86 ± 0.37	9.55 ± 0.21
P356A	3	10.39 ± 0.42	9.95 ± 0.26
E357A	4	10.44 ± 0.46	8.92 ± 0.39**
G358A	3	9.68 ± 0.31	9.46 ± 0.29
K359A	4	9.76 ± 0.12	9.94 ± 0.32
I360A	6	9.26 ± 0.14	8.40 ± 0.22**
A361L	7	9.97 ± 0.22	10.27 ± 0.22
E362A	3	9.70 ± 0.26	9.65 ± 0.33
E363A	4	10.32 ± 0.49	9.80 ± 0.32
V364A	3	9.78 ± 0.22	9.94 ± 0.15
Y365A	4	10.20 ± 0.27	10.33 ± 0.25
D366A	3	9.54 ± 0.21	9.72 ± 0.21
Y367A	4	9.64 ± 0.21	9.92 ± 0.04
I368A	5	10.13 ± 0.33	10.30 ± 0.22
M369A	4	9.92 ± 0.32	10.53 ± 0.39
H370A	5	9.42 ± 0.14	9.19 ± 0.30
I371A	4	9.96 ± 0.28	10.18 ± 0.20
L372A	3	9.93 ± 0.07	10.08 ± 0.26
M373A	3	10.51 ± 0.28	10.21 ± 0.12

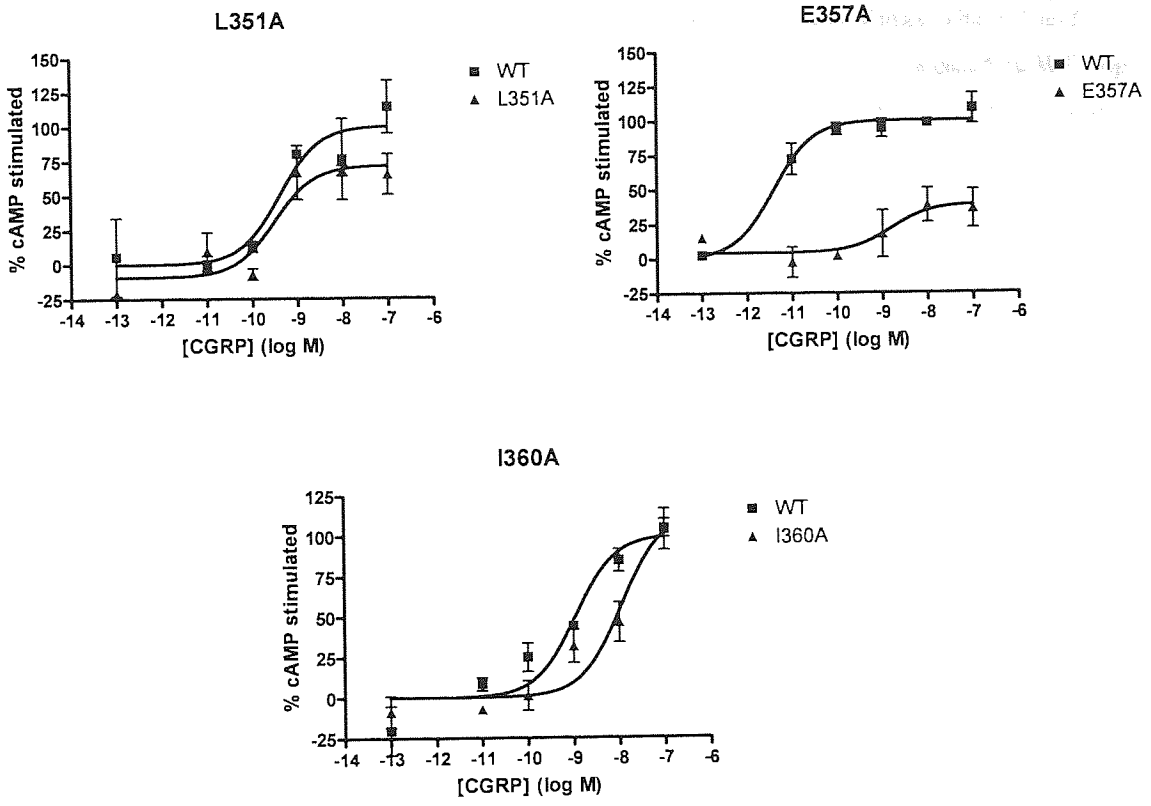
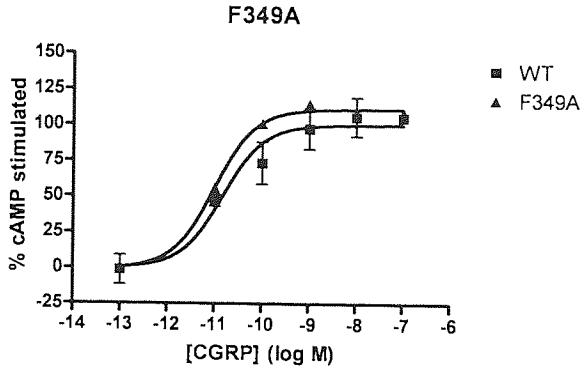
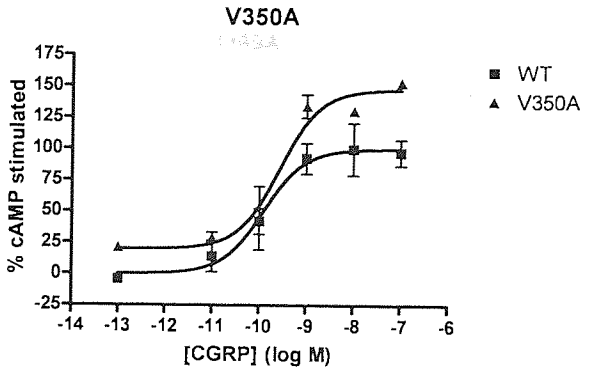


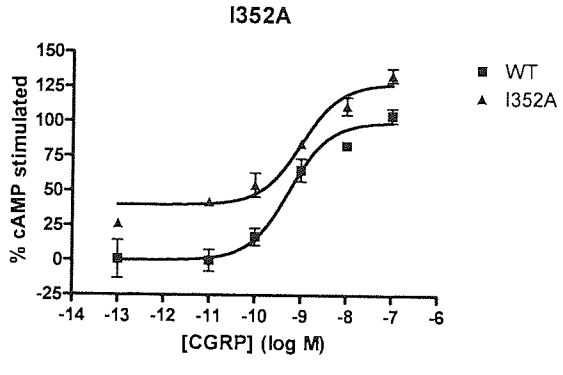
Figure 30. Representative dose-response curves of ECL3 mutants that showed a significant decrease in αCGRP potency and Emax. Sigmoidal concentration-response curves comparing the WT receptor and mutant receptors (L351A, E357A and I360A) capability of stimulating cAMP after αCGRP exposure are shown. Each WT and mutant receptor concentration-response comparison curve is a representative example of an ensemble of at least three independent experiments. Each assay point was performed in duplicate where each point on the graph represents the mean ± S.E.M. The line of best fit for each concentration-response curve was fitted with GraphPad Prism 4 see Equation 1. cAMP dose-response curves for L351A, E357A and I360A compared to WT are shown.

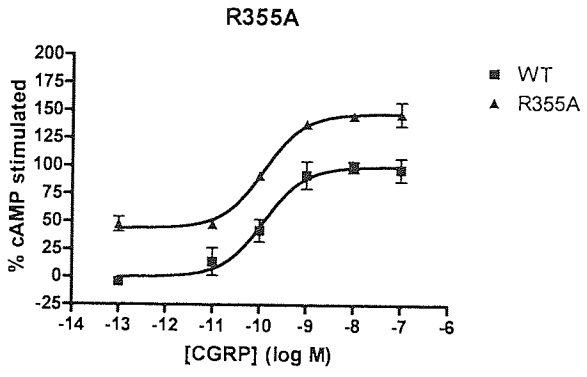
Table 17. ECL3 mutant receptors found to increase basal activity and/or Emax. The WT and mutant cAMP dose-response comparison curves were normalised from 0% to 100% based on WT Top and Bottom values generated by GraphPad Prism 4. The Emax and basal activity of each mutant was assessed by the mean of the top and bottom of each dose-response curve. The mutant receptors mean Emax and basal activity was expressed as % WT normalisation. A meaningful difference was noted if the mean size of effect differed by 20% or more. Values reported are mutant percent means ± S.E.M in relation to WT, determined from at least three independent experiments.

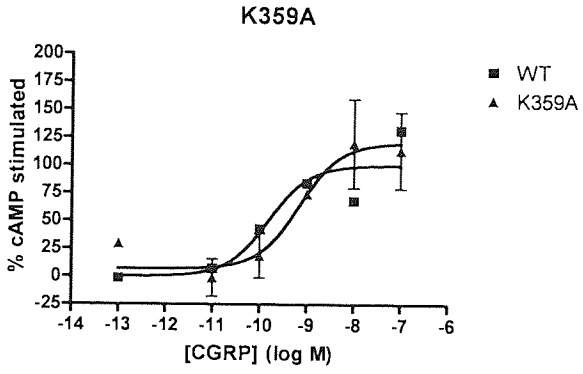
Mutant	N	Basal activity (% WT mean ± S.E.M)	Emax (% WT mean ± S.E.M)
F349A	4	10.3 ± 6.0	122.1 ± 6.8
V350A	3	19.3 ± 5.8	142.3 ± 14.6
I352A	6	25.8 ± 5.3	121.1 ± 8.9
R355A	4	39.7 ± 11.2	151.5 ± 10.9
K359A	4	12.6 ± 14.6	121.7 ± 11.1
A361L	7	15.6 ± 8.2	146.2 ± 12.7
E363A	4	7.6 ± 9.5	127.5 ± 6.5
Y365A	4	22.5 ± 1.3	119.7 ± 6.0
Y367A	4	16.9 ±11.2	123.2 ± 6.0
M369A	4	19.2 ± 6.7	139.5 ± 12.9

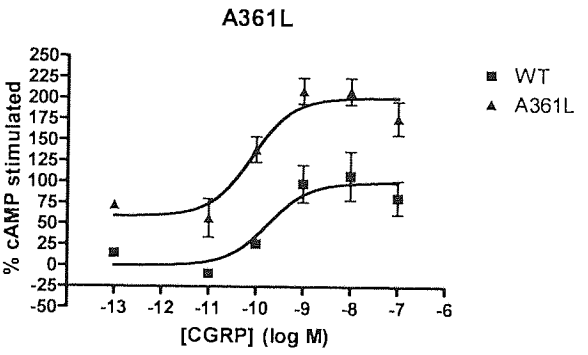

F349A
% cAMP stimulated
[CGRP] (log M)
WT
F349A


V350A
% cAMP stimulated
[CGRP] (log M)
WT
V350A


I352A
% cAMP stimulated
[CGRP] (log M)
WT
I352A


R355A
% cAMP stimulated
[CGRP] (log M)
WT
R355A


K359A
% cAMP stimulated
[CGRP] (log M)
WT
K359A


A361L
% cAMP stimulated
[CGRP] (log M)
WT
A361L

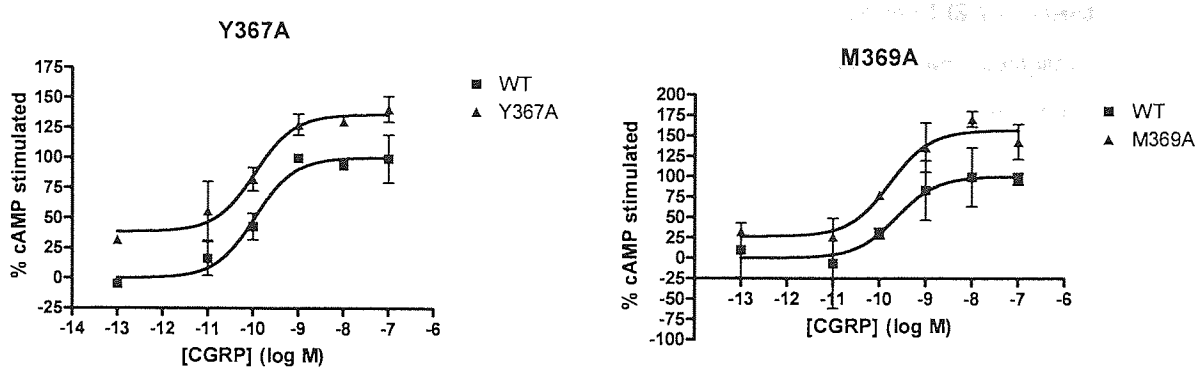


Figure 31. Representative dose-response curves of ECL3 mutants that showed an increase in basal activity and Emax. Sigmoidal concentration-response curves comparing the WT receptor and mutant receptors (F349A, V350A, I352A, R355A, K359A, A361L, Y367A and M369A) capability of stimulating cAMP after αCGRP activation are shown. Each WT and mutant receptor concentration-response comparison curve is a representative example of an ensemble of at least three independent experiments. Each assay point was performed in duplicate where each point on the graph represents the mean ± S.E.M. The line of best fit for each concentration-response curve was fitted with GraphPad Prism 4 see Equation 1.

4.2.2 Cell surface receptor expression

Expression of receptors was measured using a cell-surface ELISA (Table 18). Out of 25 mutants 12 were found to significantly reduce cell-surface expression compared to WT; L351A, W354A, R355A, E357A, G358A, K359A, E362A, Y365A, D366A, Y367A and I370A, respectively. The largest reduction was observed in E357A (a decrease of 89.0% compared to WT), closely followed by L351A (72.8% compared to WT). For the other mutant receptors highlighted above, the decrease in expression was less than 50% (see Table 18).

In contrast, an increase in cell surface expression was found for H370A (an increase of 41.1% compared to WT) and P353A (33.3% compared to WT).

Table 18. ECL3 mutant receptor cell surface expression. Cell surface expression ELISA was used to probe for the presence of the HA epitope. Mutant HA CLR/*myc* RAMP1 receptors were compared with WT HA CLR /*myc* RAMP1 receptors. 3-6 independent experiments that contained triplicate data points were used in analysis. The raw data for each independent experiment was normalised where the mean WT receptor cell-surface expression equalled a 100% and the mean negative control (*myc* RAMP1/ empty pcDNA3.1-) was equal to 0%. Values reported are mutant means ± S.E.M (% WT). Mutant cell surface expression was compared to WT receptor using a Mann-Whitney U test ($p < 0.05$ is represented by *, $p < 0.01$ represented by **, and $p < 0.001$ represented by ***).

Mutant	Cell surface expression (% WT)	Mutant	Cell surface expression (% WT)
F349A	69.0 ± 7.0*	E362A	83.8 ± 5.8*
V350A	78.9 ± 5.9	E363A	89.6 ± 11.9
L351A	27.2 ± 4.4***	V364A	135.3 ± 18.5
I352A	70.5 ± 4.6***	Y365A	60.0 ± 8.7**
P353A	133.3 ± 11.8**	D366A	85.9 ± 4.9*
W354A	88.1 ± 4.7*	Y367A	56.6 ± 4.9***
R355A	66.2 ± 5.5**	I368A	129.0 ± 11.5
P356A	85.0 ± 6.0	M369A	99.9 ± 4.8
E357A	11.0 ± 3.5***	H370A	141.1 ± 9.5***
G358A	61.4 ± 11.2*	I371A	70.5 ± 7.6**
K359A	71.8 ± 5.8***	L372A	102.3 ± 13.0
I360A	107.7 ± 8.0	M373A	95.1 ± 7.1
A361L	71.3 ± 11.0		

4.2.3 Total receptor expression

Mutant receptors that were either found to have a significantly different mean pEC50 as assessed by cAMP accumulation or a significantly different cell surface expression were further analysed for total receptor expression (Table 19). Overall the results show that the mutant cDNA was successfully transfected, transcribed and translated. Yet, a significant 44.0% increase in total expression compared to WT was observed for P353A. Furthermore, a significant but modest decrease (21.8%) in total expression was observed for D366A.

Table 19. Summary of ECL3 mutant receptor total expression. Total expression of HA-tagged receptors both mutant and WT were analysed when co-transfected with *myc* RAMP1. At least 3 independent experiments containing triplicate data points were used in analysis. Total expression in the mutant condition was normalised to the WT condition (equal to 100%) and negative control (*myc* RAMP1/empty pcDNA3.1-, after 0.1% Triton X 100, which was equal to 0%). A Mann-Whitney U test was used to assess statistical difference between WT and mutant receptors were *, **, *** represent $p < 0.05$, 0.01 and 0.001 significance levels.

Mutant	Total expression (% WT mean ± S.E.M)	Mutant	Total expression (% WT mean ± S.E.M)
L351A	111.0 ± 12.4	A361L	87.2 ± 4.7
I352A	98.9 ± 10.3	Y365A	116.0 ± 15.2
P353A	144.0 ± 12.3 ***	D366A	78.2 ± 4.8**
W354A	103.5 ± 5.3	Y367A	70.42 ± 5.8***
R355A	110.9 ± 10.9	I368A	104.1 ± 10.9
E357A	88.6 ± 7.5	H370A	144.3 ± 6.8
G358A	92.9 ± 4.2	I371A	126.9 ± 12.7
I360A	89.5 ± 2.9		

4.2.4 Agonist mediated internalisation

αCGRP mediated internalisation was severely impaired by E357A and significantly reduced in L351A, P353A and W354A. In contrast, E362A and I371A were found to internalise more readily than WT (see Table 20).

Table 20. Summary of agonist (αCGRP) mediated internalisation on ECL3 mutant receptors. Agonist mediated internalisation of the CGRP receptor (both WT and mutant receptors) was approximated by a HA epitope probing cell surface ELISA taking into account the difference in cell surface expression levels between CGRP receptors that have or have not been exposed to 100nM of human αCGRP for an 1hr. Percent mean ± S.E.M agonist mediated internalisation was determined by 3-6 independent experiments containing 3 replicates. A Mann Whitney U test was used to compare mutant and WT percent agonist internalisation values where $p < 0.05$ is respresented by *, $p < 0.01$ is represented by **, and $p < 0.001$ is represented by ***.

Mutant	WT receptor internalisation (% mean ± S.E.M)	Mutant receptor internalisation (% mean ± S.E.M)
F349A	70.16 ± 4.8	62.83 ± 3.5
V350A	61.22 ± 8.2	48.00 ± 6.7
L351A	57.34 ± 4.0	19.84 ± 7.5***
I352A	60.98 ± 3.8	55.03 ± 6.2
P353A	74.76 ± 1.8	58.80 ± 1.8***
W354A	73.25 ± 1.8	57.26 ± 2.7***
R355A	57.71 ± 6.7	60.95 ± 6.9
P356A	65.78 ± 5.7	76.63 ± 4.8
E357A	57.13 ± 3.7	8.589 ± 6.3***
G358A	60.09 ± 6.7	73.04 ± 10.2
K359A	52.46 ± 2.7	59.11 ± 4.7
I360A	71.00 ± 2.2	71.55 ± 2.2
A361L	78.82 ± 6.7	66.22 ± 10.7
E362A	49.11 ± 2.1	58.16 ± 2.3**
E363A	68.34 ± 4.2	67.59 ± 2.8
V364A	68.87 ± 9.2	59.00 ± 7.2
Y365A	68.69 ± 4.8	76.69 ± 4.0
D366A	48.38 ± 2.6	51.89 ± 5.2
Y367A	57.35 ± 4.2	60.54 ± 5.8
I368A	64.42 ± 8.3	73.82 ± 3.5
M369A	64.77 ± 4.4	67.48 ± 2.8
H370A	56.94 ± 6.3	51.04 ± 5.2
I371A	56.49 ± 6.0	82.69 ± 3.4**
L372A	53.99 ± 4.4	62.33 ± 5.4
M373A	66.09 ± 3.7	62.70 ± 2.3

4.2.5 Inhibition of ^{125}I-hCGRP radioligand binding

Inhibition of ^{125}I-hCGRP radioligand binding assays were performed on mutant receptors that were found to have significantly different mean pEC50 cAMP dose-response curves and/or agonist mediated internalisation compared to WT receptors (see Table 21 and Figure 32). The mean pIC50 values of CGRP for two mutants were significantly reduced; L351A and E357A. Note, that I360A was approaching significance (p= 0.052). Contrastingly, the mean pIC50 of I371A showed a small increase compared to WT.

Table 21. Apparent affinities of αCGRP for ECL3 mutant receptors, estimated by inhibition of radioligand binding. Mean ± S.E.M pIC50 WT and mutant values shown. An independent two-tailed t-test was used to assess statistical differences where $p < 0.05$ is represented by *. N.M.B stands for no measurable binding.

Mutant	N	Pic50 WT	pIC50 Mutant
L351A	4	9.04 ± 0.36	N.M.B
P353A	7	9.18 ± 0.33	8.65 ± 0.24
E357A	3	9.84 ± 0.03	N.M.B
I360A	3	8.77 ± 0.09	7.83 ± 0.33
W354A	4	9.01 ± 0.12	8.75 ± 0.13
E362A	3	9.05 ± 0.04	8.74 ± 0.18
I371A	3	9.84 ± 0.03	10.16 ± 0.11*

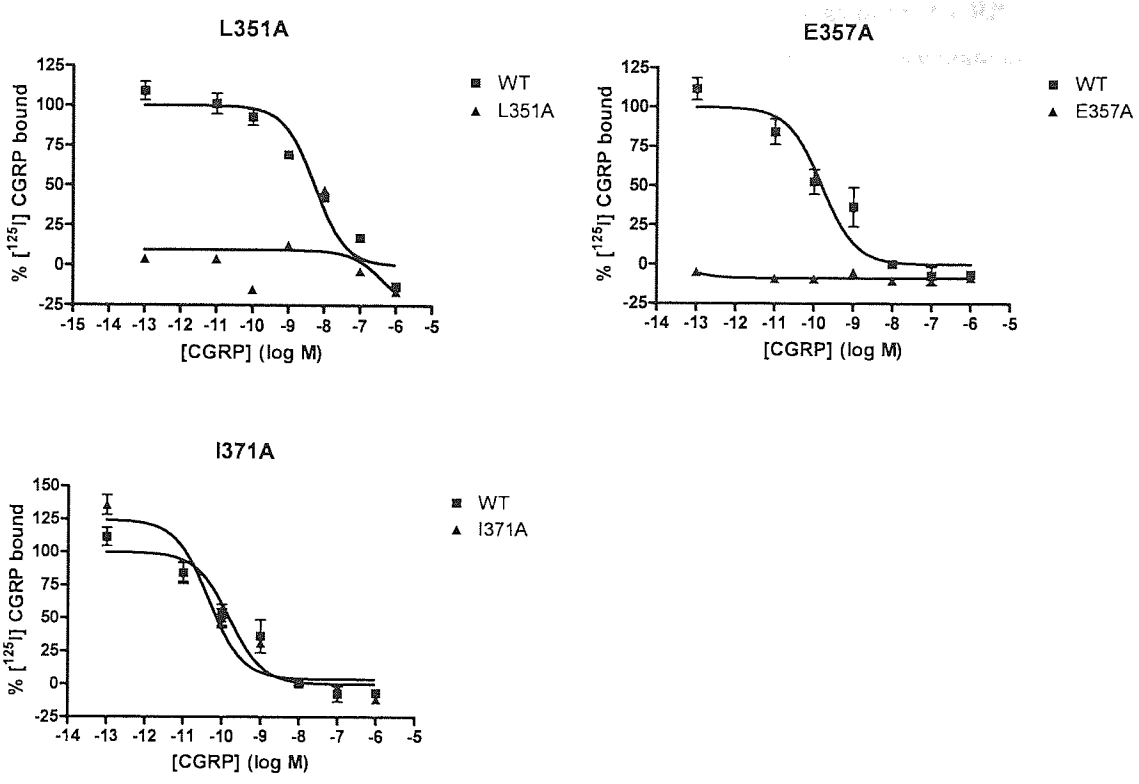


Figure 32. Inhibition of ^{125}I-hCGRP radioligand binding curves of ECL3 mutant receptors found to impair CGRP binding. Representative sigmoidal αCGRP inhibition curves of L351A, E357A and I371A compared to the WT receptor. The receptors capability of binding the radioligand ^{125}I-hCGRP (~20pM) was assessed aswell as the ability of the non-radioactive αCGRP to displace the radioactive ligand over six log concentrations. Each WT and mutant receptor curve is a representative example of an ensemble of at least three independent experiments. Each assay point was performed in duplicate where each point on the graph represents the mean ± S.E.M. The line of best fit for each comparisin curve was fitted with GraphPad Prism 4 see Equation 1. The concentration of COS-7 cell membranes was 2mg/ml.

4.2.6 Summary of important ECL3 mutations

ECL3 mutations that were found to significantly alter the potency of αCGRP to stimulate cAMP (pEC50), cell-surface expression, total expression, agonist-mediated internalisation and/or CGRP binding (pIC50) are summarised in table 22.

Table 22. The ECL1 mutations that significantly altered the pharmacology of the CGRP receptor. EC50 fold effects and IC50 fold effects were determined by sigmoidal concentration-response curves comparing the WT receptor and mutant receptors capability of either stimulating cAMP after αCGRP exposure or the inhibition of ^{125}I-hCGRP binding, respectively. Both the EC50 and IC50 fold effects represent the mean fold difference from at least three independent experiments. A two-tailed independent t-test was used to assess statistical differences. Cell surface expression, total expression and αCGRP mediated internalisation values represent the mean difference in % WT. A Mann Whitney U test was used to test for statistical differences between WT and mutant receptors on these three parameters, which were determined by at least three independent experiments. Mutants that were found not to be significantly different from WT is represented by N.S. Experiments that have not been conducted are represented by – and N.M.B stands for no measurable binding.

Mutant	EC50 fold effects after αCGRP stimulation (compared to WT receptors)	Cell surface expression (% WT mean difference)	Total expression (% WT mean difference)	αCGRP mediated internalisation (% WT mean difference)	IC50 fold effects (compared to WT receptors)
F349A	N.S.	31.0% decrease	N.S.	N.S.	-
L351A	N.S.	72.8% decrease	N.S.	65.4% decrease	N.M.B
I352A	N.S.	29.5% decrease	N.S.	N.S.	-
P353A	N.S.	33.3% increase	44.0% increase	21.3% decrease	N.S.
W354A	N.S.	11.9% increase	N.S.	21.8% decrease	N.S.
R355A	N.S.	33.3% decrease	N.S.	N.S.	-
E357A	~33 fold decrease	89.0% decrease	N.S.	85.0% decrease	N.M.B
G358A	N.S.	38.6% decrease	N.S.	N.S.	-
K359A	N.S.	28.2% decrease	N.S.	N.S.	-
I360A	~7 fold decrease	N.S.	N.S.	N.S.	N.S.
E362A	N.S.	16.2% decrease	N.S.	18.4% increase	N.S.
Y365A	N.S.	40.0% decrease	N.S.	N.S.	-
D366A	N.S.	14.1% decrease	21.8% decrease	N.S.	-
Y367A	N.S.	43.4% decrease	29.6% decrease	N.S.	-
H370A	N.S.	41.1% decrease	N.S.	N.S.	-
I371A	N.S.	29.5% decrease	N.S.	46.4% increase	~2.1 fold increase

4.3 Discussion

An alanine/leucine scan was conducted on ECL3 and the juxtamembrane regions located at the top of TM6 and TM7. Novel residues have been identified that influence CGRP receptor pharmacology. An unanticipated observation was the number of mutants that seemed to reduce cell surface expression of the CGRP receptor. Total expression of the mutant receptors is broadly comparable to WT expression levels. Therefore, factors such as transfection, translation and transcription efficiency are probably not responsible for the observed differences.

An interesting computational study conducted by Vohra *et al.,* (2007) used an evolutionary trace method to predict the dimerisation interfaces of Secretin-like GPCRs. The results suggested that TM4 and TM6 are the most likely candidates to be involved in oligomerisation. Since this publication it has become clear that TM4 is the dimerisation site in SCTR (Gao *et al.,* 2009). The stoichiometry of the CGRP receptor is ambiguous. However, if the CGRP receptor is an asymmetric complex consisting of 2 CLR:1 RAMP1 as Heroux *et al.,* (2007) suggested it is plausible to hypothesise that the CLR-CLR interface would also be TM4 since the key residues (Gly-243 and Ile-247) needed for SCTR homodimerisation are highly conserved across the family including CLR. Recently, Harikumar *et al.,* (2009) found that RAMP3 dimerised with the SCTR at TM6 and TM7. Furthermore, within this paper Harikumar and colleagues debate whether this interface could be shared in CLR but do suggest an alternative site between TM1 and TM2, although the rationale behind this alternative interface is not discussed.

Although, it remains speculative it is interesting to note that L351A, R355A, G358A, K359A, E362A, Y365A, Y367A and I371A, which were found to have a decrease in cell surface expression are all predicted to be located across TM6, ECL3 and TM7 and their side chains are predicted to face outwards toward the lipid environment (Figure 33). RAMP1 is a chaperone protein and its association to CLR is essential for CGRP receptor trafficking to the cell surface (McLatchie *et al.,* 1998). Consequently, the mutations identified could theoretically participate in a RAMP1 interface disrupting the efficiency of CLR and RAMP1 dimerisation (see Figure 33). In particular, L351 located two turns down in TM6 is essential for normal CGRP

receptor cell surface expression, which in turn when mutated to alanine reduces the Emax of the receptor and prevents high affinity αCGRP binding.

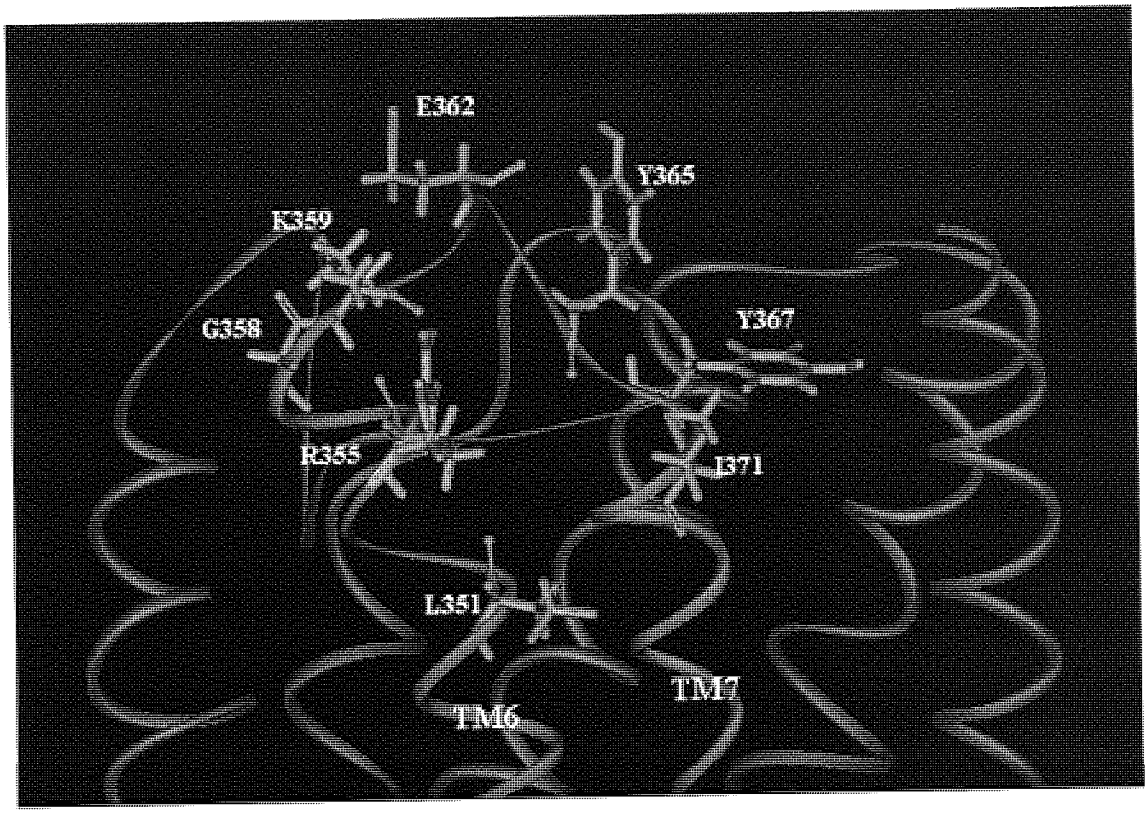


Figure 33. Residues within TM6, ECL3 and TM7 predicted to face the lipid environment and reduce cell surface expression. Side view of the inactive CLR TM bundle (cyan blue). Residues that were found to reduce cell surface expression and predicted to face the lipid environment (highlighted orange).

Mobilisation of TM6 and TM7 is thought to be required for GPCR activation. The notion that the RAMP binding site is between TM6 and TM7 is suggestive that the RAMP may assist and/or control the movement of this region. For example, selective modulation of receptor signalling has been found in VPAC1/RAMP2 complexes (Christopoulos *et al.*, 2003). Interestingly, R355A, K359A, Y365A and Y367A which were found to decrease cell surface expression by ~30% or more are also found to increase basal activity and/or Emax. It is tempting to speculate that the innate residues, which upon mutation increase the basal activity, may help constrain the CGRP receptor in an inactive conformation, a process that could be mediated between CLR and RAMP1 interactions. To validate this hypothesis further investigation is clearly needed to probe this putative interface and to rule out other possibilities for the decrease in cell surface expression observed in the mutations highlighted in Figure 30

e.g. a global decrease in CLR stability. In Secretin-like GPCRs, ECL3 and their cognate ECDs have been predicted to be in close vicinity of each other which was initially suggested to induce activation (Dong *et al.,* 2006). Yet, this association between ECL3 and the ECD may provide additional stabilising interactions needed to maintain the integrity of the receptor complex. Clearly, a multifaceted approach similar to Harikumar *et al.,* (2009), which triangulated data from a number of approaches including confocal fluorescence microscopy, peptide competition assays, BRET assays and functional assays, would be optimal in determining whether the CLR TM6/TM7 interface does accommodate RAMP1.

The mutation E357A severely impairs CGRP receptor function, primarily caused by its inability to be expressed at the cells surface. The inactive TM bundle model suggests that the side chain of E357 does not face the lipid environment but inwards toward ECL2. Specifically, E357 is predicted to make a hydrogen bond contact with T288 located in ECL2 (Figure 34). Interestingly, T288A has previously been found to reduce cAMP production and CGRP affinity (see Chapter 2). However, this interaction is not apparent in the active TM bundle. The reason for the large reduction in cell surface expression of E357A remains ambiguous, it maybe attributed to misfolding of protein or inability of the protein to insert itself into the membrane bilayer. Clearly, the putative stabilising interaction between E357 and T288 needs further investigation, initially with reciprocal mutations to see if receptor functioning is recovered.

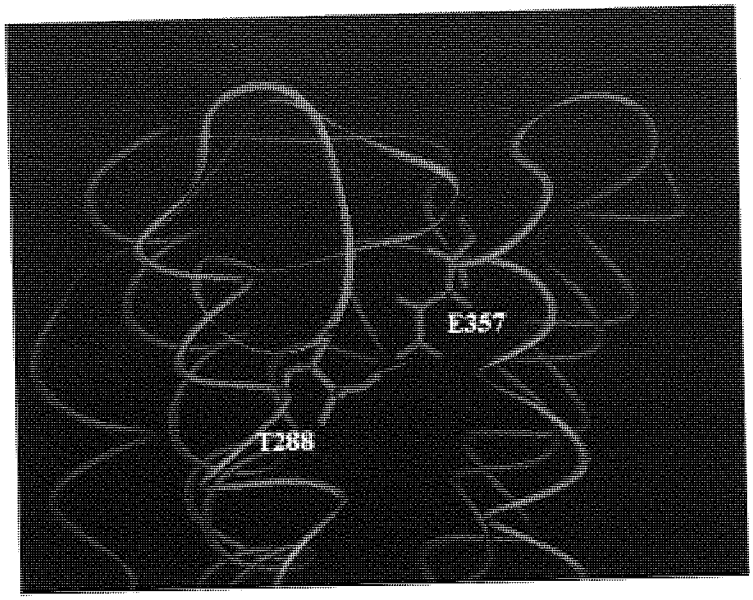


Figure 34. The putative hydrogen bond between T288 and E357. Side view of the assumed inactive CLR TM bundle based on ground-state rhodopsin (PDB accession: 1U19). T288 is red, whereas E357 is purple. The putative hydrogen bond is depicted as a green line.

The I360A significantly reduced the potency of αCGRP in evoking a cAMP response and a reduction in αCGRP affinity was observed, albeit not significant. I360 is predicted to be in the centre of ECL3 and its side chain is orientated towards ECL2. The inactive CLR TM model predicts the side chain of I360 is in close proximity to Y277 and Y278 in ECL2. Interestingly, Dr. A. Conner (see Chapter 2) found that Y277A impairs cAMP production but αCGRP binding is comparable to WT. However, Y278A significantly reduced both cAMP production and CGRP affinity, suggestive that Y277 is involved in signal transduction whereas Y278 may have a direct role in αCGRP recognition. The inactive CLR TM bundle model predicts that Y278 is orientated so it is above Y277 and Y277 faces downwards toward the TM bundle (Figure 35).

In contrast, the active CLR TM model suggests that the relative position of I360 shifts so that it is in close proximity to Y292, this hypothesis has not yet been assessed by our group (see Figure 35). It is reasonable to speculate that I360 contributes either directly to the αCGRP binding site or contributes indirectly by stabilising ECL2 (see Figure 35).

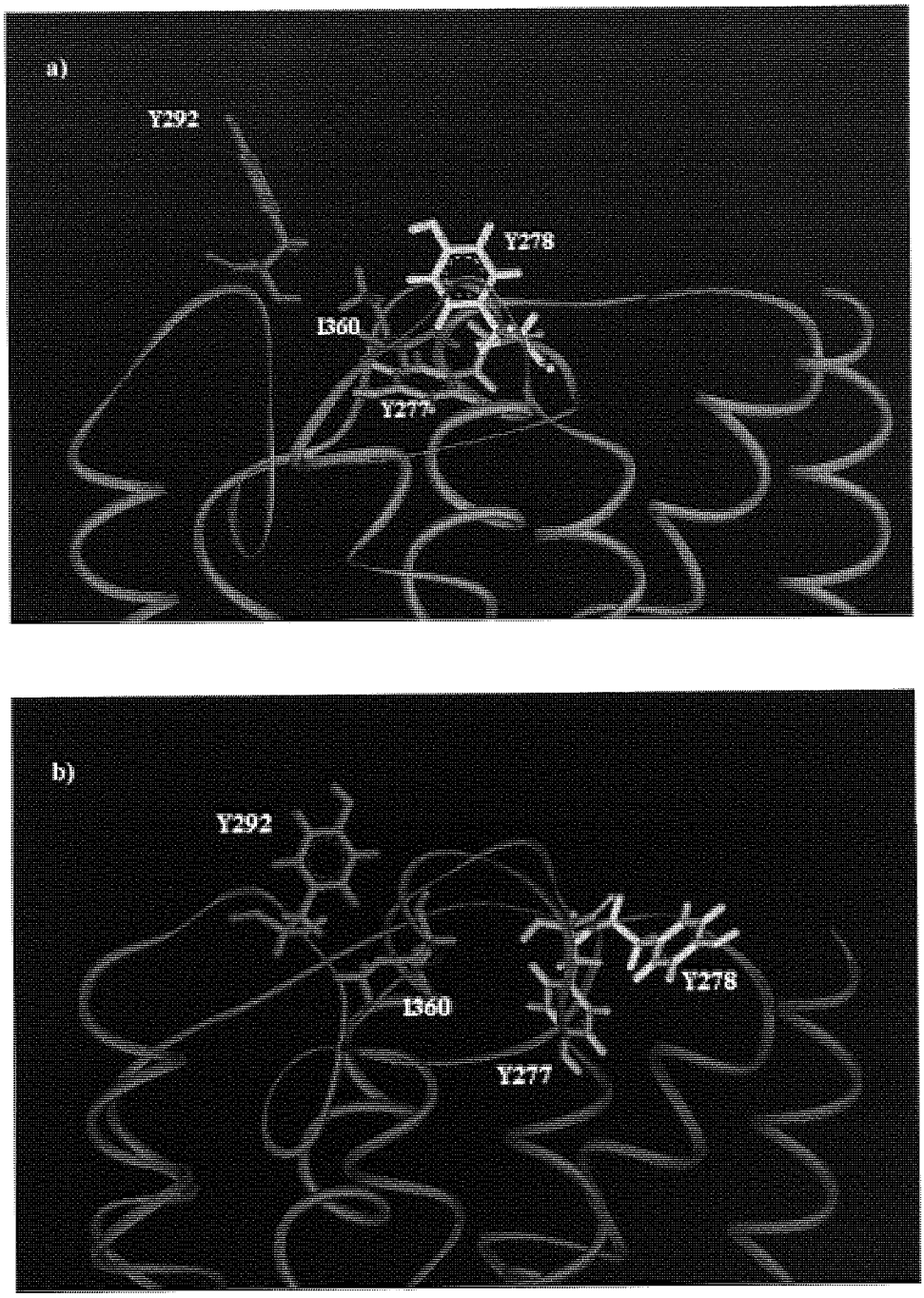


Figure 35. The relative position of I360 within the TM bundle. a) Side view of assumed inactive CLR TM bundle (cyan ribbon) based on ground-state rhodopsin (PDB accession: 1U19) showing position of I360 (red) relative to Y277 (orange), Y278 (yellow) and Y292 (green). b) Side view of assumed active CLR TM bundle (red ribbon) based on the PDB accession 3DQB showing position of I360 (blue) relative to Y277, Y278 and Y292 (colour scheme as above).

P353A and W354A were both found to significantly impair agonist-mediated internalisation, yet the size of effect was relatively small in both cases (~16%). The cell surface expression of P353A was found to be higher than WT but this seems to be

proportional to the total expression. Consequently, it could be inferred that this is not directly a pharmacological phenomenon but either a transfection, transcription or/and translation by-product.

The effects of P353A have been found and discussed previously (see Conner *et al.*, 2005). However, to add to this discussion it is interesting when a proline is identified as it can cause a kink or bend in an α-helix, an architectural feature that may be critical to receptor functioning. Monaghan *et al.*, (2008) identified four residues in the PTH1 receptor that were in close proximity to Ser-1 of PTH; L368, Y421, F424 and M425. Although, none of these residues are conserved in CLR (see Figure 36) a multiple sequence alignment reveals that P353 is in the equivalent position to M425. Moreover, F349 of CLR, whose mutation was found to have increased the Emax of the receptor but reduced cell surface expression, is in the cognate position of Y421. Finally, I352 of CLR, whose mutation was found to slightly decrease pEC50 (although not significantly) and decrease cell surface expression and increase both the Emax and basal activity, was predicted to be in the equivalent position of F424. Although, F349, I352 and P353 appear necessary for typical CGRP receptor pharmacology, only minor changes to the pEC50 and agonist mediated internalisation were observed suggestive that the mode of αCGRP binding may differ from that of the parathyroid hormone. However, given that alanine mutations can be tolerated even when made in the binding interface further cross-linking studies are required to probe the orientation of αCGRP whilst bound to its receptor.

```
                 *  **
PTHR1   MPLFGVHYIVFMATPYTE
CLR     VPLLGIEFVLIPWRPEGK

CRFR1   LPLLGITYMLFFVNPGED
CRFR2   LPLLGITYMLFFVNPGED
CTR     VPLLGIQFVVFPWRPSNK
GIPR    VPLLGVHEVVFAPVTEEQ
GLP1R   IPLLGTHEVIFAFVMDEH
GLP2R   IPLLGVHEILFSFITDDQ
GLR     IPLLGVHEVVFAFVTDEH
GHRHR   IPLFGIHYIIFNFLPDNA
SCTR    IPLFGIHYIVFAFSPEDA
VPAC2   IPLFGVHYMVFAVFPISI
VPAC1   IPLFGVHYIMFAFFPDNF
PAC1R   IPLFGIHYTVFAFSPENV
PTHR2   VLVFGVHYIVFVCLPHS-
```

Figure 36. A human Secretin-like GPCR alignment of the exofacial end of TM6 highlighting key differences between PTHR1 and CLR. T-coffee server used to generate multiple sequence alignment. * highlight Y421, F424 and M425 on the PTHR1. These residues were identified by Monaghan *et al.,* (2008) as being in close proximity to Ser-1 of PTH. These residues are not found in the CLR suggesting a different or modified mode of agonist binding.

Noteworthy, is the observation that I371A was found to internalise more readily than WT. Furthermore, the mutation significantly reduced cell surface expression but enhanced αCGRP binding. Consequently, the removal of this side chain to a smaller methyl group increases the potency of αCGRP. I371 is predicted to reside on TM7 and face into the bundle, being spatially close to TM1, and may contribute to the packing of the putative binding crevice that includes the triple cluster at the top of TM2 (L195, V198 and A199 outlined in Chapter 3) and D280 of ECL2 (see Figure 37).

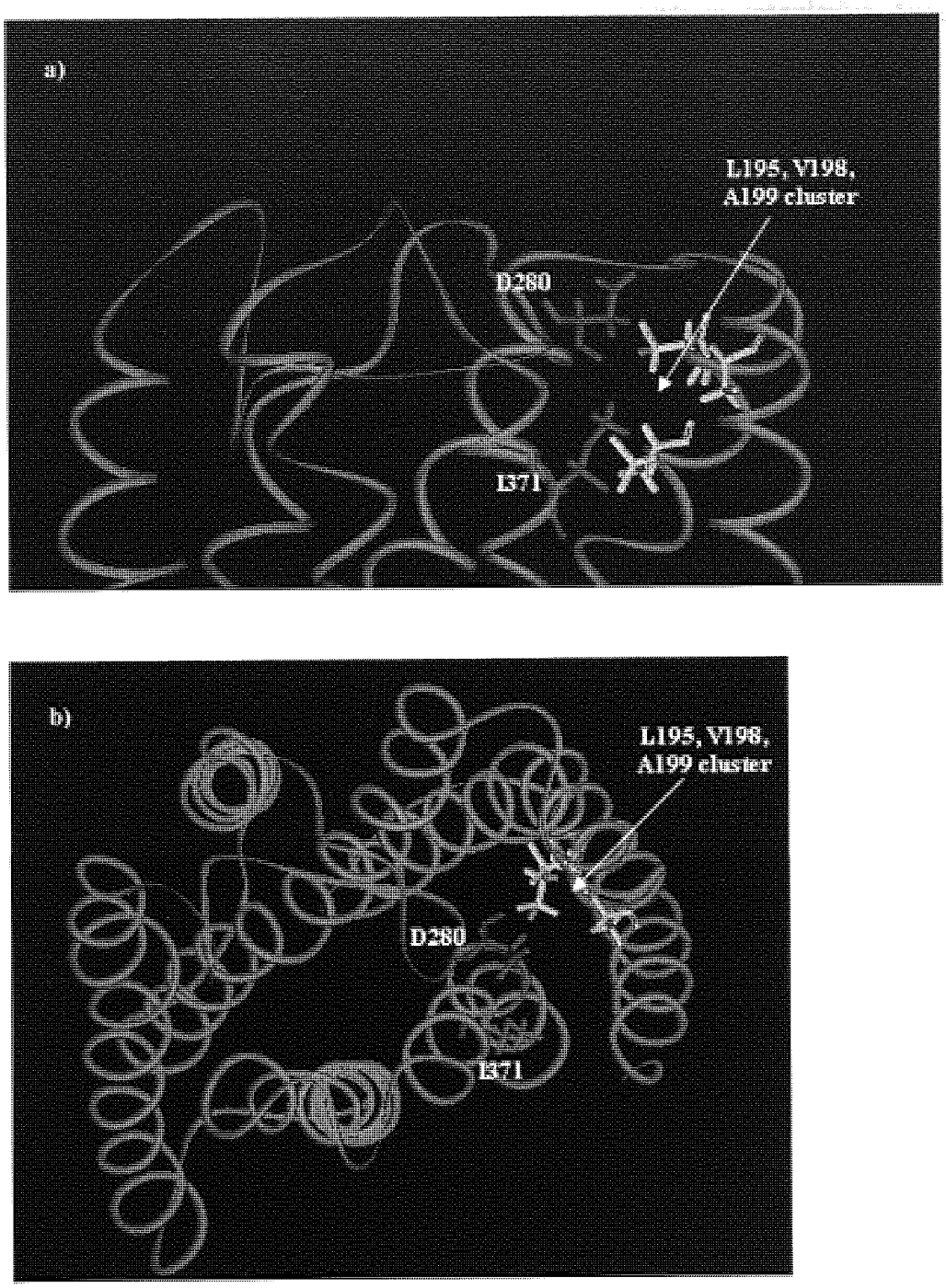


Figure 37. Position of I371 relative to D280 and the triplet cluster identified at the exofacial end of TM2. a) Side view of the assumed inactive CLR TM bundle (cyan blue ribbon) generated from ground-state rhodopsin (PDB accession: 1U19) showing the position of side chains of I371 (green), D280 (red), L195, V198 and A199 (yellow). b) Extracellular view of inactive CLR TM bundle, colour scheme as above.

In conclusion, taking into account both ECL1 and ECL3 alanine/leucine scans in conjunction with the results of Dr A.Conner's preliminary work on ECL2, the key functional epitopes within the extracellular domain of CLR are beginning to emerge. It is plausible that the αCGRP has a diffuse pharmacophore that spans across the TM

bundle. However, the orientation of the ligand still remains speculative. Antagonist binding experiments using $\alpha CGRP_{8-37}$ on identified mutants may give some insight into this issue. Alternatively, direct cross-linking experiments may be more fruitful. ECL3 and its juxtamembrane regions provide epitopes that support ECL2, which maybe essential in stabilising the protein and orthosteric binding site. TM7 may play a direct role in contributing to a CGRP binding pocket based on the predicted position of I371. More excitingly, the alanine scan outlined may provide the first experimental evidence, which indicates the elusive RAMP1 interface. However, a more direct strategy such as the TOXCAT assay (Russ and Engelman *et al.,* 1999), BRET or FRET assay involving the appropriate mutants would have to be employed to validate this hypothesis.

Chapter 5: Identifying important residues within the extreme N-terminus of CLR.

5.1 Introduction

The extreme N-terminus of CLR (defined as E23-A60) has been regarded as an important site for both αCGRP binding (Banerjee *et al.*, 2006) and RAMP1 association (Ittner *et al.*, 2005). The plethora of elucidated ECD structures of Secretin-like GPCRs with cognate ligands has suggested that this family share a common sushi domain fold (See Chapter 1 for further details). Furthermore, a general model is emerging of ligand binding to Secretin-like GPCRs, where the C-termini of the ligand interacts directly with the ECD, humorously referred to as the 'hot dog in a bun' model (Pioszak and Xu, 2008). More detailed studies suggest that the C-termini of the ligands come into contact with the extreme N-termini of the receptor, which adopts an α helical structure (Parthier *et al.*, 2007 and Pioszak and Xu, 2008). It is tempting to assume that αCGRP adopts a similar mode of binding to its receptor. However, the requirement for RAMP1 complicates matters. Consequently, it is plausible that it could have a dual role in interacting directly with αCGRP and RAMP1.

To further investigate the role of the extreme N-terminus of CLR, we report the results of an alanine scan on residues E23-A60, where native alanines have also been substituted for leucines. The results are discussed in light of recent advances in understanding the architecture of the CLR and RAMP1 interface.

5.2 Method

See Chapter 2 for general methods. Radioligand binding was conducted by Dr P. Miller (Leeds University) and the method used is described below.

5.2.1 Radioligand binding

Confluent cells from five 160-cm^2 Petri dishes (pre-coated with poly-D-lysine), were washed with PBS, followed by the addition of 15ml of ice-cold sterile double distilled water to induce cell lysis. Following 5 min incubation on ice, the ruptured cells were thoroughly washed with ice-cold PBS before being scraped from the plates and pelleted by centrifugation in a bench-top centrifuge (13,000 g for 30 min). The crude membrane pellet was resuspended in 1ml binding buffer (25 mM HEPES pH 7.4, 2.5mM $CaCl_2$, 1mM $MgCl_2$, 50mg/l bacitracin) and forced through a 23G needle. 0.1ml aliquots were snap-frozen in liquid nitrogen and stored at -70°C. Membranes were slowly thawed on ice before diluting to a concentration that gave total radioligand binding of <10% total counts added. In a reaction volume of 200µl, 75pM (~60,000 cpm) ^{125}I-CGRP with or without 1µM unlabelled CGRP and Cos7 membranes expressing the receptor of interest were combined, all diluted in binding buffer. Assays were carried out for 1h in MultiScreen 96-well Filtration Plates (Glass fibre filters, 0.65µm pore size, Millipore, Bedford, MA) pre-soaked in 1% non-fat milk/PBS. After the incubation, membrane-associated radioligand was harvested by transferring the assay mixture to the filtration plate housed in a vacuum manifold. The wells of the filtration plate were washed three times with 0.2ml PBS before harvesting the filter discs. Filter-bound radioactivity was measured in a gamma counter (RiaStar 5405 counter; PerkinElmer Life and Analytical Sciences, Waltham, MA). Total radioligand bound was <10% and non-specific binding was ~1% of total counts added.

5.3 Results

5.3.1 Stimulation of cAMP production

Each mutant was challenged with human αCGRP and cAMP production was measured (see Table 23). I41A, Q45A, C48A and Y49A all had a reduced pEC50 when compared to WT (Figure 38). A44L showed a borderline 8 fold decrease in potency, which was not found to be statistically significant. There was almost a 21 fold decrease in potency with C48A; 15 fold decrease with Y49A, 10 fold decrease seen with Q45A and a modest 6 fold decrease in I41A. Furthermore, the mean Emax of Y49A was decreased by 36.7% ± 8.5% and the mean Emax of C48A was also found to have decreased by 29.9% ± 8.5%, while the mean basal activity resembled WT.

I32A, G35A and T37A were found to have ~8-9 fold increase in potency for αCGRP compared to WT as assessed by mean pEC50 (Figure 39). Interestingly, the mean basal activity was found to have decreased by 25.81% ± 8.75% in I32A, while the mean Emax was found to have increased by 30.1% ± 9.45%. Moreover, the mean Emax of G35A was found to have increased by 23.1% ± 15.8% but the mean basal activity resembled WT. Similarly, T37A was found to have an increased mean Emax (29.1% ± 12.6%) while the mean basal activity resembles WT.

It is noteworthy to mention that K51A had a 30.2 % ± 7.0% increase in mean basal activity and a 22.7% ± 5.9% increase in Emax. Similarly, E29A was found to have a 37.1% ± 16.6% increase in mean basal activity compared to WT and the mean Emax was found to be 71.4% ± 5.1% higher compared to WT (see Figure 40). The remaining mutant dose-response curves resembled WT.

Table 23. Comparison between the mean WT and mean extreme N-terminal mutant pEC50 values. Values are pEC50 means ± S.E.M. $p < 0.05$ is represented by *. $p < 0.01$ is represented by ** and $p < 0.001$ is represented by ***. pEC50 mutant values were compared to WT using an independent two-tailed t-test.

Mutant	N	pEC50 WT	pEC50 Mutant	Mutant	N	pEC50 WT	pEC50 Mutant
E23A	4	8.98 ± 0.32	9.18 ± 0.37	M42A	6	9.81 ± 0.26	9.86 ± 0.10
L24A	4	9.09 ± 0.35	9.39 ± 0.42	T43A	3	9.77 ± 0.41	9.74 ± 0.37
E25A	3	10.32 ± 0.22	10.06 ± 0.46	A44L	5	9.32 ± 0.20	8.41 ± 0.45
E26A	5	9.33 ± 0.57	9.33± 0.21	Q45A	3	9.11 ± 0.08	8.10 ± 0.18**
S27A	4	9.26 ± 0.14	9.51 ± 0.34	Y46A	4	9.17 ± 0.19	9.12 ± 0.44
P28A	4	9.11 ± 0.33	9.90 ± 0.37	E47A	5	9.57 ± 0.21	9.40 ± 0.32
E29A	4	9.42 ± 0.06	9.26 ± 0.20	C48A	7	10.09 ± 0.23	8.77 ± 0.38**
D30A	4	9.66 ± 0.29	9.68 ± 0.41	Y49A	5	9.93 ± 0.34	8.75 ± 0.25*
S31A	7	9.11 ± 0.31	9.00 ± 0.37	Q50A	5	9.49 ± 0.10	9.61 ± 0.49
I32A	5	8.95 ± 0.11	9.90 ± 0.29*	K51A	4	9.45 ± 0.37	9.66 ± 0.25
Q33A	3	9.28 ± 0.22	9.04 ± 0.03	I52A	6	9.79 ± 0.24	9.86 ± 0.29
L34A	7	9.55 ± 0.30	9.51 ± 0.23	M53A	5	9.72 ± 0.30	10.20 ± 0.24
G35A	4	8.95 ± 0.14	9.86 ± 0.27*	Q54A	4	9.38 ± 0.15	9.60 ± 0.33
V36A	4	9.27 ± 0.18	9.90 ± 0.30	D55A	3	9.74 ± 0.27	9.91 ± 0.14
T37A	4	9.25 ± 0.18	10.22 ± 0.23*	P56A	3	9.74 ± 0.14	9.21 ± 0.18
R38A	5	9.17 ± 0.15	9.32 ± 0.35	I57A	3	9.23 ± 0.27	9.27 ± 0.33
N39A	4	9.24 ± 0.34	9.55 ± 0.42	Q58A	6	10.12 ± 0.36	10.28 ± 0.37
K40A	4	9.61 ± 0.15	10.01 ± 0.25	Q59A	5	9.68 ± 0.26	9.70 ± 0.32
I41A	7	10.02 ± 0.34	9.22 ± 0.096*	A60L	6	9.65 ± 0.44	9.56 ± 0.43

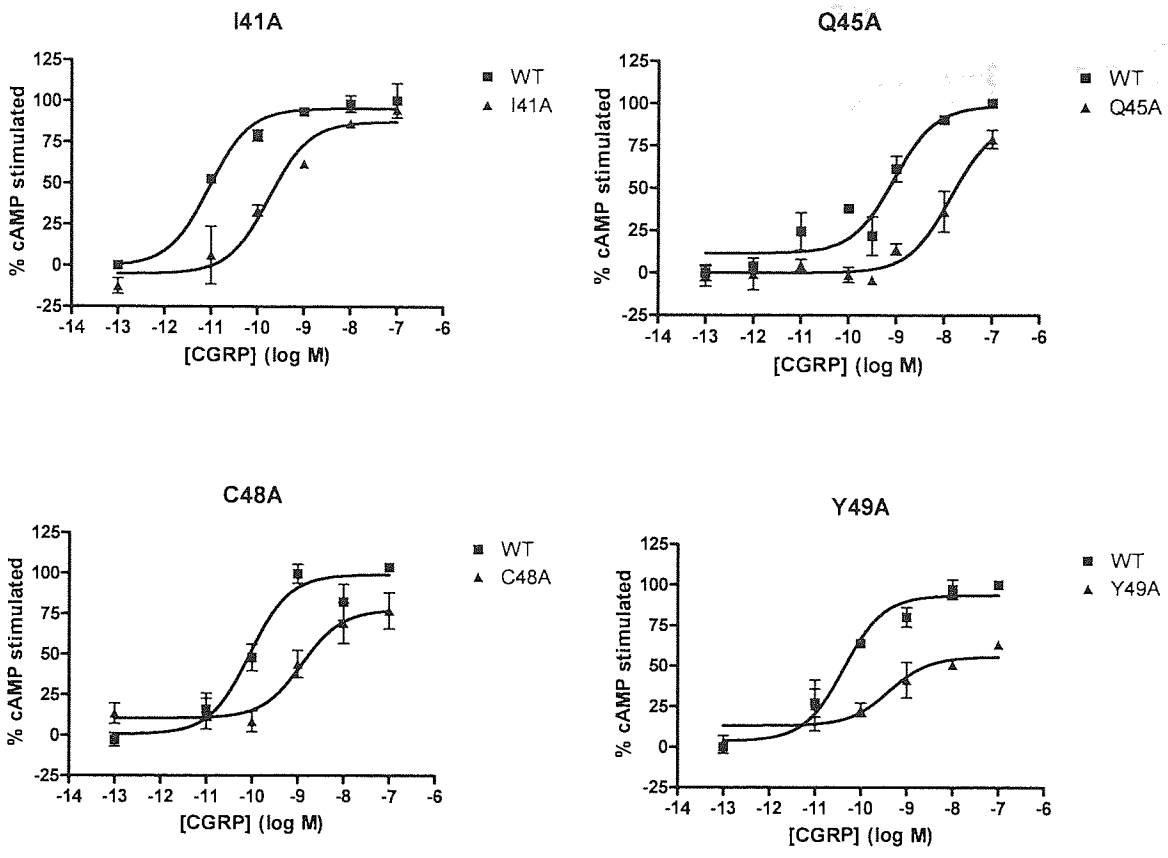


Figure 38. Representative dose-response curves of extreme N-terminus mutants that were found to significantly decrease αCGRP potency. Sigmoidal concentration-response curves comparing the WT receptor and mutant receptors (I41A, Q45A, C48A and Y49A) capability of stimulating cAMP after αCGRP activation are shown. Each WT and mutant receptor concentration-response comparison curve is a representative example of an ensemble of at least three independent experiments. Each assay point was performed in duplicate where each point on the graph represents the mean ± S.E.M. The line of best fit for each concentration-response curve was fitted with GraphPad Prism 4 see Equation 1.

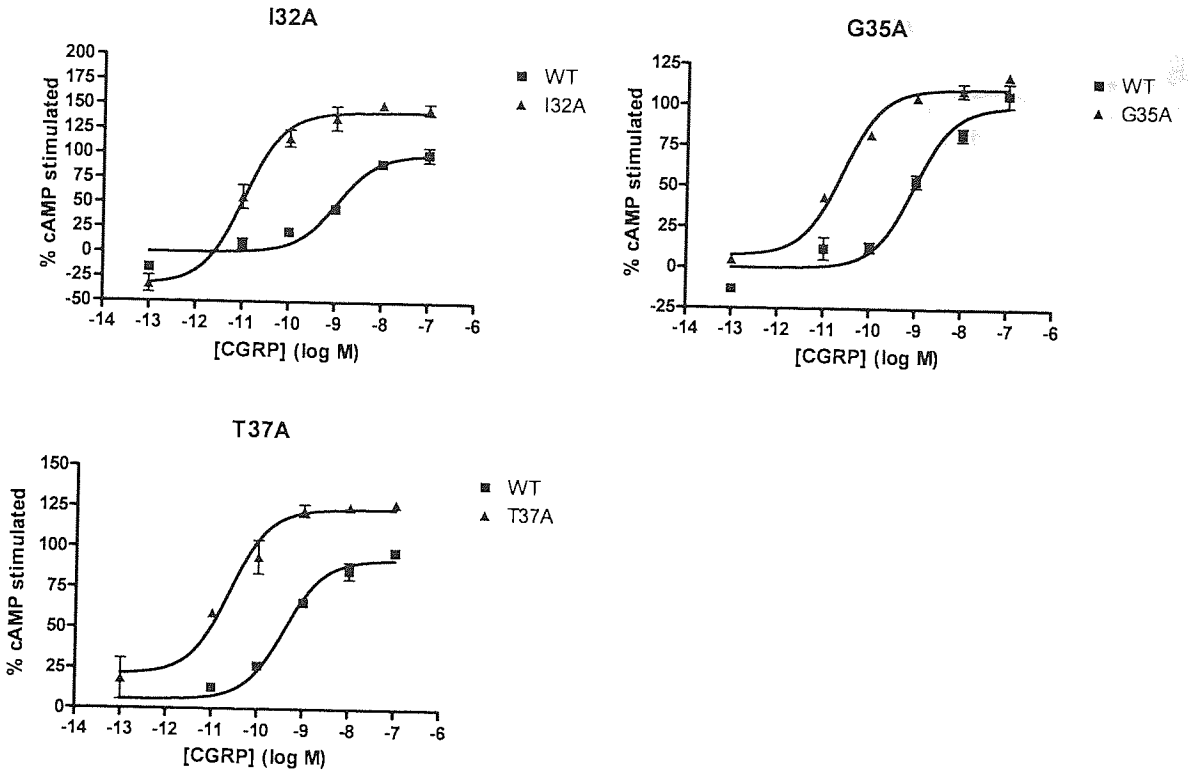


Figure 39. Extreme N-terminus mutant receptors found to significantly enhance αCGRP potency compared to WT receptors. Sigmoidal concentration-response curves comparing the WT receptor and mutant receptors (I32A, G35A and T37A) capability of stimulating cAMP after αCGRP activation are shown. Each WT and mutant receptor concentration-response comparison curve is a representative example of an ensemble of at least three independent experiments. Each assay point was performed in duplicate where each point on the graph represents the mean ± S.E.M. The line of best fit for each concentration-response curve was fitted with GraphPad Prism 4 see Equation 1.

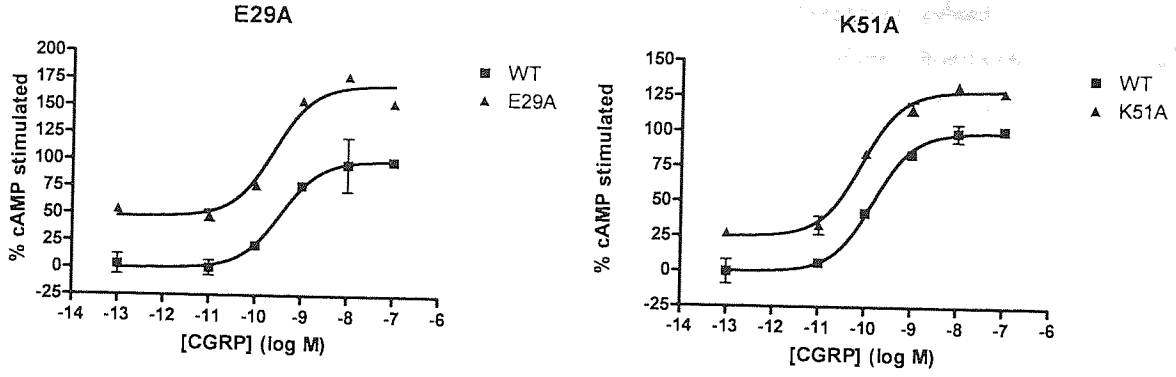


Figure 40. E29A and K51A representative dose-response curves demonstrating an increase in basal activity and Emax compared to the WT receptor. Representative sigmoidal concentration-response curves comparing E29A and K51A mutant receptors against the WT receptors capability of stimulating cAMP after αCGRP activation. Each WT and mutant receptor concentration-response comparison curve is a representative example of an ensemble of at least three independent experiments. Each assay point was performed in duplicate where each point on the graph represents the mean ± S.E.M. The line of best fit for each concentration-response curve was fitted with GraphPad Prism 4 see Equation 1.

5.3.2 Cell surface expression and radioligand binding

Cell-surface expression of all mutant receptors was measured (Table 24). Large reductions were seen for Q45A, C48A and Y49A with small reductions for T43A and A44L. Whilst nine mutants showed a significant increase in cell-surface expression, only for R38A and Y46A was this increase greater than 50% of the WT.

For four mutants, I41A, A44L, C48A and Y49A, high-affinity αCGRP binding was abolished (Table 24), consistent with the reduced potency and/or cell surface expression shown by these receptors. There was a reduction in the binding seen with Q45A, although it was not abolished.

Table 24. Cell surface expression and binding properties of extreme N-terminus mutant receptors. Cell surface expression ELISA was used to probe for the presence of the HA epitope. Mutant HA CLR/*myc* RAMP1 receptors were compared with WT HA CLR /*myc* RAMP1 receptors. 3-6 independent experiments that contained triplicate data points were used in analysis. The raw data for each independent experiment was normalised where the mean WT receptor cell-surface expression equalled a 100% and the mean negative control (*myc* RAMP1/ empty pcDNA3.1-) was equal to 0%. Values reported are mutant means ± S.E.M (% WT). Mutant cell surface expression was compared to WT receptor using a Mann-Whitney U test ($p < 0.05$ is represented by *, $p < 0.01$ represented by **, and $p < 0.001$ represented by ***). Binding shows % specific binding of ^{125}I-CGRP for each mutant receptor; WT receptor specific binding (equal to 100%) was determined by three independent experiments conducted in triplicate $^{+}$represents mutant receptors with no measurable specific binding.

Mutant	Cell surface expression (% WT mean ± S.E.M)	Specific binding (% WT mean ± S.E.M)	Mutant	Cell surface expression (% WT mean ± S.E.M)	Specific binding (% WT mean ± S.E.M)
E23A	108.7 ± 5.1	81.7 ± 4.0	M42A	83.2 ± 6.3	97.2 ± 12.0
L24A	111.1 ± 9.2	101.6 ± 3.5	T43A	83.6 ± 5.3*	87.6 ± 12.3
E25A	143.0 ± 16.4*	90.6 ± 12.5	A44L	82.7 ± 6.0*	4.3 ± 1.6^{+}
E26A	123.7 ± 8.6	71.9 ± 10	Q45A	66.6 ± 3.3***	61.4 ± 3.7
S27A	117.1 ± 7.5	116.4 ± 22.2	Y46A	167.0 ± 8.6***	71.4 ± 16.7
P28A	95.9 ± 9.9	90.2 ± 14	E47A	81.7 ± 4.7 **	90.0 ± 10.7
E29A	108.6 ± 9.7	90.5 ± 0.8	C48A	45.4 ± 3.9***	2.7 ± 1.2^{+}
D30A	123.4 ± 6.3**	74.1 ± 5	Y49A	33.8 ± 5.6***	5.5 ± 3.1^{+}
S31A	111.4 ± 9.7	98.6 ± 2.4	Q50A	98.5 ± 9.4	93.5 ± 2.5
I32A	114.3 ± 9.9	114.7 ± 13.7	K51A	82.0 ± 8.4	85.9 ± 6.7
Q33A	141.0 ± 8.3***	59.2 ± 14.2	I52A	123.9 ± 5.8*	86.2 ± 1.9
L34A	129.4 ± 14.6	103.0 ± 14.2	M53A	109.6 ± 5.1	95.8 ± 6.7
G35A	147.6 ± 9.2**	103.0 ± 8.4	Q54A	106.5 ± 9.3	103.1 ± 15.8
V36A	116.1 ± 9.3*	87.8 ± 6	D55A	100.5 ± 7.9	86.7 ± 10.4
T37A	108.5 ± 8.6	88.3 ± 8.1	P56A	131.1 ± 7.0**	83.3 ± 8.8
R38A	161.7 ± 9.0***	89.0 ± 8.6	I57A	105.4 ± 9.9	102.1 ± 4.3
N39A	147.6 ± 10.0**	94.9 ± 13.5	Q58A	88.0 ± 10.6	78.8 ± 7.9
K40A	83.9 ± 4.0	91.1 ± 9.0	Q59A	102.9 ± 7.9	82.1 ± 6.6
I41A	133.6 ± 8.7	2.0 ± 2.7^{+}	A60L	79.4 ± 7.1	104.9 ± 6

5.3.3 Total expression of receptors

Total CLR production as measured for these mutants by a whole-cell ELISA was only significantly reduced for A44L (81.7% ± 6.5% of WT) and Y49A (77.6% ± 3.8% of WT). This modest decrease suggests that the mutants were synthesised with reasonable efficiency so the reduction in cell surface expression is probably caused by defective trafficking or insertion into the membrane. In contrast nine mutants showed a significant increase in cell-surface expression. Interestingly, eight out of these nine mutations were also found to have increased cell surface expression. The largest increases were found in R38A, Y46A and I52A. Consequently, the cause in cell surface expression in these mutations could be attributed to enhanced efficiency in transfection or synthesis (see Table 25 for summary).

Table 25. Total expression of extreme N-terminus mutant receptors. Total expression of HA-tagged receptors both mutant and WT were analysed when co-transfected with *myc* RAMP1. At least 3 independent experiments containing triplicate data points were used in analysis. Total expression in the mutant condition was normalised to the WT condition (equal to 100%) and negative control (*myc* RAMP1/empty pcDNA3.1-, after 0.1% Triton X 100, which was equal to 0%).A Mann-Whitney U test was used to assess statistical difference between WT and mutant receptors were *, **, *** represent $p <$ 0.05, 0.01 and 0.001 significance levels.

Mutant	Total expression (% WT mean ± S.E.M)	Mutant	Total expression (% WT mean ± S.E.M)
E23A	91.4 ± 3.4	T43A	116.0 ± 9.4
E25A	88.7 ± 1.7	A44L	81.7 ± 6.5*
D30A	141.5 ± 11.0*	Q45A	109.1 ± 13.8
I32A	138.1 ± 7.6***	Y46A	196.6 ± 30.0***
Q33A	135.0 ± 11.9	E47A	99.5 ± 2.0
G35A	126.7 ± 5.3**	C48A	91.5 ± 3.08
T37A	112.9 ± 9.3	Y49A	77.6 ± 3.8*
R38A	158.8 ± 7.6***	I52A	176.1 ± 13.3*
N39A	130.6 ± 7.1**	P56A	111.6 ± 14.9
I41A	100.1 ± 6.7		

5.3.4 Agonist mediated internalisation

Twelve mutants showed a decrease in receptor internalisation (see Table 26). Large effects were seen with I41A, A44L, Q45A, C48A and Y49A where it was either greatly impaired or totally abolished. All these residues show impaired cAMP responsiveness and with the exception of Q45A, these mutants also showed impaired radioligand binding confirming their importance.

Table 26. Summary of extreme N-terminus mutant receptors responsiveness to αCGRP mediated internalisation. Agonist mediated internalisation of the CGRP receptor (both WT and mutant receptors) was approximated by a HA epitope probing cell surface ELISA taking into account the difference in cell surface expression levels between CGRP receptors that have or have not been exposed to 100nM of human αCGRP for an 1hr. Percent mean ± S.E.M agonist mediated internalisation was determined by 3-6 independent experiments containing 3 replicates. A Mann Whitney U test was used to compare mutant and WT percent agonist internalisation values where $p < 0.05$ is respresented by *, $p < 0.01$ is represented by **, and $p < 0.001$ is represented by ***.

Mutant	% Internalisation (mean ± S.E.M)		Mutant	% Internalisation (mean ± S.E.M)	
	WT	Mutant		WT	Mutant
E23A	60.9 ± 5.9	61.15 ± 5.45	M42A	52.15 ± 6.68	60.32 ± 5.58
L24A	57.5 ± 6.4	63.81 ± 3.77	T43A	66.03 ± 5.30	69.75 ± 3.56
E25A	57.7 ± 3.9	57.42 ± 4.45	A44L	71.16 ± 2.30	51.67 ± 3.59**
E26A	71.1 ± 5.8	71.66 ± 3.79	Q45A	54.84 ± 2.69	27.45 ± 7.51**
S27A	58.4 ± 7.0	53.36 ± 7.22	Y46A	89.19 ± 2.05	56.83 ± 3.73***
P28A	62.7 ± 5.48	58.27 ± 2.85	E47A	68.72 ± 2.89	61.44 ± 2.85
E29A	63.40 ± 5.64	60.34 ± 2.86	C48A	57.04 ± 3.36	-14.65 ± 4.2***
D30A	65.40 ± 2.68	57.55 ± 5.80*	Y49A	49.80 ± 4.81	0.00 ± 9.34***
S31A	59.86 ± 5.98	64.01 ± 5.71	Q50A	68.58 ± 3.42	65.79 ± 3.06
I32A	64.96 ± 5.86	46.42 ± 4.43*	K51A	73.14 ± 2.87	76.48 ± 3.57
Q33A	63.92 ± 4.20	52.42 ± 4.14*	I52A	75.73 ± 4.89	51.41 ± 4.06***
L34A	54.84 ± 4.78	48.11 ± 4.84	M53A	68.28 ± 4.54	65.22 ± 3.41
G35A	62.51 ± 5.34	61.18 ± 3.24	Q54A	65.97 ± 6.18	63.03 ± 4.29
V36A	55.77 ± 6.11	44.86 ± 8.79	D55A	72.80 ± 1.96	65.32 ± 2.74
T37A	61.89 ± 5.12	54.01 ± 5.50	P56A	72.16 ± 2.78	56.22 ± 2.71***
R38A	54.54 ± 5.99	33.79 ± 5.02*	I57A	68.28 ± 6.03	59.61 ± 7.04
N39A	54.44 ± 3.13	62.20 ± 5.60	Q58A	69.32 ± 4.83	75.50 ± 3.29
K40A	67.16 ± 5.22	64.61 ± 4.02	Q59A	49.42 ± 3.45	59.90 ± 3.33
I41A	50.06 ± 6.03	16.92 ± 9.94***	A60L	55.80 ± 3.23	57.09 ± 4.14

5.3.5 Summary of important extreme N-terminal mutations

Extreme N-terminal mutations that were found to significantly alter the potency of αCGRP to stimulate cAMP (pEC50), cell-surface expression, total expression, agonist-mediated internalisation and/or CGRP binding (pIC50) are summarised in table 27.

Table 27. The extreme N-terminal mutations that significantly altered the pharmacology of the CGRP receptor. EC50 fold effects were determined by sigmoidal concentration-response curves comparing the WT and mutant receptors capability of stimulating cAMP after αCGRP exposure. The EC50 mean fold differences were determined from at least three independent experiments. A two-tailed independent t-test was used to assess statistical differences. Cell surface expression, total expression and αCGRP mediated internalisation values represent the mean difference in % WT. A Mann Whitney U test was used to test for statistical differences between WT and mutant receptors on these three parameters, which were determined by at least three independent experiments. Mutants that were found not to be significantly different from WT are represented by N.S. M.B stands for measurable binding whereas N.M.B stands for no measurable binding of CGRP.

Mutant	EC50 fold effects after αCGRP stimulation (compared to WT receptors)	Cell surface expression (% WT mean difference)	Total expression (% WT mean difference)	αCGRP mediated internalisation (% WT mean difference)	Measurable CGRP specific binding
E25A	N.S.	43.0% increase	N.S.	N.S.	M.B.
D30A	N.S.	23.4% increase	41.5% increase	12.0% decrease	M.B.
I32A	~9 fold increase	N.S.	38.1% increase	28.5% decrease	M.B.
Q33A	N.S.	41.0% increase	N.S.	18.0% decrease	M.B.
G35A	~8 fold increase	47.6% increase	26.7% increase	N.S.	M.B.
T37A	~9 fold increase	N.S.	N.S.	N.S.	M.B.
R38A	N.S.	61.7% increase	58.8% increase	39.1% decrease	M.B.
N39A	N.S.	47.6% increase	30.6% increase	N.S.	M.B.
I41A	~6 fold decrease	N.S.	N.S.	66.2% decrease	N.M.B
T43A	N.S.	16.4% decrease	N.S.	N.S.	M.B.
A44L	N.S.	17.3% decrease	N.S.	27.4% decrease	N.M.B
Q45A	~10 fold decrease	33.4% decrease	N.S.	50.0% decrease	M.B.
Y46A	N.S.	67.0% increase	96.6% increase	36.2% decrease	M.B.
E47A	N.S.	18.3% increase	N.S.	N.S.	M.B.
C48A	~21 fold decrease	54.6% decrease	N.S.	125.7% decrease	N.M.B.
Y49A	~15 fold decrease	66.2% decrease	22.4% decrease	100.0% decrease	N.M.B.
I52A	N.S.	23.9% increase	76.1% increase	32.1% decrease	M.B.
P56A	N.S.	31.1% increase	N.S.	22.1% decrease	M.B.

5.4 Discussion

The results of this investigation suggest that a small number of residues in the extreme N-terminus of CLR are important for receptor function, in particular, I41, A44, Q45, C48 and Y49. Mutation of these residues either disrupts αCGRP-stimulated cAMP production, αCGRP binding or CLR expression at the cell surface when co-transfected with RAMP1. Moreover, there is also evidence for a second cluster of important residues consisting of I32, G35 and T37.

C48 is predicted to take part in one of the highly conserved disulphide bonds that characterise the ECDs of Secretin-like GPCRs and its mutation to alanine in other receptors also causes a loss of signaling (Lisenbee *et al.,* 2005). The other residues, apart from Y49 are not widely conserved and so must have receptor-specific roles. However, a prerequisite in understanding αCGRP binding is to first comprehend the architecture of the CGRP receptor ECD.

Barwell *et al.,* (2010) proposed a speculative model of the CGRP ECD made from the recently published RAMP1 crystal structure (Kusano *et al.,* 2008) and a homology based model of the CLR ECD, although no existing distance restraints had been characterised between CLR and RAMP1 at that time. Protein-Protein docking with low-resolution (homology) models is extremely challenging. In recent years, Ilya Vakser's group from the University of Kansas has been at the forefront of this field. Their program Global RAnge Molecular Matching (GRAMMv1.03) employs an empirical rigid-body geometric fit technique that performs an exhaustive 6-dimensional search through the relative translations and rotations of the molecules in an attempt to locate the area of the global minimum of intermolecular energy between the structures of interest (Katchalski-Katzir *et al.,* 1992, Vakser, 1995, Vakser *et al.,* 1999). In spite the decrease in accuracy of GRAMM when dealing with low resolution structures, it has been successfully used to determine the gross conformations of protein complexes (Lett *et al.,* 2004). A full description of the methodology of the construction of the CGRP ECD can be found in Chapter 2.

Taking into account the short comings of low resolution docking experiments the speculative model had certain strengths at the time of its publication and required careful consideration. The model was an asymmetric complex containing two CLR

ECD molecules and one RAMP1 ECD molecule, in line with Heroux *et al.,* (2007). It was predicted that the two CLR-ECDs dock together to produce a somewhat symmetrical complex. The interface between the two CLR ECDs was located in loop 4 (Parthier *et al.,* 2009 nomenclature) between P89-S99. The corresponding disordered loop in the NMR mouse CRF receptor ECD structure is predicted to be relatively stable as evidenced by very broad cross-peaks in the [^{15}N,^{1}H]-TROSY spectrum (Grace *et al.,* 2007). Consequently, this was considered to be a plausible docking region of the two molecules, which could be further stabilised by neighbouring β-strands. Moreover, the CLR ECD dimer being symmetrical (i.e. a mirror image) meant that the TM4 interface between the two receptors remained plausible.

The model suggested that the CLR N-terminal helix, which was predicted to adopt an α-helical conformation between N39 and Q54, would dock against helix 3 of RAMP1. F93, H97 and F101 located in helix 3 of RAMP1 had already postulated to participate in a CLR binding interface (Kusano *et al.,* 2008). It was plausible to foresee that a hydrogen bond could occur between Q45 of CLR-A and H97 of RAMP1 and Y49 of CLR-A and F93 of RAMP1 may pack together. Moreover, I41 appeared to be too far away from RAMP1 to make a significant hydrophobic interaction, consistent with its normal cell-surface expression (see Figure 41).

However, there were consequences if the RAMP1 resided on this side of the complex. Firstly, RAMP1 would mask loop 4 and therefore the typical 'hot dog in the bun' model of Secretin-like ligand binding would not be applicable. However, it has long been theorised that RAMP1 association may mask certain binding epitopes so the receptor is capable of accommodating a different repertoire of ligands (Hilairet *et al.,* 2001a). Moreover, there is tentative evidence to suggest that αCGRP may have a different mode of binding compared to the hydrophobic driven 'hot dog in a bun model' as Howitt *et al.,* (2003) suggested that R11 and R18 located on the hydrophilic face of the amphipathic helix of αCGRP was more likely to participate in binding compared to the hydrophobic helix face of the ligand. A second consequence of the RAMP1 molecule residing at this side of the complex is its approximate position of its TM domain. The two CLR ECDs and RAMP1 ECD are predicted to be somewhat perpendicular in relation to the supposed TM domain in this model. Taking into account the position of the C-terminus of the CLR-A ECD, which gives an indication

on the top of TM1, it would be more likely for the RAMP1 TM domain to interact with TM1, TM2 and/or TM3.

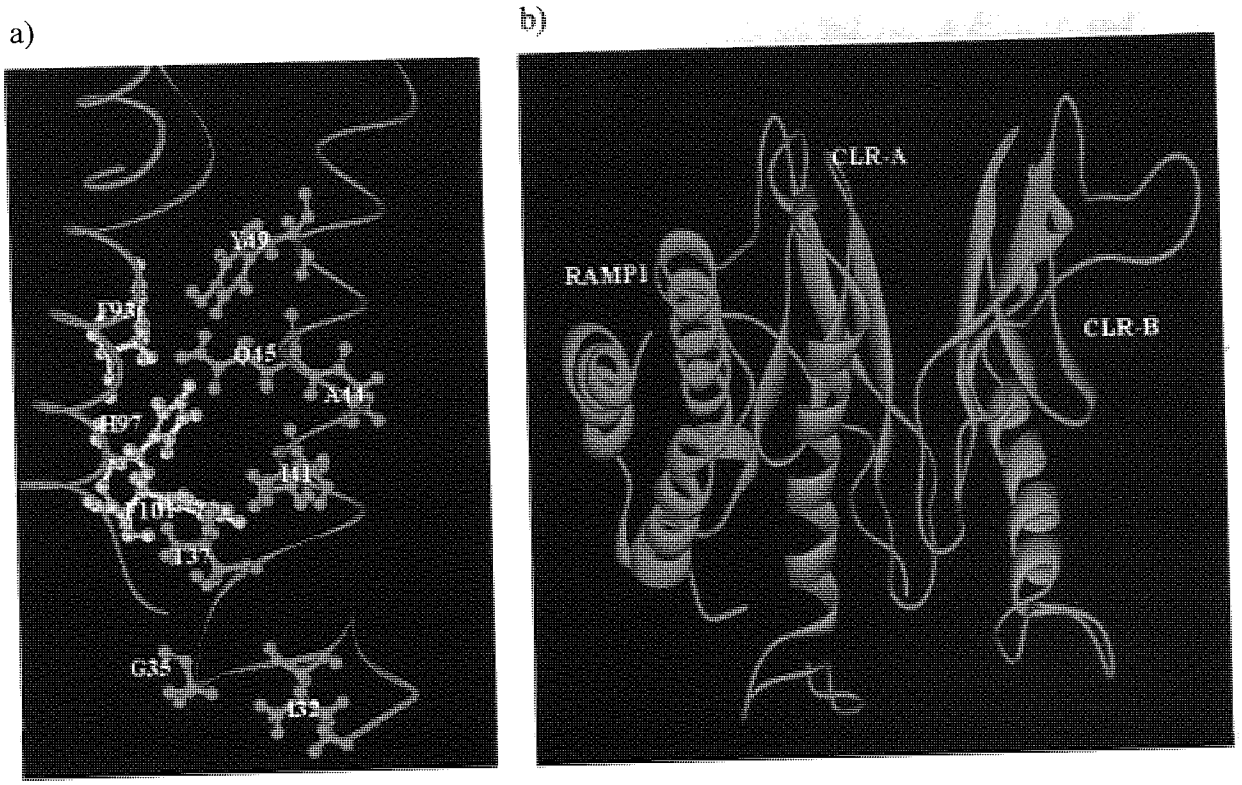


Figure 41. Speculative trimeric CGRP receptor ECD. a) Postulated interface between RAMP1 (green ribbon) and CLR-A (cyan ribbon). A comparative protein structure of ECD of CLR, from residues 23-134 was generated using Modeller9v3 (Sali and Blundell, 1993 and Fiser *et al.*, 2003). The comparative model was based on two templates GIPR ECD (Parthier *et al.*, 2007) and PTHR1 ECD (Pioszak and Xu, 2008). The RAMP1 ECD crystal structure (Kusano *et al.*, 2008) was downloaded from the RCSB Protein Data Bank (www.rcsb.org/pdb/). Y49, Q45, A44, I41, T37, G35 and I32 of CLR-A highlighted orange. F93. H97 and F101 highlighted yellow. b) Ribbon interpretation of postulated trimeric CGRP ECD, RAMP1 (green), CLR-A (cyan) and CLR-B (orange).

Shortly after the speculative model was submitted Miller *et al.*, (2010) published an interesting paper that analysed the binding of the CGRP receptor antagonists (BIBN4096BS and MK-0974) on the mutants made above. The results suggested that M42A disrupted antagonist binding. W74 located on RAMP1 has long been recognised as a key residue for BIBN4096BS affinity (Doods *et al.*, 2000 and Mallee *et al.*, 2002). Consequently, the authors speculate that M42 of CLR and W74 of RAMP1 may be in close proximity. However, the Barwell *et al.*, (2010) speculative model suggests that M42 was just too far away from W74 for both to be involved in

BIBN4096BS binding. However, W74 was located in a pivotal position in the model where its presence could potentially alter the relative position of the CLR N-terminal helix, which may have offered a tentative explanation for this novel observation. Yet, clearly Miller *et al.,* (2010) findings questioned the plausibility of the initial model suggesting a revised model would need to be constructed but still a difficult and challenging task would have remained as no clear distance restraints between the two molecules was known at this point in time.

At the end of January 2010 a key advance was made in the field when Koth *et al.,* (2010) published an exhaustive biochemical and biophysical characterisation of the CGRP receptor ECD, which was only a heterodimer (not a trimer) containing one RAMP1 ECD and just one CLR ECD. Koth and colleagues elegant work consisted of NMR structural analysis to assess the mobility of the CGRP ECD and the conformational changes induced upon BIBN4096BS binding. The results suggested the CGRP ECD was particularly flexible and the largest effects were observed in helix 2 of RAMP1 upon antagonist binding. Interestingly, the group state that they have X-ray crystallography data on the unliganded state of the CGRP ECD and various diffracting crystals with small antagonists bound. However, this data has not been published as of yet and the atomic coordinates have not been deposited into the Protein Data Bank. Yet, Koth *et al.,* (2010) provide a figure of the crystal structure of the CGRP ECD bound to BIBN4096BS in their recent publication (see Figure 42).

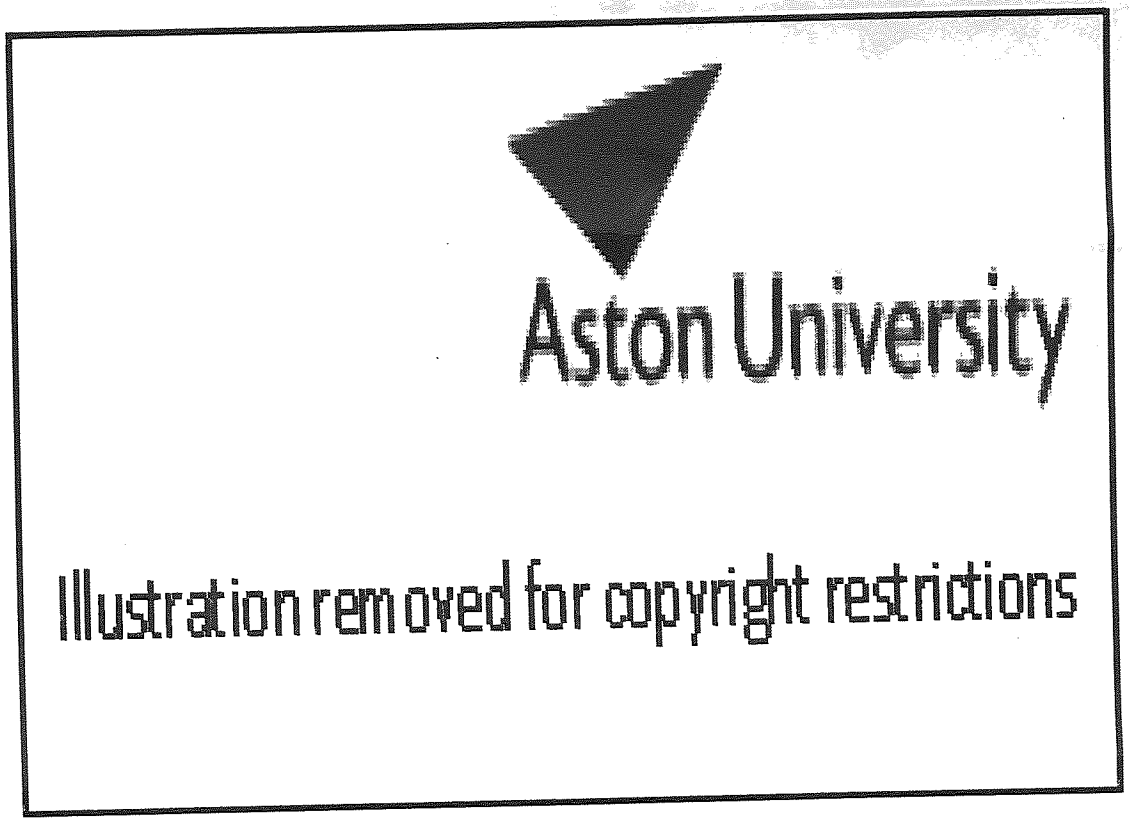


Figure 42. X-ray crystal structure of the CGRP ECD with BIBN4096BS bound. Taken from Koth *et al.*, (2010) to show perturbations observed by NMR when BIBN4096BS (purple) binds to CLR23-133 (green) and RAMP126-117 (cyan). RAMP126-117 residues exhibiting amide chemical shift changes upon addition of ligand are colored according to small (yellow), moderate (orange) and large (red) magnitude effects.Unassigned gap residues are colored grey and the sidechains of W74 and W84 are shown.

This preliminary figure clearly confirms that RAMP1 associates with the extreme N-terminus of CLR. In contrast to the Barwell *et al.*, (2010) speculative model the RAMP resides on the other side of the CLR N-terminal helix. The shortcomings in the Barwell *et al.*, model maybe attributed to the current lack in understanding the orientation of the ECD of Secretin-like GPCRs relative to their TM domains, along with the ambiguity of the stoichiometric arrangement of proteins within the CGRP receptor. The figure of the elucidated CGRP ECD suggests that it is feasible that ligand binding to this receptor may resemble the typical 'hot dog in a bun' mode of binding as loop 4 is not masked. The hypothetical model suggested that I41 and A44 may play a role in αCGRP binding, while Q45 and Y49 may be involved in RAMP1

association. These conclusions still remain feasible even though the RAMP1 is in an alternative orientation, yet further investigation is needed to support or refute this conclusion (see Figure 43).

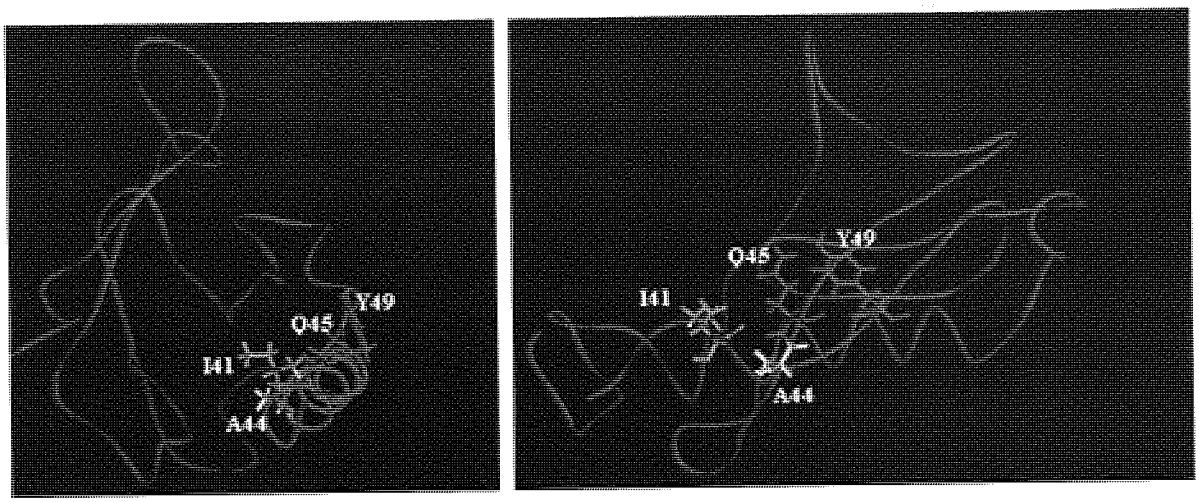


Figure 43. The predicted orientaion of I41, A44, Q45 and Y49 may provide insight into their functonal role. A comparative protein structure of ECD of CLR, from residues 23-134 was generated using Modeller9v3 (Sali and Blundell, 1993 and Fiser *et al.*, 2003). The comparative model was based on two templates GIPR ECD (Pathier *et al.*, 2007) and PTHR1 ECD (Pioszak and Xu, 2008). CLR ECD represented as a green ribbon. A44 side chain yellow, I41 side chain orange, Q45 side chain red and Y49 side chain purple.

I32A, G35A and T37A form a separate cluster of residues on the extreme N-terminus of CLR that increase αCGRP potency. I32A also reduces agonist mediated internalisation. At this part of the N-terminus, the accuracy of Secretin-like GPCRs alignments decreases considerably. Accordingly, it is difficult to use the initial model to make any useful comments and the figure provided by Koth *et al.*, (2010) gives little information on exact locations of residues. However, it suggests that there are extended contact points for αCGRP along the entire region. Deletion of the first 18 residues of this section (as far as K40) gives a receptor that cannot respond to αCGRP (Koller *et al.*, 2002) and L24A and L34A have previously been found to decrease αCGRP potency, but not binding (Banerjee *et al.*, 2006). Whilst there is agreement with the current study that this part of the N-terminus has a role in αCGRP binding, the residues identified are different. There may be cell-line specific factors at work; as authors observed a decrease in binding but not in signaling, their receptors may have been very efficiently coupled to Gs.

In conclusion, recent efforts have confirmed the importance of the extreme N-terminus of the CLR and its importance in CGRP receptor pharmacology. This specific study has highlighted specific residues that may contribute to αCGRP binding although observing the eagerly anticipated crystal structure of the CGRP heterodimer ECD may shed new light on this intriguing phenomenon.

Chapter 6: Progress to date in RAMP purification.

6.1 Introduction

Structural determination of membrane proteins has been hindered by difficulties in expression, solubilisation, purification and crystallisation (see White, 2009). Developing high throughput purification protocols that are applicable to all recombinant membrane proteins has not yet been achieved due to a plethora of pitfalls. Insufficient protein yield, the inherent instability of the protein and the limited understanding of the host system's lipid bilayer plague most attempts to produce structural biophysical data. Consequently, the choice of protein expression system and the recombinant protein are imperative. Yet, predicting the conditions that produce functional proteins at high yields has to be determined empirically (Darby *et al.*, 2010).

In spite of the difficulties in recombinant membrane protein production a protocol that can produce high protein concentrations of RAMP2 and RAMP3 is outlined here. The ability to purify the RAMP family may ultimately lead to biophysical structural data and may aid dimerisation studies. The expression system of choice is the methylotrophic yeast *Pichia pastoris* X33. There are key advantages for using this expression system. Firstly, it is relatively inexpensive when compared to certain mammalian expression systems and can be grown to much higher densities. Secondly, recombinant protein expression can be tightly regulated by the methanol-inducible alcohol oxidase I gene (AOX1). Finally, post-translational modifications are eukaryotic and *Pichia pastoris* is capable of producing disulphide bonds.

This work was done in collaboration with Dr M. Jamshad.

6.2 Results

6.2.1 Screening for colonies expressing RAMP proteins

10 colonies of each RAMP construct were screened. One colony was found to be expressing RAMP1. Two colonies were found to be expressing RAMP2 and four colonies were expressing RAMP3 (see Figure 44). Taking into consideration the yeast consensus sequence and hexa histidine tag the theoretical molecular weight of RAMP1 is 15.1kDa, RAMP2 is 17.0kDa and RAMP3 is 14.9kDa as predicted by the compute pI/Mw utility on the expasy server (Bjellqvist *et al.*, 1993 and 1994). However, these predictions do not take into account post-translational modifications such as N-glycosylation or the existence of different oligomeric states of RAMP proteins. For example, Hilairet *et al.*, (2001) suggested independent RAMP1 expression in HEK293T cells would result in homodimerisation.

Lane order:

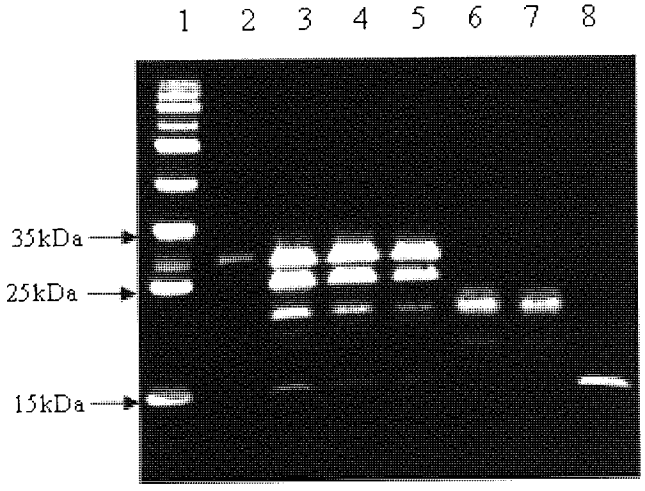


Figure 44. Western blot assessment of colonies expressing RAMP proteins. 10μg of total protein loaded per lane as assessed by NanpDrop 1000. The primary antibody used to probe the nitrocellulose membrane was the Clontech 6x HIS monoclonal antibody (Albumin free- Catalogue number: 631212). Followed by a secondary, Horse Radish Peroxidase (HRP) conjugated anti-mouse IgG antibody. The HRP signal was detected by EZ-ECL chemiluminescence. Lane 1: National Diagnostics Protometric ladder. Lane 2: RAMP3 (colony 1). Lane 3: RAMP3 (colony 2). Lane 4: RAMP3 (colony 3). Lane 5: RAMP3 (colony 4). Lane 6: RAMP2 (colony 1). Lane 7: RAMP2 (colony 2). Lane 8: RAMP1 (colony1).

6.2.2 RAMP fermentation

RAMP1 (colony 1), RAMP2 (colony 1) and RAMP3 (colony 3) were selected for fermentations. The fermentation for RAMP2 and RAMP3 were successfully completed. The overall biomass of *P.pastoris* expressing RAMP2 was 187 g/L wet cells. The overall biomass of *P.pastoris* expressing RAMP3 was 135 g/L wet cells. A western blot was conducted immediately after fermentation and confirmed RAMP2 and RAMP3 expression (see Figure 45). Unfortunately, 38 hours into the RAMP1 fermentation the methanol pump underwent a technical fault allowing too much methanol into the vessel. Presumably, the excess methanol caused toxicity and no RAMP1 expression was detected after fermentation (Figure 45). Consequently, no further work was conducted on RAMP1.

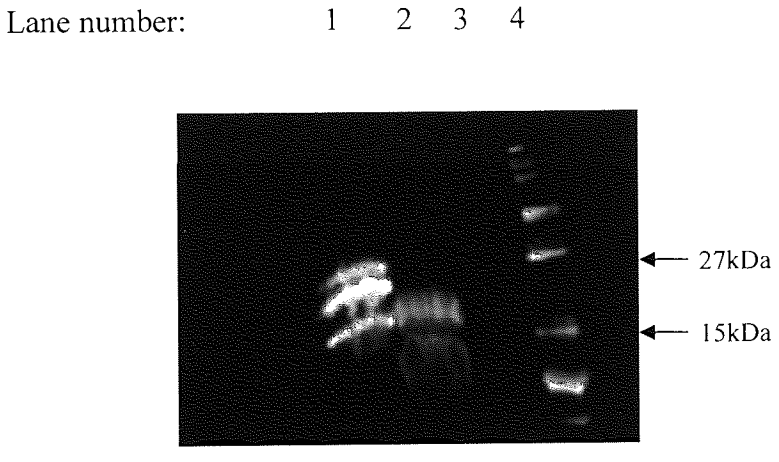


Figure 45. Western blot conducted immediately after fermentation. 10µg of total protein loaded per lane. Protein concentrations appoximated by NanoDrop 1000. The primary antibody used to probe the nitrocellulose membrane was the Clontech 6x HIS monoclonal antibody (Albumin free- Catalogue number: 631212). Followed by a secondary, Horse Radish Peroxidase (HRP) conjugated anti-mouse IgG antibody. The HRP signal was detected by chemiluminescence using the EZ-ECL chemiluminescence solution. Lane 1: RAMP3. Lane 2: RAMP2. Lane 3: RAMP1. Lane 4: plus pageRuler. The Western blot was conducted immediately after fermentation and presumably the 'blurry' blot is caused by methanol and basal salts in the samples.

6.2.3 RAMP2 and 3 solubilisation

The ProFoldin membrane protein kit was used to solubilise RAMP2 and RAMP3. It was found that Solution 8 produced the highest solubilised yields of protein assessed by Western blot analysis.

6.2.4 RAMP2 purification

RAMP2 was purified using the Qiagen 1.5ml Ni-NTA superflow column gravity flow method and the sample purity was assessed by Western blot analysis and silver staining (see Figure 46 and 47). Fractions from lane 6, 7 and 8 were pooled together, concentrated and dialysed to give a final 1.1mg/ml in a total of 2mls $NaH_2PO_4.2H_2O$/Solution 8. Taking into consideration that 20g of wet cells was cracked to produce this amount of protein; it can be approximated that ~20mg of protein could be potentially produced per 1L fermentation of RAMP2.

Figure 46. Western blot analysis of RAMP2 purification. 10µl of each sample was loaded on to the SDS polyacrylamide gel along with 1x Laemmli buffer. The primary antibody used to probe the nitrocellulose membrane was the Clontech 6x HIS monoclonal antibody (Albumin free- Catalogue number: 631212). Followed by a secondary, Horse Radish Peroxidase (HRP) conjugated anti-mouse IgG antibody. The HRP signal was detected by EZ-ECL chemiluminescence. Lane 1: Protometric ladder. Lane 2: flow-through. Lane 3: Wash 1. Lane 4: Wash 2. Lane 5: Elution 1 (100mM imidazole). Lane 6: Elution 2 (150mM imidazole). Lane 7: Elution 3 (200mM imidazole). Lane 8: Elution 4 (250mM imidazole). Lane 9: Elution 5 (300mM imidazole).

Lane number:

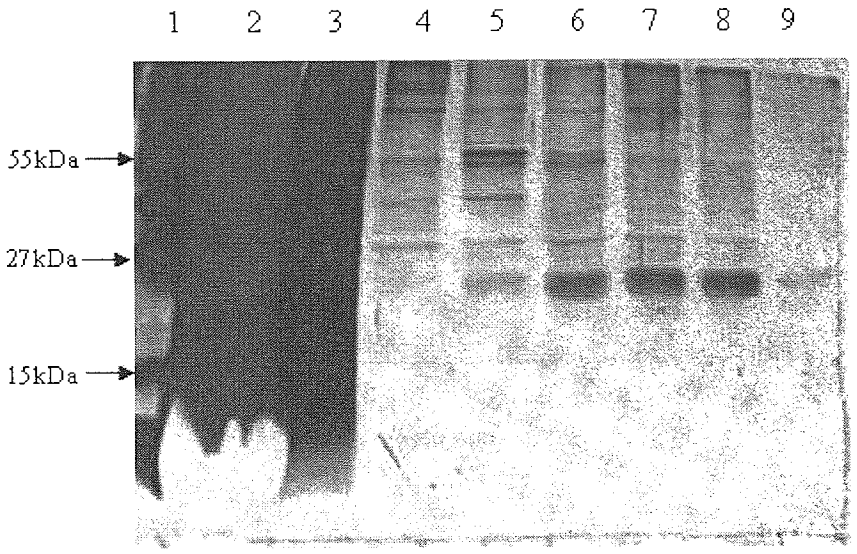


Figure 47. Silver stain analysis of RAMP2 purification. 10μl of each sample was loaded on to the SDS polyacrylamide gel along with 1x Laemmli buffer. Lane 1: Protometric ladder. Lane 2: flow-through. Lane 3: Wash 1. Lane 4: Wash 2. Lane 5: Elution 1 (100mM imidazole). Lane 6: Elution 2 (150mM imidazole) Lane 7: Elution 3 (200mM imidazole). Lane 8: Elution 4 (250mM imidazole). Lane 9: Elution 5 (300mM imidazole).

6.2.5 RAMP2 circular dichroism

Circular dichroism (CD) spectroscopy is a biophysical technique that measures the difference in the absorption of left-handed polarized light against right-handed polarized light. An ordered structure is typically asymmetrical and will result in positive and negative signals within the ultra-violet spectrum. Consequently, CD analysis can be used to determine whether a protein is folded. Taking the RAMP1 crystal structure into consideration (Kusano *et al.*, 2008) it is assumed that RAMP2 should predominately be α-helical. The CD spectra of an α-helix is characterized by a negative peak with a separate maxima of similar magnitude at 208nm and 222nm. Furthermore, a broad positive peak should be present between 190-200nm. CD analysis on RAMP2 suggests that the protein is folded and follows a typical α-helical read out (see Figure 48).

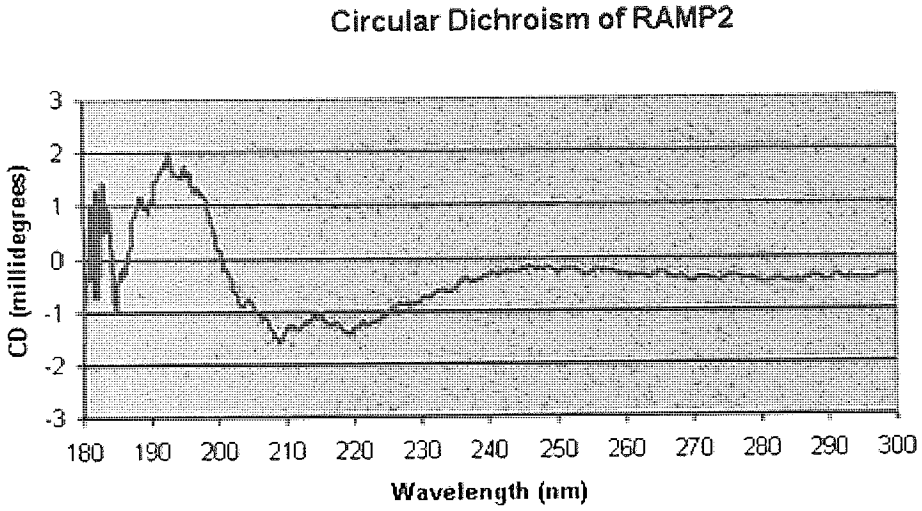


Figure 48. CD analysis of RAMP2. 50µl of dialysed RAMP2 sample (0.4 mg/ml) as assessed by the NanoDrop 1000 was used in CD analysis. The CD spectrum was collected using a cell pathlength of 0.1mm using the Jasco J-715 spectropolarimeter. The CD spectrum was an average of 96 scans, which was buffered corrected against 50mM $NaH_2PO_4.2H_2O$ containing 1% Soultion 8 pH 7.2. A negative signal is located at 208 nm and 222 nm.

6.2.6 RAMP3 purification

A Ni-NTA agarose batch method was used for RAMP3 purification. The method utilised a manual gradient elution protocol. Silver stain analysis demonstrates that this procedure was able to isolate RAMP3 successfully (see Figure 49).

Lane order:

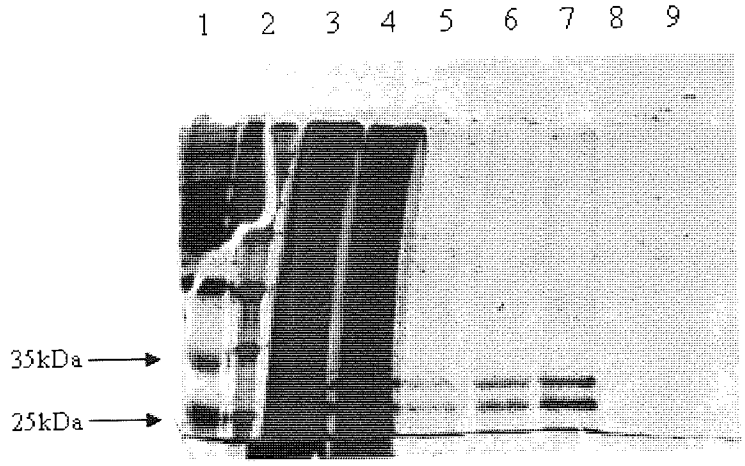


Figure 49. Silver stain of RAMP3 purification. 10µl of each sample was loaded on to the SDS polyacrylamide gel along with 1x Laemmli sample buffer. Lane 1: Protometric ladder. Lane 2: plus pageRuler. Lane 3: flow-through. Lane 4: wash 1 (40mM imidazole). Lane 5: Wash 2 (40mM imidazole wash). Lane 6: Elution 1 (80mM imidazole). Lane 7: Elution 2 (150mM imidazole). Lane 8: Elution 3 (150mM imidazole). Lane 9: Elution 4 (300mM imidazole).

Elution 1 and 2 (lane 6 and 7, respectively) were pooled together and concentrated to 0.75mg/ml in a total volume of 3ml. The sample underwent dialysis to remove unwanted imidazole. Taking into consideration that 20g (wet weight) of RAMP3 yeast pellet was used as the starting material for Emulsiflex-C3 cell disruption; it can be approximated that a 1L fermentation could yield ~15mg of protein. The concentrated RAMP3 sample was assessed with a silver stain, coomassie stain and Western blot to evaluate the purity and state of the protein after the concentration and dialysis steps (see Figure 50).

Figure 50. Concentrated RAMP3 silver stain, coomassie stain and Western blot assessment.

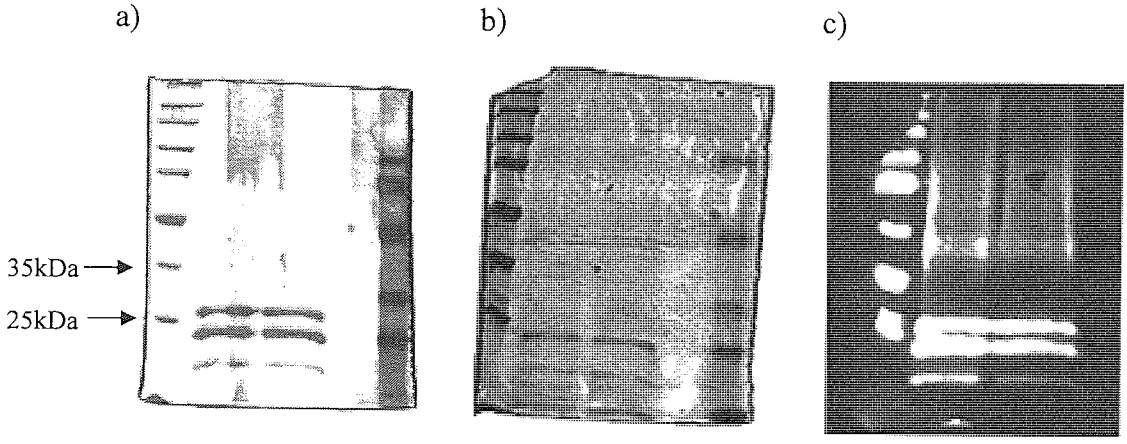


Figure 50. 10µg of total protein loaded per lane as assessed by NanpDrop 1000. a) Represents a silver stain, where the lane order from left to right includes the protometric ladder, concentrated RAMP3 (sample 1), concentrated RAMP3 (sample 2), plus pageRuler ladder. b) Brilliant blue coomassie stain where the lane order the same as in a). c) Western Blot analysis. The primary antibody used to probe the nitrocellulose membrane was the Clontech 6x HIS monoclonal antibody (Albumin free- Catalogue number: 631212). Followed by a secondary, Horse Radish Peroxidase (HRP) conjugated anti-mouse IgG antibody. The HRP signal was detected by EZ-ECL chemiluminescence. Western blot lane order was the same as in a).

A second purification of RAMP3 was conducted in an attempt to produce more protein to test the solubility of RAMP3 at higher concentrations. The second purification began with 36g of yeast pellet and different concentrations of imidazole in the elutions were used. Interestingly, 100mM imidazole was required for RAMP3 elution compared to 80mM imidazole in the first purification. The results are summarised by a silver stain (see Figure 51).

Lane order:

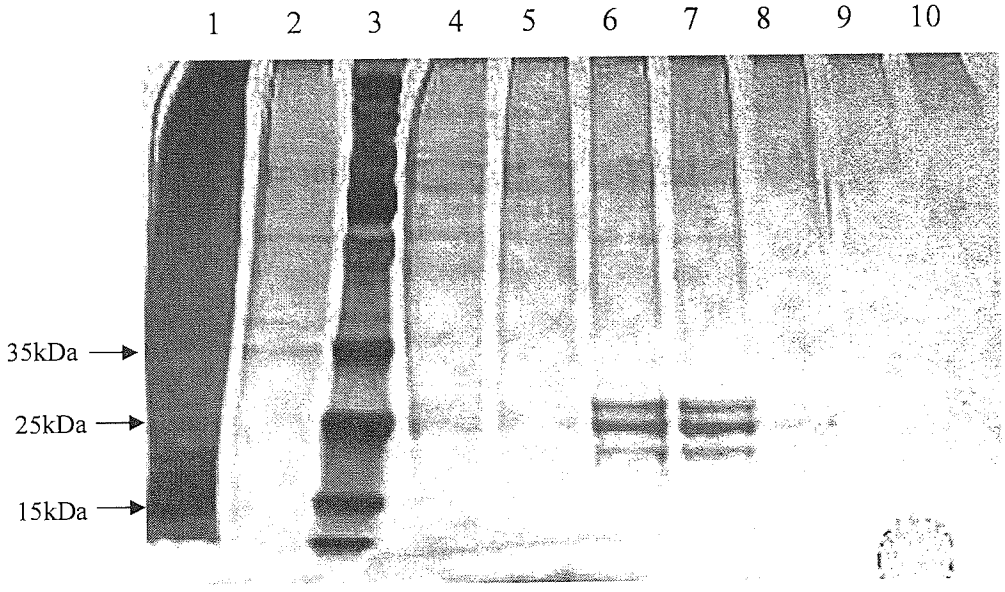


Figure 51. Silver stain of second RAMP3 purification. 10µl of each sample was loaded on to the SDS polyacrylamide gel along with 1x Laemmli sample buffer. Lane 1: flow-through. Lane 2: wash 1 (10mM imidazole). Lane 3: National diagnostic protometric ladder. Lane 4: wash 2 (40mM imidazole). Lane 5: wash 3 (40mM imidazole). Lane 6: elution 1 (100mM imidazole). Lane 7: elution 2 (250mM imidazole). Lane 8: elution 3 (350mM imidazole). Lane 9: elution 4 (500mM imidazole). Lane 10: elution 5 (700mM imidazole).

Elution 1, 2 and 3 (fractions from lane 6, 7 and 8 respectively) were pooled together and concentrated to 3.0mg/ml in a total volume of 1.2 ml. The concentrated RAMP3 sample was assessed by a silver stain and Western blot to ensure dialysis and concentration steps were successsful (see Figure 52).Although, RAMP3 has been successfully retained throughout these procedural steps (as assessed by hexa histidine epitope probing) the Western blot analysis does show hexa histidine epitope detection of proteins well above ~25kDa. This may provide early evidence for RAMP3 aggregation. However, more sophisticated experiments are required to assess the propensity for RAMP3 to aggregate (see discussion).

Lane order:

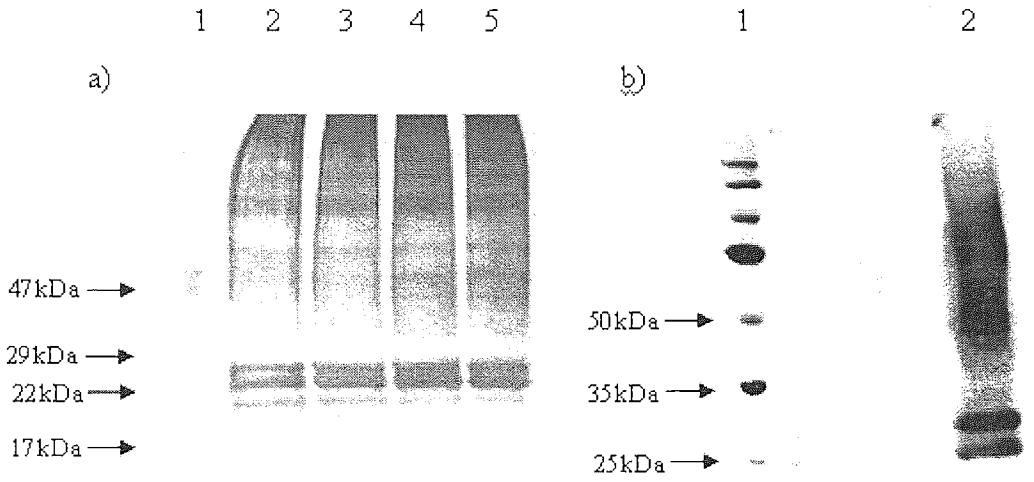


Figure 52. Assessment of 3.0mg/ml concentrated RAMP3 sample. Protein concentrations assessed by NanpDrop 1000 a) Represents silver stain analysis. Lane 1: GeneFlow wide range protomarker. Lane 2: 6µg of RAMP3. Lane 3: 8µg of RAMP3. Lane 4: 18µg of RAMP3. Lane 5: 24µg of RAMP3. b) Western blot background reversed to white. The primary antibody used to probe the nitrocellulose membrane was the Clontech 6x HIS monoclonal antibody (Albumin free- Catalogue number: 631212). Followed by a secondary, Horse Radish Peroxidase (HRP) conjugated anti-mouse IgG antibody. The HRP signal was detected by EZ-ECL chemiluminescence. Lane 1: protometric ladder. Lane 2: 30µg of RAMP3 sample.

6.2.7 RAMP3 circular dichroism

CD analysis was conducted on the second concentrated RAMP3 purification sample. The CD spectrum suggests that RAMP3 is folded and broadly follows an α-helical trace (see Figure 53).

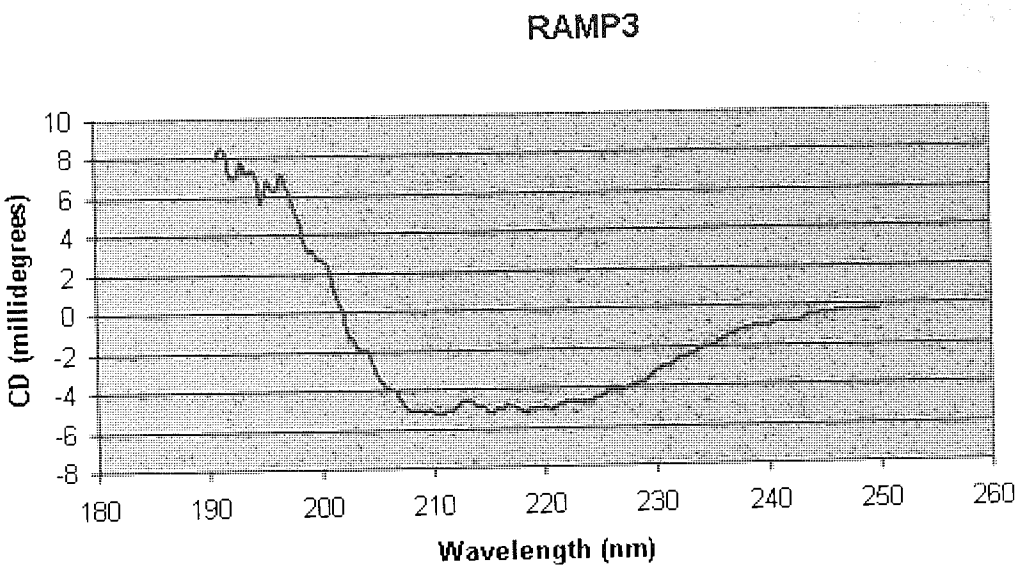


Figure 53. CD analysis of RAMP3. 50µl of dialysed RAMP3 sample (2.6 mg/ml; as assessed by the NanoDrop 1000) was used in CD analysis. The CD spectrum was collected using a cell pathlength of 0.1mm using the Jasco J-715 spectropolarimeter. The CD spectrum was an average of 4 scans, which was buffered corrected against 50mM $NaH_2PO_4.2H_2O$ containing 1% Soultion 8 pH 7.2.

6.3 Discussion

Each member of the human RAMP family was C-terminally tagged with a hexa histidine tag and successfully cloned into the *P.pastoris* expression vector pPICZB downstream of the AOX1 promoter. Furthermore, it was found that the wild type X33 strain of *P.pastoris* was capable of expressing all three recombinant RAMP proteins in successful transformants. Moreover, RAMP2 and RAMP3 expressing colonies were amenable to 1L fermentations and despite a technical fault there is no reason at present why RAMP1 expressing colonies could not be cultivated using this approach to generate large amounts of biomass prior to protein purification. RAMP2 and RAMP3 were solubilised and purified using a manual approach which was capable of yielding a CD spectrum that suggested that both RAMP2 and RAMP3 were folded.

In spite of the initial encouraging results there are many practical problems that need to be resolved if the RAMP proteins are to be to fully characterised. At present, the cellular localisation of RAMP1, 2 and 3 has not been thoroughly investigated. Only the membrane fraction of *P.pastoris* has been probed for the presence of RAMPs, based on a tentative assumption that the protein is more likely to be folded when

inserted into a membrane. However, taking into consideration that RAMP2 and RAMP3 are being expressed independently without a partner GPCR it would be of interest to determine where the proteins reside in this particular host system. An initial but crude experiment would be to conduct a western blot analysis comparing the relative concentrations of RAMP in the membrane fraction to the non-membrane fraction. However, to further understand the protein distribution within the cell more sophisticated analysis is required. Laser confocal microscopy combined with immunofluorescence could potentially be implemented to visualise the distribution of RAMP protein in fixed cells. However, the incorporation of a fluorophore e.g. Green fluorescent protein to replace the C-terminal histidine tag in the RAMP constructs could lead to quantitative time-lapse imaging, which is capable of providing kinetic information on protein migration in a live cell (see Presley *et al.,* 2005, for a review on live cell optical techniques).

The choice of detergent is essential as membrane proteins have a tendency to non-specifically aggregate over-time (Gutmann *et al.,* 2007). Non-specific irregular protein aggregation in solution can have adverse effects on future experiments for instance it hinders crystal lattice formation (Gutmann *et al.,* 2007). Iwata (2003) noted that the detergent used for extraction and purification will not necessarily be optimal in promoting the formation of diffracting crystals. Newstead *et al.,* (2008) analysed the crystallisation conditions of 121 α-helical membrane proteins and found that the most successful group of detergents were alkyl maltopyranosides, specifically n-Dodecyl-β-D-maltopyranoside (DDM). DDM ($C_{24}H_{46}O_{11}$, M_w = 510.62) is a relatively large detergent that forms large micelles which is beneficial since it reduces the likelihood of protein aggregation but its disadvantage is that the large micelles reduce the amount of exposed protein needed to form protein–protein interactions that ultimately facilitate protein crystal lattice formation. Newstead *et al.,* (2008) noted that the second most successful detergent family was the alkyl glucopyranosides. A popular member of this family of detergents is n-Octyl-β-D-glucopyranoside (β-OG). β-OG ($C_{14}H_{28}O_6$, M_w = 292.37) is smaller than DDM and therefore forms smaller micelles. However, Newstead *et al.,* (2008) found β-OG was more successful in crystalising protein channels. The reason for this is obscure but Newstead *et al.,* (2008) suggested that channels were inherently more stable than other membrane proteins within the analysis inferring that a protein has to be particularly stable to tolerate β-OG micelles.

RAMP2 and RAMP3 have been solubilised in the presence of ProFoldin solution 8 (a non-denaturing zwitterionic detergent). Dr Mohammed Jamshad is currently conducting sedimentation velocity analytical ultracentrifugation (AUC) experiments on the RAMP samples at Birmingham University in Dr Tim Dafforn's Laboratory. This approach can give estimations of the molecular mass of the protein and inferences on the proteins oligomeric state. Moreover, the amount of protein aggregation can be estimated, which is of particular interest given that protein/detergent complexes have to be in a monodisperse state for crystallisation.

The ability of ProFoldin solution 8 to facilitate protein crystal formation has not been assessed neither has the solubility of RAMP proteins over 3mg/ml. These are key factors that have to be verified to evaluate whether the RAMP protein constructs and/or the purification procedure are sufficient for crystallisation trials.

However, taking into consideration that the approach outlined here is modular, additional steps can be added or existing steps modified. For example, if protein aggregation becomes problematic at higher RAMP concentrations preventative steps or aggregation removal steps could be integrated into the current methodology. The addition of a kosmotrope, which contributes to the stability and structure of water-water interactions, can enhance the native protein stability reducing aggregation. Strong kosmotropes include $MgSO_4$ and Na_2SO_4 and weaker kosmotropes include NaCl and KCl (Bondos and Bicknell 2002). Alternatively the addition of a choatrope, such as $CaCl_2$ and $MgCl_2$, can destabilise the protein aggregate directly. Other preventative aggregation solutes include sugars (e.g. glucose and sucrose), polyhydric alcohols (e.g. glycerol) and amino acids (e.g. glycine and L-arginine)- see Bondos and Bicknell (2002). A common removal step for protein aggregates is size exclusion chromatography. However, this approach is low throughput and typically requires large amounts of protein and buffer containing detergent and can ultimately be very expensive (Gutmann *et al.*, 2007). Inevitably protein loss will occur at each step of a purification process and this has to be carefully considered if increasing the number of steps in the procedure.

Homogeneity of a protein sample is another factor that can have implications on structural studies. RAMP3 has four Asn-X-Ser/Thr sequons (X can be any amino acid

except a proline) suggestive that it could potentially have four N-glycosylation sites. RAMP2 has one Asn-X-Ser/Thr sequon and RAMP1 has none. The host system *Pichia pastoris* is able to N-glycosylate proteins. Although, N-glycosylation is primarily associated with secreted proteins (Daly and Hearn, 2005) it may explain the heterogeneity in the protein population in the RAMP2 and RAMP3 samples as assessed by SDS PAGE. To test this hypothesis the purified RAMP samples could be treated with an exo-glycosidase (e.g. endo-β-*N*-acetylglucosaminidases) and re-analysed.

The general dogma suggests that the standard practice is to submit a protein for crystallisation trials when the protein is ~90% pure as assessed by a Coomassie stain. However, even when the advances in affinity chromatography are taken into consideration the removal of all impurities in a protein sample remains extremely challenging.

Kors *et al.,* (2009) analysed the effects of impurities on membrane protein crystallisation using the photosynthetic reaction centre of *Rhodobacter sphaeroides* as a test case. Kors *et al.,* (2009) assessed three common crystallisation methodologies including sitting-drop vapour diffusion, microbatch plug and lipidic cubic phase. The results suggested that the sitting-drop vapour diffusion and microbatch plug based method was sensitive to impurities as it hindered crystal nucleation. Yet crystals were still obtained using these methods at purities as low as ~75% in vapour diffusion method and ~60% in the microbatch plug method. However, the lipidic cubic phase crystallisation approach could produce high quality diffracting crystals (3.5 Å) when only 50% of the target protein was in the sample, which contained both protein and lipid impurities. However, when the target protein dropped to ~43% crystallisation was no longer achievable. Kors *et al.,* (2009) advised that lipidic cubic phase crystallisation should be the method of choice for membrane protein crystallisation trials. However, the findings generally suggest that crystallisation is very robust and is not particularly sensitive to impurities.

RAMP2 was purified using the Qiagen 1.5ml Ni-NTA superflow column gravity flow approach while RAMP3 was purified using a batch method. The main difference between the two approaches is the time given for the target protein to equilibrate with the Ni-NTA resin. The RAMP3 batch approach allows the target protein to equilibrate

with the resin overnight whereas the target protein only interacts with the resin in the RAMP2 column gravity flow approach as the lysate drains through the column. However, a direct assessment between the two approaches where the yield of each elute and the amount of target protein in flow-through (i.e. amount of target protein unable to bind to the resin) has not been conducted and therefore it remains difficult to directly compare the two approaches. However, on face value it appears the RAMP3 purification method seems more successful. Future aspirations would be to design an automated purification procedure using fast protein liquid chromatography. Theoretically, this approach is superior to the manual approach since the concentration of imidazole can gradually be increased and the target protein elution would be expected to follow a Gaussian distribution whereby the purest fractions could then be selected.

The purity of the RAMP3 samples by the batch manual approach suggests the impurity of the sample will not be a bottleneck factor in crystallisation trials. However, given the recent findings by Kors *et al.,* (2009) and the recent increase in the number of membrane protein structures being elucidated by lipidic cubic phase crystallisation a change in strategy may have to be proposed to get full RAMPs to crystallisation trials. Grabe *et al.,* (2003) produced a theoretical analysis of how the lipidic cubic phase method facilitated crystal nucleation and growth, when the cubic matrix is made from monoolein lipids. Briefly, when a protein is presented to the cubic structure it embeds itself and proteins tend to cluster together this induces strain to the surrounding bilayer (a process driven by elastostatic forces). In attempt to relieve this strain the matrix flattens into lamellar stacks that eventually form arrays that can undergo X-ray analysis. However, it was noted by Grabe *et al.,* (2003) that larger proteins prefer the lamellar stack formation but smaller proteins do not have enough elastic driving force to induce the shift to lamellar stacks. Consequently, Grabe *et al.,* (2003) went onto suggest that this crystallisation approach is only amenable to membrane proteins that have five or more TM regions.

Taking into consideration that RAMPs only contain a single TM region it maybe beneficial to express the RAMP proteins with an associated GPCR, such as CLR. Attempting high-resolution structural studies on a CLR and RAMP complex has many practical obstacles and obviously if achieved would be regarded as a significant milestone in Secretin-like GPCR research. Yet, the recent success in Rhodopsin-like

GPCR structural research has provided a benchmark that can be emulated. The pharmacological profile of CLR and RAMP1, 2 and 3 co-expression has already been characterised in the yeast *Saccharomyces cerevisiae* and was found to resemble mammalian profiles (Miret *et al.,* 2002) suggestive that a yeast host could produce functional CGRP and AM receptors. Once solubilised the thermostability of CGRP or AM receptors could be assessed by a radioligand based assay incubated at different temperatures, an approach used by Serrano-Vega (2008) that led to the crystallisation of the turkey adrenergic β1 receptor. Although, differential scanning calorimetry is reported as the most frequent approach used to assess protein stability (Cueto *et al.,* 2002). Such an approach would give novel insight into the inherent thermostability of the CGRP and AM receptors.

Moreover, ICL3 is predicted to only be 8 residues in length in the CLR. Therefore, it would not be expected to be as flexible and problematic as in Rhodopsin-like GPCRs. Consequently, the development of fusion proteins and ICL3 specific antibodies may not be necessary. However, if the intracellular face of CLR was inherently too flexible and hindered crystallisation the addition of RCP, an endogenous protein shown to interact with CLR, may provide additional structure. However, this hypothesis would have to be empirically tested and problems with purifying RCP would have to be overcome (Tolun *et al.,* 2007).

However, being able to analyse independent RAMP constructs should not be underestimated. The production of independent RAMP structures could potentially provide assays that will aid in understanding RAMP homo and hetero-dimerisation. For example, surface plasmon resonance is an optical method that is able to measure the refractive index near a sensor. Consequently, it can assess protein interactions. The process involves immobilising a protein onto the surface sensor while the binding protein (often referred to as the analyte) is in solution and continuously flowed over the surface sensor, while changes in refractive index are being assessed. Theoretically, the RAMP proteins could be immobilised on to Ni-NTA BIAcore sensor chips (see Nieba *et al.,* 1997). This would be advantageous as the hexa histidine tag is located on the C-terminus of RAMP constructs and therefore the orientation of the RAMP protein would be consistent. Then purified ECDs of Secretin-like GPCRs could be used as the analyte. This could be used to assess which Secretin-like ECDs interact

with a RAMP and the extent to which this interaction is governed by the ECD of the Secretin-like ECD.

In conclusion, the RAMP family are able to be expressed using *P. pastoris* as a host system and a degree of success has been accomplished in purifying RAMP2 and 3. Yet, this body of work remains in its infancy and extensive work is required before high resolution structural studies could be a feasible goal. However, the limitations, methodology and ideas discussed here could be useful to guide future work in this area.

Chapter 7: General discussion and future considerations

The CGRP receptor is a complex macromolecule and its tertiary structure remains elusive. Consequently, fundamental questions remain in understanding how this receptor facilitates agonist binding and mediates signal transduction. The primary focus of this thesis was to identify 'hot spots' in the extracellular face of the CGRP receptor. This pragmatic approach was used in an attempt to gain novel insight into the role of specific amino acids within the CGRP receptor. To this end the results were interpreted in light of contemporary views and concepts that are applicable to Secretin-like GPCRs.

Ligand binding to the CGRP receptor remains a thorny issue. The recent advances by Koth *et al.,* (2010) have shed light on non-peptide antagonist binding which appears to be entirely governed by the ECD of the CGRP receptor. However, agonist binding is more complicated as the C-terminus of αCGRP, which is required for high affinity binding, is predicted to dock against the ECD of the CGRP receptor whereas the N-terminus of the ligand, required for efficacy, binds to the TM domain of the receptor. Taking into account Koth *et al.,* (2010) recent findings it is most probable that the C-terminus of αCGRP adheres to the 'hot dog in a bun' model as loop 4 of the CLR ECD is accessible. Interesting, an alanine/leucine scan of the extreme N-terminus of the CLR (discussed in Chapter 5) identified novel residues that were important for CGRP receptor pharmacology, most notably I41A, A44L, Q45A, C48A and Y49A. Given that I41A, Q45A and Y49A show periodicity it is assumed that the residues are part of the predicted N-terminal helix and may participate in the RAMP1 interface or/and αCGRP binding.

However, it remains ambiguous how the N-terminus of αCGRP penetrates into and activates the TM domain. Chapter 3 and 4 discuss an alanine/leucine scan of ECL1 and ECL3 and its associated TM regions. The results taken together with Dr. A Conner's preliminary work on ECL2 provides a novel insight into the important residues on the extracellular surface of the CGRP receptor. In turn, a speculative model of the active and inactive state was constructed to aid visualisation of the TM domain. Investigation into ECL1, along with the top of TM2 and TM3, found that C212 was required for normal CGRP receptor functioning, presumably because it forms a disulphide bond with C282, an interaction that is thought to be preserved

across all GPCR families. A novel triplet cluster of residues (L195, V198 and A199) that resides at the top of TM2 was identified, which may directly contribute to the αCGRP orthosteric binding site. However, further investigation is needed to validate this assumption. H219A, L220A and L222A alter cAMP production and given that all three residues are predicted to be located in the middle of TM3 it is plausible that the residues are involved in signal transduction and/or contribute to TM packing. Furthermore, A203L and A206L that are predicted to be located in ECL1 were found to increase cAMP production. Although, the underling mechanism for this remains elusive in may give an insight into αCGRP docking, albeit inhibition binding assays did not reveal significant effects.

Investigation into ECL3 and the corresponding TM regions also revealed interesting and novel results. In particular, L351 is located two turns down in TM6 and predicted to face the lipid environment. Mutating this residue to alanine impaired the potency and the binding of CGRP along with decreasing cell surface expression. Recently, Harikumar *et al.,* (2009) has suggested that RAMP3 binds to TM6 and TM7 in secretin receptors. Consequently, it is tempting to suggest that the effects of L351A and the number of mutations in this region that were found to reduce cell surface expression could be attributed to RAMP1 association. However, to establish this further research is required. Harikumar *et al.,* (2009) primarily used a TM peptide BRET competition assay to determine which TM regions within the secretin receptor allow RAMP3 association. This approach in conjunction with cysteine mapping mutagenesis strategy outlined by Gao *et al.,*(2009) could establish whether RAMP1 does dock against TM6 and TM7 of CLR and give insight into the participating residues. An alternative strategy is to use the TOXCAT assay (Russ and Engelman, 1998), which evaluates direct associations between TM regions.

In spite of initial encouraging results fundamental questions regarding the orientation of the ligand in relation to the TM domain remain. Neumann *et al.,* (2008) suggested that Secretin-like GPCR ligands all share a helix N-capping motif and postulated that this motif could induce activation in all cognate receptors. However, Runge *et al.,* (2003) work from the glucagon and glucagon like receptors suggested that the N-terminus of the agonist was located between TM1, TM2 and TM7 but the Monaghan *et al.,* (2008) model of parathyroid hormone docking suggested the N-terminus of the ligand ends up near TM5 and TM6. It is perfectly feasible that the two ligands do

have distinct binding sites. Moreover, as Secretin-like GPCR ligands are relatively large peptides it is understandable that the pharmacophore is diffuse and that the ligands may span the entire extracellular TM domain. It is tempting to speculate that there could in fact be two molecular switches located on the extracellular face of Secretin-like GPCRs; one located between TM1, TM2 and TM7 and the other located between TM6 and TM7 which could work in unison or synergistically. This provocative (and note untested hypothesis) could also shed light on receptors that were found to be activated by the WDN/G motif as it was suggested that this 'internal agonist' may induce activation *via* TM6/ECL3 (Dong *et al.,* 2006). Maybe secretin is able to fully activate the TM1, TM2 and TM7 switch but needs further assistance when activating the TM5 and TM6 switch.The plausibility of an internal agonist within Secretin-like GPCRs remains unclear but it is interesting to note that another well conserved WD epitope is found in most ECL2 domains of Secretin-like GPCRs.

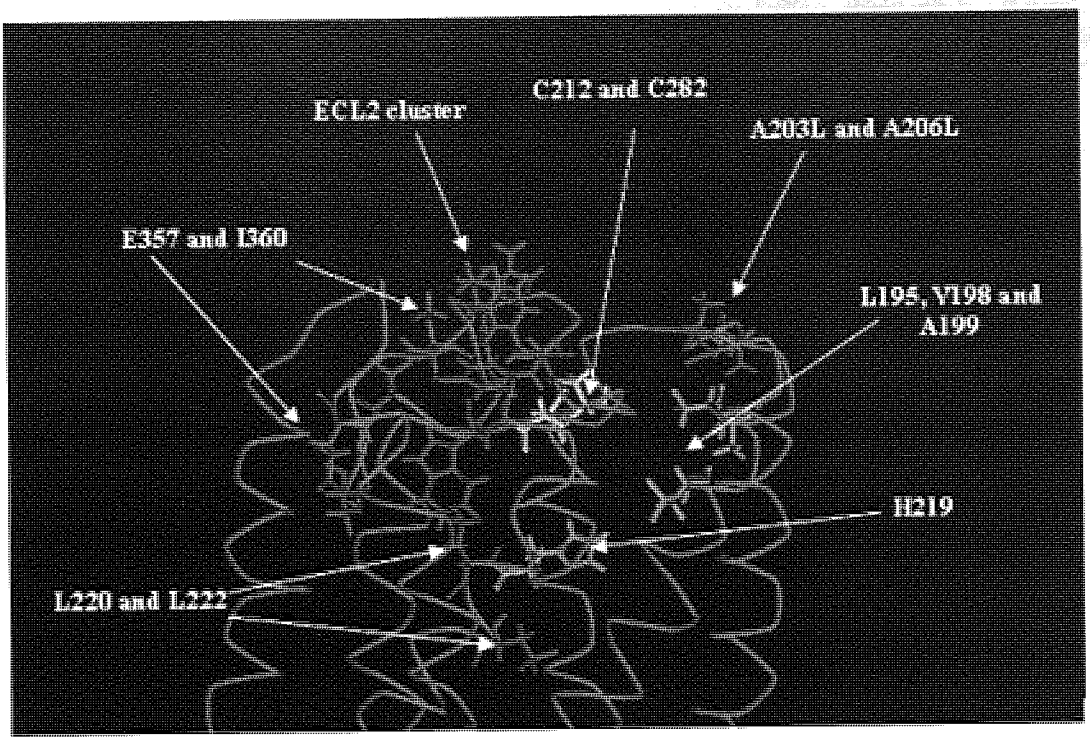


Figure 54. Summary of residues within the extracellular face of CLR that when mutated affect CGRP receptor pharmacology. Assumed inactive CLR TM domain (cyan ribbon) generated from ground-state rhodopsin (PDB accession: 1U19). Residues found to decrease the potency of αCGRP in stimulating cAMP when mutated to alanine/leucine are highlighted orange when associated with ECL1, red when associated with ECL2 and bright pink when associated with ECL3. The disulphide bond between C212 and C282 highlighted yellow. Residues found to increase the potency of αCGRP in stimulating cAMP are highlighted purple.

A hypothesis of this nature is currently difficult to assess as deciphering the relative position of the N-terminal ECD in relation to the TM bundle is extremely challenging, since the two domains are linked by a flexible loop. Vilardaga *et al.*, (1997) investigation into the role of cysteine residues in the rat secretin receptor using serine mutations provides the only empirical distance restraint between the two domains as a disulphide bond was suggested to be present between Cys-11 (located in the extreme N-terminus of the receptor) and Cys-186 (located at the top of TM2). However, even with one distance restraint it still remains difficult to predict the relative ECD orientation. In respect to solving this problem an approach similar to Thomas *et al.*, (2009) could be employed. Thomas *et al.*, (2009) used an engineered disulphide bond formation strategy to study conformational changes in PTHR1, where a Factor Xa cleavage site was incorporated in ICL3.If an engineered disulphide bond had formed successfully between TM2 and TM7 the receptor would remain the same molecular weight after Factor Xa treatment. However, instead of analysing distance restraints

between TM domains, disulphide bonds could be introduced between the ECD and TM domain in Secretin-like GPCRs.

Yet, the position of the αCGRP N-terminus within the TM domain is still unclear. An initial study to investigate this would be to perform binding experiments or Schild plots using the antagonist $\alpha CGRP_{8-37}$ on mutations that have already been found to alter normal αCGRP binding. This may give an indication on the location of the active N-terminal epitope. However, more sophisticated experiments employing photaffinity cross-liking studies or cysteine trapping experiments may be more informative.

Chapter 6 discusses a procedure used to purify the human RAMP family. The initial results are positive but taking into account the high concentrations needed for biophysical characterisation, such as X-ray crystal, protein aggregation maybe a future concern. Parthier *et al.,* (2009) suggested the next milestone in Secretin-like GPCR research is the elucidation of the whole GPCR or at least the TM domain. Consequently, the 'divide and conquer' approach used here may want to be revised and attempts on gathering high resolution data for the entire CGRP receptor may want to be attempted. However, given that no high resolution data is currently available for Secretin-like TM domains only speculative models based on presumed structure similarity can be generated. An inactive and active CLR TM bundle was generated using an alignment put forward by Professor Christopher Reynold's group. The main limitation in developing such alignments is being able to determine the exact start and end of each TM helix, in particular in this case TM5. SCAM is an experimental technique that can be used with a range of thiol-specific reagents that vary in size, charge and membrane permeability, which could be used to detect the beginning of TM regions and identify periodicity in the CLR TM bundle (see Bogdanov *et al.,* 2004).

In conclusion, a full understanding of receptor activation is being able to identify different states of the receptor along with understanding the molecular process that governs the trajectory between such states. Currently, understanding the CGRP receptor on this mechanistic level remains in its infancy. The lack of high resolution structural data on the TM domain regions coupled with discrepancies in the stoichiometric arrangement of the receptor needs to be addressed. However, this is no mean feat. The stoichiometry of the CGRP receptor could be governed by a plethora

of variables including cell-line specific factors, cell compartmentalisation and lipid bilayer lateral movement. However, research in this area has remained vibrant, which in part maybe ascribed to the potential pharmaceutical benefit, but also the CGRP receptor encompasses many novel and challenging aspects of GPCR research such as dimerisation, accessory proteins and elucidation of large diffuse pharmacophores. Consequently, the pursuit in discovering the mechanism/s that govern CGRP receptor activation remains both novel and exciting.

References

Ahuja S, Hornak V, Yan EC, Syrett N, Goncalves JA, Hirshfeld A, Ziliox M, Sakmar TP, Sheves M, Reeves PJ, Smith SO, Eilers M. (2009a) Helix movement is coupled to displacement of the second extracellular loop in rhodopsin activation. Nat Struct Mol Biol 16(2):168-75

Ahuja S, Crocker E, Eilers M, Hornak V, Hirshfeld A, Ziliox M, Syrett N, Reeves PJ, Khorana HG, Sheves M, Smith SO. (2009b) Location of the retinal chromophore in the activated state of rhodopsin*. J Biol Chem 284 (15):10190-201

Ahuja S and Smith SO (2009) Multiple switches in G protein-coupled receptor activation. Trends Pharmacol Sci 30(9):494-502

Aiyar N, Daines RA, Disa J, Chambers PA, Sauermelch CF, Quiniou M, Khandoudi N, Gout B, Douglas SA, Willette RN (2001) Pharmacology of SB-273779, a nonpeptide calcitonin gene-related peptide 1 receptor antagonist. J Pharmacol Exp Ther 296:768-775

Aldecoa A, Gujer R, Fischer JA, 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 471:156-160

Altenbach C, Kusnetzow AK, Ernst OP, Hofmann KP, Hubbell WL (2008) High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. Proc Natl Acad Sci U S A 105:7439-7444

Amara SG, Jonas V, Rosenfeld MG, Ong ES, Evans RM (1982) Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. Nature 298:240-244

Andrew ER, Bradbury A, and Eades, RG (1959) Removal of Dipolar Broadening of Nuclear Magnetic Resonance Spectra of Solids by Specimen Rotation. Nature 183: 1802.

Arnis S, Fahmy K, Hofmann KP, Sakmar TP (1994) A conserved carboxylic acid group mediates light-dependent proton uptake and signaling by rhodopsin. J Biol Chem 269:23879-23881

Arulmani U, Maassenvandenbrink A, Villalon CM, Saxena PR (2004) Calcitonin gene-related peptide and its role in migraine pathophysiology. Eur J Pharmacol 500:315-330

Assil-Kishawi I, Abou-Samra AB (2002) Sauvagine cross-links to the second extracellular loop of the corticotropin-releasing factor type 1 receptor. J Biol Chem 277:32558-32561

Banerjee S, Evanson J, Harris E, Lowe SL, Thomasson KA, Porter JE (2006) Identification of specific calcitonin-like receptor residues important for calcitonin gene-related peptide high affinity binding. BMC Pharmacol 6:9

Bas DC, Rogers DM , and Jensen JH (2008) Very fast prediction of pKa values for protein-ligand complexes. Proteins, 73, 765-783

Barwell J, Miller PS, Donnelly D, Poyner DR (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 31:170-176

Behar V, Nakamoto C, Greenberg Z, Bisello A, Suva LJ, Rosenblatt M, Chorev M (1996) Histidine at position 5 is the specificity "switch" between two parathyroid hormone receptor subtypes. Endocrinology 137:4217-4224

Bellot G, Granier S, Bourguet W, Seyer R, Rahmeh R, Mouillac B, Pascal R, Mendre C, Demene H (2009) Structure of the third intracellular loop of the vasopressin V2 receptor and conformational changes upon binding to gC1qR. J Mol Biol 388:491-507

Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340:783-795

Benitez-Paez A, Cardenas-Brito S (2008) Dissection of functional residues in receptor activity-modifying proteins through phylogenetic and statistical analyses. Evol Bioinform Online 4:153-169

Bergwitz C, Jusseaume SA, Luck MD, Juppner H, Gardella TJ (1997) Residues in the membrane-spanning and extracellular loop regions of the parathyroid hormone (PTH)-2 receptor determine signaling selectivity for PTH and PTH-related peptide. J Biol Chem 272:28861-28868

Bisello A, Adams AE, Mierke DF, Pellegrini M, Rosenblatt M, Suva LJ, Chorev M (1998) Parathyroid hormone-receptor interactions identified directly by photocross-linking and molecular modeling studies. J Biol Chem 273:22498-22505

Bissantz C, Logean A, Rognan D (2004) High-throughput modeling of human G-protein coupled receptors: amino acid sequence alignment, three-dimensional model building, and receptor library screening. J Chem Inf Comput Sci 44:1162-1176

Bogdanov M, Zhang W, Xie J, Dowhan W (2005) Transmembrane protein topology mapping by the substituted cysteine accessibility method (SCAM(TM)): application to lipid-specific membrane protein topogenesis. Methods 36:148-171

Bomberger JM, Parameswaran N, Hall CS, Aiyar N, Spielman WS (2005a) Novel function for receptor activity-modifying proteins (RAMPs) in post-endocytic receptor trafficking. J Biol Chem 280:9297-9307

Bomberger JM, Spielman WS, Hall CS, Weinman EJ, Parameswaran N (2005b) Receptor activity-modifying protein (RAMP) isoform-specific regulation of

adrenomedullin receptor trafficking by NHERF-1. J Biol Chem 280:23926-23935

Bondos SE, Bicknell A (2003) Detection and prevention of protein aggregation before, during, and after purification. Anal Biochem 316:223-231

Bouschet T, Martin S, Henley JM (2005) Receptor-activity-modifying proteins are required for forward trafficking of the calcium-sensing receptor to the plasma membrane. J Cell Sci 118:4709-4720

Brain SD, Cox HM (2006) Neuropeptides and their receptors: innovative science providing novel therapeutic targets. Br J Pharmacol 147 Suppl 1:S202-211

Brain SD, Grant AD (2004) Vascular actions of calcitonin gene-related peptide and adrenomedullin. Physiol Rev 84:903-934

Breeze AL, Harvey TS, Bazzo R, Campbell ID (1991) Solution structure of human calcitonin gene-related peptide by 1H NMR and distance geometry with restrained molecular dynamics. Biochemistry 30:575 - 582

Bruccoleri RE, Karplus M (1987) Prediction of the folding of short polypeptide segments by uniform conformational sampling. Biopolymers 26:137-168

Buhlmann N, Leuthauser K, Muff R, Fischer JA, Born W (1999) A receptor activity modifying protein (RAMP)2-dependent adrenomedullin receptor is a calcitonin gene-related peptide receptor when coexpressed with human RAMP1. Endocrinology 140:2883-2890

Bywater RP (2005) Location and nature of the residues important for ligand recognition in G-protein coupled receptors. J Mol Recognit 18:60-72

Carpenter KA, Schmidt R, von Mentzer B, Haglund U, Roberts E, Walpole C (2001) Turn structures in CGRP C-terminal analogues promote stable arrangements of key residue side chains. Biochemistry 40:8317-8325

Chakravarty S, Godbole S, Zhang B, Berger S, Sanchez R (2008) Systematic analysis of the effect of multiple templates on the accuracy of comparative models of protein structure. BMC Structural Biology 8:31

Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science 318:1258-1265

Chiba T, Yamaguchi A, Yamatani T, Nakamura A, Morishita T, Inui T, Fukase M, Noda T, Fujita T (1989) Calcitonin gene-related peptide receptor antagonist human CGRP-(8-37). Am J Physiol 256:E331-335

Christopoulos A, Christopoulos G, Morfis M, Udawela M, Laburthe M, Couvineau A, Kuwasako K, Tilakaratne N, Sexton PM (2003) Novel receptor partners and function of receptor activity-modifying proteins. J Biol Chem 278:3293-3297

Chugunov AO, Simms J, Poyner DR, Dehouck Y, Rooman M, Gilis D, Langer I (awaiting publication) Evidence that interaction between conserved residues in transmembrane helices 2,3 and 7 are crucial for human VPAC1 receptor activation.

Chugunov AO, Novoseletsky VN, Arseniev AS, Efremov RG (2007) A novel method for packing quality assessment of transmembrane alpha-helical domains in proteins. Biochemistry (Mosc) 72:293-300

Clackson T, Wells JA (1995) A hot spot of binding energy in a hormone-receptor interface. Science 267:383-386.

Claros MG, von Heijne G (1994) TopPred II: an improved software for membrane protein structure predictions. Comput Appl Biosci 10:685-686

Colquhoun D (2006) Agonist-activated ion channels. Br J Pharmacol 147 Suppl 1:S17-26

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 30:451-455

Conner AC, Hay DL, Simms J, Howitt SG, Schindler M, Smith DM, Wheatley M, Poyner DR (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 67:20-31

Conner AC, Simms J, Conner MT, Wootten DL, Wheatley M, Poyner DR (2006) Diverse functional motifs within the three intracellular loops of the CGRP1 receptor. Biochemistry 45:12976-12985

Conner AC, Simms J, Hay DL, Mahmoud K, Howitt SG, Wheatley M, Poyner DR (2004) Heterodimers and family-B GPCRs: RAMPs, CGRP and adrenomedullin. Biochem Soc Trans 32:843-846

Conner M, Hicks MR, Dafforn T, Knowles TJ, Ludwig C, Staddon S, Overduin M, Gunther UL, Thome J, Wheatley M, Poyner DR, Conner AC (2008) Functional and biophysical analysis of the C-terminus of the CGRP-receptor; a family B GPCR. Biochemistry 47:8434-8444

Couvineau A, Ceraudo E, Tan YV, Laburthe M (2009) VPAC1 receptor binding site: Contribution of photoaffinity labeling approach. Neuropeptides

Couvineau A, Fabre C, Gaudin P, Maoret JJ, Laburthe M (1996) Mutagenesis of N-glycosylation sites in the human VIP 1 receptor. Ann N Y Acad Sci 805:558-562

Crocker E, Eilers M, Ahuja S, Hornak V, Hirshfeld A, Sheves M, Smith SO (2006) Location of Trp265 in metarhodopsin II: implications for the activation mechanism of the visual receptor rhodopsin. J Mol Biol 357:163-172

Cserzo M, Wallin E, Simon I, von Heijne G, Elofsson A (1997) Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. Protein Eng 10:673-676

Cueto M, Dorta MJ, Munguia O, Llabres M (2003) New approach to stability assessment of protein solution formulations by differential scanning calorimetry. Int J Pharm 252:159-166

Cuthbert AW (2006) A brief history of the British Pharmacological Society. Br J Pharmacol 147 Suppl 1:S2-8

Cuthbertson JM, Doyle DA, Sansom MS (2005) Transmembrane helix prediction: a comparative evaluation and analysis. Protein Eng Des Sel 18:295-308

Daly R, Hearn MT (2005) Expression of heterologous proteins in Pichia pastoris: a useful experimental tool in protein engineering and production. J Mol Recognit 18:119-138

Darby RAJ, Jamshad M, Grgic L, Holmes WJ, Bill RM (2010) Production of Recombinant G protein-coupled Receptor in Yeast for structural and Functional Analysis. Cited from Poyner DR and Wheatley M (2010) G Protein-Coupled Recptors Essential Methods. Wiley and Blackwell Oxford,

de Prado BM, Russo AF (2006) CGRP receptor antagonists: A new frontier of anti-migraine medications. Drug Discov Today Ther Strateg 3:593-597

DePristo MA, de Bakker PI, Lovell SC, Blundell TL (2003) Ab initio construction of polypeptide fragments: efficient generation of accurate, representative ensembles. Proteins 51:41-55

Ding WQ, Cheng ZJ, McElhiney J, Kuntz SM, Miller LJ (2002) Silencing of secretin receptor function by dimerization with a misspliced variant secretin receptor in ductal pancreatic adenocarcinoma. Cancer Res 62:5223-5229

Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, Baker NA (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. Nucleic Acids Res 35:W522-525

Dong M, Cox RF, Miller LJ (2009) Juxtamembranous region of the amino terminus of the family B G protein-coupled calcitonin receptor plays a critical role in small-molecule agonist action. J Biol Chem 284:21839-21847

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, Lam PC, Pinon DI, Sexton PM, Abagyan R, Miller LJ (2008) Spatial approximation between secretin residue five and the third extracellular loop of its receptor provides new insight into the molecular basis of natural agonist binding. Mol Pharmacol 74:413-422

Dong M, Pinon DI, Asmann YW, Miller LJ (2006) Possible endogenous agonist mechanism for the activation of secretin family G protein-coupled receptors. Mol Pharmacol 70:206-213

Donnelly D (1997) The arrangement of the transmembrane helices in the secretin receptor family of G-protein-coupled receptors. FEBS Lett 409:431-436

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 129:420-423

Du P, Alkorta I (1994) Sequence divergence analysis for the prediction of seven-helix membrane protein structures: I. Comparison with bacteriorhodopsin. Protein Eng 7:1221-1229

Edvinsson L (2004) Blockade of CGRP receptors in the intracranial vasculature: a new target in the treatment of headache. Cephalalgia 24:611-622

Evans BN, Rosenblatt MI, Mnayer LO, Oliver KR, Dickerson IM (2000) CGRP-RCP, a novel protein required for signal transduction at calcitonin gene-related peptide and adrenomedullin receptors. J Biol Chem 275:31438-31443

Farrens DL, Altenbach C, Yang K, Hubbell WL, Khorana HG (1996) Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. Science 274:768-770

Fiser A, Do RK, Sali A (2000) Modeling of loops in protein structures. Protein Sci 9:1753 - 1773

Fiser A, Sali A (2003) Modeller: generation and refinement of homology-based protein structure models. Methods Enzymol 374:461-491

Fitzsimmons TJ, Zhao X, Wank SA (2003) The extracellular domain of receptor activity-modifying protein 1 is sufficient for calcitonin receptor-like receptor function. J Biol Chem 278:14313-14320

Fluhmann B, Muff R, Hunziker W, Fischer JA, Born W (1995) A human orphan calcitonin receptor-like structure. Biochem Biophys Res Commun 206:341-347

Foord SM, Wise A, Brown J, Main MJ, Fraser NJ (1999) The N-terminus of RAMPs is a critical determinant of the glycosylation state and ligand binding of calcitonin receptor-like receptor. Biochem Soc Trans 27:535-539

Fraser NJ, Wise A, Brown J, McLatchie LM, Main MJ, Foord SM (1999) The amino terminus of receptor activity modifying proteins is a critical determinant of glycosylation state and ligand binding of calcitonin receptor-like receptor. Mol Pharmacol 55:1054-1059

Fredriksson R, Lagerstrom MC, Lundin LG, Schioth 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

Fujiyoshi Y (1998) The structural study of membrane proteins by electron crystallography. Adv Biophys 35:25-80

Gaddum, JH (1937) The quantitative effects of antagonistic drugs, *J Physiol (Lond)* 89:7–9.

Gao F, Harikumar KG, Dong M, Lam PC, Sexton PM, Christopoulos A, Bordner A, Abagyan R, Miller LJ (2009) Functional importance of a structurally distinct homodimeric complex of the family B G protein-coupled secretin receptor. Mol Pharmacol 76:264-274

Gardella TJ, Luck MD, Jensen GS, Usdin TB, Juppner H (1996) Converting parathyroid hormone-related peptide (PTHrP) into a potent PTH-2 receptor agonist. J Biol Chem 271:19888-19893

Geraud G, Keywood C, Senard JM (2003) Migraine headache recurrence: relationship to clinical, pharmacological, and pharmacokinetic properties of triptans. Headache 43:376-388

Gether U (2000) Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. Endocr Rev 21:90-113

Ginalski K, Rychlewski L (2003) Protein structure prediction of CASP5 comparative modeling and fold recognition targets using consensus alignment approach and 3D assessment. Proteins 53 Suppl 6:410-417

Gkountelias K, Tselios T, Venihaki M, Deraos G, Lazaridis I, Rassouli O, Gravanis A, Liapakis G (2009) Alanine scanning mutagenesis of the second extracellular loop of type 1 corticotropin releasing factor receptor revealed residues critical for peptide binding. Mol Pharmacol

Grabe M, Neu J, Oster G, Nollert P (2003) Protein interactions and membrane geometry. Biophys J 84:854-868

Grace CR, Perrin MH, DiGruccio MR, Miller CL, Rivier JE, Vale WW, 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 101:12836-12841

Grace CR, Perrin MH, Gulyas J, Digruccio MR, Cantle JP, Rivier JE, Vale WW, 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

Gutmann DA, Mizohata E, Newstead S, Ferrandon S, Postis V, Xia X, Henderson PJ, van Veen HW, Byrne B (2007) A high-throughput method for membrane protein solubility screening: the ultracentrifugation dispersity sedimentation assay. Protein Sci 16:1422-1428

Hakala JM, Vihinen M (1994) Modelling the structure of the calcitonin gene-related peptide. Protein Eng 7:1069-1075

Hanson MA, Cherezov V, Griffith MT, Roth CB, Jaakola VP, Chien EY, Velasquez J, Kuhn P, Stevens RC (2008) A specific cholesterol binding site is established by the 2.8 A structure of the human beta2-adrenergic receptor. Structure 16:897-905

Harikumar KG, Happs RM, Miller LJ (2008a) Dimerization in the absence of higher-order oligomerization of the G protein-coupled secretin receptor. Biochim Biophys Acta 1778:2555-2563

Harikumar KG, Morfis MM, Sexton PM, Miller LJ (2008b) Pattern of intra-family hetero-oligomerization involving the G-protein-coupled secretin receptor. J Mol Neurosci 36:279-285

Hashimoto H, Ogawa N, Hagihara N, Yamamoto K, Imanishi K, Nogi H, Nishino A, Fujita T, Matsuda T, Nagata S, Baba A (1997) Vasoactive intestinal

polypeptide and pituitary adenylate cyclase-activating polypeptide receptor chimeras reveal domains that determine specificity of vasoactive intestinal polypeptide binding and activation. Mol Pharmacol 52:128-135

Hauser M, Kauffman S, Lee BK, Naider F, Becker JM (2007) The first extracellular loop of the Saccharomyces cerevisiae G protein-coupled receptor Ste2p undergoes a conformational change upon ligand binding. J Biol Chem 282:10387-10397

Hay DL, Christopoulos G, Christopoulos A, Sexton PM (2006) Determinants of 1-piperidinecarboxamide, N-[2-[[5-amino-l-[[4-(4-pyridinyl)-l-piperazinyl]carbonyl]pentyl]amino]-1- [(3,5-dibromo-4-hydroxyphenyl)methyl]-2-oxoethyl]-4-(1,4-dihydro-2-oxo-3(2 H)-quinazolinyl) (BIBN4096BS) affinity for calcitonin gene-related peptide and amylin receptors--the role of receptor activity modifying protein 1. Mol Pharmacol 70:1984-1991

Hebert TE, Moffett S, Morello JP, Loisel TP, Bichet DG, Barret C, Bouvier M (1996) A peptide derived from a beta2-adrenergic receptor transmembrane domain inhibits both receptor dimerization and activation. J Biol Chem 271:16384-16392

Henke H, Sigrist S, Lang W, Schneider J, Fischer JA (1987) Comparison of binding sites for the calcitonin gene-related peptides I and II in man. Brain Res 410:404-408

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

Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J. Chem. Theory Comput. 4:435-447

Hilairet S, Belanger C, Bertrand J, Laperriere A, Foord SM, Bouvier M (2001a) Agonist-promoted internalization of a ternary complex between calcitonin receptor-like receptor, receptor activity-modifying protein 1 (RAMP1), and beta-arrestin. J Biol Chem 276:42182-42190

Hilairet S, Foord SM, Marshall FH, 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 276:29575-29581

Hill AV (1909) The mode of action of nicotine and curari, determined by the form of the contraction curve and the method of temperature coefficients. J Physiol 39:361-373

Hill SJ (2006) G-protein-coupled receptors: past, present and future. Br J Pharmacol 147 Suppl 1:S27-37

Hirokawa T, Boon-Chieng S, Mitaku S (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics 14:378-379

Hofmann K and Stoffel W (1993) TMbase - A database of membrane spanning proteins segments Biol. Chem. Hoppe-Seyler 374,166

Holst B, Zoffmann S, Elling CE, Hjorth SA, Schwartz TW (1998) Steric hindrance mutagenesis versus alanine scan in mapping of ligand binding sites in the tachykinin NK1 receptor. Mol Pharmacol 53:166-175

Holtmann MH, Ganguli S, Hadac EM, Dolu V, Miller LJ (1996a) Multiple extracellular loop domains contribute critical determinants for agonist binding and activation of the secretin receptor. J Biol Chem 271:14944-14949

Holtmann MH, Hadac EM, Ulrich CD, Miller LJ (1996b) Molecular basis and species specificity of high affinity binding of vasoactive intestinal polypeptide by the rat secretin receptor. J Pharmacol Exp Ther 279:555-560

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 138:325-332

Hubbard JA, Martin SR, Chaplin LC, Bose C, Kelly SM, Price NC (1991) Solution structures of calcitonin-gene-related-peptide analogues of calcitonin-gene-related peptide and amylin. Biochem J 275 (Pt 3):785-788

Hubbell WL, Altenbach C, Hubbell CM, Khorana HG (2003) Rhodopsin structure, dynamics, and activation: a perspective from crystallography, site-directed spin labeling, sulfhydryl reactivity, and disulfide cross-linking. Adv Protein Chem 63:243-290

Hulme EC, Lu ZL, Ward SD, Allman K, Curtis CA (1999) The conformational switch in 7-transmembrane receptors: the muscarinic receptor paradigm. Eur J Pharmacol 375:247-260

Ittner LM, Koller D, Muff R, Fischer JA, Born W (2005) The N-terminal extracellular domain 23-60 of the calcitonin receptor-like receptor in chimeras with the parathyroid hormone receptor mediates association with receptor activity-modifying protein 1. Biochemistry 44:5749-5754

Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, Ijzerman AP, Stevens RC (2008) The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. Science 322:1211-1217

Jackson RM, Gabb HA, Sternberg MJ (1998) Rapid refinement of protein interfaces incorporating solvation: application to the docking problem. J Mol Biol 276:265-285

Jacobson MP, Pincus DL, Rapp CS, Day TJ, Honig B, Shaw DE, Friesner RA (2004) A hierarchical approach to all-atom protein loop prediction. Proteins 55:351-367

Jacoby E, Bouhelal R, Gerspacher M, Seuwen K (2006) The 7 TM G-protein-coupled receptor target family. ChemMedChem 1:761-782

Johnson MS, Overington JP (1993) A structural basis for sequence comparisons. An evaluation of scoring methodologies. J Mol Biol 233:716-738

Jojart B, Kiss R, Viskolcz B, Keseru GM (2008) Activation mechanism of the human histamine H4 receptor--an explicit membrane molecular dynamics simulation study. J Chem Inf Model 48:1199-1210

Jones DT (2007) Improving the accuracy of transmembrane protein topology prediction using evolutionary information. Bioinformatics 23:538-544

Juhasz G, Zsombok T, Modos EA, Olajos S, Jakab B, Nemeth J, Szolcsanyi J, Vitrai J, Bagdy G (2003) NO-induced migraine attack: strong increase in plasma calcitonin gene-related peptide (CGRP) concentration and negative correlation with platelet serotonin release. Pain 106:461-470

Juretic D, Zoranic L, Zucic D (2002) Basic charge clusters and predictions of membrane protein topology. J Chem Inf Comput Sci 42:620-632

Kandt C, Ash WL, Tieleman DP (2007) Setting up and running molecular dynamics simulations of membrane proteins. Methods 41:475-488

Kapas S, Clark AJ (1995) Identification of an orphan receptor gene as a type 1 calcitonin gene-related peptide receptor. Biochem Biophys Res Commun 217:832-838

Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA (1992) Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. Proc Natl Acad Sci U S A 89:2195-2199

Kenakin T (2007) Collateral efficacy in drug discovery: taking advantage of the good (allosteric) nature of 7TM receptors. Trends Pharmacol Sci 28:407-415

Kennedy SP, Sun D, Oleynek JJ, Hoth CF, Kong J, Hill RJ (1998) Expression of the rat adrenomedullin receptor or a putative human adrenomedullin receptor does not correlate with adrenomedullin binding or functional response. Biochem Biophys Res Commun 244:832-837

Klein P, Kanehisa M and DeLisi C (1985) The detection and classification of membrane-spanning proteins, Biochim. Biophys. Acta 815 ,468–476

Kobilka BK (2007) G protein coupled receptor structure and activation. Biochim Biophys Acta 1768:794-807

Kolakowski LF, Jr. (1994) GCRDb: a G-protein-coupled receptor database. Receptors Channels 2:1-7

Koller D, Born W, Leuthauser K, Fluhmann B, McKinney RA, Fischer JA, 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 Lett 531:464-468

Kors CA, Wallace E, Davies DR, Li L, Laible PD, Nollert P (2009) Effects of impurities on membrane-protein crystallization in different systems. Acta Crystallogr D Biol Crystallogr 65:1062-1073

Koth CM, Abdul-Manan N, Lepre CA, Connolly PJ, Yoo S, Mohanty AK, Lippke JA, Zwahlen J, Coll JT, Doran JD, Garcia-Guzman M, Moore JM (2010) Refolding and Characterization of a Soluble Ectodomain Complex of the Calcitonin Gene-Related Peptide Receptor. Biochemistry

Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567-580

Kusano S, Kukimoto-Niino M, Akasaka R, Toyama M, Terada T, Shirouzu M, Shindo T, Yokoyama S (2008) Crystal Structure of the Human Receptor Activity-Modifying Protein 1 Extracellular Domain. Protein Sci

Kuwasako K, Cao YN, Chu CP, Iwatsubo S, Eto T, Kitamura K (2006) Functions of the cytoplasmic tails of the human receptor activity-modifying protein components of calcitonin gene-related peptide and adrenomedullin receptors. J Biol Chem 281:7205-7213

Kuwasako K, Kitamura K, Nagoshi Y, Cao YN, Eto T (2003) Identification of the human receptor activity-modifying protein 1 domains responsible for agonist binding specificity. J Biol Chem 278:22623-22630

Laburthe M, Couvineau A, Tan V (2007) Class II G protein-coupled receptors for VIP and PACAP: structure, models of activation and pharmacology. Peptides 28:1631-1639

Langmuir, I (1918) The adsorption of gases on plane surfaces of glass, mica and platinum, *J. Am. Chem. Soc.* 40:1362–1403.

Laskowski RA (2001) PDBsum: summaries and analyses of PDB structures. Nucleic Acids Res 29:221 - 222

Laskowski RA, Macarthur MW, Moss DS, Thornton JM (1993) {PROCHECK}: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. 26:283-291

Lassen LH, Haderslev PA, Jacobsen VB, Iversen HK, Sperling B, Olesen J (2002) CGRP may play a causative role in migraine. Cephalalgia 22:54-61

Lawson Z, Wheatley M (2004) The third extracellular loop of G-protein-coupled receptors: more than just a linker between two important transmembrane helices. Biochem Soc Trans 32:1048-1050

Leach K, Valent C, Sexton PM, Christopoulos A Measurement of Ligand-G protein coupled receptor interactions. Found in Poyner DR and Wheatley M (2010) G Protein-Coupled Receptors Essential methods. Wiley-Blackwell, A John Wiley and Sons, Ltd. Publication.

Leach K, Sexton PM, Christopoulos A (2007) Allosteric GPCR modulators: taking advantage of permissive receptor pharmacology. Trends Pharmacol Sci 28:382-389

Lee C, Luck MD, Juppner H, Potts JT, Jr., Kronenberg HM, Gardella TJ (1995) Homolog-scanning mutagenesis of the parathyroid hormone (PTH) receptor reveals PTH-(1-34) binding determinants in the third extracellular loop. Mol Endocrinol 9:1269-1278

Lett D, Hsing M, Pio F (2004) Interaction profile-based protein classification of death domain. BMC Bioinformatics 5:75

Leuthauser K, Gujer R, Aldecoa A, McKinney RA, Muff R, Fischer JA, Born W (2000) Receptor-activity-modifying protein 1 forms heterodimers with two G-protein-coupled receptors to define ligand recognition. Biochem J 351 Pt 2:347-351

Lisenbee CS, Dong M, Miller LJ (2005) Paired cysteine mutagenesis to establish the pattern of disulfide bonds in the functional intact secretin receptor. J Biol Chem 280:12330-12338

Li H, Robertson AD, and Jensen JH (2005) Very fast empirical prediction and interpretation of protein pKa values. Proteins, 61, 704-721.

Liu B, Li S, Hu J (2004) Technological advances in high-throughput screening. Am J Pharmacogenomics 4:263-276

Liu Y, Engelman DM, Gerstein M (2002) Genomic analysis of membrane protein families: abundance and conserved motifs. Genome Biol 3:research0054

Lodowski DT, Angel TE, Palczewski K (2009) Comparative analysis of GPCR crystal structures. Photochem Photobiol 85:425-430

Luca S, White JF, Sohal AK, Filippov DV, van Boom JH, Grisshammer R, Baldus M (2003) The conformation of neurotensin bound to its G protein-coupled receptor. Proc Natl Acad Sci U S A 100:10706-10711

Luck MD, Carter PH, Gardella TJ (1999) The (1-14) fragment of parathyroid hormone (PTH) activates intact and amino-terminally truncated PTH-1 receptors. Mol Endocrinol 13:670-680

Lynch B, Kaiser ET (1988) Biological properties of two models of calcitonin gene related peptide with idealized amphiphilic alpha-helices of different lengths. Biochemistry 27:7600 - 7607

MacKenzie KR, Prestegard JH, Engelman DM (1997) A transmembrane helix dimer: structure and implications. Science 276:131-133

Maggi CA, Rovero P, Giuliani S, Evangelista S, Regoli D, Meli A (1990) Biological activity of N-terminal fragments of calcitonin gene-related peptide. Eur J Pharmacol 179:217 - 219

Mallee JJ, Salvatore CA, LeBourdelles B, Oliver KR, Longmore J, Koblan KS, Kane SA (2002) Receptor activity-modifying protein 1 determines the species selectivity of non-peptide CGRP receptor antagonists. J Biol Chem 277:14294-14298

Manning MC (1989) Conformation of the alpha form of human calcitonin gene-related peptide (CGRP) in aqueous solution as determined by circular dichroism spectroscopy. Biochem Biophys Res Commun 160:388-392

Martin VT, Goldstein JA. (2005) Evaluating the safety and tolerability profile of acute treatments for migraine. *Am J Med.* 118 (Suppl 1):S36–44.

McGuffin LJ, Bryson K, Jones DT (2000) The PSIPRED protein structure prediction server. Bioinformatics 16:404-405

McLatchie LM, Fraser NJ, Main MJ, Wise A, Brown J, Thompson N, Solari R, Lee MG, Foord SM (1998) RAMPs regulate the transport and ligand specificity of the calcitonin-receptor-like receptor. Nature 393:333-339

Michalsky E, Goede A, Preissner R (2003) Loops In Proteins (LIP)--a comprehensive loop database for homology modelling. Protein Eng 16:979-985

Miller PS, Barwell J, Poyner DR, Wigglesworth MJ, Garland SL, Donnelly (2010) D Non-peptidic antagonists of the CGRP receptor, BIBN4096BS and MK-0974, interact with wthe calcitonin receptor-like receptor via methionine-42 and RAMP1 via tryptophan-74. Biochem Biophys Res Commun 391:437-442

Miret JJ, Rakhilina L, Silverman L, Oehlen B (2002) Functional expression of heteromeric calcitonin gene-related peptide and adrenomedullin receptors in yeast. J Biol Chem 277:6881-6887

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

Moont G, Gabb HA, Sternberg MJ (1999) Use of pair potentials across protein interfaces in screening predicted docked complexes. Proteins 35:364-373

Moult J, James MN (1986) An algorithm for determining the conformation of polypeptide segments in proteins by systematic search. Proteins 1:146-163

Muff R, Born W, Fischer JA (2001) Adrenomedullin and related peptides: receptors and accessory proteins. Peptides 22:1765-1772

Muff R, Buhlmann N, Fischer JA, Born W (1999) An amylin receptor is revealed following co-transfection of a calcitonin receptor with receptor activity modifying proteins-1 or -3. Endocrinology 140:2924-2927

Mulderry PK, Ghatei MA, Spokes RA, Jones PM, Pierson AM, Hamid QA, Kanse S, Amara SG, Burrin JM, Legon S, et al. (1988) Differential expression of alpha-CGRP and beta-CGRP by primary sensory neurons and enteric autonomic neurons of the rat. Neuroscience 25:195-205

Murakami M, Kouyama T (2008) Crystal structure of squid rhodopsin. Nature 453:363-367

Nakai K, Kanehisa M (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics 14:897-911

Neubig RR, Spedding M, Kenakin T, Christopoulos A (2003) International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. XXXVIII. Update on terms and symbols in quantitative pharmacology. Pharmacol Rev 55:597-606

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

Newstead S, Ferrandon S, Iwata S (2008) Rationalizing alpha-helical membrane protein crystallization. Protein Sci 17:466-472

Nieba L, Nieba-Axmann SE, Persson A, Hamalainen M, Edebratt F, Hansson A, Lidholm J, Magnusson K, Karlsson AF, Pluckthun A (1997) BIACORE analysis of histidine-tagged proteins using a chelating NTA sensor chip. Anal Biochem 252:217-228

Njuki F, Nicholl CG, Howard A, Mak JC, Barnes PJ, Girgis SI, Legon S (1993) A new calcitonin-receptor-like sequence in rat pulmonary blood vessels. Clin Sci (Lond) 85:385-388

Nyblom M, Oberg F, Lindkvist-Petersson K, Hallgren K, Findlay H, Wikstrom J, Karlsson A, Hansson O, Booth PJ, Bill RM, Neutze R, Hedfalk K (2007)

Exceptional overproduction of a functional human membrane protein. Protein Expr Purif 56:110-120

Okada T, Fujiyoshi Y, Silow M, Navarro J, Landau EM, Shichida Y (2002) Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. Proc Natl Acad Sci U S A 99:5982-5987

Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V (2004) The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure. J Mol Biol 342:571-583

Olde B, Sabirsh A, Owman C (1998) Molecular mapping of epitopes involved in ligand activation of the human receptor for the neuropeptide, VIP, based on hybrids with the human secretin receptor. J Mol Neurosci 11:127-134

Ostermeier C, Michel H (1997) Crystallization of membrane proteins. Curr Opin Struct Biol 7:697-701

Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Le Trong I, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M (2000) Crystal structure of rhodopsin: A G protein-coupled receptor. Science 289:739-745

Parthier C, Kleinschmidt M, Neumann P, Rudolph R, Manhart S, Schlenzig D, Fanghanel J, Rahfeld JU, Demuth HU, Stubbs MT (2007) Crystal structure of the incretin-bound extracellular domain of a G protein-coupled receptor. Proc Natl Acad Sci U S A 104:13942-13947

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

Paton WD, Rang HP (1965) The uptake of atropine and related drugs by intestinal smooth muscle of guinea-pig in relation to acetylcholine receptors. Proc R Soc Lond B Biol Sci 163:1-44

Persson B, Argos P (1997) Prediction of membrane protein topology utilizing multiple sequence alignments. J Protein Chem 16:453-457

Pham V, Dong M, Wade JD, Miller LJ, Morton CJ, Ng HL, Parker MW, Sexton PM (2005) Insights into interactions between the alpha-helical region of the salmon calcitonin antagonists and the human calcitonin receptor using photoaffinity labeling. J Biol Chem 280:28610-28622

Pham V and Sexton PM (2003) Review: Photoaffinity Scanning in mapping of the the peptide interface of Class II G protein-coupled receptors. J Peptide Sci 10:179-203

Pham V, Wade JD, Purdue BW, Sexton PM (2004) Spatial proximity between a photolabile residue in position 19 of salmon calcitonin and the amino terminus of the human calcitonin receptor. J Biol Chem 279:6720-6729

Pioszak AA, Parker NR, Suino-Powell K, Xu HE (2008) Molecular recognition of corticotropin-releasing factor by its G-protein-coupled receptor CRFR1. J Biol Chem 283:32900-32912

Pioszak AA, Xu HE (2008) Molecular recognition of parathyroid hormone by its G protein-coupled receptor. Proc Natl Acad Sci U S A 105:5034-5039

Plati J, Tsomaia N, Piserchio A, Mierke DF (2007) Structural features of parathyroid hormone receptor coupled to Galpha(s)-protein. Biophys J 92:535-540

Poirot O, O'Toole E, Notredame C (2003) Tcoffee@igs: A web server for computing, evaluating and combining multiple sequence alignments. Nucleic Acids Res 31:3503-3506

Poyner DR, Sexton PM, Marshall I, Smith DM, Quirion R, Born W, Muff R, Fischer JA, Foord SM (2002) International Union of Pharmacology. XXXII. The mammalian calcitonin gene-related peptides, adrenomedullin, amylin, and calcitonin receptors. Pharmacol Rev 54:233-246

Poyner DR, Soomets U, Howitt SG, Langel U (1998) Structural determinants for binding to CGRP receptors expressed by human SK-N-MC and Col 29 cells: studies with chimeric and other peptides. Br J Pharmacol 124:1659-1666

Presley JF (2005) Imaging the secretory pathway: the past and future impact of live cell optical techniques. Biochim Biophys Acta 1744:259-272

Pusey ML, Liu ZJ, Tempel W, Praissman J, Lin D, Wang BC, Gavira JA, Ng JD (2005) Life in the fast lane for protein crystallization and X-ray crystallography. Prog Biophys Mol Biol 88:359-386

Qi T, Christopoulos G, Bailey RJ, Christopoulos A, Sexton PM, Hay DL (2008) Identification of N-terminal receptor activity-modifying protein residues important for calcitonin gene-related peptide, adrenomedullin, and amylin receptor function. Mol Pharmacol 74:1059-1071

Rang HP (2006) The receptor concept: pharmacology's big idea. Br J Pharmacol 147 Suppl 1:S9-16

Rasmussen SG, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VR, Sanishvili R, Fischetti RF, Schertler GF, Weis WI, Kobilka BK (2007) Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. Nature 450:383-387

Reggio PH (2006) Computational methods in drug design: modeling G protein-coupled receptor monomers, dimers, and oligomers. AAPS J 8:E322-336

Rist B, Lacroix JS, Entzeroth M, Doods HN, Beck-Sickinger AG (1999) CGRP 27-37 analogues with high affinity to the CGRP1 receptor show antagonistic properties in a rat blood flow assay. Regul Pept 79:153-158

Robinson SD, Aitken JF, Bailey RJ, Poyner DR, Hay DL (2009) Novel peptide antagonists of adrenomedullin and calcitonin gene-related peptide receptors: identification, pharmacological characterization, and interactions with position

74 in receptor activity-modifying protein 1/3. J Pharmacol Exp Ther 331:513-521

Roh J, Chang CL, Bhalla A, Klein C, Hsu SY (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 279:7264-7274

Rohl CA, Strauss CE, Chivian D, Baker D (2004) Modeling structurally variable regions in homologous proteins with rosetta. Proteins 55:656-677

Rosenbaum DM, Rasmussen SG, Kobilka BK (2009) The structure and function of G-protein-coupled receptors. Nature 459:356-363

Rost B (1996) PHD: predicting one-dimensional protein structure by profile-based neural networks. Methods Enzymol 266:525-539

Rost B, Fariselli P, Casadio R (1996) Topology prediction for helical transmembrane proteins at 86% accuracy. Protein Sci 5:1704-1718

Rovero P, Giuliani S, Maggi CA (1992) CGRP antagonist activity of short C-terminal fragments of human alpha CGRP, CGRP(23-37) and CGRP(19-37). Peptides 13:1025 - 1027

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

Runge S, Schimmer S, Oschmann J, Schiodt CB, Knudsen SM, Jeppesen CB, Madsen K, Lau J, Thogersen H, Rudolph R (2007) Differential structural properties of GLP-1 and exendin-4 determine their relative affinity for the GLP-1 receptor N-terminal extracellular domain. Biochemistry 46:5830-5840

Runge S, Thogersen H, Madsen K, Lau J, Rudolph R (2008) Crystal structure of the ligand-bound glucagon-like peptide-1 receptor extracellular domain. J Biol Chem 283:11340-11347

Russ WP, Engelman DM (1999) TOXCAT: a measure of transmembrane helix association in a biological membrane. Proc Natl Acad Sci U S A 96:863-868

Sakmar TP, Franke RR, Khorana HG (1989) Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin. Proc Natl Acad Sci U S A 86:8309-8313

Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779-815

Salvatore CA, Hershey JC, Corcoran HA, Fay JF, Johnston VK, Moore EL, Mosser SD, Burgey CS, Paone DV, Shaw AW, Graham SL, Vacca JP, Williams TM, Koblan KS, Kane SA (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 324:416-421

Sambrook J and Russell DW (2001) Molecular Cloning: A Laboratory manual (3rd edition). Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York

Samudrala R, Moult J (1998) An all-atom distance-dependent conditional probability discriminatory function for protein structure prediction. J Mol Biol 275:895-916

Scheerer P, Park JH, Hildebrand PW, Kim YJ, Krauss N, Choe HW, Hofmann KP, Ernst OP (2008) Crystal structure of opsin in its G-protein-interacting conformation. Nature 455:497-502

Schild HO (1997) pA, a new scale for the measurement of drug antagonism. 1947. Br J Pharmacol 120:29-46; discussion 27-28

Schioth HB, Fredriksson R (2005) The GRAFS classification system of G-protein coupled receptors in comparative perspective. Gen Comp Endocrinol 142:94-101

Schwartz TW, Frimurer TM, Holst B, Rosenkilde MM, Elling CE (2006) Molecular mechanism of 7TM receptor activation--a global toggle switch model. Annu Rev Pharmacol Toxicol 46:481-519

Schwartz TW, Rosenkilde MM (1996) Is there a 'lock' for all agonist 'keys' in 7TM receptors? Trends Pharmacol Sci 17:213-216

Seck T, Baron R, Horne WC (2003) Binding of filamin to the C-terminal tail of the calcitonin receptor controls recycling. J Biol Chem 278:10408-10416

Sexton PM, Poyner DR, Simms J, Christopoulos A, Hay DL (2009) Modulating receptor function through RAMPs: can they represent drug targets in themselves? Drug Discov Today

Shacham S, Marantz Y, Bar-Haim S, Kalid O, Warshaviak D, Avisar N, Inbal B, Heifetz A, Fichman M, Topf M, Naor Z, Noiman S, Becker OM (2004) PREDICT modeling and in-silico screening for G-protein coupled receptors. Proteins 57:51-86

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 274:17033-17041

Shen MY, Sali A (2006) Statistical potential for assessment and prediction of protein structures. Protein Sci 15:2507-2524

Shi L, Javitch JA (2004) The second extracellular loop of the dopamine D2 receptor lines the binding-site crevice. Proc Natl Acad Sci U S A 101:440-445

Shimamura T, Hiraki K, Takahashi N, Hori T, Ago H, Masuda K, Takio K, Ishiguro M, Miyano M (2008) Crystal structure of squid rhodopsin with intracellularly extended cytoplasmic region. J Biol Chem 283:17753-17756

Sine SM, Wang HL, Bren N (2002) Lysine scanning mutagenesis delineates structural model of the nicotinic receptor ligand binding domain. J Biol Chem 277:29210-29223

Soto CS, Fasnacht M, Zhu J, Forrest L, Honig B (2008) Loop modeling: Sampling, filtering, and scoring. Proteins 70:834-843

Steenbergh PH, Hoppener JW, Zandberg J, Lips CJ, Jansz HS (1985) A second human calcitonin/CGRP gene. FEBS Lett 183:403-407

Steiner S, Muff R, Gujer R, Fischer JA, 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 41:11398-11404

Stephenson RP (1956) A modification of receptor theory. Br J Pharmacol Chemother 11:379-393

Sun C, Song D, Davis-Taber RA, Barrett LW, Scott VE, Richardson PL, Pereda-Lopez A, Uchic ME, Solomon LR, Lake MR, Walter KA, Hajduk PJ, Olejniczak ET (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 104:7875-7880

Tanford C (1978) The hydrophobic effect and the organization of living matter. Science 200:1012-1018

Taylor WR, Munro RE, Petersen K, Bywater RP (2003) Ab initio modelling of the N-terminal domain of the secretin receptors. Comput Biol Chem 27:103-114

Thomas BE, Woznica I, Mierke DF, Wittelsberger A, Rosenblatt M (2008) Conformational changes in the parathyroid hormone receptor associated with activation by agonist. Mol Endocrinol 22:1154-1162

Tolun AA, Dickerson IM, Malhotra A (2007) Overexpression and purification of human calcitonin gene-related peptide-receptor component protein in Escherichia coli. Protein Expr Purif 52:167-174

Topiol S, Sabio M (2009) X-ray structure breakthroughs in the GPCR transmembrane region. Biochem Pharmacol 78:11-20

Torres J, Stevens TJ, Samso M (2003) Membrane proteins: the 'Wild West' of structural biology. Trends Biochem Sci 28:137-144

Tusnady GE, Simon I (1998) Principles governing amino acid composition of integral membrane proteins: application to topology prediction. J Mol Biol 283:489-506

Tusnady GE, Simon I (2001) The HMMTOP transmembrane topology prediction server. Bioinformatics 17:849-850

Tycko R (2001) Biomolecular solid state NMR: advances in structural methodology and applications to peptide and protein fibrils. Annu Rev Phys Chem 52:575-606

Udawela M, Christopoulos G, Tilakaratne N, Christopoulos A, Albiston A, Sexton PM (2006) Distinct receptor activity-modifying protein domains differentially modulate interaction with calcitonin receptors. Mol Pharmacol 69:1984-1989

Vakser IA (1995) Protein docking for low-resolution structures. Protein Eng 8:371-377

Vakser IA (1996) Low-resolution docking: prediction of complexes for underdetermined structures. Biopolymers 39:455-464

Vakser IA (1997) Evaluation of GRAMM low-resolution docking methodology on the hemagglutinin-antibody complex. Proteins Suppl 1:226-230

Vakser IA, Matar OG, Lam CF (1999) A systematic study of low-resolution recognition in protein--protein complexes. Proc Natl Acad Sci U S A 96:8477-8482

van Rossum D, Hanisch UK, Quirion R (1997) Neuroanatomical localization, pharmacological characterization and functions of CGRP, related peptides and their receptors. Neurosci Biobehav Rev 21:649-678

Vilardaga JP, Di Paolo E, Bialek C, De Neef P, Waelbroeck M, Bollen A, Robberecht P (1997) Mutational analysis of extracellular cysteine residues of rat secretin receptor shows that disulfide bridges are essential for receptor function. Eur J Biochem 246:173-180

Vincze T, Posfai J, Roberts RJ (2003) NEBcutter: A program to cleave DNA with restriction enzymes. Nucleic Acids Res 31:3688-3691

Vohra S, Poyner D, Millineaux PM, Reeves PJ, Upton G, Reynolds CA (in preparation) Aplant GPCR unlocks the class A -class B alignment: application to CGRP

Vohra S, Chintapalli SV, Illingworth CJ, Reeves PJ, Mullineaux PM, Clark HS, Dean MK, Upton GJ, Reynolds CA (2007) Computational studies of Family A and Family B GPCRs. Biochem Soc Trans 35:749-754

Wallin E, von Heijne G (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. Protein Sci 7:1029-1038

Wand AJ, Ehrhardt MR, Flynn PF (1998) High-resolution NMR of encapsulated proteins dissolved in low-viscosity fluids. Proc Natl Acad Sci U S A 95:15299-15302

Warne T, Serrano-Vega MJ, Baker JG, Moukhametzianov R, Edwards PC, Henderson R, Leslie AG, Tate CG, Schertler GF (2008) Structure of a beta1-adrenergic G-protein-coupled receptor. Nature 454:486-491

White SH (2009) Biophysical dissection of membrane proteins. Nature 459:344-346

Wimalawansa SJ (1996) Calcitonin gene-related peptide and its receptors: Molecular genetics, physiology, pathophysiology, and therapeutic potentials. Endocr Rev 17:533 - 585

Wisskirchen FM, Doyle PM, Gough SL, Harris CJ, Marshall I (2000) Bioactive beta-bend structures for the antagonist halpha CGRP(8 - 37) at the CGRP1 receptor of the rat pulmonary artery. Br J Pharmacol 129:1049 - 1055

Xiang Z, Soto CS, Honig B (2002) Evaluating conformational free energies: the colony energy and its application to the problem of loop prediction. Proc Natl Acad Sci U S A 99:7432-7437

Yang Y, Zhou Y (2008) Ab initio folding of terminal segments with secondary structures reveals the fine difference between two closely related all-atom statistical energy functions. Protein Sci 17:1212-1219

Yeagle PL, Albert AD (2007) G-protein coupled receptor structure. Biochim Biophys Acta 1768:808-824

Yohannan S, Yang D, Faham S, Boulting G, Whitelegge J, Bowie JU (2004) Proline substitutions are not easily accommodated in a membrane protein. J Mol Biol 341:1-6

Zhang C, Liu S, Zhou Y (2004) Accurate and efficient loop selections by the DFIRE-based all-atom statistical potential. Protein Sci 13:391-399

Zhou H, Zhou Y (2002) Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. Protein Sci 11:2714-2726

Zhu K, Pincus DL, Zhao S, Friesner RA (2006) Long loop prediction using the protein local optimization program. Proteins 65:438-452

Appendix

Sequencing primers for HA CLR pcDNA3.1 (-)

T7 primer

Forward: 5' TAATACGACTCACTATAGGGAACCC 3'

TM2 primer

Forward: 5' ATCTGTTCTTCTCATTTGTTTGTAACT 3'

TM4 primer

Reverse: 5' CCTTCAGGTCGCCATGGAATCAGCAC 3'

BGH primer

Reverse: 5' TAGAAGGCACAGTCGAGGCTG 3'

Oligonucleotide primers for generation of mutants

E23A

Forward: 5' CGAGGGATCCGCAGCCTTAGAAGAGAGTC 3'
Reverse: 5' GACTCTCTTCTAAGGCTGCGGATCCCTCG 3'

L24A

Forward: 5' GAGGGATCCGCAGAAGCCGAAGAGAGTCCTGAG 3'
Reverse: 5' CTCAGGACTCTCTTCGGCTTCTGCGGATCCCTC 3'

E25A

Forward: 5' GATCCGCAGAATTAGCCGAGAGTCCTGAGGAC 3'
Reverse: 5' GTCCTCAGGACTCTCGGCTAATTCTGCGGATC 3'

E26A

Forward: 5' CCGCAGAATTAGAAGCCAGTCCTGAGGACTC 3'
Reverse: 5' GAGTCCTCAGGACTGGCTTCTAATTCTGCGG 3'

S27A

Forward: 5' CCGCAGAATTAGAAGAGGCCCCTGAGGACTCAATTC 3'
Reverse: 5' GAATTGAGTCCTCAGGGGCCTCTTCTAATTCTGCGG 3'

P28A

Forward: 5' GCAGAATTAGAAGAGAGTGCCGAGGACTCAATTCAGTTG 3'
Reverse: 5' CAACTGAATTGAGTCCTCGGCACTCTCTTCTAATTCTGC 3'

E29A

Forward: 5' GAATTAGAAGAGAGTCCTGCCGACTCAATTCAGTTGGGAG 3'
Reverse: 5' CTCCCAACTGAATTGAGTCGGCAGGACTCTCTTCTAATTC 3'

D30A

Forward: 5' GAGAGTCCTGAGGCCTCAATTCAGTTGG 3'
Reverse: 5' CCAACTGAATTGAGGCCTCAGGACTCTC 3'

S31A

Forward: 5' GAAGAGAGTCCTGAGGACGCCATTCAGTTGGGAGTTAC 3'
Reverse: 5' GTAACTCCCAACTGAATGGCGTCCTCAGGACTCTCTTC 3'

I32A

Forward: 5' GAGTCCTGAGGACTCAGCCCAGTTGGGAGTTACTAG 3'
Reverse: 5' CTAGTAACTCCCAACTGGGCTGAGTCCTCAGGACTC 3'

Q33A

Forward: 5' GAAGAGAGTCCTGAGGACTCAATTGCCTTGGGAGTTACTAGAAATAAAATC 3'
Reverse: 5' GATTTTATTTCTAGTAACTCCCAAGGCAATTGAGTCCTCAGGACTCTCTTC 3'

L34A

Forward: 5' GTCCTGAGGACTCAATTCAGGCCGGAGTTACTAGAAATAAAATC 3'
Reverse: 5' GATTTTATTTCTAGTAACTCCGGCCTGAATTGAGTCCTCAGGAC 3'

G35A

Forward: 5' CTGAGGACTCAATTCAGTTGGCCGTTACTAGAAATAAAATCATG 3'
Reverse: 5' CATGATTTTATTTCTAGTAACGGCCAACTGAATTGAGTCCTCAG 3'

V36A

Forward: 5' GACTCAATTCAGTTGGGAGCCACTAGAAATAAAATCATGAC 3'
Reverse: 5' GTCATGATTTTATTTCTAGTGGCTCCCAACTGAATTGAGTC 3'

T37A

Forward: 5' GACTCAATTCAGTTGGGAGTTGCCAGAAATAAAATCATGACAGC 3'
Reverse: 5' GCTGTCATGATTTTATTTCTGGCAACTCCCAACTGAATTGAGTC 3'

R38A

Forward: 5' GACTCAATTCAGTTGGGAGTTACTGCCAATAAAATCATGACAGCTCAATATG 3'
Reverse: 5' CATATTGAGCTGTCATGATTTTATTGGCAGTAACTCCCAACTGAATTGAGTC 3'

N39A

Forward: 5' CAATTCAGTTGGGAGTTACTAGAGCCAAAATCATGACAGCTCAATATG 3'
Reverse: 5' CATATTGAGCTGTCATGATTTTGGCTCTAGTAACTCCCAACTGAATTG 3'

K40A

Forward:
5 'CAATTCAGTTGGGAGTTACTAGAAATGCCATCATGACAGCTCAATATGAATGTTAC 3'
Reverse:
5' GTAACATTCATATTGAGCTGTCATGATGGCATTTCTAGTAACTCCCAACTGAATTG 3'

I41A

Forward: 5' GTTGGGAGTTACTAGAAATAAAGCCATGACAGCTCAATATGAATG 3'
Reverse: 5' CATTCATATTGAGCTGTCATGGCTTTATTTCTAGTAACTCCCAAC 3'

M42A

Forward: 5' GGGAGTTACTAGAAATAAAATCGCCACAGCTCAATATGAATGTTACC 3'
Reverse: 5' GGTAACATTCATATTGAGCTGTGGCGATTTTATTTCTAGTAACTCCC 3'

T43A

Forward: 5' GAGTTACTAGAAATAAAATCATGGCCGCTCAATATGAATGTTACCAAAAG 3'
Reverse: 5' CTTTTGGTAACATTCATATTGAGCGGCCATGATTTTATTTCTAGTAACTC 3'

A44L

Forward:
5' GAGTTACTAGAAATAAAATCATGACACTGCAATATGAATGTTACCAAAAGATTATG 3'
Reverse:
5' CATAATCTTTTGGTAACATTCATATTGCAGTGTCATGATTTTATTTCTAGTAACTC 3'

Q45A

Forward: 5' CTAGAAATAAAATCATGACAGCTGCCTATGAATGTTACCAAAAGATTATG 3'
Reverse: 5' CATAATCTTTTGGTAACATTCATAGGCAGCTGTCATGATTTTATTTCTAG 3'

Y46A

Forward: 5' GAAATAAAATCATGACAGCTCAAGCCGAATGTTACCAAAAGATTATGC 3'
Reverse: 5' GCATAATCTTTTGGTAACATTCGGCTTGAGCTGTCATGATTTTATTTC 3'

E47A

Forward:

5' GAAATAAAATCATGACAGCTCAATATGCCTGTTACCAAAAGATTATGCAAGACCC 3'

Reverse:

5' GGGTCTTGCATAATCTTTTGGTAACAGGCATATTGAGCTGTCATGATTTTATTTC 3'

C48A

Forward:

5'

GAAATAAAATCATGACAGCTCAATATGAAGCCTACCAAAAGATTATGCAAGACCCCATTC

3'

Reverse:

5' GAATGGGGTCTTGCATAATCTTTTGGTAGGCTTCATATTGAGCTGTCATGATTTTATTTC

3'

Y49A

Forward: 5' GACAGCTCAATATGAATGTGCCCAAAAGATTATGCAAGAC 3'

Reverse: 5' GTCTTGCATAATCTTTTGGGCACATTCATATTGAGCTGTC 3'

Q50A

Forward: 5' GACAGCTCAATATGAATGTTACGCCAAGATTATGCAAGACCCCATTC 3'

Reverse: 5' GAATGGGGTCTTGCATAATCTTGGCGTAACATTCATATTGAGCTGTC 3'

K51A

Forward: 5' GCTCAATATGAATGTTACCAAGCCATTATGCAAGACCCCATTCAAC 3'

Reverse: 5' GTTGAATGGGGTCTTGCATAATGGCTTGGTAACATTCATATTGAGC 3'

I52A

Forward: 5' CAATATGAATGTTACCAAAAGGCCATGCAAGACCCCATTCAACAAG 3'

Reverse: 5' CTTGTTGAATGGGGTCTTGCATGGCCTTTTGGTAACATTCATATTG 3'

M53A

Forward: 5' CAATATGAATGTTACCAAAAGATTGCCCAAGACCCCATTCAACAAGCAGAAG 3'
Reverse: 5' CTTCTGCTTGTTGAATGGGGTCTTGGGCAATCTTTTGGTAACATTCATATTG 3'

Q54A

Forward: 5' GAATGTTACCAAAAGATTATGGCCGACCCCATTCAACAAGCAGAAG 3'
Reverse: 5' CTTCTGCTTGTTGAATGGGGTCGGCCATAATCTTTTGGTAACATTC 3'

D55A

Forward: 5' CAAAAGATTATGCAAGCCCCCATTCAACAAGC 3'
Reverse: 5' GCTTGTTGAATGGGGGCTTGCATAATCTTTTG 3'

P56A

Forward: 5' CAAAAGATTATGCAAGACGCCATTCAACAAGCAGAAG 3'
Reverse: 5' CTTCTGCTTGTTGAATGGCGTCTTGCATAATCTTTTG 3'

I57A

Forward: 5' GATTATGCAAGACCCCGCCCAACAAGCAGAAGGC 3'
Reverse: 5' GCCTTCTGCTTGTTGGGCGGGGTCTTGCATAATC 3'

Q58A

Forward: 5' GATTATGCAAGACCCCATTGCCCAAGCAGAAGGCGTTTAC 3'
Reverse: 5' GTAAACGCCTTCTGCTTGGGCAATGGGGTCTTGCATAATC 3'

Q59A

Forward: 5' CAAGACCCCATTCAAGCAGCAGAAGGCGTTTAC 3'
Reverse: 5' GTAAACGCCTTCTGCTGCTTGAATGGGGTCTTG 3'

A60L

Forward: 5' CAAGACCCCATTCAACAACTGGAAGGCGTTTACTGCAAC 3'

Reverse: 5' GTTGCAGTAAACGCCTTCCAGTTGTTGAATGGGGTCTTG 3'

L195A

Forward: 5' GTAACAATCATTCACGCCACTGCAGTGGCCAAC 3'
Reverse: 5' GTTGGCCACTGCAGTGGCGTGAATGATTGTTAC 3'

T196A

Forward: 5' GTAACAATCATTCACCTCGCCGCAGTGGCCAACAACCAG 3'
Reverse: 5' CTGGTTGTTGGCCACTGCGGCGAGGTGAATGATTGTTAC 3'

A197L

Forward: 5' CAATCATTCACCTCACTCTGGTGGCCAACAACCAGG 3'
Reverse: 5' CCTGGTTGTTGGCCACCAGAGTGAGGTGAATGATTG 3'

V198L

Forward: 5' CACCTCACTGCAGCCGCCAACAACCAG 3'
Reverse: 5' CTGGTTGTTGGCGGCTGCAGTGAGGTG 3'

A199L

Forward: 5' CACCTCACTGCAGTGCTGAACAACCAGGCCTTAG 3'
Reverse: 5' CTAAGGCCTGGTTGTTCAGCACTGCAGTGAGGTG 3'

N200A

Forward: 5' CACTGCAGTGGCCGCCAACCAGGCCTTAG 3'
Reverse: 5' CTAAGGCCTGGTTGGCGGCCACTGCAGTG 3'

N201A

Forward: 5' CTGCAGTGGCCAACGCCCAGGCCTTAGTAG 3'
Reverse: 5' CTACTAAGGCCTGGGCGTTGGCCACTGCAG 3'

Q202A

Forward: 5' GCAGTGGCCAACAACGCCGCCTTAGTAGCCAC 3'
Reverse: 5' GTGGCTACTAAGGCGGCGTTGTTGGCCACTGC 3'

A203L

Forward: 5' CAGTGGCCAACAACCAGCTGTTAGTAGCCACAAATC 3'
Reverse: 5' GATTTGTGGCTACTAACAGCTGGTTGTTGGCCACTG 3'

L204A

Forward: 5' GCCAACAACCAGGCCGCCGTAGCCACAAATCC 3'
Reverse: 5' GGATTTGTGGCTACGGCGGCCTGGTTGTTGGC 3'

V205A

Forward: 5' CAACCAGGCCTTAGCCGCCACAAATCCTG 3'
Reverse: 5' CAGGATTTGTGGCGGCTAAGGCCTGGTTG 3'

A206L

Forward: 5' CAACAACCAGGCCTTAGTACTGACAAATCCTGTTAGTTGC 3'
Reverse: 5' GCAACTAACAGGATTTGTCAGTACTAAGGCCTGGTTGTTG 3'

T207A

Forward: 5' CCAGGCCTTAGTAGCCGCCAATCCTGTTAGTTGC 3'
Reverse: 5' GCAACTAACAGGATTGGCGGCTACTAAGGCCTGG 3'

N208A

Forward: 5' GGCCTTAGTAGCCACAGCCCCTGTTAGTTGCAAAG 3'
Reverse: 5' CTTTGCAACTAACAGGGGCTGTGGCTACTAAGGCC 3'

P209A

Forward: 5' CCTTAGTAGCCACAAATGCCGTTAGTTGCAAAGTGTCC 3'
Reverse: 5' GGACACTTTGCAACTAACGGCATTTGTGGCTACTAAGG 3'

V210A

Forward: 5' GTAGCCACAAATCCTGCCAGTTGCAAAGTGTCC 3'
Reverse: 5' GGACACTTTGCAACTGGCAGGATTTGTGGCTAC 3'

S211A

Forward: 5' GTAGCCACAAATCCTGTTGCCTGCAAAGTGTCCCAGTTC 3'
Reverse: 5' GAACTGGGACACTTTGCAGGCAACAGGATTTGTGGCTAC 3'

K213A

Forward: 5' CAAATCCTGTTAGTTGCGCCGTGTCCCAGTTCATTC 3'
Reverse: 5' GAATGAACTGGGACACGGCGCAACTAACAGGATTTG 3'

V214A

Forward: 5' CAAATCCTGTTAGTTGCAAAGCCTCCCAGTTCATTCATCTTTAC 3'
Reverse: 5' GTAAAGATGAATGAACTGGGAGGCTTTGCAACTAACAGGATTTG 3'

S215A

Forward: 5' GTTAGTTGCAAAGTGGCCCAGTTCATTCATC 3'
Reverse: 5' GATGAATGAACTGGGCCACTTTGCAACTAAC 3'

Q216A

Forward: 5' CTGTTAGTTGCAAAGTGTCCGCCTTCATTCATCTTTACCTGATG 3'
Reverse: 5' CATCAGGTAAAGATGAATGAAGGCGGACACTTTGCAACTAACAG 3'

F217A

Forward: 5' GTTGCAAAGTGTCCCAGGCCATTCATCTTTACCTG 3'
Reverse: 5' CAGGTAAAGATGAATGGCCTGGGACACTTTGCAAC 3'

I218A

Forward: 5' CAAAGTGTCCCAGTTCGCCCATCTTTACCTGATGGG 3'
Reverse: 5' CCCATCAGGTAAAGATGGGCGAACTGGGACACTTTG 3'

L220A

Forward: 5' CAAAGTGTCCCAGTTCATTCATGCCTACCTGATGGGCTGTAATTAC 3'
Reverse: 5' GTAATTACAGCCCATCAGGTAGGCATGAATGAACTGGGACACTTTG 3'

L222A

Forward: 5' CCCAGTTCATTCATCTTTACGCCATGGGCTGTAATTACTTTTG 3'
Reverse: 5' CAAAAGTAATTACAGCCCATGGCGTAAAGATGAATGAACTGGG 3'

M223A

Forward: 5' CAGTTCATTCATCTTTACCTGGCCGGCTGTAATTACTTTTGGATG 3'
Reverse: 5' CATCCAAAAGTAATTACAGCCGGCCAGGTAAAGATGAATGAACTG 3'

F349A

Forward: 5' CATTGCTTGGCATTGAAGCCGTGCTGATTCCATGGC 3'
Reverse: 5' GCCATGGAATCAGCACGGCTTCAATGCCAAGCAATG 3'

V350A

Forward: 5' CTTGGCATTGAATTTGCCCTGATTCCATGGCGAC 3'
Reverse: 5' GTCGCCATGGAATCAGGGCAAATTCAATGCCAAG 3'

L351A

Forward: 5' CTTGGCATTGAATTTGTGGCCATTCCATGGCGACCTGAAG 3'
Reverse: 5' CTTCAGGTCGCCATGGAATGGCCACAAATTCAATGCCAAG 3'

I352A

Forward: 5' CATTGAATTTGTGCTGGCCCCATGGCGACCTGAAG 3'
Reverse: 5' CTTCAGGTCGCCATGGGGCCAGCACAAATTCAATG 3'

R355A

Forward: 5' GAATTTGTGCTGATTCCATGGGCCCCTGAAGGAAAGATTGCAGAG 3'
Reverse: 5' CTCTGCAATCTTTCCTTCAGGGGCCCATGGAATCAGCACAAATTC 3'

P356A

Forward: 5' GCTGATTCCATGGCGAGCCGAAGGAAAGATTGCAG 3'
Reverse: 5' CTGCAATCTTTCCTTCGGCTCGCCATGGAATCAGC 3'

E357A

Forward: 5' CCATGGCGACCTGCCGGAAAGATTGCAG 3'
Reverse: 5' CTGCAATCTTTCCGGCAGGTCGCCATGG 3'

G358A

Forward: 5' CATGGCGACCTGAAGCCAAGATTGCAGAGGAG 3'
Reverse: 5' CTCCTCTGCAATCTTGGCTTCAGGTCGCCATG 3'

K359A

Forward: 5' CCATGGCGACCTGAAGGAGCCATTGCAGAGGAGGTATATG 3'
Reverse: 5' CATATACCTCCTCTGCAATGGCTCCTTCAGGTCGCCATGG 3'

A361L

Forward: 5' CGACCTGAAGGAAAGATTCTGGAGGAGGTATATGACTAC 3'
Reverse: 5' GTAGTCATATACCTCCTCCAGAATCTTTCCTTCAGGTCG 3'

E362A

Forward: 5' CTGAAGGAAAGATTGCAGCCGAGGTATATGACTACATC 3'
Reverse: 5' GATGTAGTCATATACCTCGGCTGCAATCTTTCCTTCAG 3'

E363A

Forward: 5' GAAGGAAAGATTGCAGAGGCCGTATATGACTACATCATG 3'
Reverse: 5' CATGATGTAGTCATATACGGCCTCTGCAATCTTTCCTTC 3'

V364A

Forward: 5' GAAAGATTGCAGAGGAGGCCTATGACTACATCATGCAC 3'
Reverse: 5' GTGCATGATGTAGTCATAGGCCTCCTCTGCAATCTTTC 3'

Y365A

Forward: 5' GAAAGATTGCAGAGGAGGTAGCCGACTACATCATGCACATCC 3'
Reverse: 5' GGATGTGCATGATGTAGTCGGCTACCTCCTCTGCAATCTTTC 3'

D366A

Forward: 5' CAGAGGAGGTATATGCCTACATCATGCACATC 3'
Reverse: 5' GATGTGCATGATGTAGGCATATACCTCCTCTG 3'

Y367A

Forward: 5' GAGGAGGTATATGACGCCATCATGCACATCC 3'
Reverse: 5' GGATGTGCATGATGGCGTCATATACCTCCTC 3'

I368A

Forward: 5' GGAGGTATATGACTACGCCATGCACATCCTTATGC 3'
Reverse: 5' GCATAAGGATGTGCATGGCGTAGTCATATACCTCC 3'

M369A

Forward: 5' GAGGTATATGACTACATCGCCCACATCCTTATGCACTTC 3'
Reverse: 5' GAAGTGCATAAGGATGTGGGCGATGTAGTCATATACCTC 3'

I371A

Forward: 5' GACTACATCATGCACGCCCTTATGCACTTCC 3'
Reverse: 5' GGAAGTGCATAAGGGCGTGCATGATGTAGTC 3'

L372A

Forward: 5' CTACATCATGCACATCGCCATGCACTTCCAGGGTC 3'
Reverse: 5' GACCCTGGAAGTGCATGGCGATGTGCATGATGTAG 3'

M373A

Forward: 5' CTACATCATGCACATCCTTGCCCACTTCCAGGGTCTTTTG 3'
Reverse: 5' CAAAAGACCCTGGAAGTGGGCAAGGATGTGCATGATGTAG 3'