



The
University
Of
Sheffield.

Expression of a functional adrenomedullin type 2 receptor extracellular domain

By:

Joseph Lewis Holmes

**A thesis submitted in partial fulfillment of the requirements for the
degree of Doctor of Philosophy**

**The University of Sheffield
Faculty of Medicine, Dentistry and Health
Department of Oncology and Metabolism**

Submission Date:

September 2020

Acknowledgements

I would firstly like to say thanks to my excellent PhD supervisors, Professor Tim Skerry and Dr Gareth Richards. Both have been incredible during my post graduate degree and I would not be in this position without your support and help over the years. Your support and supervision were not only vital for the completion of my project, but also critical for my personal development. I would not be the all-rounded scientist I am today without your guidance and the opportunities you have given me to work in such a great research team, which I will be forever grateful for. I would like to extend a special thanks to Gareth for not only being a great supervisor but also for being a great friend, on the good and the most challenging days.

I also could not have completed this PhD without the support from Dr Paris Avgoustou. You were always at hand when I needed your guidance and your excellent knowledge during the small molecule screening analysis was much appreciated. A special thanks also to Dr Ameera Jailani for your help during statistical analysis and support when carrying out tricky calculations. I would also like to thank fellow lab mates Kamilla Bigos and Ewan Lilley for your support over the years and making the whole experience as fun as it could have been. I would like to express my thanks to Dr Karan Shah for always going above and beyond to help those who require it, and also for the many drinks and shots we have shared together. Those times will live long in the memory.

A very special thanks goes out to the Wellcome Trust for funding my PhD and to Dr Jean-Olivier Zirimwabagabo and Professor Joe Harrity for synthesising the novel small molecules for my project. A thanks also goes out to Professor John Rafferty for your help and support during the protein crystallography screening. Your guidance and detailed explanations of the process was much appreciated.

Last but most definitely not least, I wish to wholeheartedly say thanks to my family, friends and my girlfriend Emily for your love and unceasing encouragement over the years and now you finally have a record of what I actually do! Finally, I would like to express my gratitude to anyone who may have directly, or indirectly, had a positive impact on my research experience and thank the whole of the department for creating a friendly and welcoming working environment.

Here is not just to making great work colleagues, but also to making great friends.

Abstract

Introduction: The Calcitonin like-receptor (CLR) is a class B GPCR that couples with Receptor Activity Modifying Protein-1 (RAMP1), RAMP2 or RAMP3 to form the CGRP, AM₁ and AM₂ receptors respectively. Novel small molecules and fragments which selectively antagonise the CGRP and/or AM₂ receptors have been developed by utilising structure-based drug design, but the binding pocket occupancies of these compounds and fragments are unknown.

Hypothesis: Functional RAMP-CLR fusion proteins can be expressed and crystallised to outline the binding pocket occupancy of novel antagonists on the CGRP and AM₂ receptors.

Methods: The extracellular domains (ECD) of the RAMP-CLR protein tagged with a 10xhistidine tag and maltose binding protein were expressed in CHO-K1 cell media. Proteins were purified using immobilised metal affinity chromatography. Protein functionality and small molecule association to the receptors was determined by using an HTRF assay, displacing AM₂ receptor antagonists in single (10 µM fragments) or dose-response manner. The purified protein was used in crystallography screening trials via the sitting drop vapour diffusion method.

Results: Proteins were purified to >95% purity. Peptide displacement on both receptors corresponded with previous literature and purified proteins held accurate affinities for both high and low affinity compounds. Fragments targeting the CLR domain of the receptors have different potencies between the CGRP and AM₂ receptors ($p < 0.0471$ - 0.0001), with small molecule displacement up to 50% higher on the AM₂ receptor. Fragments with selectivity for RAMP3 have been identified ($p < 0.0121$ - 0.0001) and could be applied to any GPCR that associates with RAMP3. The highest affinity CLR targeting fragment (62% displacement) and RAMP3 targeting fragment (38% displacement) were combined to produce full length antagonists. Preliminary data shows the compound linker is vital for high antagonist potency. Commercial antagonists of the CGRP receptor displace novel compounds at lower potencies. Crystals have been generated of the RAMP3-CLR ECD bound to lead small molecule antagonist, SHF-1036.

Conclusion: RAMP-CLR ECD fusion proteins can be successfully expressed, purified, and utilised for drug screening analysis. RAMP proteins could be allosterically modulating the CLR receptor to enable peptide and compound selectivity. The RAMP3-CLR ECD structure will be resolved using X-ray diffraction.

Contents

Acknowledgements	ii
Abstract	iii
Contents	iv
List of Figures	ix
List of Tables	xiv
Abbreviations	xvi
Chapter 1: General Introduction	1
1.1 G-Protein Coupled Receptors	2
1.2 GPCR Signalling	5
1.3 Calcitonin Receptor-Like Receptor (CLR) Family	5
1.4 Receptor Activity Modifying Proteins (RAMPs) Effect on GPCR Coupling.	6
1.5 RAMP1 Extracellular Domain Complex	10
1.6 RAMP2 extracellular domain complex	12
1.7 Predicted structure of the RAMP3 extracellular domain	13
1.8 CLR extracellular domain complex and the binding pocket	16
1.9 Electrostatic Interactions between CLR and RAMP1	18
1.10 Electrostatic Interaction Between CLR and RAMP2	20
1.11 Electrostatic Interaction Between CLR and RAMP3	22
1.12 Comparison of non-bound CGRP and AM₁ extracellular domain structures	23
1.13 CGRP peptide binding to the CGRP receptor ECD	25
1.14 AM peptide binding to the AM₁ Receptor ECD	28
1.15 Residues Enabling Ligand Specificity for the CGRP and AM₁ Receptors	31
1.16 Predicted peptide binding on the AM₂ receptor extracellular domain	33
1.17 Project Rationale, aims and objectives	36
Chapter 2: Construct Development and Protein Expression	38
2.1 Introduction	39
2.1.1 Prokaryote Protein Expression System	39
2.1.2 Glycan Addition to Membrane Proteins	41
2.1.3 Mammalian Protein Expression System	42
2.1.4 Addition of Protein Tags	43
2.1.5 Aim and Objectives	45
2.2 Methods	47
2.2.1 Transformation	47
2.2.2 Plasmid Maxiprep DNA Extraction	47

2.2.3 RAMP-CLR ECD Construct Restriction Reactions and Ligations	48
2.2.4 Signal Peptide Restriction Reactions and Ligation into the MBP vector	49
2.2.5 Miniprep Plasmid Extraction	51
2.2.6 Vector Sequencing.....	52
2.2.7 Tissue Culture	52
2.2.8 Transfection of the SP-MBP-RAMP1/2/3-CLR Construct into CHO-K1 Cells	53
2.2.9 Cell Lysate Collection.....	55
2.2.10 Protein quantification: DC Protein Assay	55
2.2.11 Denaturing Protein Deglycosylation.....	56
2.2.12 Western Blot: Gel Electrophoresis and Membrane Transfer	57
2.2.13 Western Blot: Blocking and Antibody Probing	58
2.2.14 Western Blot: Detection.....	58
2.3 Results	59
2.3.1 CMV Promoter- Signal Peptide Sequence	59
2.3.2 RAMP1/2/3-CLR ECD-Histag Cloning Sequence	60
2.3.3 RAMP1/2/3-CLR Fusion protein expression	64
2.3.4 Expression of RAMP-CLR Fusion proteins in Cell Lysates	65
2.3.5 Probing for the Fusion Protein Histidine Tag	66
2.3.6 Deglycosylation of the RAMP1/2/3-CLR fusion proteins.....	67
2.4 Discussion	68
2.4.1 Protein Construct Generation and Rationale	69
2.4.2 RAMP-CLR Fusion Protein Expression in CHO cells.	69
2.4.3 Glycosylation of the RAMP-CLR Fusion Proteins.....	70
2.5 Conclusion	71
Chapter 3: Purification of the RAMP3-CLR Fusion Proteins	72
3.1 Introduction	73
3.1.1 Affinity Chromatography.....	74
3.1.2 Size Exclusion Chromatography	77
3.1.3 Aim and Objectives.....	79
3.2 Methods	80
3.2.1 Tissue Culture	80
3.2.2 Media Sample and RAMP-CLR Fusion Protein Concentration	81
3.2.3 Fusion Protein Purification	84
3.2.4 Coomassie Blue SDS-PAGE Staining	87
3.2.5 Native PAGE.....	87
3.2.6 Western Blot analysis	88

3.2.7	Histag ELISA for Protein quantification.....	88
3.3	Results	91
3.3.1	Purification Using the MBP Tag	91
3.3.2	Initial RAMP3-CLR Histidine Tag Purification Attempts.....	92
3.3.3	Effect of a Longer Histidine Tag on Protein Purification.....	97
3.3.4	RAMP2-CLR Fusion Protein Purification.....	98
3.3.5	RAMP3-CLR Fusion Protein Purification.....	101
3.3.6	Purified RAMP-CLR Fusion Protein Native PAGE Analysis	103
3.4	Discussion	104
3.4.1	RAMP-CLR Column Chromatography.....	105
3.4.2	MBP purification	107
3.4.3	RAMP-CLR Native PAGE Analysis	108
3.5	Conclusion	108
Chapter 4: Fusion Protein Functionality and Binding Pocket Occupancy of Novel CGRP and AM₂ Antagonists.....		109
4.1	Introduction	110
4.1.1	Olcegepant Binding to the CGRP Receptor	110
4.1.2	Telcagepant Association to the CGRP receptor.....	113
4.1.3	Development of MK-3207.....	116
4.1.4	The Development of Ubrogapant and Atogepant	118
4.1.5	Aim and Objectives	123
4.2	Methods	124
4.2.1	Homogenous Time Resolved Fluorescence (HTRF) Assay: Saturation binding..	124
4.2.2	HTRF Assay: Drug Displacement Response.....	127
4.2.3	Native Protein Deglycostlation	128
4.2.4	Hisbead purification.....	128
4.2.5	Data analysis	130
4.3	Results	131
4.3.1	Spectral Scan of HTRF Response.....	131
4.3.2	Saturation Binding Cruves of the RAMP1/3-CLR Fusion Proteins.....	133
4.3.3	Displacement of SHF-1257 Using the Calcitonin Family of Peptides.....	134
4.3.4	Displacement of SHF-1257 Using Novel Compounds.	137
4.3.5	SHF-1257 Displacement Using Previously Developed CGRP Antagonists.....	142
4.3.6	Effect on Glycan Addition on Receptor Binding Capabilities	144
4.4	Discussion	148
4.4.1	Peptide dissociation of SHF-1257	150

4.4.2 Olcegepant and Telcagepant Association to the RAMP1-CLR Fusion Protein ...	151
4.4.3 Olcegepant and Telcagepant Association to the RAMP3-CLR Fusion Protein ...	152
4.4.5 Comparison of SHF-638 Enantiomers and SHF-1036 Displacement	154
4.4.6 Effect of Glycans on Antagonist Binding	156
Chapter 5: Fragment Screening and Compound Development Against the RAMP-CLR receptors	158
5.1 Introduction	159
5.1.1 Structural Comparison of Previous CGRP Antagonists	159
5.2 Aim and Objectives	162
5.2 Materials and Methods	163
5.2.1 HTRF Binding Assay: Single Dose Response of RAMP/CLR Fragments	163
5.2.2 cAMP Assay: Peptide Stimulation on Overexpressing Cells	163
5.2.3 cAMP Assay: Competition Assay between fragments and peptides	166
5.2.3 Data analysis	167
5.3 Results	169
5.3.1 RAMP Fragment Single Dose Displacement	169
5.3.2 CLR Fragments Single Point Displacement	176
5.3.3 Fragment Inhibition on Receptor Overexpressing Cell lines	184
5.4 Discussion	188
5.4.1 RAMP Fragment Association to the RAMP1/3-CLR Fusion Proteins	188
5.4.2 CLR Fragment Displacement on RAMP-CLR Fusion Proteins	196
5.4.3 CLR Fragment development	198
5.4.4 Combining RAMP and CLR Fragments for Full Length Compounds	200
5.4.6 Effect on cAMP Inhibition	207
Chapter 6: Crystallography Screening of the RAMP-CLR ECD	209
6.1 Introduction	210
6.1.1 Formation of Protein Crystals	211
6.1.2 Imaging the RAMP-CLR ECD structures	214
6.1.5 Aims and Objectives	217
6.2 Methods	218
6.2.1 Protein Concentration and Buffer Exchange	218
6.2.2 Pre-incubation with CGRP and AM ₂ novel antagonist	218
6.2.3 Protein Crystallography Screening	219
6.3 Results	221
6.3.1 Crystal Screening	221
6.4 Discussion and Future Work	224

6.4.1 RAMP2-CLR Crystal Screening.....	224
6.4.2 RAMP1-CLR and RAMP3-CLR Fusion Protein Crystals.....	224
6.5 Conclusion	227
Chapter 7: General Discussion	228
7.0 General Discussion	229
7.1 The RAMP-CLR Fusion Protein can be Purified in a 1-Step Approach.	230
7.2 HTRF Assay Development	231
7.3 Displacement of SHF-1257 Shows a Functional Fusion Protein Which Retains Precise Affinity for Novel Compounds	233
7.4 RAMP Fragments Show Selectivity Between RAMP-CLR Fusion Protein Receptors.	234
7.5 CLR Fragment Displacement Supports Theory of Allosteric Modulation of CLR by RAMPs	235
7.6 Testing of Fragment Molecules Leads to the Production of Full-Length Antagonists.	237
7.7 cAMP Detection in Fragment/Compound-Receptor Inhibition	238
7.8 Crystallography Modelling and Future Work.....	239
7.9 Conclusion	242
References	243
S. Supplementary Section	258
S.1 Supplementary Section 1: Signal peptide, MBP and RAMP-CLR ECD construct Sequences	258
S.2. Supplementary Section 2: pSF-CMV-Puro-NH ₂ -MBP Vector Map	261

List of Figures

Chapter 1

Figure 1.1 The general structure of a G-protein coupled receptor.....	3
Figure 1.2 Basic response from GPCR activation.....	5
Figure 1.3: Schematic of the RAMP-CLR receptor.....	10
Figure 1.4: The ECD structure of the RAMP1 protein	11
Figure 1.5: The ECD structure of the RAMP2 protein.	13
Figure 1.6: The extracellular domain amino-acid sequence overlay of the RAMP1, RAMP2 and RAMP3 proteins.	14
Figure 1.7: The ECD structure of the RAMP3 protein	16
Figure 1.8: Structure of the RAMP1-CLR ECD complex.	17
Figure 1.9: The interface between RAMP1 and CLR ECD.....	20
Figure 1.10: The interface between RAMP2 and the CLR ECD.....	22
Figure 1.11: Structural comparison of the ligand free CGRP and AM ₁ receptor ECDs.....	25
Figure 1.12: The comparison of the ligand-bound and ligand-free structures of the CGRP receptor ECD	28
Figure 1.13: Structural comparison of the ligand-bound and ligand free AM ₁ receptor ECD	31
Figure 1.14: Overlay of the CGRP and AM peptides on the ligand-free CGRP and AM ₁ receptors.	33
Figure 1.15: Surface view comparison of the CGRP, AM ₁ and AM ₂ receptors.....	35

Chapter 2

Figure 2.1: Schematic of the RAMP-CLR fusion protein generated via cloning methods.....	49
Figure 2.2: Schematic of the BM40 secretory peptide cloning process.....	51
Figure 2.3: Example of a standard line for the DC BioRad protein assay.....	56
Figure 2.4: Sanger sequencing of the signal peptide cloned into the MBP vector..	60
Figure 2.5: Sanger sequencing of the RAMP1-CLR-10xHistag into the SP-MBP vector	61
Figure 2.6: Sanger sequencing of the RAMP2-CLR-10xHistag into the SP-MBP vector	62
Figure 2.7: Sanger sequencing of the RAMP3-CLR-10xHistag into the SP-MBP vector	63

Figure 2.8: Western Blot analysis of the RAMP1/2/3-CLR fusion protein expression in CHO-K1 cells	64
Figure 2.9: Western Blot analysis of the RAMP1/2/3-CLR fusion protein expression in CHO-K1 cell lysates.	65
Figure 2.10: Western Blot analysis of the RAMP1/2/3-CLR fusion proteins probing for their 10xhistidine tags	66
Figure 2.11: Western Blot analysis of the deglycosylated RAMP1/2/3-CLR fusion proteins probing for the proteins MBP tag.	67

Chapter 3

Figure 3.1: The association of a histidine tag to nickel ions during IMAC	75
Figure 3.2: Size Exclusion Chromatography process.....	78
Figure 3.3: Comparison of conventional direct flow filtration and tangential flow filtration (TFF).	82
Figure 3.4: Schematic of the TFF system	83
Figure 3.5: Schematic diagram of the AKTA start purification system.	84
Figure 3.6: Schematic diagram of the Histag ELISA for protein quantification.....	89
Figure 3.7: Standard curve for the CREB1 Recombinant Human Protein N-His.MBP Tag ELISA.....	90
Figure 3.8: Western blot of the RAMP3-CLR fusion protein purification using its MBP tag.	91
Figure 3.9: RAMP3-CLR fusion protein purity using a 6x histidine tag and FCS supplemented cell growth media	94
Figure 3.10: RAMP3-CLR fusion protein purity using 300 mM NaCl and FCS supplemented media	95
Figure 3.11: Purification of the RAMP3-CLR fusion protein utilising a 6x histidine tag in CD FortiCHO media.....	96
Figure 3.12: Purification of the RAMP1-CLR fusion protein utilising a 10x histidine tag.....	98
Figure 3.13: Purification of the RAMP2-CLR fusion protein from stable expression in CHO-K1 cells	100
Figure 3.14: Purification of the RAMP2-CLR fusion protein from transiently transfected CHO-K1 cells.....	101
Figure 3.15: Purification of the RAMP3-CLR fusion protein from utilising a 10x histidine tag.	102

Figure 3.16: Native gel electrophoresis of purified RAMP-CLR fusion proteins.....	103
---	-----

Chapter 4

Figure 4.1: Chemical structure of the CGRP antagonist olcegepant	111
Figure 4.2: Surface and ribbon representation of olcegepant bound to the RAMP1-CLR ECD (PDB: 3N7S)	112
Figure 4.3: The chemical structure of the CGRP antagonist telcagepant	113
Figure 4.4: Comparison of CGRP receptor binding pocket occupancy by telcagepant and olcagepant.....	116
Figure 4.5: The development of CGRP antagonist MK-3207.....	117
Figure 4.6: Comparison of MK-3207 and telcagepant binding to the CGRP receptor	118
Figure 4.7: Structural comparison of ubrogepant and atogepant	119
Figure 4.8: Comparison of ubrogepant and telcagepant binding pocket occupancy in the CGRP receptor	121
Figure 4.9: Schematic of the HTRF assay.....	124
Figure 4.10: Spectral view of terbium cryptate and BODIPY-FL.....	125
Figure 4.11: Well layout for the HTRF assay	127
Figure 4.12: Spectral scan of the HTRF assay	132
Figure 4.13: HTRF saturation binding curves of SHF-1257 on the RAMP1/3-CLR fusion proteins	133
Figure 4.14: Ligand displacement curves	135
Figure 4.15: Dose response curve of SHF-1036 displacing SHF-1257 on the RAMP1/3-CLR fusion protein receptors.....	138
Figure 4.16: Structures of CGRP and AM ₂ receptor antagonists.....	139
Figure 4.17: Dose response curves of SHF-638, SHF-770 and SHF-771 on the RAMP1 and RAMP3-CLR fusion proteins	141
Figure 4.18: Olcegepant and Telcagepant HTRF displacement curves	143
Figure 4.19: HTRF response post deglycosylation of the RAMP1-CLR fusion protein	145
Figure 4.20: HTRF response post deglycosylation of the RAMP3-CLR fusion protein	146
Figure 4.21: Structural comparison of SHF-1036, SHF-770 and SHF-771.....	155
Figure 4.22: Superimposition of the SHF-770 and SHF-771 onto the CGRP receptor.....	156

Chapter 5

Figure 5.1: The identification of the CLR bind region, the interface and the CLR/RAMP binding region	161
Figure 5.2: Schematic of the cAMP assay used to test CGRP, AM ₁ and AM ₂ receptor inhibition	164
Figure 5.3: Single dose (10 μM) displacement of SHF-1257 from RAMP1 and RAMP3-CLR fusion proteins using RAMP fragments	170
Figure 5.4: Single dose (10 μM) displacement of SHF-1257 from RAMP1 and RAMP3-CLR fusion proteins using RAMP fragments	171
Figure 5. 5 Structural comparison of RAMP fragments SHF-971 and SHF-968.....	172
Figure 5.6: Structural comparison of RAMP fragments SHF-418, SHF-209, SHF-709, SHF-1309, SHF-972	173
Figure 5.7: Structural comparison of RAMP fragments SHF-1140, SHF-1141 and SHF-1142	175
Figure 5.8: Single dose (10 μM) displacement of SHF-1257 from RAMP1 and RAMP3-CLR fusion proteins using CLR fragments	177
Figure 5.9: Single dose (10 μM) displacement of SHF-1257 from RAMP1 and RAMP3-CLR fusion proteins using CLR fragments	178
Figure 5.10: Structural comparison of CLR fragments SHF-1088, SHF-1351, SHF-952, SHF-695, SHF-985, SHF-1346.....	180
Figure 5. 11: Structural comparison of CLR fragments SHF-969, SHF-356 and SHF-207	182
Figure 5.12: Structural comparison of CLR fragments SHF-359, SHF-1314 and SHF-1335.	183
Figure 5.13: Agonist stimulation of 1231N1 O/E cell lines.	184
Figure 5.14: Activity of small molecule fragment antagonists against the CGRP, AM ₁ and AM ₂ receptors	186
Figure 5.15: Region of RAMP fragment targeting during FBDD.....	189
Figure 5.16: Development of the RAMP fragment	192
Figure 5.17: Development of the RAMP fragment and possible changes for receptor affinity	194
Figure 5.18: The final adaptation of the RAMP fragment.....	195
Figure 5.19: Region of CLR fragment targeting during FBDD.....	197
Figure 5.20: Development of the CLR fragment.	199
Figure 5.21: Development of the CLR fragment using SHF-1335.	200

Figure 5.22: Developed compounds for targeted AM ₂ selectivity after combing RAMP and CLR fragments	203
Figure 5.23: Developed compounds for targeted CGRP and AM ₂ receptors after combing RAMP and CLR fragments.....	205
Figure 5.24: Developed compounds for targeted for CGRP or AM ₂ selectivity using the CLR fragment SHF-1335	207

Chapter 6

Figure 6.1: Summary of the crystallography process.....	210
Figure 6.2: X-ray diffraction of a protein crystal	211
Figure 6.3: Phase diagram of protein supersaturation	213
Figure 6.4: Typical outcomes from crystallography experiment.....	214
Figure 6.5: In phase and out of phase diffraction waves.	215
Figure 6.6: Schematic of the Mosquito Liquid Dispensing system.....	220
Figure 6.7: RAMP3-CLR fusion protein crystal with SHF-1036 novel antagonist	222
Figure 6.8: Deglycosylated RAMP3-CLR fusion protein crystal with SHF-1036 novel antagonist.....	222
Figure 6.9: Glycosylated RAMP3-CLR fusion protein crystal without addition of novel antagonist.....	223

Chapter 7

Figure 7.1: Potential sites of antagonist binding on the ligand free CGRP and AM ₁ receptors ECD.	241
---	-----

List of Tables

Chapter 1

Table 1.1: Summary of the GPCR classification and their family characteristics.....	4
Table 1.2: Rank order of potency of agonists against to calcitonin family of receptors	8
Table 1.3: Interactions between RAMPs and GPCR proteins. Adapted from (Hay and Pioszak, 2016)	9
Table 1.4: Amino-acid residues which comprise the CGRP and AM ₁ receptor binding pockets.....	17
Table 1.5: Key residues involved in the RAMP1-CLR Interface	19
Table 1.6: Summary of the Interactions in the RAMP1-CLR interface.....	19
Table 1. 7: Summary of the key residues involved in association of RAMP2 and CLR proteins	21
Table 1.8: Summary of the interactions which make up the RAMP2-CLR interface	21
Table 1.9: Key interactions between the CGRP/AM peptides and their receptors	30

Chapter 2

Table 2.1: Protein standard concentrations for the BCA assay	55
--	----

Chapter 3

Table 3.1: Dilution of stock protein for the ELISA standard curves	90
--	----

Chapter 4

Table 4.1: Key CGRP receptor residues involved in hydrogen bonding to olcegepant and telcagepant	116
Table 4.2: Pharmacological Properties of CGRP antagonists.....	122
Table 4.3: Structures and binding affinities of SHF-1036 and SHF-1257.	125
Table 4.4: Parameters used during HTRF plate reading	127
Table 4.5: Values of SHF-1257 saturation binding to the RAMP1/3-CLR fusion proteins ..	134

Table 4.6: Parameters of ligand displacement curves on the RAMP1-CLR and RAMP3-CLR fusion proteins	136
Table 4.7: Values of SHF-1036 displacing SHF-1257 on the RAMP1/3-CLR fusion proteins	138
Table 4.8: IC50 values for SHF-638, SHF-770 and SHF-771 on the RAMP1-CLR fusion protein	142
Table 4.9: IC50 values for SHF-638, SHF-770 and SHF-771 on the RAMP3-CLR fusion protein	142
Table 4.10: SHF-1257 association values to glycosylated and deglycosylated RAMP1-CLR fusion.....	146
Table 4.11: SHF-1257 association values to the glycosylated and deglycosylated RAMP3-CLR fusion protein	147

Chapter 5

Table 5.1: Ligand dilutions for cAMP stimulation	165
Table 5.2: Plate reader settings for the cAMP assay.....	166
Table 5.3: Fragment Dilutions from cAMP assay	167
Table 5.4: Pharmacological profiles of RAMP fragments on displacing SHF-1257	176
Table 5.5: Pharmacological profiles of CLR fragments on displacing SHF-1257	183
Table 5.6: Best fit values of ligand stimulation of CGRP, AM ₁ and AM ₂ O/E cells	185
Table 5.7: Best fit values from fragment inhibition of the CGRP, AM ₁ and AM ₂ receptors	187

Chapter 6

Table 6.1: Screening conditions which resulted in protein crystals.....	223
---	-----

Abbreviations

Ala: Alanine

AM: Adrenomedullin

AM₁: Adrenomedullin Receptor 1

AM₂: Adrenomedullin 2 Peptide

AM₂: Adrenomedullin Receptor 2

AMY: Amylin

AMY₁: Amylin 1 Receptor

AMY₂: Amylin 2 Receptor

AMY₃: Amylin 3 Receptor

Arg: Arginine

Asn: Asparagine

Asp: Aspartic acid

Bmax: Binding Maximum

Bp: base pairs

cAMP: Cyclic Adenosine Monophosphate

CGRP: Calcitonin Gene Related Peptide

CHAPS: 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate

CHO-K1: Chinese Hamster Ovary Cells

CLR: Calcitonin-Like Receptor

CT: Calcitonin

CTR: Calcitonin Receptor

CV: Column Volume

Cys: Cysteine

DNA: Deoxyribonucleic Acid

DsbA: Disulphide Bond A Isomerase

DsbC: Disulphide Bond C Isomerase

DTT: dithiothreitol

E. coli: Escherichia coli

EC50: Half Maximal Excitatory Concentration

ECD: Extracellular Domain

ECL: Extracellular Loop

EDTA: Ethylenediaminetetraacetic acid

ELISA: Enzyme linked immunosorbent assay

ER: Endoplasmic Reticulum

EU/SA-cAMP: Europium/Streptavidin-Labelled cAMP

FBDD: Fragment Based Drug Design

FBS: Foetal Bovine Serum

FCS: Foetal Calf Serum

FITC: Fluorescein Isothiocyanate

FRET: Fluorescence Resonance Energy Transfer

FT: Flow Through

GDP: Guanine 5'-diphosphate

Gln: Glutamine

Glu: Glutamic acid

Gly: Glycine

GPCR: G-Protein Coupled Receptor

GST: Glutathione S-Transferase

GTP: Guanine Triphosphate

HEK: Human Embryonic Kidney

His: Histidine

HisTag: Histidine Tag

HRP: Horse Radish Peroxidase

HTRF: Homogenous Time Resolved Fluorescence

HYPERFlasks: High Yield Performance Flasks

IBMX: 3-isobutyl-1-methylxanthine

IC50: Half Maximal Inhibitory Concentration

ICL: intracellular Loop

Ile: Isoleucine

IMAC: Immobilised Metal Affinity Chromatography

IMD: Intermedin/adrenomedullin 2

K_d : Dissociation Constant

Leu: Leucine

Lys: Lysine

MBP: Maltose Binding Protein

Met: Methionine

MW: Molecular Weight

MWCO: Molecular Weight Cut Off

NTA: nitrilotriacetate

O/E: Overexpressing

PAGE: Polyacrylamide Gel Electrophoresis

PBS: Phosphate Buffered Saline

PDB: Protein Data Bank

PEG: polyethylene glycol

PEI: Polyethylenimine

Phe: Phenylalanine

PP: Pre-purified

Pro: Proline

RAMP: Receptor Activity Modifying Protein

rSAP: Shrimp Alkaline Phosphatase

SBDD: Structure Based Drug Design

SD: Standard Deviation

SDS: Sodium Dodecyl Sulphate

SEC: Size Exclusion Chromatography

Ser: Serine

SOC: Super Optimal Broth with Catabolite Suppression

SP: Signal Peptide

TBS-T: Tris Buffered Saline with Tween-20

TFF: Tangential Flow Filtration

Thr: Threonine

TM: Transmembrane

TMB: 3,3',5,5'-tetramethylbenzidine

TR-FRET: Time Resolved Fluorescence Resonance Energy Transfer

Tris: Trisma

Trp: Tryptophan

Tyr: Tyrosine

Val: Valine

WT: Wild Type

α : Alpha

β : Beta

γ : Gamma

Chapter 1: General Introduction

1.1 G-Protein Coupled Receptors

G-protein Coupled Receptors are a superfamily of integral membrane proteins which consists of over 800 members and are encoded by ~5% of the human genome (Zhang, Devries and Skolnick, 2006). The receptor proteins are identified by their seven-transmembrane (7TM) helices which are connected by 3 extracellular (ECL 1-3) and 3 intracellular loops (ICL 1-3), with these loops and N-terminus collectively known as the juxtamembrane region (figure 1.1). GPCRs can be divided into 5 families in humans, creating the GRAFs system of classification (Glutamate, Rhodopsin, Adhesion, Frizzled/Taste2 and Secretin) (Schiöth and Fredriksson, 2005). These include the Class A rhodopsin family, the Class B secretin family, the adhesion family, the Class C glutamate family and the Class F frizzled/taste family (Lundin, Oth and Breeding, 2003)(Zhang, Zhao and Wu, 2015). The distinguishable characteristics and properties of each family is summarised in table 1.1

GPCRs recognise a variety of extracellular stimuli including hormones, chemokines, ions and neurotransmitters by the receptors extracellular loops and occasionally the N-terminus of the receptor (extracellular domain) (Zhang, Zhao and Wu, 2015). Upon their stimulation, the signal is transduced and the GPCR receptor undergoes a series of conformational changes which ultimately releases the G-protein from the receptor C-terminus. The G-protein consists of an α (alpha), β (beta) and γ (gamma) subunit with a guanine 5'-diphosphate (GDP) molecule bound in its inactive form and is responsible for the downstream signalling cascade of the receptor (figure 1.1).

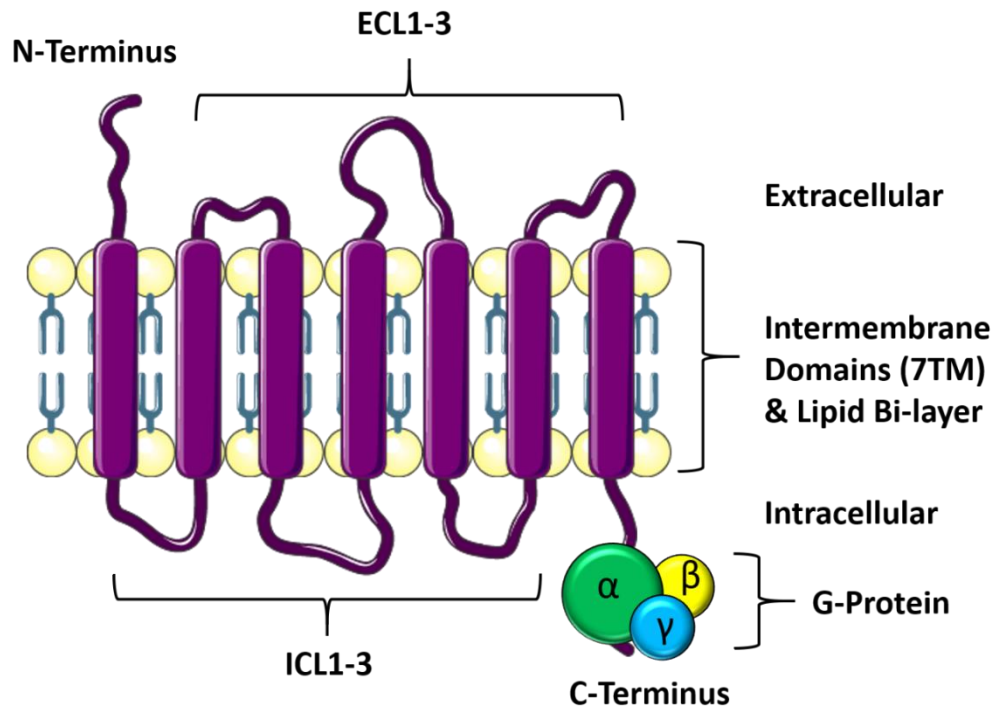


Figure 1.1 The general structure of a G-protein coupled receptor. Outlined is the intermembrane domain with intra/extracellular loops and the heterotrimeric G-protein. The extracellular N-terminus and intracellular C-terminus is annotated. Image created using Servier Medical Art under the creative commons license.

The most significant structural differences between GPCRs which results in varying binding properties are found in the receptors extracellular domain (ligand binding domain and the extracellular loops) (Rosenbaum, Rasmussen and Kobilka, 2014). These structural differences allow the receptors to bind to different external stimuli and produce an intracellular response. The differences in their structure allows the receptors to be categorised into their specific class of GPCR. Examples in structural variances include the ECL2 of rhodopsin (Class A) forming a short β -sheet which interacts with its ligand (*Cis*-retinal) while the ECL2 of the β -adrenergic receptor (Class B) forms a short α -helices structure which is stabilised by intra- and inter- molecular disulphide bonds (Rosenbaum *et al.*, 2007)(Cherezov *et al.*, 2008)(Okada *et al.*, 2004).

Table 1.1: Summary of GPCR classification and their family characteristics

GPCR Family	Receptor Characteristics
Rhodopsin Receptor (Class A)	<p>The largest family of GPCRs which are receptors for peptides and small ligands.</p> <p>Can be further divided:</p> <ul style="list-style-type: none"> - Group 1: GPCRs for small ligands and binding site is in the 7TM - Group 2: Peptide binding site is in the N-terminus and extracellular loops - Group 3: Characterised by large extracellular loops with peptide binding domain on the N-terminus and at least one contact is formed with ECL1 and 3. (Tuteja, 2009).
Secretin Family (Class B)	<p>Characterised by a large N-terminal domain. Share a similar morphology with group 3 in the rhodopsin receptors but do not share any sequence homology and bind high molecular weight ligands/peptides (Tuteja, 2009)</p>
Glutamate Receptor (Class C)	<p>These hold large ectodomains and consist of metabotropic glutamate, Ca²⁺ sensing, taste, and olfactory receptors. Most ligand binding pockets are conserved in the extracellular region but some receptors have shown allosteric binding sites within the 7TM bundle (Gacasan, Baker and Parrill, 2018)</p>
Adhesion Receptors	<p>Consists of highly variable N-terminal regions in number of amino acids, ranging from 200-2800 in length. These receptors bind extracellular molecules rather than peptide hormones and are involved in cell proliferation and migration (Langerstrom and Helgi, 2008)(Gacasan, Baker and Parrill, 2018)</p>
Frizzled Receptor Family (Family F)	<p>These receptors are primarily involved in embryonic development and are known to bind secreted glycoproteins. Their N-terminus is cysteine rich and is conserved in the receptor family (Tuteja, 2009)(Gacasan, Baker and Parrill, 2018) (Langerstrom and Helgi, 2008).</p>

1.2 GPCR Signalling

Once external stimuli have bound to the GPCR receptor, the receptor induces intracellular signalling through the heterotrimeric G-protein. The GPCR undergoes a conformational change which enables the GDP molecule to be exchanged for guanine triphosphate (GTP), prompting the $G\alpha$ subunit to be released from the dimeric $G\beta\gamma$ subunit (Tuteja, 2009). The $G\alpha$ unit can initiate the $G\alpha$ -s signalling pathway which increases concentrations of secondary messenger cyclic adenosine monophosphate (cAMP), which ultimately results in gene transcription of specific proteins (figure 1.2). This signalling pathway can be hindered if the $G\alpha$ -s pathway is activated, inhibiting cAMP production. An additional activation pathway includes the $G\alpha$ -q pathway. This pathway activates phospholipase C which ultimately leads to Ca^{2+} ion increase intracellular and activates ion channels in the cell. These receptor pathways are summarised in figure 1.2.

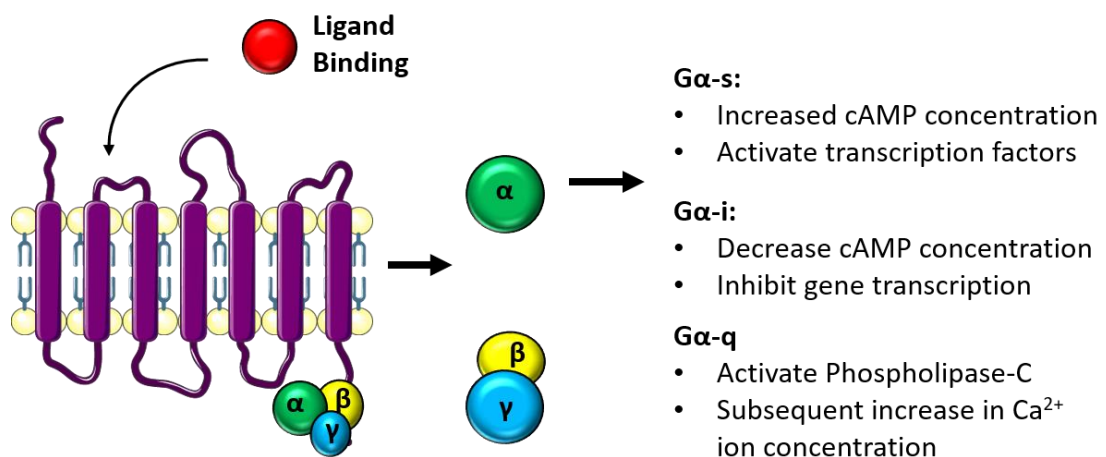


Figure 1.2 Basic response from GPCR activation. Activating GPCRs induces the $G\alpha$ subunit to dissociate from the $G\beta\gamma$ subunits and induce the signalling cascade. The type of signals produced is summarised and dependant on the receptor activated.

1.3 Calcitonin Receptor-Like Receptor (CLR) Family

The Calcitonin Like-Receptor (CLR) is a 461-amino acid GPCR protein which was identified as a novel receptor being closely related to the calcitonin receptor (CTR), sharing ~56% sequence homology with the human receptors (Fluhmann *et al.*, 1995). The 2 proteins are a part of the Class B Secretin family of GPCRs and both possess large N-terminal extracellular domains (ECD) which are vital for peptide binding. The class B GPCR family have large N-

terminal domains which are essential for peptide binding, with ligands associating to the receptors by the 'two domain model' (Hoare, 2005). The model describes the C-terminus of the peptide first being captured by the N-terminus of the receptor. This allows the ligand to effectively anchor itself to the receptor and bind to the upper TM domains and ECLs (Hoare, 2005). This allows receptor activation and has been confirmed by numerous studies using chimeric receptors and their peptides (Barwell *et al.*, 2012) (Holtmann, Hadac and Miller, 1995)(Moore *et al.*, 1995). The CLR and CTR receptors also share the same N-terminal structure, consisting of 2 anti-parallel β -sheets and α -helices, but the detailed structure of CLR ECD is discussed later in this chapter (Section 1.8).

CTR is primarily stimulated by the calcitonin peptide (CT) which belongs to the calcitonin family of peptides, together with amylin (AMY), adrenomedullin (AM), intermedin (IMD, also referred to as adrenomedullin 2) and the calcitonin gene related peptide (CGRP). Although their peptide homology in their primary sequences are weak, they are grouped due to their strong similarities in their secondary structure (Poyner *et al.*, 2002). At the point of the CLR receptor discovery, its native peptide was not distinguished with both the CT peptide and the closely related neuropeptide CGRP, not stimulating the CLR receptor when it was expressed in COS-7 cells (Fluhmann *et al.*, 1995). It wasn't until the CLR receptor was expressed in human embryonic kidney cells (HEK293) that a 60-fold increase in cAMP levels was observed upon cell stimulation with the CGRP peptide (Aiyar *et al.*, 1996). This led to the discovery that receptor activity modifying proteins (RAMPs), proteins which are found endogenously in HEK293 cells, were important regulators in trafficking the CLR receptor to the cell surface and altering ligand pharmacology (McLatchie *et al.*, 1998) .

1.4 Receptor Activity Modifying Proteins (RAMPs) Effect on GPCR Coupling.

RAMP proteins are single transmembrane domain proteins which can modulate the function and binding properties of GPCRs upon their association. To date, there are 3 known RAMPs in humans (RAMP1, 2 and 3) which hold short C-terminal domains (~9 amino-acids) and N-terminal domains of ~90-100 amino acids long, depending on the RAMP subtype (Sexton *et al.*, 2001). All 3 RAMPs have been seen to associate to the CLR receptor from biosynthetic compartments in the cell (golgi apparatus and endoplasmic reticulum) and 3 distinct receptor subtypes are formed upon their association to the CLR receptor (McLatchie *et al.*, 1998)(Harikumar *et al.*, 2010). RAMP proteins have also been seen to associate with other

GPCRs, including the CTR receptor to aid their cell surface expression and form other distinguishable receptor subtypes, with table 1.3 summarising these findings (Hay and Pioszak, 2016).

If RAMP1 associates to the CLR receptor it forms the CGRP receptor, which is primarily stimulated by the CGRP peptide, although the receptor does hold some affinity for other calcitonin family peptides (table 1.2) (McLatchie *et al.*, 1998). CGRP is a very potent vasodilator, believed to have a strong role in blood flow regulation, recruitment of inflammatory cells and pain modulation with both itself and the CGRP receptor being expressed in trigeminal neurones which form C- and A- fibres (Ceppa, Lennerz and Ru, 2008)(Eftekhari *et al.*, 2010)(Russell *et al.*, 2014). There is strong evidence which proposes that CGRP receptor stimulation by CGRP has a primary role in migraine pathophysiology, despite years of previous research into the link between the vasodilator substance P and pain modulation (Diener, 2003)(Hansen *et al.*, 2010). CGRP has a direct link to pain pathophysiology, with studies showing intravenous injection of CGRP inducing migraine-like attacks in those who have previously suffered migraines (Uddman *et al.*, 1985). Other studies show the normalisation of CGRP levels after the administration of antimigraine drug sumatriptan (Goadsby and Edvinsson, 1993). There are several CGRP antagonists which can treat migraine pain by selectively binding to the CGRP receptor. The specific binding interactions of these antagonists to the CGRP receptor are discussed further in chapter 4.

When RAMP2 interacts with the CLR receptor, it forms the adrenomedullin 1 receptor (AM₁) which is primarily stimulated by the AM peptide but does have affinity for CGRP, IMD, AMY and CT (table 1.2). AM is a 52 amino-acid multifunctional peptide which was first identified in the pheochromocytoma of the human adrenal gland but is expressed in numerous organs and tissues such as cardiac and skeletal muscle, kidneys and liver tissues (Blom *et al.*, 2012)(Kitamura *et al.*, 1993). AM signalling through the AM₁ receptor induces angiogenesis, promotes proper vascular integrity and prevents cell apoptosis (Oehler *et al.*, 2001)(Koyama *et al.*, 2013). Although this signalling purpose is a protective response, AM is upregulated in response to hypoxia and therefore has been implied to be responsible for some types of tumour development (Garayoa *et al.*, 2000)(Oehler *et al.*, 2001)(Oehler *et al.*, 2003) (Zudaire *et al.*, 2006) (Keleg *et al.*, 2007)(Deng *et al.*, 2012)(Berenguer-daiz *et al.*, 2013) (Aggarwal *et al.*, 2013) (Larráyoiz *et al.*, 2014).

Association of RAMP3 with the CLR receptor results in the adrenomedullin 2 receptor (AM₂) which shares equal potency for the AM and IMD peptide, but also has partial affinity for the

CGRP peptide (table 1.2). Although the AM₂ receptor shares equal potency with AM and IMD, it is important to note that only the AM peptide acts as the full agonist and IMD acts as a partial agonist (Weston *et al.*, 2016). AM peptide signalling through the AM₂ receptor exhibits similar effects to that of AM₁ receptor signalling. AM₂ receptor stimulation induces angiogenesis but is thought to have a lesser role in promoting proper vascular integrity to AM₁ receptors signalling (Yamauchi *et al.*, 2014). The rank order of ligands to each receptor subtype can be seen in table 1.2.

Table 1.2: Rank order of potency of agonists against to calcitonin family of receptors

Receptor	Composition	Rank Order of agonist potency
CGRP Receptor	CLR and RAMP1	CGRP > IMD=AM ≥ AMY ≥ CT
AM₁	CLR and RAMP2	AM > CGRP > IMD > AMY ≥ CT
AM₂	CLR and RAMP3	AM = IMD > CGRP > AMY > CT
AMY1	CTR and RAMP1	AMY ≥ CGRP > CT > AM
AMY2	CTR and RAMP2	Poorly Characterised
AMY3	CTR and RAMP3	AMY > CGRP > AM

Table 1.3: Interactions between RAMPs and GPCR proteins. Adapted from (Hay and Pioszak, 2016)

GPCR Receptor	GPCR Class	RAMP Interacting	Result of the Interaction	Reference
GPR30 (oestrogen receptor)	A	RAMP3	Receptor trafficking	(Lenhart <i>et al.</i> , 2014)
CLR receptor (CGRP, AM ₁ and AM ₂)	B	RAMP1-3	Receptor trafficking and altering ligand pharmacology	(McLatchie <i>et al.</i> , 1998)
Calcitonin Receptor (AMY ₁ , AMY ₂ , AMY ₃)	B	RAMP1-3	Modulated ligand pharmacology	(Poyner <i>et al.</i> , 2002)
Parathyroid hormone receptor 1	B	RAMP2	Unknown	(Christopoulos <i>et al.</i> , 2003)
Parathyroid hormone receptor 2	B	RAMP3	Unknown	(Christopoulos <i>et al.</i> , 2003)
Vasoactive Intestinal Polypeptide Receptor 1	B	RAMP1-3	Modulates Receptor signalling/pharmacology	(Christopoulos <i>et al.</i> , 2003)
Vasoactive Intestinal Polypeptide Receptor 2	B	RAMP1-3	Modulates Receptor signalling/pharmacology	(Wootten <i>et al.</i> , 2013)
Corticotropin-releasing hormone receptor 1	B	RAMP2	Traffics the receptor to the cell surface and modulates receptor signalling	(Wootten <i>et al.</i> , 2013)
Glucagon Receptor	B	RAMP2	Unknown	(Christopoulos <i>et al.</i> , 2003)
Secretin Receptor	B	RAMP3	Receptor Trafficking to the cell surface	(Harikumar <i>et al.</i> , 2010)
Atypical chemokine	B	RAMP3	Aid ligand pharmacology and signalling	(Mackie <i>et al.</i> , 2019)
Calcium Sensing receptor	C	RAMP1 and RAMP3	Traffics the receptor to the cell surface	(Desai <i>et al.</i> , 2014)

1.5 RAMP1 Extracellular Domain Complex

CGRP receptor stimulation results in migraine pain and signalling via the AM₁ and AM₂ receptors promotes angiogenesis and aids cancer cell survival, meaning these receptors are clinically applicable. As the N-terminus of the secretin family (class B) of GPCR receptors is essential for peptide binding, their structures have been heavily investigated. The N-terminal structures of both the CLR and RAMP proteins have been developed to aid drug design, in attempts to antagonise the receptor and block the key interactions of their stimulants. The N-terminal ligand binding site is outlined in figure 1.3.

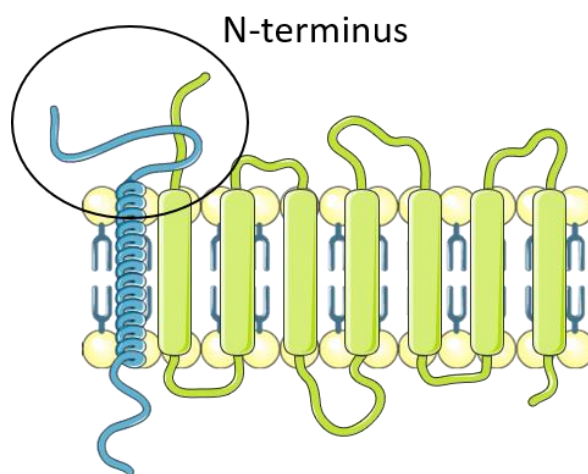


Figure 1.3: Schematic of the RAMP-CLR receptor. The RAMP protein (blue) and CLR receptor (green) coupled at the cell membrane. Outlined is the N-terminus of the receptor which is vital for peptide binding.

Kusano *et al.*, 2008 were the first group to generate an extracellular domain (ECD) crystal model of the RAMP1 protein (PDB: 2YX8). More recent studies have developed the structure of the RAMP1 protein linked with the CLR receptor, with the general RAMP1 structure being replicated between each image (ter Haar *et al.*, 2010)(Booe *et al.*, 2015)(Liang *et al.*, 2019). The RAMP1 ECD (residues Cys27-Ser107) has a 3 α -helices bundle (α 1-3) with α 2 aligned in an anti-parallel manner in comparison to α 1 and 3, which was predicted in previous mutagenesis studies (Simms *et al.*, 2006). There are 3 disulphide bonds within the triple helices, located on residues Cys27-Cys82, Cys40-Cys72 and Cys57-Cys104, with mutagenesis of these residues suggesting the Cys40-72 and Cys57-104 disulphide bonds are vital for the proper cell surface expression of the CLR-RAMP1 heterodimer (Kusano *et al.*, 2008).

A noticeable characteristic of RAMP1 is the 'kink' at Leu39 in the α 1 helix (figure 1.4), which

appears to be stabilised by the Cys40-Cys72 disulphide bond (Kusano *et al.*, 2008). The Leu39 kink has had a subsequent effect on displacing the hydrogen bonds between the Leu39 carbonyl group and Gln43 amino group. Further abnormal hydrogen bonds include those formed between the Leu36 carbonyl group and Leu41 amino-group rather than the amino group of Cys40, and an hydrogen bond between the carbonyl group of Arg37 and amino-group of Thr42, rather than the amino group of Leu41 (Kusano *et al.*, 2008). Within the alpha helix bundle of the RAMP1 ECD, there are numerous hydrophobic interactions between the $\alpha 1$ and $\alpha 3$ helices, involving the side chains of $\alpha 1$ amino-acid residues Tyr32, Leu35, Leu41, Phe44, Met48 and Val51, and the $\alpha 3$ amino acid residues Val89, Phe92, Val96, Try100 and Phe101 (Kusano *et al.*, 2008). These interactions will likely play a role in the protein structure and conformation.

The electrostatic potential distribution of RAMP1 shows the presence of a hydrophobic patch in the shallow concave area formed between the $\alpha 2$ and $\alpha 3$ helices (figure 1.4), with residues in this being highly conserved between the three RAMP subtypes. This hydrophobic patch is the region of the RAMP1-CLR interface, which is discussed further in section 1.10 (Kuwasako, Kitamura, Nagoshi, *et al.*, 2003).

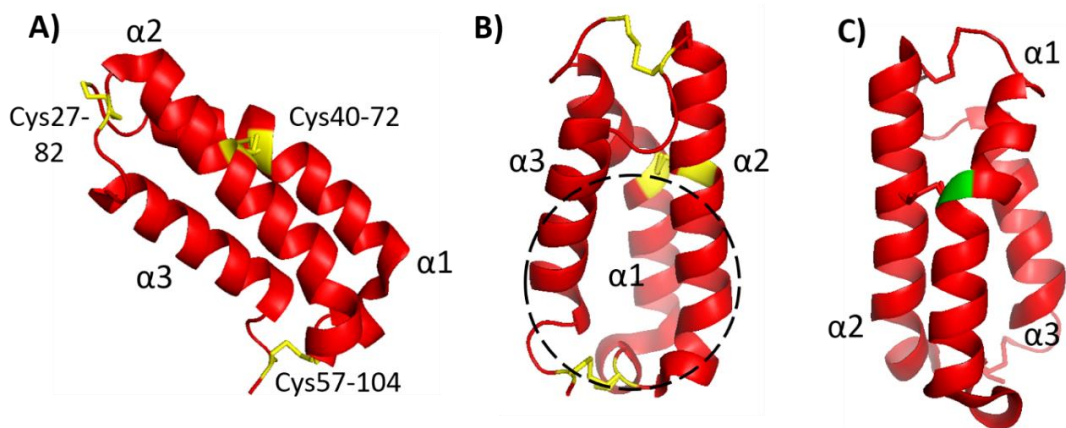


Figure 1.4: The ECD structure of the RAMP1 protein. **A)** The RAMP1 ECD (red) (PDB: 2YX8) comprises 3 α -helices with $\alpha 2$ aligned in an anti-parallel manner. The 3 cysteine bonds within the RAMP1 ECD structure are highlighted in yellow. **B)** The hydrophobic path between the $\alpha 2$ and $\alpha 3$ helices is annotated by the black dashed ring. **C)** The reverse view (behind) of the RAMP1 triple helices with Leu39 kink annotated in green. Image created using PyMOL Molecular Graphics System, Version 1.7.4.5- Educational Product, Schrodinger, LLC.

1.6 RAMP2 extracellular domain complex

Kusano *et al.*, 2012 were the first to describe the extracellular domain structure of the RAMP2 protein but more recent studies replicated and confirmed the protein structure (Booe *et al.*, 2015)(Liang *et al.*, 2020). The RAMP2 ECD (PDB: 3AQF) was found to be structurally similar to RAMP1 despite only a 30% sequence homology between all the three RAMP subtypes and RAMP2 ECD being 27 amino acids longer than both RAMP1 and 3 (figure 1.5) (Klein, Matson and Caron, 2016). The RAMP2 ECD (residues Gly56-Ser139) adopts the same three-helix bundle, with the $\alpha 2$ helices positioned in an anti-parallel manner in comparison to the $\alpha 1$ and 3 helices (figure 1.5). The identifiable difference between the RAMP1 and RAMP2 proteins is the kink in the α -helices 1 of RAMP1. The Leu39 kink is not present in RAMP2 (leucine is replaced by phenylalanine in RAMP2) and a straight $\alpha 1$ is observed, allowing easier structural identification between the two protein subtypes (Kusano *et al.*, 2012) (figure 1.5).

The RAMP2 structure contains 2 disulphide bonds which connect the $\alpha 1$ and $\alpha 2$ helices (Cys68-Cys99) and the loop between the $\alpha 1$ -2 helices and the C-terminal tail (Cys84-Cys 131) (figure 1.5). RAMP2 does not have the Cys27-Cys82 disulphide bridge that connects the loop between $\alpha 2$ -3 and the N-terminal of the RAMP sequence which is present in the RAMP1 structure (Kusano *et al.*, 2012)(Kusano *et al.*, 2008). Mutagenesis of all 4 cysteine residues in RAMP2 report reduced association of AM to the AM₁ receptor, suggesting the disulphide bonds contribute to the stability and structural integrity of the RAMP2 ECD structure (Kuwasako, Kitamura, Uemura, *et al.*, 2003).

The RAMP2 ECD also has a hydrophobic patch similar to that of the RAMP1 EDC, between the $\alpha 2$ and 3 helices (figure 1.5). This is also the region which forms the RAMP2-CLR interface and is discussed in section 1.11 (Kusano *et al.*, 2012). The top (latter) end of the $\alpha 2$ helices and the $\alpha 2$ -3 helices loop on RAMP2 holds residues which are key for AM peptide binding and selectivity, with these contacts being discussed further in sections 1.14 and 1.15.

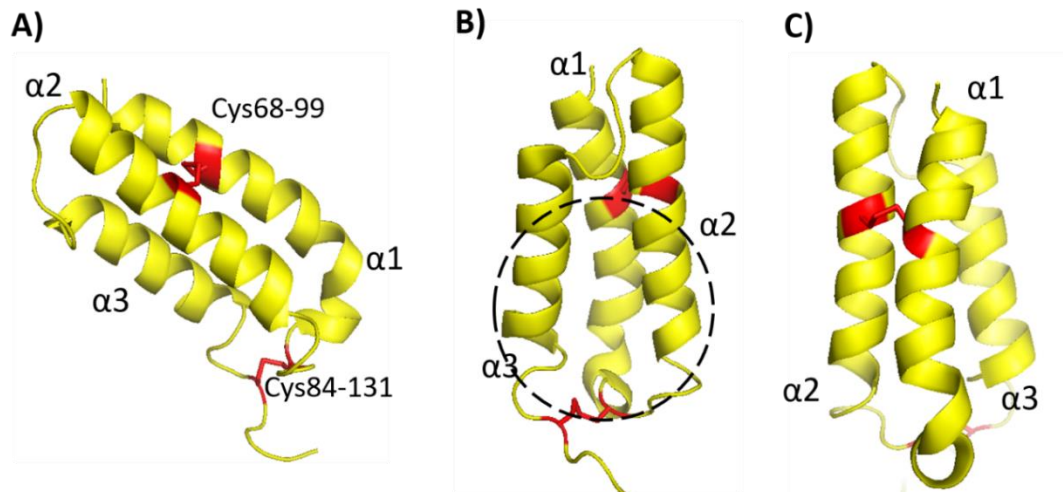


Figure 1.5: The ECD structure of the RAMP2 protein. **A)** The RAMP2 ECD (yellow) (PDB: 3AQF) comprises 3 α -helices with α 2 aligned in an anti-parallel manner. The 2 cysteine bonds within the RAMP2 ECD structure are highlighted in red. **B)** The hydrophobic path between the α 2 and α 3 helices is annotated by the black dashed ring. **C)** The reverse view of the RAMP2 structure. Image created using PyMOL Molecular Graphics System, Version 1.7.4.5- Educational Product, Schrodinger, LLC.

1.7 Predicted structure of the RAMP3 extracellular domain

The full-length structures for the CGRP, AM₁ and AM₂ receptors have been published which show all 3 receptors coupled to their G-protein subunits (Liang *et al.*, 2019)(Liang *et al.*, 2020). However, the overall electron density map of the AM₂ receptor ECD was poor, and therefore the article was focused primarily on the intermembrane segments and how the N-terminus of the peptides interact with the receptor. The specific structure of the RAMP3 ECD is therefore unclear. The basic structure however shows the RAMP3 protein adopting the same triple helices structure, with the α 2 helices being positioned in an anti-parallel manner in comparison to the α 1 and 3 helices (figure 1.7) (Liang *et al.*, 2020).

To potentially aid the structural prediction of the RAMP3 ECD, sequence overlays of the three RAMP subtypes can be utilised to predict the key structural elements of the receptor (figure 1.6). As the RAMP proteins form similar structures and the RAMP3 protein holds affinity for the CGRP and AM peptides, it is likely that RAMP3 will form a combination of the RAMP1 and RAMP2 structures which can facilitate the binding of both peptides. This means an overlay of sequences is feasible to predict the structural properties by noticing key changes and similarities between the sequences. Figure 1.6 shows the overlay of the three

amino acid (protein) sequences and figure 1.7 shows the structure of the poorly diffracted RAMP3 ECD protein from Liang *et al.*, 2020.

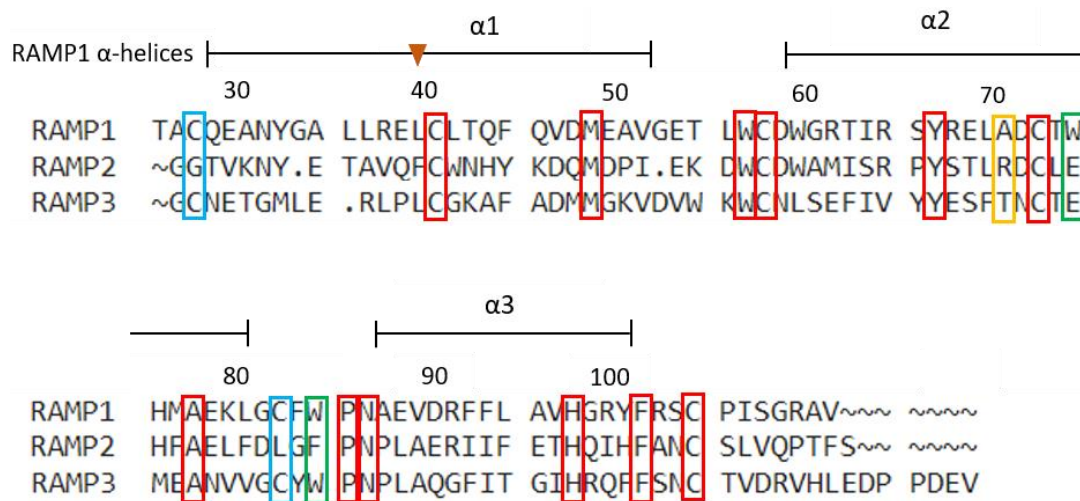


Figure 1.6: The extracellular domain amino-acid sequence overlay of the RAMP1, RAMP2 and RAMP3 proteins. Sequences are outlined in comparison to the RAMP1 alpha helices. Residues outlined in red boxes indicate a fully conserved residue. Those in blue indicate cysteine residues for disulphide bond formation conserved between RAMP1 and RAMP3 and green indicates key residues for peptide selectivity to the RAMP proteins. Residues outlined in yellow shows residues which enables antagonist selectivity for the CGRP receptor. The brown arrow indicates the position of the Leu39 kink found in the RAMP1 α1 helices. Sequences aligned on Clustal Omega, multiple sequence alignment comparison tool.

Figure 1.6 shows the sequence overlay with fully conserved amino acids between all three proteins outlined in red boxes. The most noticeable of these are 4 cysteines residues which are found in all 3 RAMP sequences. In the RAMP1 and RAMP2 receptors, the residues are responsible for forming cysteine bridges between the α1 and α2 helices (Cys40-Cys72 in RAMP1 and Cys68-Cys99 in RAMP2), and between the α1- α2 loop and the C-terminal tail (Cys57-Cys104 in RAMP1 and Cys84-Cys131 in RAMP2) (figures 1.4 and 1.5) (Kusano *et al.*, 2008, 2012). All these disulphide bonds formed for the RAMP1 and RAMP2 proteins have been shown to be important for structural integrity and peptide binding from previous mutagenesis studies (Kuwasako, Kitamura, Uemura, *et al.*, 2003; Simms *et al.*, 2006, 2009). The RAMP2 protein was found to not hold the 3rd cystine bond in comparison to RAMP1 (Cys27-Cys82) and does not have the cystine residue there to form this interaction (Kusano *et al.*, 2012). Sequence overlay however shows the cystine residues which form the 3rd cystine bridge is conserved between the RAMP1 and RAMP3 sequences, which means this additional cystine bridge is likely to form in the RAMP3 protein. This disulphide bond has

been seen to stabilise the loop between the $\alpha 2$ and $\alpha 3$ helices where key interactions between RAMP1 and the CGRP peptide are found (Booe *et al.*, 2015). A similar interaction is therefore likely to be found in RAMP3. While the published RAMP3 structure shows the formation of cystine bonds between the $\alpha 1$ -2 helices, it does not contain the amino acid sequence which would allow the disulphide bond to form between the N-terminal sequence of the RAMP3 $\alpha 1$ helices and the $\alpha 2$ -3 loop (figure 1.7). It is therefore unclear whether a bond will form in this position.

The noticeable difference between the RAMP1 and RAMP2 protein structures is the kink at the Leu39 residue of the RAMP1 $\alpha 1$ helices. As previously described, the RAMP1 $\alpha 1$ helix contains a kink at Leu39 and enables the correct positioning of residues for peptide association to the receptor complex (Booe *et al.*, 2015). The same kink is not observed in the RAMP2 ECD helices. Sequence overlay (figure 1.6) shows the leucine is conserved between the RAMP1 and RAMP3 proteins, which may result in the same kink being observed in the RAMP3 protein. The kink is likely to reposition the $\alpha 1$ helices so the disulphide bond (equivalent to Cys27-Cys82 in RAMP1) can form, as previously stated. As the $\alpha 1$ helices on the published RAMP3 protein contains a shorter sequence however, it is unclear whether this kink is formed in the RAMP3 ECD (figure 1.7) (Liang *et al.*, 2020).

In some contrast to this, a proline residue is also found in the RAMP3 $\alpha 1$ (Pro38), located 1 residue upstream from the leucine residue. This may influence structural confirmation of the protein. Due to the nature of the R-group in proline, the amine group of the amino acid cannot form a hydrogen bond and can affect structural integrity of α -helices. The proline side chain is likely to sterically interfere with the backbone of the helices, which can force a bend of $\sim 30^\circ$. This therefore is most likely to cause a disruption of the helices and reposition downstream loops and residues which aid AM and CGRP peptide binding. As mentioned with the Leu39 positioning however, the $\alpha 1$ helices in the published RAMP3 ECD uses a shorter amino acid sequence than described and therefore is unclear how this may influence the protein structure (Liang *et al.*, 2020).

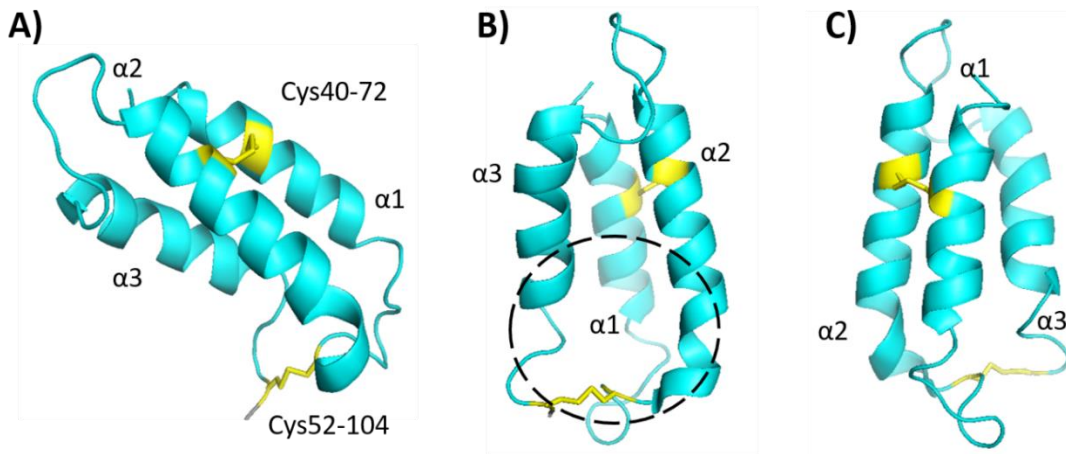


Figure 1.7: The ECD structure of the RAMP3 protein. **A)** The RAMP3 ECD (magenta) (PDB: 6E3Y) comprises 3 α -helices with α 2 aligned in an anti-parallel manner. The 2 cysteine bonds within the RAMP3 ECD structure are highlighted in yellow. **B)** The hydrophobic path between the α 2 and α 3 helices is annotated by the black dashed ring. **C)** The reverse view (behind) of the RAMP3 structure. Image created using PyMOL Molecular Graphics System, Version 1.7.4.5- Educational Product, Schrodinger, LLC.

1.8 CLR extracellular domain complex and the binding pocket

The CLR ECD and RAMP ECD have been co-crystallised to investigate the structure of the CLR receptor and the ECD binding pocket (ter Haar *et al.*, 2010) (Kusano *et al.*, 2012) (Booe *et al.*, 2015). The structure of the CLR receptor ECD is outlined in figure 1.8. The CLR receptor (PDB:3N7P) initially forms an α -helices structure from its N-terminus (residues 35-53), which is aligned roughly perpendicular to the RAMP ECD protein, specifically to the RAMP α helices 2 and 3 (adjacent to the hydrophobic pocket) forming the RAMP-CLR interface (figure 1.8) (ter Haar *et al.*, 2010). Following from the α -helices is an irregular (residues 55-64) secondary structure loop that extends into a finger-like structural motif involving residues 65-81, leading into 2 antiparallel β -sheet structures (figure 1.8). This motif has 2 disulphide bonds, Cys48-74 and Cys65-105, which are responsible for the structural integrity and stability of the protein (Watkins *et al.*, 2014). At the tip of this motif is the CLR Trp72 residue, creating the 'Trp shelf', and is key for peptide and antagonist binding to the receptor (ter Haar *et al.*, 2010)(Booe *et al.*, 2015).

With the Trp72 bulge of CLR dividing the two regions, the receptor holds a hydrophobic 'patch' which is formed from the base of CLR loop 4 to loop 3, and a binding 'pocket' which includes CLR loop 2 extending to the base of loop 4 (figure 1.8) (Booe *et al.*, 2015). The pocket is made up of CLR residues Asp70, Gly71, Trp121, Thr122, Tyr124 and Trp72, while the patch comprised CLR residues Phe92, Phe95, Tyr124 and Trp72 (Booe *et al.*, 2015).

Additional RAMP residues comprise the pocket, which include RAMP1 residues Trp84 and Pro85 or RAMP2 residues Arg97, Glu101, Glu105 and Pro112 (Booe *et al.*, 2015). Key residues that comprise the patch and pocket are summarised in table 1.4. Mutagenesis of RAMP1 Trp84 shows clear involvement of the Trp84 in peptide binding (Watkins *et al.*, 2014). RAMP2 Glu101 on the AM₁ receptor has also been seen to be vital for AM binding, with previous mutagenesis studies showing a 26-fold reduced potency of the AM peptide to the receptor (Booe *et al.*, 2015). This implies that the AM and CGRP peptides occupy the same site on the receptor binding pocket, with small changes in amino acid interactions and the RAMP protein dictating the selectivity of the peptides (Booe *et al.*, 2010).

Table 1.4: Amino-acid residues which comprise the CGRP and AM₁ receptor ECD binding pockets

Protein Receptor	Residues in Binding Pocket
CGRP Receptor	CLR Pocket: Asp70, Gly71, Trp72 Trp121, Thr122, Tyr124
	CLR Patch: Phe92, Phe95, Tyr124
	RAMP1 Pocket: Trp84 and Pro85
AM ₁ Receptor	CLR Pocket: Asp70, Gly71, Trp72 Trp121, Thr122, Tyr124
	CLR Patch: Phe92, Phe95, Tyr124
	RAMP2 Pocket: Arg97, Glu101, Glu105 and Pro112

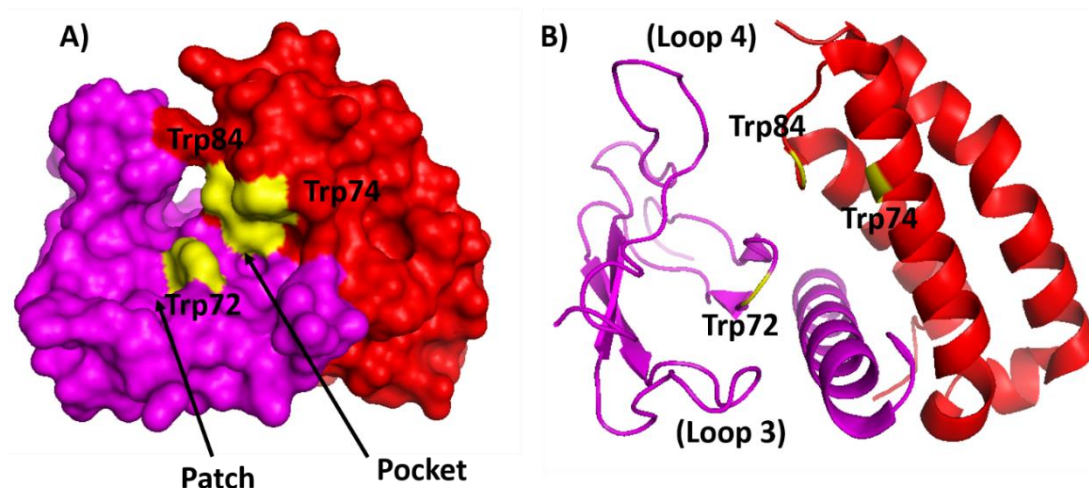


Figure 1.8: Structure of the RAMP1-CLR ECD complex. **A)** The surface representation of the ligand-free RAMP1-CLR ECD complex (PDB: 3N7P). RAMP1 ECD is coloured red and CLR ECD is coloured magenta. RAMP1 Trp84 and Trp74 and CLR Trp72 are coloured yellow throughout **B)** A ribbon diagram of the RAMP1 and CLR ECD structures. RAMP1 is positioned adjacent to the α helices of CLR. Images were generated using PyMOL Molecular Graphics System, Version 1.7.4.5-Educational Product, Schrodinger, LLC.

1.9 Electrostatic Interactions between CLR and RAMP1

There are numerous electrostatic and hydrophobic interactions that occur between the extracellular domains, predominantly between the CLR α -helices and RAMP1/RAMP2 α -helices 2 and 3 due to their tight positioning (figure 1.9). There are 2 residues in the CLR ECD that could prove to be essential for structural stability and surface expression when interacting with the RAMP1 protein. The carbonyl side group of Gln54 on the CLR α -helices interacts with RAMP1 Arg102 and Cys104 via their amine groups (ter Haar *et al.*, 2010). Noticeably, the previous Kusano *et al.*, 2008 study shows Cys104 mutagenesis on the RAMP1 ECD impairs the RAMP1-CLR dimer and reduces cell surface expression. This implies this interaction is key for the chaperone ability of the RAMP1 protein with the CLR receptor (ter Haar *et al.*, 2010).

The carbonyl group of CLR Gln50 interacts with the His97, Phe101 and Pro105 residues in the RAMP1 ECD, which corresponds with previous mutagenesis data (Kusano *et al.*, 2008)(ter Haar *et al.*, 2010). This indicates the electrostatic interaction between CLR Gln54 and RAMP1 His97 and Phe101 could be essential for cell surface expression and stability of CGRP receptor binding pocket. The mutagenesis of RAMP1 Trp84 also leads to a reduction of RAMP1-CLR cell surface expression, implying the residue has a dual role in peptide binding and the stabilisation of the RAMP1-CLR heterodimer (Moore *et al.*, 2010a). The key interactions between the RAMP1 and CLR receptor and their effect on protein integrity are summarised in table 1.5. Additional electrostatic interactions between the RAMP1 and CLR receptor occur between CLR Tyr49 and Gln45 and RAMP1 Asp90 (α -helices 3) and Tyr66 (α -helices 2) respectively (ter Haar *et al.*, 2010). CLR Tyr46 also makes two contacts with RAMP1 Trp59 and Phe101 (ter Haar *et al.*, 2010). The hydrophobic interactions and hydrogen bonding between RAMP1 and CLR are summarised in table 1.6 and a visual representation of the interface is shown in figure 1.9.

Table 1.5: Key residues involved in the RAMP1-CLR Interface

Inter-residue binding between RAMP1 and CLR	Effect After mutagenesis
CLR Gln54 to RAMP1 Arg102 and Cys104	Reduced dimer surface expression Reduced chaperone activity of RAMP1 on CLR
CLR Gln50 to RAMP1 His97, Phe101 and Pro105	Reduced surface expression Impaired CGRP binding to the RAMP1 – CLR receptor

Table 1.6: Summary of the Interactions in the RAMP1-CLR interface

CLR residue	RAMP1 residue and interaction
Arg38	Salt Bridge: Asp71
Met42	Hydrophobic: Tyr66, Ala70, Arg67
Thr43	Hydrogen: Asp90
Gln45	Hydrogen: Tyr66
Tyr46	Hydrophobic: Trp59, Phe101, Tyr66
Tyr49	Hydrophobic: Phe93, Leu94, His97
Gln50	Hydrogen: His97, Phe101, Pro105
Met53	Hydrophobic: Leu94, Gly98, Arg102
Gln54	Hydrogen Bonds: Cys104, Arg102

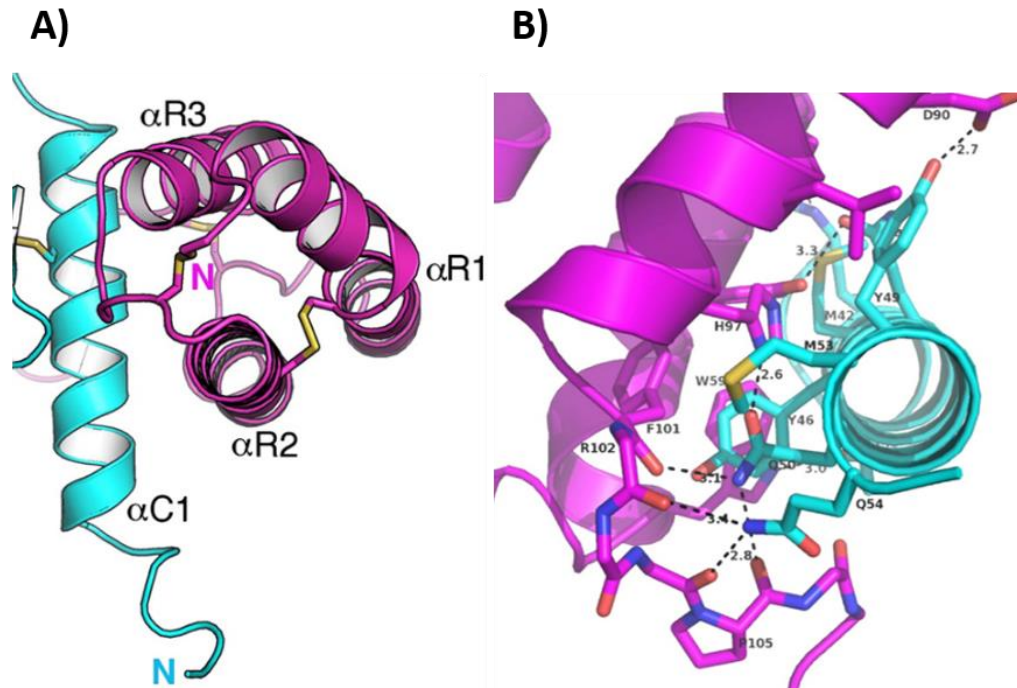


Figure 1.9: The interface between RAMP1 and CLR ECD. Key interactions between the RAMP1 ECD (magenta) and CLR receptor ECD (cyan) interface. **A)** Shows the RAMP1-CLR ECD from an aerial view. **B)** The interface between the CLR α - helices and the RAMP1 $\alpha 2$ and $\alpha 3$ helices. Hydrogen bonds are annotated using dashed lines. Image from ter Haar et al 2010 with permission from Elviseer.

1.10 Electrostatic Interaction Between CLR and RAMP2

CLR interacts with RAMP2 in a similar manner, with numerous hydrophobic interactions between RAMP2 and the CLR receptor, in-particular CLR residues Gln45, Gln50 and Gln54 (Kusano *et al.*, 2012). These are the same CLR amino-acids which associate to the RAMP1 protein, implying those amino acids are essential for all RAMP-CLR receptor subtypes and hold a similar interface. Contacts between RAMP2 and CLR are summarised in table 1.7.

Mutagenesis shows deletion of RAMP2 $\alpha 2$ residues Trp86-Pro92 significantly reduces cell surface expression (Kuwasako *et al.*, 2001)(Kusano *et al.*, 2012). Although this mutagenesis decreased cell surface expression, Kusano *et al.*, 2012 do not describe any contacts or interactions between the RAMP2 and the CLR receptor from these residues. This therefore implies these residues alter the RAMP2 structure and conformation which affects its ability to form the interface with the CLR receptor.

Mutagenesis shows that the combined alanine substitution of RAMP2 His124Ala and

His127Ala, amino acids which interact with CLR Gln48 and Gln50 (figure 1.10), reduces the chaperone ability of RAMP2 and effects AM binding to the receptor (Kuwasako *et al.*, 2008). RAMP2 His124 is located in a similar position as RAMP1 His97 (responsible for RAMP1-CLR surface expression and CGRP binding) and RAMP3 His97, but RAMP3 His97Ala mutagenesis does not have any significant effect on CLR cell surface expression but does impair AM binding (Kuwasako *et al.*, 2008). This implies that the histidine residues have differing roles in trafficking the CLR to the cell surface but are all important in the binding of CGRP or AM. It could also give an insight into the slight structural difference of the RAMP3 receptor in comparison to the RAMP1/2 receptor if this histidine residue does not interact with the CLR receptor. Key RAMP2 residues which associate to CLR are summarised in table 1.8.

Table 1. 7: Summary of the key residues involved in association of RAMP2 and CLR proteins

CLR Amino Acids	RAMP2 amino acid associations
Met42	Hydrophobic: Tyr 93, Arg97, Ser94
Gln45	Hydrogen Bond: Tyr93
Tyr46	Hydrophobic: Trp86, Phe128, Tyr93, His124
Tyr49	Hydrophobic: Ile120, Phe121, His123 Hydrogen Bond: Glu117
Gln50	Hydrogen Bonds: His124, Phe128, Ser132
Met53	Hydrophobic: Phe121, Gln125, Ala129
Gln54	Hydrogen Bonds: Cys131, Ala129
Gly71	Hydrogen Bond: Arg97

Table 1.8: Summary of the interactions which make up the RAMP2-CLR interface

RAMP2 Amino Acid residues	Effect After Mutagenesis
Trp86-Pro92	Significantly reduces CLR cell surface expression
His127 and His124	Reduces cell surface expression Effects AM binding to the AM ₁ receptor

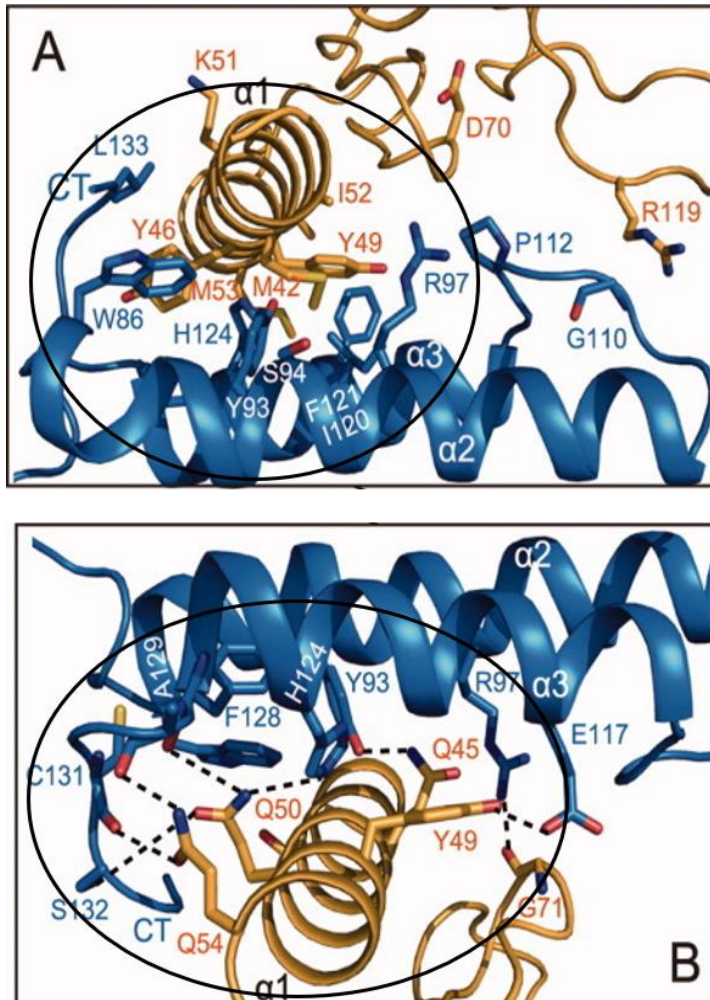


Figure 1.10: The interface between RAMP2 and the CLR ECD. Key interactions are shown between RAMP2 (blue) and the CLR (orange) which include hydrophobic interactions **(A)** and hydrogen bonds **(B)**. Key residues outlined are important for receptor trafficking which include CLR Gln45, 50 and 54, and RAMP2 Trp86-Pro92, His127 and His127 (Kusano et al., 2012). Image taken from Kusano et al 2012 with permission from John Wiley and Sons

1.11 Electrostatic Interaction Between CLR and RAMP3

Despite there currently being no developed crystal image of the AM₂ receptor ECD, predictions of amino-acid residues that make up the RAMP3-CLR interface can be speculated from previous mutagenesis data. RAMP3 residues 59-68 were investigated which is the equivalent region to RAMP2 86-92. These residues in the RAMP2 protein saw significant reduction in cell surface expression (Kuwasaki *et al.*, 2001). Upon RAMP3 59-68 mutagenesis, there was a significant reduction in AM potency and specific AM binding when co-expressed with the CLR receptor, but how much of this is due to reduced cell surface expression is unclear (Kuwasaki *et al.*, 2001). The location of these residues is likely to be at

the beginning of the RAMP3 $\alpha 2$ helices and unless there is complete repositioning of the CGRP or AM peptide during their association, they are not likely to form contacts with the peptide (Booe *et al.*, 2015).

RAMP3 His97Ala reduces potency and specific AM binding to the AM₂ receptor but is unclear whether this was due to decreased cell surface expression (Kuwasako *et al.*, 2008). The RAMP3 His97 residue however is the equivalent residue of RAMP1 His97, which was seen to form key bonds in the RAMP1-CLR ECD interface and therefore it is likely to be conserved in the AM₂ receptor ECD (ter Haar *et al.*, 2010).

RAMP3 Trp84Ala mutagenesis shows a significant decrease in cell surface expression when co-expressed with the CLR receptor, and is a conserved residue between the RAMP1 and RAMP3 proteins (Watkins *et al.*, 2014). However, the equivalent residue in RAMP2 (Phe111), does not show a decrease in cell surface expression (Moore *et al.*, 2010a). This is implying that the interface changes between each RAMP subtype and some residues may have a dual role of aiding peptide binding and stabilising of the RAMP-CLR heterodimers (Moore *et al.*, 2010a)

1.12 Comparison of non-bound CGRP and AM₁ extracellular domain structures

The ligand free structures of the CGRP receptor ECD (PDB: 3N7P) and the AM₁ receptor ECD (PDB: 3AQF) have been previously published (ter Haar *et al.*, 2010; Kusano *et al.*, 2012). While there are distinguishable differences between the RAMP1 and RAMP2 protein structures, it has not been previously stated how RAMPs may alter CLR conformation in the ECD of the receptor to aid ligand binding/selectivity. As the CLR is genetically the same between each receptor, any structural conformation differences must be due to RAMPs allosterically modulating its structure.

When overlaying the structures of the receptor proteins, small distinct differences can be observed. From the CLR α -helices, an irregular loop extends to a finger like motif (which holds the Trp shelf or Trp72 bulge) leading into 2 antiparallel β -sheet structures. These loops form key binding sites for both the CGRP and AM peptides which are discussed further in section 1.15. Structures of the finger-like motif and loop 3 of the CGRP and AM₁ ECDs (compared in figure 1.11) appear almost identical, with key amino acids identified for peptide binding by mutagenesis studies (CLR Trp72, Tyr91, Phe92, Asp94 and Phe95)

appearing to be structurally aligned (Booe *et al.*, 2015). The patch and pocket of the receptor could therefore be separated in a similar manner in all 3 receptor subtypes.

However, the structures differ at the loop 4 of the CLR receptor which could give structural aid to selectively bind different peptides between the receptors. Although the CLR loop 4 of each structure is very similar it does not align, which gives repositioning of specific amino acids, including CLR Arg119 (figure 1.11). The amino acid terminus of CLR Arg119 is positioned higher, out of the binding pocket in the AM₁ receptor ECD in comparison to the CGRP receptor ECD, where the CLR Arg119 terminus is more downward pointing into the binding pocket (towards CLR Trp72) (figure 1.11). As CLR Arg119 has been implied to be important in binding the CGRP peptide in the CGRP receptor (56-fold decrease in peptide potency after CLR Arg119 mutagenesis) and much less important for AM potency to the AM₁ receptor (4-fold decrease in peptide potency after CLR Arg119 mutagenesis), it is implying the structural positioning of this residue is aiding ligand selectivity between the 2 receptor subtypes (Booe *et al.*, 2015). This could support the concept that RAMPs allosterically modulate CLR receptors to aid peptide selectivity (Booe *et al.*, 2018). The conformation the CLR receptor adopts to contact and aid peptide binding in the AM₂ receptor is unclear, but as CLR Arg119 plays a larger role in CGRP binding, it may be likely to adopt a structure similar to that of the CGRP receptor (Booe *et al.*, 2015).

Further differences are observed at the latter end of the α -helices of the CLR receptor (figure 1.11). While this region of the receptor does not appear to be involved in any direct contact with the peptides, the region forms the RAMP-CLR interface which aids its trafficking to the cell surface, as stated in section 1.9 and 1.10 (ter Haar *et al.*, 2010; Kusano *et al.*, 2012). If this α -helices is positioned differently in relation to RAMP1 and RAMP2, it could alter the positioning of the RAMP on the receptor and therefore influence where key contact points are made in the binding pocket. While this is speculative, it is important to note the CLR protein residues 23-133 were expressed in the CGRP receptor ECD image and CLR residues 56-139 were expressed in the AM₁ receptor ECD model (ligand-free) (ter Haar *et al.*, 2010) (Kusano *et al.*, 2012). This would therefore produce a shorter helices structure and could therefore affect structural alignment between the 2 images.

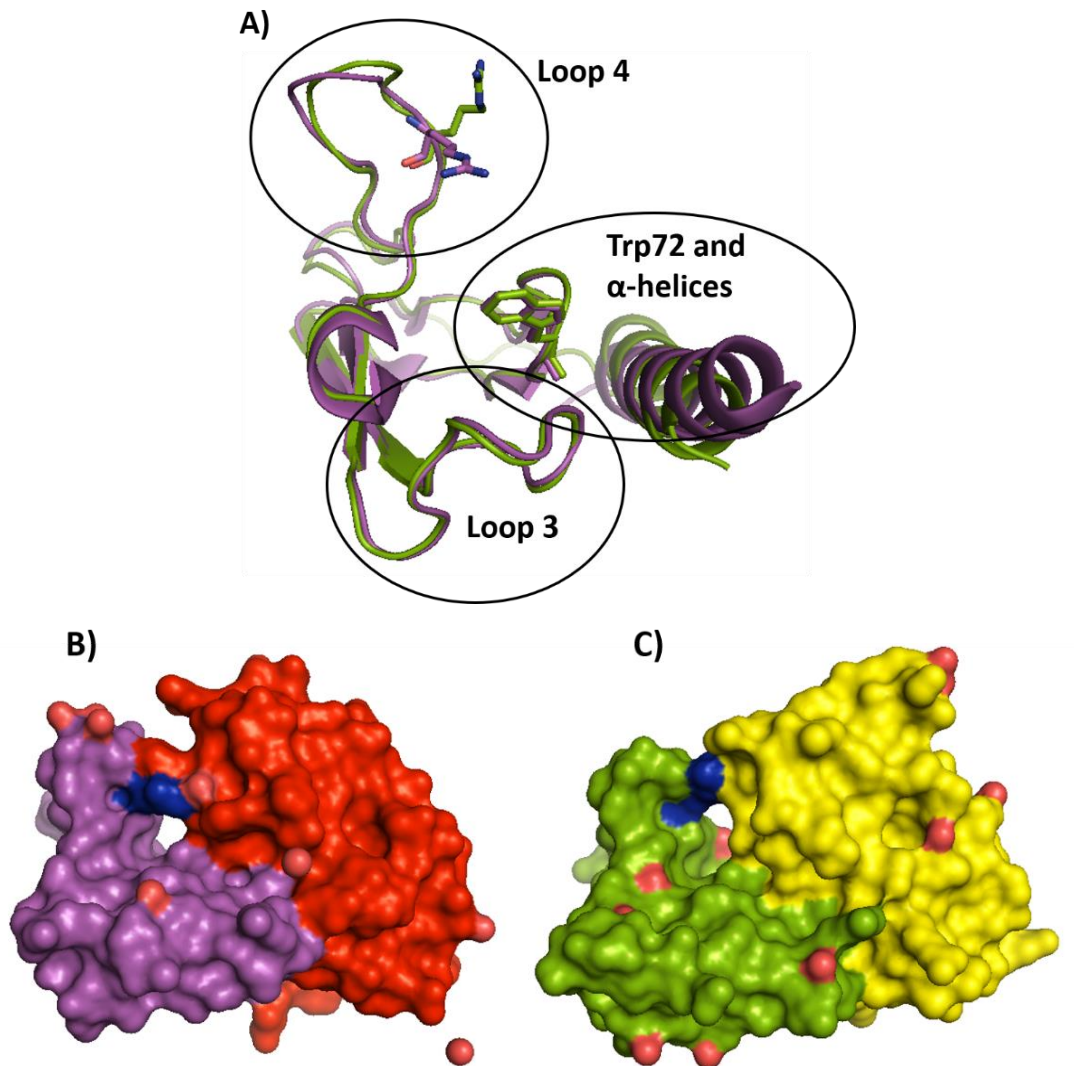


Figure 1.11: Structural comparison of the ligand free CGRP and AM₁ receptor ECDs. (A). Cartoon structural overlay of CLR protein from the CGRP receptor ECD (PDB: 3N7P) (magenta) and AM₁ receptor ECD (PDB: 3AQF) (green). Colours are consistent through the figure. CLR α -helices and the CLR loop 3/4 are annotated. CLR Trp72 and Arg119 are represented as 'stick' figures. (B) The surface representation of the CGRP receptor ECD with RAMP1 coloured red. CLR Arg119 is coloured blue. (C) The surface representation of the AM₁ receptor ECD with the RAMP2 protein coloured yellow. CLR Arg119 is coloured blue. Images were generated using PyMOL Molecular Graphics System, Version 1.7.4.5- Educational Product, Schrodinger, LLC.

1.13 CGRP peptide binding to the CGRP receptor ECD

Ligand bound structures have been generated to assess key contact points between peptides and the receptor. Booe *et al.*, 2015 were able to generate crystal models of the CGRP and AM₁ receptor ECDs which had bound CGRP₍₂₇₋₃₇₎ and AM₍₃₅₋₅₂₎ peptides. CGRP extends along the CLR patch and pocket forming connections with CLR Trp72, Phe92, Asp94,

Phe95 in loop 3 of the CLR ECD and forms further connections with His114, Arg119, Trp121, Thr122 and Tyr124 around loop 4 of the CLR ECD. These connections were confirmed with alanine scanning, resulting in a reduction in CGRP potency to the CGRP receptor (Booe *et al.*, 2015). A β -turn on CGRP allows the peptide to pack against the CLR protein and make contacts with CLR Thr122 as the C-termini of the peptide extends to form hydrophobic contacts with CLR Gln71, Trp72 and RAMP1 Trp84 via its Phe37 phenyl ring (figure 1.12) (Booe *et al.*, 2015). Further hydrogen bond contacts are made between CGRP Val32 contacting the Trp72 bulge and CGRP Thr30 forming main-chain and side-chain connections with Asp94 on CLR loop 3 (Booe *et al.*, 2015).

This data shows only 1 single contact from CGRP to RAMP1 (Trp84) which aligns with previous mutagenesis data, which states RAMP1 Trp84 is key for CGRP binding (Qi and Hay, 2010). All contacts between CGRP and the CGRP receptor are summarised in figure 1.12. Interestingly, there appears to be conformational differences in the receptor between the ligand free and ligand bound structure (figure 1.12) (Booe *et al.*, 2015). A 'clamp' like movement is observed between the CLR loop 3 and 4 upon ligand association to the receptor. The CLR loop 3 could be re-positioned due to the contact from CGRP Thr30 and CLR Asp94, and loop 4 is clamped inwards due to the β -turn of CGRP contacting CLR Ser117 via its Phe35 residue (Booe *et al.*, 2015).

Upon superimposing the ligand free (PDB: 3N7P) and ligand bound structures (PDB: 4RWG), RAMP1 Phe83 rotates away from the CLR loop 4 and CLR Arg119 shifts to form a hydrogen bond with the side chain of RAMP1 Trp84 to accommodate the CGRP peptide (figure 1.12) (Booe *et al.*, 2015). Mutagenesis studies show that the CLR Arg119 and RAMP1 Trp84 are important for peptide potency to the CGRP receptor, with a 56-fold and 30-fold decrease upon the mutations respectively (Moore *et al.*, 2010b; Booe *et al.*, 2015). This implies the hydrogen bonding between CLR Arg119 and the CGRP peptide is important for peptide binding, and the subsequent reposition and hydrogen bonding of CLR Arg119 to RAMP1 Trp84, could be important for signal transduction. This hydrogen bond formation may stabilise the shift, resulting in subsequent conformational change in the full-length receptor for receptor activation. A shift in the RAMP1 α 1-3 helices is also observed when comparing the structures, which could be an effect of the CLR Arg119 hydrogen bond and rotation of RAMP1 Phe83 (figure 1.12). Despite this shift leading to the residue movement in the helices, the RAMP1 Trp84 residue remains relatively unmoved between the 2 structures.

The hydrophobic contact made by CGRP Phe37 ring and RAMP1 Trp84 also appears to be

important for peptide binding. Mutagenesis of CLR Arg119 or RAMP1 Trp84 does not completely diminish peptide potency at the receptor (Moore *et al.*, 2010a; Booe *et al.*, 2015). CLR Arg119 appears to have no effect on the positioning on RAMP1 Trp84 and therefore if CLR Arg119 was edited via mutagenesis, the Phe37 ring of CGRP could still form hydrophobic contacts with RAMP1 Trp84. It may imply that the contact at RAMP1 Trp84 is the instigator of the CLR Arg119 shift. The contact between the CLR Asp119 and RAMP1 Trp84 may 'lock' the bound peptide to induce the signal with both interactions playing a strong role in CGRP potency to the CGRP receptor. Due to the lack of published evidence, it is unclear whether a dual mutation of CLR Arg119 and RAMP1 Trp84 would completely diminish peptide binding.

Other key amino acids include the interaction between CLR Asp94 with CGRP Thr30 residue. Mutagenesis of CLR Asp94 completely diminishes potency for the receptor implying it is vital for CGRP binding and signal transduction (Booe *et al.*, 2015). CLR Asp94 could be important for stabilising further interactions which occur beyond the receptors ECD (Liang *et al.*, 2019). This may include 'anchoring' the peptide to the surface of the receptor, to allow the extension of the peptide to form interactions with the full-length receptor. Other residues include CLR Trp121 which completely diminished CGRP binding upon its mutagenesis (Booe *et al.*, 2015). CLR Trp121 is a residue which hydrogen bonds to CGRP Gly33, implying this interaction is vital for peptide potency. This is also a region of CGRP antagonist site occupancy, implying this site occupation is key for antagonist potency on the receptor (ter Haar *et al.*, 2010).

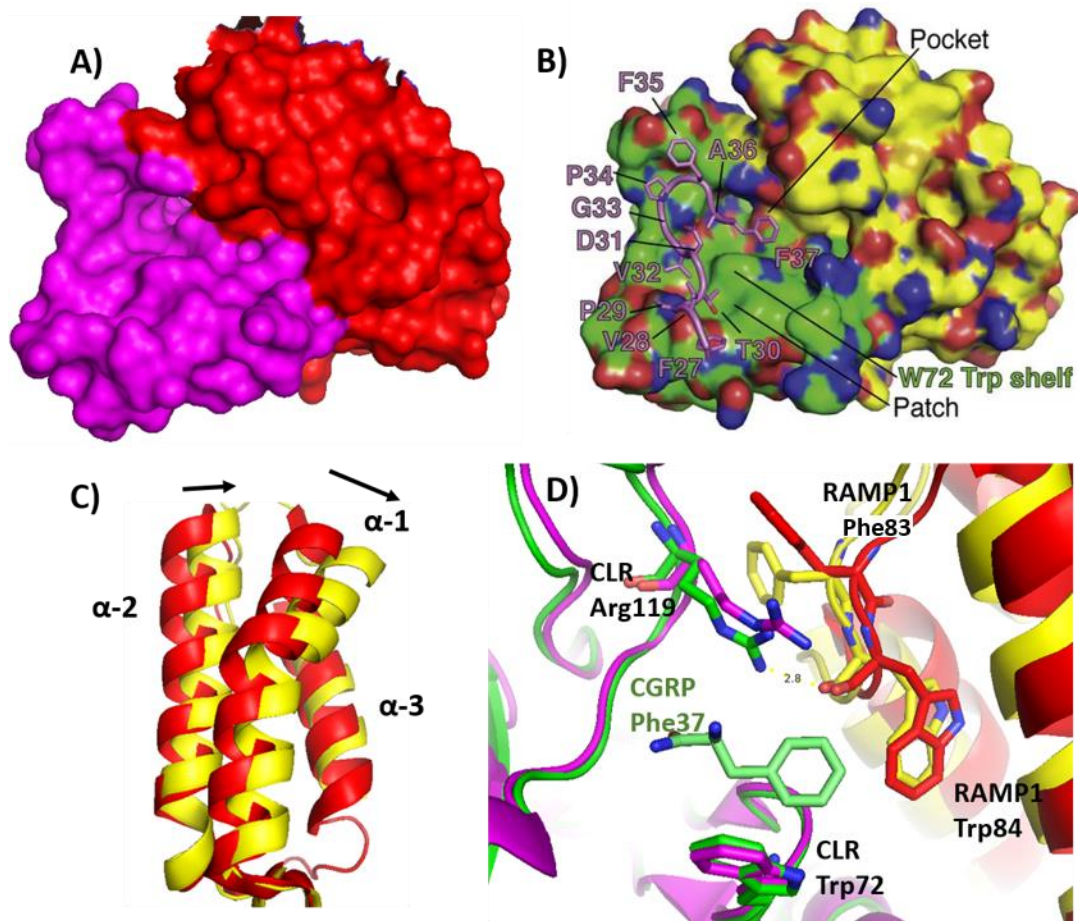


Figure 1.12: The comparison of the ligand-bound and ligand-free structures of the CGRP receptor ECD. (A) Surface representation of the ligand-free CGRP receptor ECD (PDB: 3N7R) structure. CLR is coloured magenta and the RAMP1 protein is coloured red. These colours are consistent through the figure for the ligand-free structure. (B) Surface representation of the ligand-bound CGRP receptor ECD (PDB: 4RWG). The CLR receptor is coloured green and the RAMP1 is coloured yellow. These colours are consistent for the bound CGRP receptor ECD. The CGRP peptide is coloured magenta. Image from Booe et al., 2015 with permission from Elviser (creative commons attribution license). (C) Ribbon representation of the alignment of ligand-free and ligand-bound CGRP receptor ECD structures. Arrows indicate the shift in the helices after the peptide has bound. (D) The differences in residue positioning in the binding pocket for the ligand-bound and ligand-free CGRP receptor ECD. The key residues of the RAMP1 and CLR proteins which are involved in peptide binding are annotated with the Phe37 ring of CGRP shown (light green). Outlined is the shift of the CLR Arg119 and RAMP1 Phe111 upon peptide binding. RAMP1 Trp84 and CLR Trp72, both important for peptide binding are shown. Structures aligned using PyMOL Molecular Graphics System, Version 1.7.4.5- Educational Product, Schrodinger, LLC.

1.14 AM peptide binding to the AM₁ Receptor ECD

AM is seen to contact the CLR receptor in a similar manner as CGRP. Alanine scanning and the ligand bound crystal model (PDB: 4RWF) determined CLR residues Trp72, Phe92, Phe95, Trp121 and Tyr124 were essential for AM binding, seeing a >40-fold decrease in AM binding/potency upon their mutagenesis (Booe *et al.*, 2015). Alanine scanning of CLR Asp94,

His114, Arg119 and Thr122 has less effect on AM binding, with a 4 to 9-fold decrease in AM potency (Booe *et al.*, 2015). This contrasts with CGRP binding to the CLR receptor, which saw a ≥ 20 -fold decrease in peptide potency on all the residues but noticeably at CLR Asp94, which completely diminished CGRP potency (Booe *et al.*, 2015).

The β turn on the AM peptide allows AM Tyr52 and Lys46 residues to contact Arg97, Glu101 and Glu105 on the $\alpha 2$ of RAMP2, and the extension of a single helical turn allows AM Lys64 to contact the CLR Trp72 bulge (figure 1.13). This in turn allows AM to pack against the Trp shelf via its Y52 residue in an edge to face manner (figure 1.13) and enables contacts from AM Phe43 and Ala42 to the CLR patch. This corresponds partly with previous mutagenesis data which shows evidence that RAMP2 residues Glu101 and Phe111 are key for AM binding (Watkins *et al.*, 2014)(Qi *et al.*, 2011)(Moore *et al.*, 2010a). While Phe111 doesn't directly interact with AM, it is positioned closely to Glu101 and Glu105, implying it could be important in implicating an electrostatic charge, affecting residual binding to AM or is important for structural stability (figure 1.13). A summary of the key interactions are summarised in table 1.9. Despite AM holding contacts with more residues with RAMP2 than CGRP with RAMP1, there is no helices shift like the one observed in the RAMP1 structure (figure 1.13)

In contrast to CGRP association to the receptor, CLR Trp72 plays a much larger role in AM association to the AM₁ receptor, observing a 126-fold decrease upon its mutagenesis in comparison to the 44-fold difference observed on the CGRP receptor (Booe *et al.*, 2015). This directly correlated to the structural data, which shows increased hydrophobic interactions made between AM and CLR Trp72, resulting in a larger loss of potency upon its mutagenesis (Booe *et al.*, 2015). When superimposing the bound and non-bound AM₁ structures, there appears to be the slight rotation of the CLR Trp72 residue (the Trp shelf). This could be due to the contacts made at CLR Trp72, which could enable the better edge-to-face stacking off the AM Tyr52 ring (figure 1.13).

In addition to this, there are less conformational differences between the ligand bound and ligand free images in comparison to the CGRP receptor. Structural alignment shows minimal differences with the exception of CLR loop 3 (figure 1.12). While there are no major shifts in key contact amino acids in that loop (CLR Phe95 or Phe92) there is the rotation (away from the binding pocket) of CLR Glu90, which appears to give way for the AM peptide and allow better access to the binding pocket (figure 1.13). There were no reported interactions between CLR Glu90 and the AM peptide in the AM₁ ECD receptor (Booe *et al.*, 2015). There

also appears to be a slight rotation of RAMP2 Glu105, which could allow the contacts between AM Lys46 via hydrogen bonding.

Interestingly, there appears to be a complete 180° turn of RAMP2 His102 (α 2 helices) when the ligand is bound to the receptor (figure 1.13). While this is a significant change in residue positioning, it is unclear why this rotation occurs with no observed contacts with the AM peptide and mutagenesis shows it has no effect on cell surface expression, AM potency or specific peptide binding (Kuwasako *et al.*, 2008). As the residue is located in close proximity of RAMP2 Glu101, a key binding residue of the AM peptide, it could be rotating to give way and enable binding of the AM peptide (Booe *et al.*, 2015). Key interactions between CGRP and AM peptides to their requisite receptors are summarised in table 1.9.

Table 1.9: Key interactions between the CGRP/AM peptides and their receptors

Peptide	Key RAMP and CLR residues for peptide binding
Calcitonin-Gene Related Peptide (CGRP)	CLR- residues Trp72, Phe92, Asp94, Phe95, His114, Arg119, Trp121, Thr122 and Tyr124 RAMP1- residues Trp84
Adrenomedullin (AM)	CLR- (essential residues) - Trp72, Phe92, Phe95, Trp121 and Tyr124 CLR- (non-essential) – Asp94, His114, Arg119 and residues Thr122 RAMP2 - (essential residues) – Glu101, Phe111 RAMP2 - (nonessential residues) – Arg97, Glu105

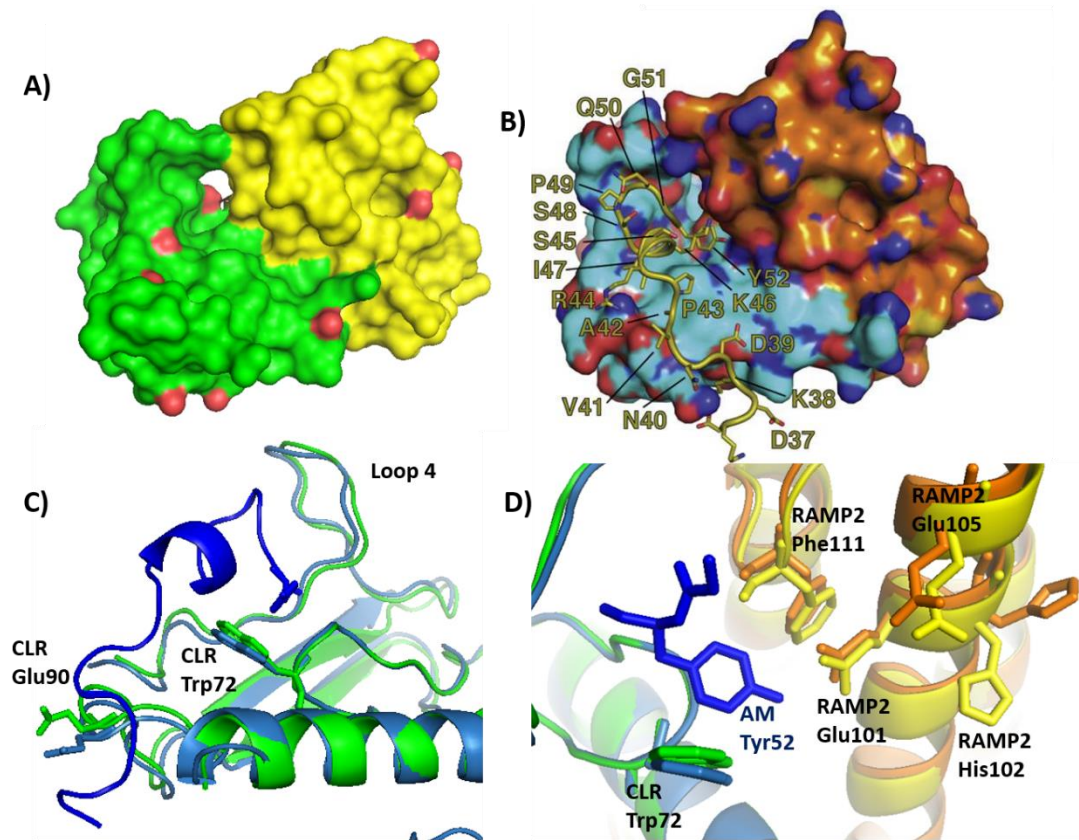


Figure 1.13: Structural comparison of the ligand-bound and ligand free AM₁ receptor ECD. (A) Surface representation of the ligand-free AM₁ receptor ECD (PDB: 3AQF). The RAMP2 protein is coloured yellow and CLR is coloured green. These colours are consistent through the figure for the ligand-free receptor. (B) The surface representation of the bound AM₁ receptor ECD (PDB: 4RWF). The CLR protein is coloured light blue and the RAMP2 protein coloured orange. These colours are consistent through the figure for the ligand-bound structure. The AM peptide is coloured yellow. Image from Booe et al., 2015 with permission from Elvise (creative commons attribution license). (C) Ribbon structure alignment of the ligand-free and ligand bound CLR receptor ECD structure. The AM peptide is coloured dark blue. Outlined is the positioning of the CLR loop 4 and Trp72 which form key contact points with the AM peptide. CLR Glu90 is shown which appears to shift upon peptide binding. (D) Structural alignment of the binding pocket between the ligand-free and ligand-bound AM₁ receptor ECD. Outlined are the residues important for AM binding to the receptor and their positioning in the binding pocket. Structures aligned using PyMOL Molecular Graphics System, Version 1.7.4.5- Educational Product, Schrodinger, LLC.

1.15 Residues Enabling Ligand Specificity for the CGRP and AM₁ Receptors

The key difference between CGRP and AM selectivity on their requisite receptors is the positioning and size of key residues within the receptor binding pocket, particularly those from the RAMP protein. The Phe37 phenyl ring of CGRP forms hydrophobic connections with RAMP1 Trp84 but the equivalent RAMP2 residue, Phe111, is smaller and despite structural similarities between RAMP1 and RAMP2, the contact between RAMP2 Phe111 and the

Phe37 ring of CGRP would be lost (figure 1.14) (Booe *et al.*, 2015). This is due to the fact phenylalanine contains only one benzyl ring in its structure whereas tryptophan contains an indole group (six-membered benzene ring fused to a five-membered pyrrole ring). This means RAMP1 Trp84 will be extending further into the binding pocket than RAMP2 Phe111, forming contacts with CGRP Phe37 (figure 1.14).

A further example is the bond between AM Lys46 and Tyr52 to RAMP2 Glu101. AM Lys46 and Tyr52 would not be able to contact the equivalent residue in RAMP1 (Trp74), giving AM selectivity for RAMP2 over RAMP1. This aligns with previous mutagenesis data which shows RAMP1 Trp74Glu has no effect on CGRP potency but an increase in AM potency for the CGRP receptor, stating the importance in the residues role in ligand selectivity (Qi *et al.*, 2008)(Qi *et al.*, 2011).

Despite the bonding not being significantly important for AM potency on the AM₁ receptor, AM Try52 contacts RAMP2 Arg97 via hydrogen bonding. The equivalent residue in RAMP1 (Ala70), would not be able to form these contacts (figure 1.14). Additional to this, RAMP2 Glu105 forms ionic interactions with AM Lys46 and the glutamic acid residue is conserved in the RAMP1 protein (RAMP1 Glu78). However, structure overlay shows a repositioning of the residue in RAMP1 and therefore AM Lys46 would not be able to form these contacts in RAMP1 (figure 1.14).

Further differences may include the points of contact between peptides and CLR Arg119 (figure 1.14). It has previously been implicated that RAMP proteins are allosterically modulating the CLR receptor, which may alter their conformation and aid ligand selectivity between the receptors (Booe *et al.*, 2018)(Liang *et al.*, 2020). The overlay of the CGRP peptide on the non-bound AM₁ receptor shows that the peptide could not form contacts with the CLR Arg119 in the AM₁ receptor, reducing peptide potency on the receptor (figure 1.14). The positioning of CLR Arg119 on the AM₁ receptor would also sterically clash with RAMP1 Phe83 which is located in the α 2-3 loop of the RAMP1 protein (figure 1.14). This may therefore suggest that RAMP1 Phe83 is responsible for the repositioning of the CLR Arg119, which allows contacts with the CGRP peptide (figure 1.14). RAMP2 holds a glycine residue in the equivalent position as RAMP1 Phe83, which is smaller and therefore would not interfere with CLR Arg119 (figure 1.14). CLR Arg119 on the AM₁ receptor is not in the correct position to form contacts with CGRP or hydrogen bond with RAMP1 Trp84 (Booe *et al.*, 2015). The combined changes of the RAMP1 Trp84 to RAMP2 Phe111, RAMP1 Trp74 to RAMP2 Glu101 and the positioning of CLR Arg119 give selectivity for the peptides on their requisite

receptors (Booe *et al.*, 2015).

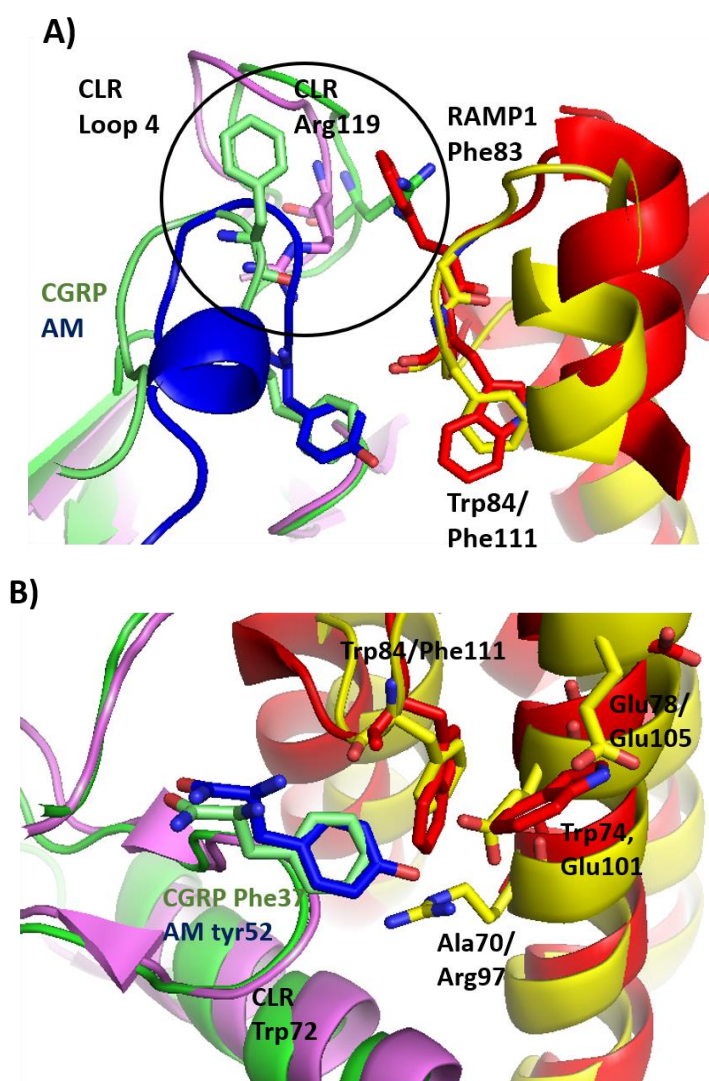


Figure 1.14: Overlay of the CGRP and AM peptides on the ligand-free CGRP and AM₁ receptor ECD. (A) Ribbon representation/overlay of the ligand free CGRP (PDB: 3n7p) and AM₁ (PDB: 3aqf) receptors with peptides. Shows the residue positioning differences at CLR loop 4 and the α 2-3 loop on the RAMP1/2 proteins. RAMP1 is coloured red and RAMP2 is coloured yellow. The CLR of the CGRP receptor ECD is coloured magenta and CLR of the AM₁ receptor ECD is coloured green. The CGRP peptide is coloured pale green and the AM peptide blue. These colours are consistent through the figure. The positioning of CLR Arg119 at the loop 4 and how it forms contacts for CGRP but not AM is outlined. It is further shown how CLR Arg119 on the AM₁ receptor ECD would clash with RAMP1 Phe83. **(B)** Outlays the binding pocket of the 2 receptors and the differing positioning of key residues enable ligand selectivity. Structures aligned using PyMOL Molecular Graphics System, Version 1.7.4.5- Educational Product, Schrodinger, LLC.

1.16 Predicted peptide binding on the AM₂ receptor extracellular domain

The extracellular domain of the published AM₂ structure was poorly established, and while

the image had the AM peptide bound to the structure, it was difficult to investigate the positioning of key amino acids which aid ligand binding (Liang *et al.*, 2020). It is therefore problematic to use this model to investigate how the peptide is interacting with the receptors N-terminus. The AM peptide appears to be adopting a similar position in the binding pocket but the interactions between itself and the receptor are unclear (Liang *et al.*, 2020). To overcome this, it may be more suitable to use a homology model to predict the key binding sites between the AM and CGRP peptides to the AM₂ receptors ECD.

A homology model was generated by inducing mutagenesis in the CGRP receptor ECD structure (figure 1.15) to the key peptide binding residues of the RAMP3 protein which included RAMP3 Glu74 (equivalent to RAMP1 Trp74). It is predicted that RAMP3 Glu74, the equivalent residue of RAMP2 Glu101, will interact via hydrogen bonding with AM Tyr52 (Booe *et al.*, 2015). This could be confirmed with the mutagenesis of Glu74Trp of RAMP3 showing a reduced potency for AM and no change in CGRP potency (Qi and Hay, 2010). As RAMP3 Glu74 is located in the same position as RAMP1 Trp74, it corresponds with data that shows CGRP is not binding to RAMP1 Trp74 (or RAMP3 Glu74) and holds a similar site of occupancy across both of the receptors.

It may also give a further insight into how RAMP3 forms the receptor binding pocket, as RAMP3 Glu74 is in the equivalent position to RAMP2 Glu101, a residue which sits deeper into the AM₁ binding pocket (in comparison to RAMP1 Trp74) and is key for AM binding to the AM₁ receptor ECD (Booe *et al.*, 2015). This implies that the RAMP3-CLR receptor may have a binding pocket which is structurally similar to the AM₁ receptor (Watkins *et al.*, 2014).

RAMP3 Trp84 which is the equivalent of RAMP1 Trp84, will contact the CGRP Phe37 phenyl ring and AM Tyr52 residues, both of which are seen to be key in peptide binding to the receptors (Booe *et al.*, 2015). This could explain previous alanine scanning data which shows diminished potency of both peptides on the AM₂ receptor upon RAMP3 Trp84Ala mutagenesis (Watkins *et al.*, 2014). It would further evaluate why CGRP has a much higher affinity for the AM₂ receptor in comparison to the AM₁ receptor, and also why the AM₂ receptor has preferred binding to AM due to more contact points. Despite this, no distinguishable AM₂ receptor ECD image has yet been developed and these interactions are yet to be determined.

As seen in the CGRP and AM₁ receptor structures, RAMP1 Phe83 is seen to rotate away which allows CLR Arg119 to make way for the ligand in the CGRP receptor binding pocket. It was also noted that this RAMP1 Phe83 could be responsible for the repositioning of CLR Arg119

to enable bonding between itself and the CGRP peptide (Booe *et al.*, 2015). It is yet to be determined if this mechanism is prevalent in the AM₂ receptor, but sequence comparison shows that RAMP3 holds a tyrosine residue in place of the RAMP1 Phe83. Tyrosine is structurally similar to phenylalanine but holds an additional hydroxyl group on the benzyl ring (phenol group). The residue therefore has the ability to act in the same manner as RAMP1 Phe83 but whether the hydroxyl group of the RAMP3 Tyr83 holds any additional contacts is unclear. It could therefore be predicted that the CLR loop 4 on the AM₂ receptor ECD holds a similar conformality as the CLR loop 4 on the CGRP receptor ECD. The surface representation of the CGRP, AM₁ and AM₂ receptors are shown in figure 1.15.

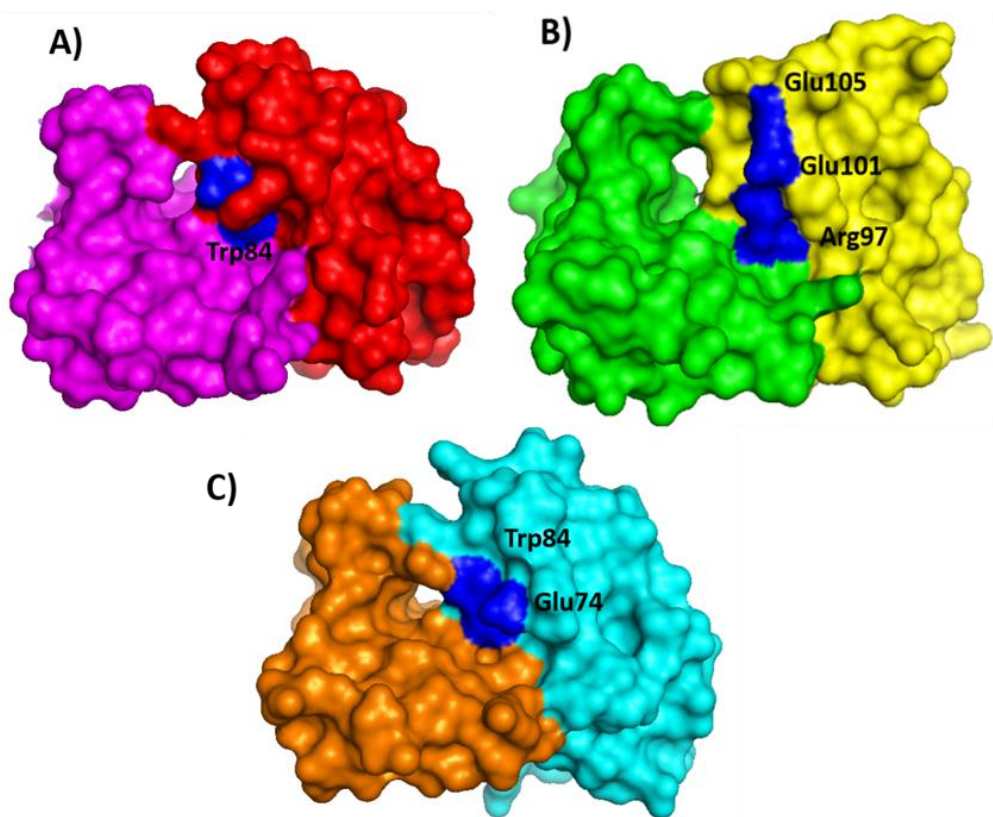


Figure 1.15: Surface view comparison of the CGRP, AM₁ and AM₂ receptor ECD. A) The surface of the ligand-free CGRP receptor (PDB: 3N7P) with RAMP1 coloured red and the CLR receptor coloured magenta. Residues which are key for peptide binding on the RAMP1 receptor are highlighted in blue (Trp84) B) The surface representation of the ligand-free AM₁ receptor (PDB: 3AQF) with the CLR coloured green and RAMP2 coloured yellow. Residues important for AM association to the receptor are highlighted (blue). C) The surface representation of the AM₂ model with positioning of suspected amino acids highlighted (blue). The model was generated by inducing mutagenesis on key residues of the ligand free CGRP receptor image (PDB: 3N7P). The CLR receptor is coloured orange and the RAMP3 protein coloured cyan. The RAMP3 Glu74, which is in the equivalent position of RAMP1 Trp74 and RAMP2 Glu101 that may have a role in peptide binding, is highlighted (blue). The RAMP3 Trp84 residue is also highlighted (blue). Images generated using PyMOL Molecular Graphics System, Version 1.7.4.5- Educational Product, Schrodinger, LLC.

1.17 Project Rationale, aims and objectives

Our lab group is currently committed to the development of CGRP and AM₂ receptor selective antagonists designed to treat migraine pain and hinder cancer progression (primarily pancreatic cancer). As there is currently no defined structure for the AM₂ receptors ECD, designing compounds and fragments which selectively antagonise the CGRP and AM₂ receptors has been difficult. Despite this, structural data from ligand free, ligand bound and compound bound CGRP/AM₁ receptors was utilised to generate lead novel compounds and small molecules which show high selectivity to the N-terminus of the CGRP and AM₂ receptors (Avgoustou *et al.*, 2020). However, their binding site occupancy for both CGRP and AM₂ receptors remain unknown. The principle aims of this study therefore, is to develop a functional RAMP-CLR ECD fusion protein for structural analysis and to investigate the binding pocket occupancy of the novel compounds. This will involve testing novel CGRP and AM₂ receptor antagonists to predict binding pocket occupancy and protein crystallisation to determine the ECD structure of the AM₂ receptor.

Hypothesis:

- Functional RAMP-CLR fusion proteins can be expressed and crystallised to outline the binding pocket occupancy of novel antagonists on the CGRP and AM₂ receptor ECDs.
- The structure and specific positioning of amino-acid residues of the receptor activity modifying proteins 1, 2 and 3 will be detrimental to antagonist selectivity on the CGRP, AM₁ and AM₂ receptor ECDs.

Objective 1: Develop RAMP-CLR ECD fusion protein constructs and express in mammalian cell culture media.

Objective 2: Purify large quantities of the RAMP-CLR fusion proteins using affinity and/or size exclusion chromatography.

Objective 3: Test the functionality of the RAMP-CLR fusion proteins and investigate the binding affinities of previously developed CGRP antagonists and novel CGRP and AM₂ selective compounds.

Objective 4: Screen fragments designed to target the RAMP or CLR proteins by utilising the RAMP-CLR fusion proteins. High affinity fragments should be developed into compounds

which could selectively antagonise the CGRP and AM₂ receptors. A cAMP assay will be developed to determine how effectively these generated compounds antagonise the receptors.

Objective 5: Crystallise the RAMP-CLR fusion proteins with/without compounds to investigate their binding pocket occupancy and determine the N-terminus structure of the AM₂ receptor.

Chapter 2: Construct Development and Protein Expression

2.1 Introduction

Recombinant protein technology is a progressive area of research which can provide an insight into receptor binding specifics and molecule interactions. The process involves the expression and purification of a recombinant protein which can be utilised to aid drug development and monitor compound/ligand interactions. To investigate the binding properties of molecules and compounds against their target receptors, significant amounts of pure, correctly folded, and functional protein is required. GPCRs are a notoriously difficult receptor family to express and purify due to their highly hydrophobic transmembrane domains leading to protein instability and aggregation when extracted/expressed out of the cell membrane. Designing the correct protein construct along with choosing the most appropriate host expression system is therefore an essential step to express a functional recombinant protein.

2.1.1 Prokaryote Protein Expression System

Bacteria is a commonly chosen expression system for large scale recombinant protein expression and purification of proteins. Using bacteria, specifically *E. coli*, as an expression host gives well known advantages in comparison to other systems. Fast growth kinetics, cheap culture media and ease of vector introduction to the bacteria enables high density cultures to be formed in a short period of time, with correct salts temperature/environment shortening doubling times to ~20 minutes (Sezonov and Ari, 2007).

When mammalian GPCRs are expressed in *E. coli*, they can often be expressed functionally in the bacterial inner cell membrane when fused with endogenously expressed bacterial proteins. Examples include β -galactosidase or the maltose binding protein (MBP) which have been proven to efficiently traffic proteins into the bacterial cell membrane (Wei and Grisshammer, 2002). If the protein is non-functional when expressed in the inner membrane, it may be expressed in inclusion bodies. Inclusion bodies are enclosed capsules of stable, aggregated proteins formed typically when recombinant proteins are expressed in large amounts and have been previously utilised during early purification studies when expressing RAMP/CLR proteins in *E. coli* (Moad and Pioszak, 2013). The formation of inclusion bodies can aid recombinant protein purification steps, with the encapsulation of predominantly the recombinant protein only, reducing non-specific capture (Moad and Pioszak, 2013). Further to this, inclusion bodies are resistant to proteolytic digestion which

can prevent potential cell toxicity and recombinant protein breakdown (Milic and Veprintsev, 2015). While this may be advantageous, proteins retained in inclusion bodies are often misfolded and inactive aggregates. The proteins therefore must be solubilised and folded into their native states along with potential refolding to resolve any misplaced disulphide bonds. This can lead to a heterogenous mix of proteins which can be difficult to separate from each other, affecting downstream applications (Hill and Pioszak, 2013).

A further disadvantage of using *E. coli* as an expression host is their inability to form proper disulphide bridges. As bacteria do not possess an endoplasmic reticulum (ER), they cannot form disulphide bonds between cysteine residues which are vital for structural robustness and function in mammalian proteins, particularly those that are membrane bound. A prime example are relevant mutagenesis studies that edit cysteine residues in RAMP1 and RAMP2 ECDs which form disulphide bonds in their final structures. These mutations resulted in decreased peptide specificity, reduced potency of CGRP and AM peptides and decreased cell surface expression in comparison to the unmodified wild type protein (Kuwasako, Kitamura, Nagoshi, *et al.*, 2003; Kuwasako, Kitamura, Uemura, *et al.*, 2003; Simms *et al.*, 2006, 2009 and Qi and Hay, 2010).

In an attempt to combat this, *E. coli* strains have been developed which have a *gor trxB* mutation (origami strain) to create a more oxidative cytoplasmic environment. This origami strain is a double strain mutation which effects the thioredoxin/glutaredoxin reductase pathway. Thioredoxin and glutaredoxin are both proteins which act as antioxidants and facilitate the reduction of proteins by cysteine thiol-disulphide exchange and their pathway removal would result in a more oxidative environment. This environment would be more compatible with the formation of correct cystine-cystine disulphide bridges (Marco, 2009)(Seras-franzoso *et al.*, 2012). A plasmid encoding disulphide bond C isomerase (DsbC) to form correct cysteine bonds may be used with the origami strain of *E. coli*. DsbC may not necessarily be required for disulphide bond formation but acts as an isomerase, reconfiguring the bonds if they are mis-positioned, commonly from those formed between adjacent cysteines in the amino-acid sequence (Gleiter and Bardwell, 2008). This can however lead to a heterogenous mix of misfolded proteins if these disulphide bonds are misplaced. With no 'protein translational check' in the golgi apparatus, this can often occur, as seen in early RAMP-CLR fusion protein expression tests (Hill and Pioszak, 2013).

2.1.2 Glycan Addition to Membrane Proteins

Protein glycosylation is a post-translational modification which involves the attachment of carbohydrates to proteins in a non-template based manner (Chandler and Costello, 2017). In the lumen of the ER, glycans can be N-linked, which attach to the nitrogen group of asparagine amino-acid side chains. The asparagine C-terminus must be located next to any amino acid (except proline), followed by a serine or threonine amino-acid (consensus: N-X-S/T). Glycans can be O-linked by attaching to hydroxyl groups on serine or threonine side chains in the golgi apparatus. In contrast to *N*-glycans, there is currently no identified consensus sequence for the addition of O-glycans and are thought to have lesser importance in protein functionality and ligand binding. The addition of O-linked glycans is thought to have a greater role in protein recognition by immune cells and protein trafficking (Ellies *et al.*, 1998).

N-glycans are responsible for increasing protein solubility and recruiting two lectin-like chaperones in the ER (calnexin and calreticulin) which retain the protein in the ER until proper protein folding is achieved (Varki *et al.*, 2009). This process can therefore be vital for receptor functionality and ligand binding. *E. coli* do not possess the ability to add glycans to proteins due to the lack of an ER, which can lead to non-functional and incorrectly folded proteins, particularly membrane bound GPCR receptors (Hill and Pioszak, 2013) (Booe *et al.*, 2015). A key example is a study by Booe *et al.*, 2015 who generated functional CGRP and AM₁ receptor proteins but could not express a functional AM₂ receptor fusion protein in *E. coli*. A functional AM₂ receptor could only be achieved by expressing the construct in mammalian cells (HEK 293) to gain these post-translational modifications, including protein glycosylation (Hill and Pioszak, 2013) (Booe *et al.*, 2015) (Roehrkasse *et al.*, 2018)

Advancements are currently being made in bacterial expression systems to add glycans to proteins. *Campylobacter jejuni* is a known bacterium which has an N-linked glycosylation pathway and can produce functional N-glycan linked proteins (Wacker *et al.*, 2002). However, the applicability and the assessment of its use to produce correctly folded mammalian proteins and its similarity to the mammalian glycosylation pathway is yet to be fully assessed, with bacterial N-glycan addition differing from their eukaryotic counterpart (Wacker *et al.*, 2002). Although it generated the potential to transfer this system to more the commonly used *E. coli* host, it has yet to be fully established.

2.1.3 Mammalian Protein Expression System

Mammalian cells are being progressively utilised over bacteria for recombinant protein expression and purification, particularly of GPCRs. Mammalian cells can express a correctly folded recombinant protein with precise post translational modifications and protein assembly. Chinese hamster ovary cells (CHO-K1) and human embryonic kidney (HEK293) cells are two commonly used mammalian cell lines for recombinant protein production and purification due to their high protein expression characteristics (Hunter *et al.*, 2018). HEK 293 cells have been previously utilised for expressing and purifying RAMP1-CLR, RAMP2-CLR and RAMP3-CLR fusion proteins which retained affinity for their requisite peptides (Roehrkasse *et al.*, 2018).

Despite these advantages, culturing mammalian cells is less cost effective in comparison to bacterial cell culture, due to costlier cell culture media and expensive reagents required to introduce vector DNA into the cells. Mammalian cells are often more sensitive to environmental factors, including pH and temperature changes which can negatively impact protein expression (Wulhfard *et al.*, 2008)(Lin *et al.*, 2015). However, in some cases, inducing mild hypothermia during mammalian cell culture can have a positive impact on recombinant protein expression. HEK293 cells have been cultured at 30°C during the expression of a RAMP3-CLR fusion protein, which has a positive impact on the recombinant protein expression (Roehrkasse *et al.*, 2018). Although the mechanism for the increase in protein expression under mild hypothermia is not fully understood, it is thought to increase mRNA stability by promoting the accumulation of cells in the G₀/G₁ phase of the cell cycle as well as altering cell metabolism and gene expression (Baik *et al.*, 2006)(Fox *et al.*, 2005). Mammalian cells also require a constant supply of CO₂ which makes their culture much more expensive in comparison to *E. coli*.

Although proteins are more likely to be correctly folded, protein expression and subsequent yield after purification is much lower in mammalian cells in comparison to bacterial expression. This means if a large amount of purified protein is required for structural studies, bacterial expression may be used if a functional protein can be generated.

2.1.4 Addition of Protein Tags

2.1.4.1 MBP Tag

Along with the expression system, protein tags used in recombinant protein production are important for expression, yield and purification. The maltose binding protein (MBP) is a 42.5 kDa periplasmic protein which is part of the maltodextrin pathway in *E. coli*. The protein is often used during large scale protein expression studies in eukaryotes, with first reports dating back to 1998 when the human T cell leukaemia virus type 1 GP21 protein was linked to the MBP protein for crystallisation (Howlett, Kemp and Pountourios, 1998). Other successful crystallography attempts with proteins tagged to an MBP protein have since been described, including those which have expressed functional RAMP-CLR fusion proteins (Booe *et al.*, 2015)(Pioszak and Xu, 2008)(Patrick *et al.*, 2013)(Pan, Nakatsu and Kato, 2013). The MBP protein is a popular choice of protein tag as it is evidenced to increase protein turnover by accelerating the folding process through the endoplasmic reticulum in comparison to other fusion protein tags, although the mechanism for this is still unclear (Reuten *et al.*, 2016). Due to its large hydrophilic surface, the MBP protein has been evidenced to dramatically increase the solubility of its fusion partner in comparison to other protein tags (Kapust and Waugh, 1999). This is vital during protein expression studies to prevent aggregation of the recombinant protein, which can become problematic when expressing highly hydrophobic GPCR receptors. The tag has also been incorporated into previous RAMP-CLR fusion protein purification steps as a secondary purification method, with the MBP protein having a high affinity for amylose (Hill and Pioszak, 2013)(Booe *et al.*, 2015).

2.1.4.2 GST Tag

Glutathione S-transferase (GST) is an alternative protein tag to the MBP protein. It is a 26 kDa metabolic isozyme from parasitic helminth, *Schistosoma japonicum*, involved in the conjugation of reduced glutathione to xenobiotic substrates. GST can be easily incorporated into fusion protein purification steps (column chromatography) due to its high affinity for glutathione and simply eluted using 10 mM reduced glutathione. With the tag being smaller than the MBP, it could be advantageous as it may not interfere with downstream protein:protein or protein:peptide interaction assays, meaning it's more likely that the fusion protein retains proper functionality. However, GST has proven to be more

problematic than the MBP protein. When tagged with larger proteins, specifically those that are highly hydrophobic or larger than 100 kDa, they have the tendency to form large insoluble aggregates and inclusion bodies (Bell *et al.*, 2013). To combat this, the detergents Triton-X and CHAPS have been incorporated into the purification steps but proteins then require refolding which can give an incorrectly folded protein and render them inactive (Deceglie *et al.*, 2012).

A further issue is that the GST protein has the tendency to dimerise. As the native protein exists as a homodimer, the protein can form large complexes which bind very strongly to the glutathione resin during purification, resulting in difficulties in elution (Maru *et al.*, 1996). Addition of the strong reducing agent, dithiothreitol, has been used to prevent dimerization but this can interfere with the fusion protein structure, which could compromise future work if the research is investigating the structural properties of the protein (Maru *et al.*, 1996)

2.1.4.3 Histidine Tag

A polyhistidine affinity tag typically consists of 6 or 10 consecutive histidine amino-acid residues on either terminus of a recombinant protein and was first used to purify galactose dehydrogenase (Lilius *et al.*, 1991). While the tag offers no favourable characteristics for protein expression, it can be easily used for protein purification processes due to the tags high affinity for metal ions during immobilised metal affinity chromatography (IMAC). The tag is small, resulting in minimal interference on protein folding and functionality in comparison to other larger recombinant protein tags and has successfully been incorporated into purification studies, including those involving RAMP1-3-CLR fusion proteins, and can give purities of >80% (Hill and Pioszak, 2013)(Booe *et al.*, 2015)(Kimple, Brill and Pasker, 2015)(Roehrkasse *et al.*, 2018).

Although the tag can be easily incorporated into the protein construct, it relies on the tag being accessible. If the recombinant protein folds in such a way that the tag is not accessible, it cannot bind to the metal ions during IMAC purification and therefore will be less affective. This could be resolved by redesigning the protein construct so the histidine tag is at the other construct terminus, but this can be time consuming and expensive with gene/DNA vector synthesis and additional cloning experiments required. In previous studies, the histidine tag has been successfully incorporated into a RAMP1/2/3-CLR fusion protein and applied to IMAC without affecting protein functionality (Booe *et al.*, 2015)(Roehrkasse *et al.*,

2018)(Liang *et al.*, 2019)(Liang *et al.*, 2020). The tag could therefore be easily incorporated into future studies involving purification of the receptor

2.1.5 Aim and Objectives

Literature has previously reported that functional ECDs of the RAMP1/2-CLR fusion proteins for AM and CGRP receptors have been expressed and purified in bacteria but the generation of a functional RAMP3-CLR fusion protein was not achieved (Booe *et al.*, 2015)(ter Haar *et al.*, 2010). More recently, a study by Roehrkasse *et al.*, 2018 expressed RAMP1/2/3-CLR fusion proteins in HEK293T cells, stating their affinity for AM, CGRP analogues and intermidin (Roehrkasse *et al.*, 2018). Based on these previous efforts, gene sequences encoding the RAMP1/2/3-CLR extracellular domains will be fused together using a (gly-ser-glu)₃ linker. The construct will hold a C-terminus 10x Histidine tag with a 6x glycine linker linked to the CLR ECD and will be cloned into a vector at the C-terminus of the gene encoding for the MBP protein. Additional to this, a signal peptide (SP) will be cloned onto the N-terminus of the MBP protein which will allow the recombinant fusion protein to be secreted into cell media. Amino acids 22-111, 55-140 and 27-119 were used for RAMP1, RAMP2 and RAMP3 proteins respectively and residues 22-144 for the CLR protein. After cloning into an expression vector, it would generate the following constructs. Specific gene sequences can be found in the supplementary data.

- N-SP-MBP-RAMP1₍₂₂₋₁₁₁₎-(Gly-Ser-Glu)₃-CLR₍₂₂₋₁₄₄₎-(Gly)₆-10xHistag-C
- N-SP-MBP-RAMP2₍₅₅₋₁₄₀₎-(Gly-Ser-Glu)₃-CLR₍₂₂₋₁₄₄₎-(Gly)₆-10xHistag-C
- N-SP-MBP-RAMP3₍₂₇₋₁₁₉₎-(Gly-Ser-Glu)₃-CLR₍₂₂₋₁₄₄₎-(Gly)₆-10xHistag-C

The generated construct will be transfected into CHO-K1 cells and expression will be monitored over 1-5 days post transfection using western blot analysis. To validate construct expression, antibodies targeting the two tags of the protein (MBP and histidine tag) will be used. A stable cell line will also be generated via antibiotic selection, which will constantly secrete the fusion protein and allow easier purification set-up in later studies. The glycosylation patterns of the fusion proteins will also be investigated.

Hypothesis

1. RAMP1/2/3-CLR ECD fusion proteins will be successfully expressed and secreted from CHO-K1 cells

Objectives

1. To clone the SP gene and the RAMP1/2/3-CLR ECD-Histag construct into a vector to generate a *N*-SP-MBP-RAMP1/2/3 ECD-CLR ECD-Histag-C construct
2. Validate the expression of the fusion proteins post transfection in CHO-K1 cells
3. Investigate the glycosylation patterns of the RAMP1/2/3-CLR fusion proteins.

2.2 Methods

2.2.1 Transformation

Transformation allows the entry of plasmid DNA into bacteria for high throughput *in vitro* plasmid production. It provides an easy and cheap way of gaining large quantities of plasmid DNA for future experimental procedures. All DNA, including plasmids and gene strings were synthesised using GeneART Gene Synthesis and Services (ThermoFisher Scientific) unless otherwise stated. The DNA encoding the ECD of RAMP1/2/3-CLR with C-terminus 10x histidine tag were synthesised into Invitrogen TOPO Plasmids (ThermoFisher Scientific). See supplementary data 1 for gene sequence design. 1 ng of the RAMP-CLR TOPO plasmids and pSF-CMV-Puro-NH₂-MBP plasmid (MBP Plasmid, Oxford Genetics cat. OG3220) were transformed separately into 20 µL One Shot TOP 10 chemically competent *E. coli* (Invitrogen) by placing samples on ice for 30 minutes followed by heat shock in a water bath at 42°C for 30 seconds. See supplementary section 2 for the pSF-CMV-Puro-NH₂-MBP vector map. 250 µL super optimal broth with catabolite suppression (SOC media, Invitrogen) was added to the plasmid/bacteria mix and samples were incubated at 37°C with shaking (220 rpm) for 1 hour. Samples were spread on agar plates with appropriate antibiotic selection (50 µg/mL kanamycin or 100 µg/mL ampicillin, Sigma Aldrich) and left overnight to grow at 37°C. Resulting colonies were picked and grown for Maxiprep DNA extraction (see section 2.2.2)

2.2.2 Plasmid Maxiprep DNA Extraction

Plasmid maxiprep DNA extraction allows quick isolation and purification of large quantities of plasmid DNA by utilising DNAs ability to strongly associate to silica. Bacterial colonies were picked using a sterile 10 µL pipette tip and placed in a baffled flask containing 250 mL sterile LB broth (Sigma) supplemented with antibiotics (50 µg/mL Kanamycin or 100 µg/mL Ampicillin, Sigma Aldrich) and grown overnight at 37°C with shaking (220 rpm). Samples were spun at 7000 x g for 15 minutes prior to the application of the maxiprep kit with the subsequent supernatant being discarded. Plasmid DNA was extracted from the resulting cell pellet by using the PureYield Plasmid Maxiprep System (Promega) in accordance with the manufactures instructions. A Stuart rotator was used to aid the bacterial lysis (10 minutes at 9 rpm) and neutralisation steps (15 minutes 12 rpm). Plasmid DNA was eluted from the column using 1 mL DNase/RNase free H₂O. Plasmid DNA was quantified using a Nanodrop

2000 spectrophotometer at 260 nm and 280 nm (ThermoFisher)

2.2.3 RAMP-CLR ECD Construct Restriction Reactions and Ligations

2.2.3.1 RAMP1/2/3-CLR Restrictions

Restriction enzymes are mainly produced by bacteria and can be utilised to cleave specific sections of DNA. This enables the overlap of 2 different complimentary fragments of DNA from different sources (plasmids) to enable them to join after cleavage. 2 µg of the MBP plasmid and TOPO Invitrogen plasmids (RAMP1/2/3-CLR-HisTag plasmids) were restricted using 20 units EcoRI-HF and KpnI-HF endonucleases with 1x NEB buffer (New England Biolabs). Final reaction volumes were diluted to 50 µL using DNA/RNase free H₂O. Reactions were left for 3 hours at 37°C and heat inactivated at 80°C for 20 minutes. This inactivates the enzyme meaning it will not interfere with downstream reactions. The MBP plasmid was then dephosphorylated using 5 units shrimp alkaline phosphatase (rSAP, New England Biolabs) at 37°C for 30 minutes followed by heat inactivation at 80°C for 5 minutes. The dephosphorylation step is used to cleave phosphate ends of the MBP plasmid to prevent self-ligation during downstream ligation reactions. Restricted MBP and RAMP-CLR plasmids were then purified by utilising the E-gel Electrophoresis System (Invitrogen). This system allows faster and more efficient DNA fragmentation isolation with the ability to retrieve the required DNA fragment, purified from the recovery well. It means there is no additional loss of DNA during gel extractions and clean-up steps. The restricted DNA samples were loaded into CloneWell II Agarose Gel 0.8% (ThermoFisher Scientific) and ran at 100 watts until the sample entered the collection wells where 50 µL RNA/DNAase free water was added and DNA fragments were collected. DNA concentration was quantified using a Nanodrop 2000 spectrophotometer at 260 nm and 280 nm (ThermoFisher).

2.2.3.2 Ligation Reaction of RAMP-CLR ECD-HisTag into the MBP Plasmid

Ligation reactions can utilise T4 DNA ligase from the T4 bacteriophage to facilitate the joining of 2 strands of DNA. The enzyme does so through the formation of phosphodiester bonds on the DNA backbone. In this instance, the RAMP-CLR ECD DNA would join to the MBP plasmid, to form a completed circular vector. 30 ng of restricted MBP vector was ligated with

the RAMP-CLR ECD construct at a 1:6 (MBP vector:RAMP-CLR ECD) ratio, using 400 units T4 DNA Ligase with 1x NEB 2.1 buffer (New England Biolabs). Final sample volume was diluted to 20 μ L using DNase/RNase free H₂O. Samples were left overnight to react at 4°C followed by heat inactivation at 65°C for 10 minutes. Ligated samples were transformed into TOP 10 Chemically Competent *E. coli* (Invitrogen) as previously stated (section 2.2.1) by mixing the entire 20 μ L reaction sample with 20 μ L *E. coli*. Resulting colonies were picked and grown overnight in 5 mL antibiotic supplemented LB broth and plasmids extracted using the Plasmid Miniprep Extraction kit (Invitrogen) as stated in section 2.2.5. Positive colonies were initially identified by re-restricting the plasmid DNA with the restriction enzymes stated and excised DNA was visualised by running the DNA on a Tris-borate EDTA (89 mM Tris, 89 mM boric acid, 2 mM EDTA) 1.5% agarose gel supplemented with 5 μ L/100mL ethidium bromide (Sigma). Positive vectors were also sent for sequencing to confirm the correct insertion/alignment of the construct.

2.2.4 Signal Peptide Restriction Reactions and Ligation into the MBP vector

A BM40 signal peptide (SP) was cloned on the N-terminus of the MBP gene in the plasmid. The successful cloning of the signal peptide would give the following construct (figure 2.1).



Figure 2.1: Schematic of the RAMP-CLR fusion protein generated via cloning methods.

2 μ g of the MBP vector was restricted with 20 units NotI-HF and EcoRI-HF endonucleases with 1x NEB buffer (New England Biolabs). Sample reactions were made up to a final volume of 50 μ L using DNA/RNase free H₂O (ThermoFisher Scientific). Samples were left to react at 37°C for 3 hours and heat inactivated at 80°C for 20 minutes, giving a gene product of 1133 bp and an open cut vector size of 6798 bp. Both the gene product and open circular vector were collected during DNA purification using the E-gel Electrophoresis System (Invitrogen) as previously stated in section 2.2.3.1. The open circular vector was dephosphorylated as previously stated in section 2.2.3.1.

The purified gene product from the MBP plasmid restriction (1133 bp), and the BM40 signal

peptide gene string (generated using GeneART Gene Synthesis and Services, ThermoFisher Scientific, see supplementary section for sequence), were restricted with 20 units BsaI-HF and NotI-HF endonucleases with 1x NEB buffer (New England Biolabs). Sample reactions were made up to a final volume of 50 μ L using DNA/RNase free H₂O. Reactions were left at 37°C for 3 hours and further heat inactivated for 20 minutes at 80°C. The DNA samples were purified using the E-gel Electrophoresis system (Invitrogen) as previously stated. The BM40 gene fragment was ligated with the purified MBP gene product at a 1:1 ratio using 400 units T4 DNA Ligase with 1x NEB 2.1 buffer (New England Biolabs) using 50 ng of MBP gene product. Final reaction volumes were made to 20 μ L using DNA/RNase Free H₂O. Reactions were left overnight at 4°C, giving a combined DNA fragment length of 1306 bp, with the reaction being heat inactivated at 65°C for 10 minutes. The resulting 1306 bp DNA fragment was ligated into the dephosphorylated open circular MBP vector at a 1:3 ratio of vector to insert DNA overnight at 4°C, using 30 ng of vector DNA and 400 units of T4 DNA ligase (New England Biolabs). Final reaction volumes were made to 20 μ L with DNA/RNase free H₂O. The reaction was heat inactivated at 65°C for 10 minutes. See figure 2.2 for the resulting end cuts and ligation process. The ligation reaction was transformed into bacteria as previously stated and left overnight to grow at 37°C on 50 μ g/mL kanamycin resistance agar plates (Sigma).

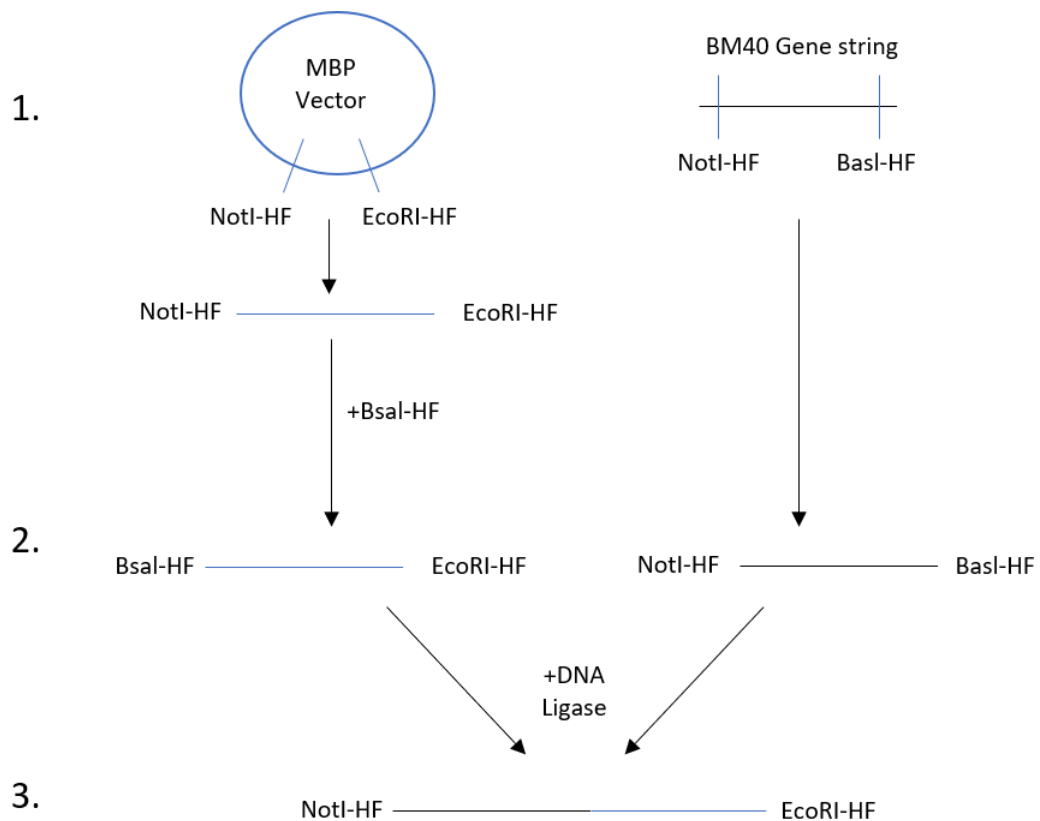


Figure 2.2: Schematic of the BM40 secretory peptide cloning process. 1. Indicates the initial cuts of both the BM40 Gene string and the DNA fragment from the MBP vector, with resulting fragment and cut sites (5'end is to the left of the fragments and the 3' to the right of the fragments). 2. The resulting restricted ends of both the DNA fragment from the MBP vector and the BM40 signal peptide after the second restriction cut. 3. The final construct length to be cloned

2.2.5 Miniprep Plasmid Extraction

Plasmid miniprep DNA extraction is a quick, small scale method of extracting plasmid DNA from bacteria. It utilises DNAs ability to bind to silica and can be eluted through a small change in pH. Bacterial colonies grown on antibiotic resistant agar plates were picked using a sterile 10 μ L pipette tip and placed in a 15 mL tube containing 5 mL sterile LB Broth (sigma) supplemented with antibiotics (50 μ g/mL kanamycin or 100 μ g/mL ampicillin, Sigma) and grown overnight at 37°C with shaking (220 rpm). Bacteria was sedimented at 3300 x g for 30 minutes and supernatant discarded. Plasmid DNA was extracted from the resulting pellet using the PureLink HiPure Plasmid Miniprep Kit (Invitrogen) in accordance to manufactures instructions. Plasmids were eluted from the silica column using 75 μ L pre-heated (65°C) TE Buffer provided by the kit.

2.2.6 Vector Sequencing

Vectors which could contain the gene encoding for the RAMP-CLR ECDs with histidine tag and signal peptide were validated via sanger sequencing to confirm correct alignment and gene incision. Sequencing was performed by the University of Sheffield Core Genomic Facility using the BigDye Terminator v3.1 Cycle Sequence Kit (ThermoFisher) and run on an Applied BioSystems 3730 DNA Analyser. The generated sequences were analysed using Seq Scanner 2 (ThermoFisher) and sequences were compared to the known synthesised DNA sequence (supplementary section 1) using the Align Sequences Nucleotide BLAST tool (National Library of Medicine, NIH). Vectors were provided at 100 ng/ μ L with the forward primers diluted to 1 pmol. The following primers were used to confirm successful cloning of DNA.

MBP Forward Primer (Detecting RAMP-CLR ECD-Histag): 5'-AGCGCTGAAGTCTTACGAGG-3'

CMV Forward Primer (Detecting Signal Peptide insert): 5'-ATGGGCGGTAGGCGTGTACG-3'

2.2.7 Tissue Culture

2.2.7.1 Maintenance of Cell Lines

CHO-K1 cells were maintained in T-175 cm² culture flasks (ThermoFisher Scientific) with Khans F-12K Medium (1X) (ThermoFisher Scientific) supplemented with 1% 10,000 units of penicillin, 10 mg/mL streptomycin (Sigma Aldrich) and 10% Foetal Bovine Serum (FBS) (full growth serum media) (Gibco- ThermoFisher Scientific) at 37°C in a 5% CO₂ incubator. Cells were passaged at 70% confluency and media changed at least twice a week.

2.2.7.2 Passage of Cells

Confluent cells were washed 2x with sterile phosphate buffered saline (PBS pH 7.4, Gibco) and incubated with 4 mL 1x TrypLE Express Dissociation Enzyme (Gibco) for 10-15 minutes at 37°C in a 5% CO₂ incubator or until all cells had detached from the flask. If necessary, cell detachment was aided with gentle agitation. Pre-heated full growth serum media was added to each flask at a 1:1 ratio to the TrypLE reagent and subsequently spun at 1000 x *g* for 5 minutes in a 15 mL tube. The supernatant was discarded, and pellets were resuspended in

10 mL growth media. 1 mL of cells were added to a new T-175 cm² culture flask containing 24 mL fresh media.

2.2.7.3 Cell counting

Cells were detached as stated above. 10 µL of cell suspensions were mixed with 10 µL 0.4% Trypan Blue Solution (Sigma) and placed into a Countess Chamber Slide (Invitrogen). The slide was inserted into a Countess Automated Cell Counter (Invitrogen) and the mean number of cells was calculated from two cell counts (duplicates).

2.2.8 Transfection of the SP-MBP-RAMP1/2/3-CLR Construct into CHO-K1 Cells

2.2.8.1 Polyethylenimine Reagent

The plasmid was transiently transfected into CHO-K1 cells using PEI (Polyethylenimine) reagent (PolyPlus). PEI introduces the DNA into the cell by binding to and condensing DNA into a positively charged particle. This allows the DNA:PEI complex to associate to the anionic cell surface, allowing it to be endocytosed into the cells. All cell culture media was pre-heated to 37°C in a water bath prior to starting the experiment. 1x10⁷ CHO-K1 cells were seeded into a T175 cm² culture flask and grown overnight in full growth serum media. 24 hours later or until cells reached 70-80% confluency, cells were washed 3x with 10 mL PBS and 20 mL opti-MEM reduced serum media [+] HEPES [+] 2.4 g/L sodium bicarbonate [+] L-glutamine (ThermoScientific) was added to each culture flask. The flasks were placed back into an incubator at 37°C and 5% CO₂ for at least 30 minutes to allow media and cells to stabilise/equilibrate prior to the addition of the transfection reagents. 50 µL PEIpro reagent was added to 1 mL opti-MEM reduced serum media, gently vortexed, and left at room temperature for 5 minutes prior to its addition to the plasmid DNA. 50 µg plasmid DNA was added to 1 mL opti-MEM reduced serum media and gently vortexed. The entire DNA:opti-MEM solution was added to the PEIpro:opti-MEM solution and was immediately vortexed gently. The resulting solution was incubated for 20 minutes at room temperature to allow the formation of the DNA:PEI complex. The entire volume of the DNA:PEI complex was added to the T175 flask containing cells with opti-MEM media and incubated at 37°C and 5% CO₂ for 24 hours. Media was replaced with CD FortiCHO Medium (ThermoFisher Scientific)

with 1x GlutaMAX and HT Supplement (ThermoFisher Scientific). Media samples were collected when appropriate.

2.2.8.2 Lipofectamine 2000

Plasmid DNA was transfected into CHO-K1 cell lines using lipofectamine 2000 (ThermoScientific). Lipofectamine is a cationic liposome which complexes with the plasmid DNA, forming a positively charged liposome. The liposome fuses to the negatively charged cell membrane allowing entry of plasmid DNA into the cell. All cell culture media was pre-heated to 37°C in a water bath prior to starting the experiment. 5×10^5 CHO-K1 cells were seeded into a 6-well plate 24 hours prior to transfection and cultured in 2 mL full growth serum media overnight at 37°C and 5% CO₂. 24 hours later or at 70-80% confluency, cells were washed gently 3x with 1 mL PBS and 2 mL opti-MEM reduced serum media [+] HEPES [+] 2.4g/L sodium bicarbonate [+] L-glutamine (ThermoScientific) was added to each well. The plate was placed back into the incubator at 37°C and 5% CO₂ for at least 30 minutes prior to the addition of transfection reagents. 7.5 µL lipofectamine 2000 reagent was added to 150 µL opti-MEM media and incubated for 5 minutes at room temperature. 4 µg SP-MBP-RAMP1/2/3-CLR-Histidine tag plasmid was mixed with 300 µL opti-MEM media and the entire solution was added to the lipofectamine reagent/opti-MEM solution and mixed gently. The lipofectamine/DNA mix was incubated for 20 minutes at room temperature to allow the formation of the liposome. PBS was used on the 'negative' samples containing lipofectamine only to supplement the DNA addition. The entire lipofectamine/DNA complex was added dropwise into the 6-well plate containing CHO-K1 cells with the solution gently stirred to mix the DNA/Lipofectamine complex into the media. The plate was incubated for 24 hours at 37°C and 5% CO₂. The transfection media was then removed and washed gently 2x with 1 mL PBS before the addition of full growth Ham's F-12K media with 10 µg/mL puromycin antibiotic. The media was changed after 3 days to maintain puromycin concentration for selection of positively transfected cells. Once selected colonies had been established 2 weeks after transfection, puromycin concentration was lowered to 1 µg/mL for a selection maintenance.

2.2.9 Cell Lysate Collection

Cell lysates were collected to detect the presence of any fusion protein being retained intracellularly. Unless stated, the following was done on ice or at 4°C. Cells transfected using the PEIpro transfection method (see chapter 2, section 2.2.8.1) were detached from the flasks as previously stated (chapter 2, section 2.2.7.2). Cells were spun at 1000 x *g* for 5 minutes and subsequent supernatant discarded. Lysis buffer (20 mM Tris-HCl pH 7.4, 300 mM NaCl and 1 mM EDTA) was added at 20 µL/20 mg of cell pellet and cells were sonicated in 30 second bursts 4x using a rod sonicator. Resulting lysates were spun at 13,000 x *g* for 20 minutes with the resulting supernatant collected. Protein samples were stored at -20°C until use. Protein concentrations in the lysates were quantified using the DC BioRad Protein Assay.

2.2.10 Protein quantification: DC Protein Assay

Protein lysates and cell media were quantified using the DC protein assay (BioRad). This is a colourimetric assay that determines protein concentration following detergent solubilisation. The assay is based on a reaction between alkaline copper tartrate and folin solution. The protein reacts with the copper in the alkaline medium which subsequently leads to the reduction of folin by 1, 2 or 3 oxygen atoms, developing a blue colour which is directly proportional to the amount of folin reduction. This holds a maximum absorbance at 750 nm. A working solution was made by adding 20 µL reagent S per 1 mL of alkaline copper tartrate solution (reagent A). Bovine serum albumin was used as a standard and was diluted in the appropriate media (lysis buffer or fresh cell media) to produce a standard curve (Table 2.1)

Table 2.1: Protein standard concentrations for the DC Biro-Rad assay

BSA Standards (mg/mL)	Buffer Amount (µL)	Bovine Serum Albumin Amount (µL)
0	20	0
0.294	16	4
0.588	12	8
0.882	8	12
1.17	4	16
1.47	0	20

Samples, including standards, were laid out in triplicates in a 96-well plate and protein samples added neat or diluted 1:10. 25 μ L of working reagent was added to each well followed by 200 μ L of reagent B. The samples were placed briefly on a plate shaker at 600 rpm before being left to react at room temperature for 30 minutes. The absorbance was read using a Perkin Elmer Ensign Plate Reader (Perkin Elmer) at 750nm. A standard line was plotted on GraphPad Prism 7 and unknown values were interpolated from the standard line (linear). An example of a BSA standard line can be found in figure 2.3.

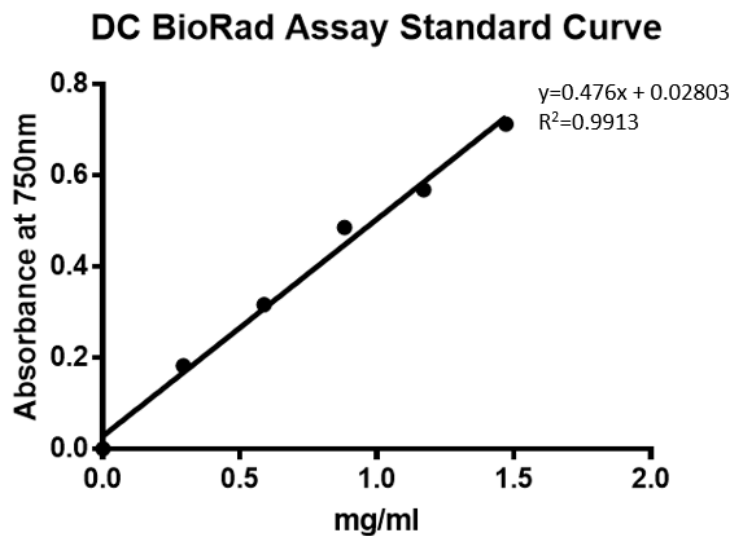


Figure 2.3: Example of a standard line for the DC BioRad protein assay

2.2.11 Denaturing Protein Deglycosylation

Protein in the cell media was deglycosylated using the GlycoProfile™ II Enzymatic In-Solution N-Deglycosylation Kit (Sigma Aldrich) according to manufactures instructions. The kit first denatures the protein and PNGase F, an enzyme which cleaves the innermost GlcNAc to asparagine, is added to remove N-linked glycans from the protein sample. This would remove sugars added during protein glycosylation, giving an insight into the extent of glycan addition to the recombinant protein. If necessary, the kit can also remove the glycans under non-denaturing conditions.

5 μ L denaturant solution containing 2% octyl β -D-glucopyranoside and 100 mM 2-mecaptoethanol was added to 90 μ L of media and incubated at 100°C for 10 minutes. The

media sample was briefly cooled at room temperature before the addition of 5 μ L reaction buffer (200mM ammonium bicarbonate). The media was separated equally into 2 separate tubes to act as a -ve and +ve enzyme reaction. 5 μ L of PNGase F Enzyme solution (500 units/mL) was added to the +ve and 5 μ L H₂O to the -ve control. The reaction was incubated at 37°C for 2 hours followed by 4°C overnight. The reaction was stopped by incubating the solution at 100°C for 10 minutes.

2.2.12 Western Blot: Gel Electrophoresis and Membrane Transfer

Western blots are commonly used as a semi-quantitative technique to monitor the expression of a protein within a cell lysate or in cell culture media. Samples can be diluted in buffer and denatured via heat with the addition of a reducing agent to reduce disulphide bonds in a protein structure. Proteins within the sample can be separated using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and subsequently be transferred from the gel to a membrane, which can be blocked to prevent non-specific antibody interactions. A primary antibody which can target the protein of interest is then applied to the membrane before the addition of a secondary antibody, which is used to bind to the primary antibody. The secondary antibody is often conjugated to horse radish peroxidase (HRP) which allows detection of protein bands via chemiluminescence.

5 μ g of quantified protein were diluted in 25 μ L Lamelli buffer (Bio-rad) (62.5 mM Tris-HCl, pH 6.8, 25% glycerol, 2% SDS and 0.01% Bromophenol Blue) and 5 μ L 500 mM dithiothreitol (DTT) (ThermoFisher). DNase/RNase Free H₂O (Invitrogen) was added to the samples to make a final sample volume of 50 μ L. Positive control samples included either the MBP protein expressed without the RAMP-CLR ECD or the recombinant CREB Human Protein, N-His.MBP tag (ThermoFisher Scientific). Samples were heated at 95°C for 10 minutes and allowed to cool at room temperature. The entire sample was loaded into a 4-20% Mini-PROTEAN TGX Precast Protein Gel (Bio-Rad) with 10 μ L Precision Plus Protein Dual Colour Standards (Bio-Rad) for size comparison. Gels were placed into a Vertical Electrophoresis Cell (Bio-Rad) for SDS-PAGE, with the chamber filled with 1 x Tris/Glycine/SDS buffer (25 mM Tris, 192 mM Glycine, 0.1% SDS, pH 8.3) (Bio-Rad). The samples were run at 150 mV for ~50 minutes. Gels were removed from the cassette and transferred onto a Trans-Blot Turbo Mini 0.2 μ m PVDF membrane (BioRad) by using a Trans-Blot Turbo Transfer System (Bio-rad).

2.2.13 Western Blot: Blocking and Antibody Probing

2.2.13.1 Anti-MBP antibody

Membranes were placed in blocking solution (5% skimmed milk and 0.1% Tween-20 in PBS) for 1 hour at room temperature on a plate shaker at 750 rpm. The membrane was placed in 4 mL blocking solution containing 1 µg (1:2000 dilution) primary HRP conjugated mouse Anti-MBP antibody (Abcam) and left on a stuart roller at room temperature for 1 hour. The membrane was washed 3x for 5 minutes in TBS-T (50 mM Tris-HCl pH 7.4, 150 mM NaCl and 0.1% tween) followed by 3x washes for 5 minutes in distilled water.

2.2.13.2 HisTag Antibody

Membranes were placed in blocking solution (5% skimmed milk and 0.1% tween-20 in PBS) for 1 hour at room temperature on a plate shaker at 750 rpm. The membrane was placed in 4 mL blocking solution containing 0.1 µg (1:4000 dilution) primary MAb mouse Anti-HisTag Antibody (Abcam) and left on a stuart roller overnight at 4°C. Membranes were washed 3x for 10 minutes in TBS-T before being transferred into blocking solution containing 3 µg (1:5000) HRP Conjugated Goat Anti-Mouse IgG (Agilent Dako) and left on a stuart roller for 1 hour at room temperature. Membranes were washed 3x for 5 minutes in TBS-T followed by 3x 5 minutes in distilled H₂O.

2.2.14 Western Blot: Detection

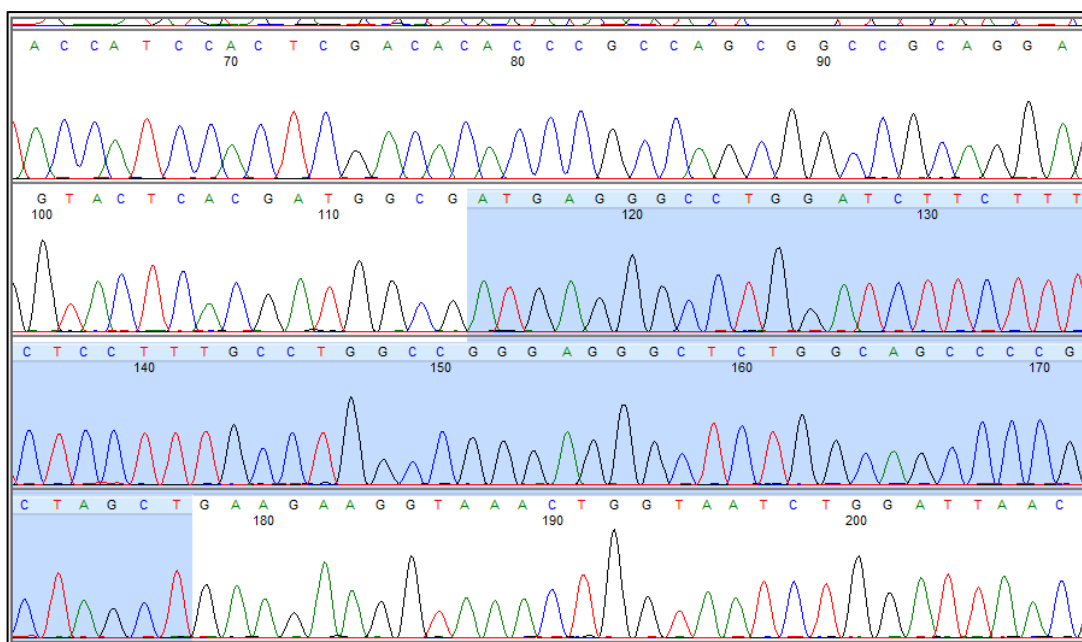
Protein on the membrane was detected using the SuperSignal West Dura Extended Duration Substrate (Thermo Scientific). 250 µL of luminol solution was mixed with 250 µL hydrogen peroxide solution and placed onto the membrane with both sides being covered. In the presence of peroxide buffer and HRP (from the antibody), luminol will undergo oxidation into an excited state which emits light and can be visualised with a photo imaging system. The membrane with the detection buffer is then placed into a transparent plastic film and visualised in the Bio-Rad ChemiDoc XR+ System (Bio-Rad).

2.3 Results

2.3.1 CMV Promoter- Signal Peptide Sequence

As previously stated, both the genes encoding for the signal peptide and RAMP-CLR ECDs were to be cloned into a vector containing the gene for the MBP protein (MBP vector). The signal peptide was to be cloned to the N-terminus of the MBP gene and the RAMP-CLR ECD gene was to be cloned to the C-terminus of the MBP gene.

First, the signal peptide was cloned into the vector. The vector containing the signal peptide and MBP genes would then be applied to clone in the RAMP-CLR constructs. Colonies which grew on antibiotic selection plates were further cultured and subsequent plasmids extracted. The plasmids were sent for sanger sequencing to confirm the successful incision of the signal peptide. Sequencing results were aligned using Align Sequences Nucleotide Basic Local Alignment Search Tool (BLAST) to confirm correct sequence in comparison to the known sequence generated. Figure 2.4 shows the sanger sequencing results of the signal peptide into the MBP vector. The Sanger sequence was aligned with the 'actual' synthesised BM40 signal peptide sequence and showed 100% identity comparison (figure 2.4).



Score	Expect	Identities	Gaps	Strand
122 bits(66)	3e-34	66/66(100%)	0/66(0%)	Plus/Plus
Query 1	ATGAGGGCCTGGATCTTCTTTCTCCTTTGCCTGGCCGGGAGGGCTCTGGCAGCCCCGCTA	60		
Sbjct 1	ATGAGGGCCTGGATCTTCTTTCTCCTTTGCCTGGCCGGGAGGGCTCTGGCAGCCCCGCTA	60		
Query 61	GCTGAA	66		
Sbjct 61	GCTGAA	66		

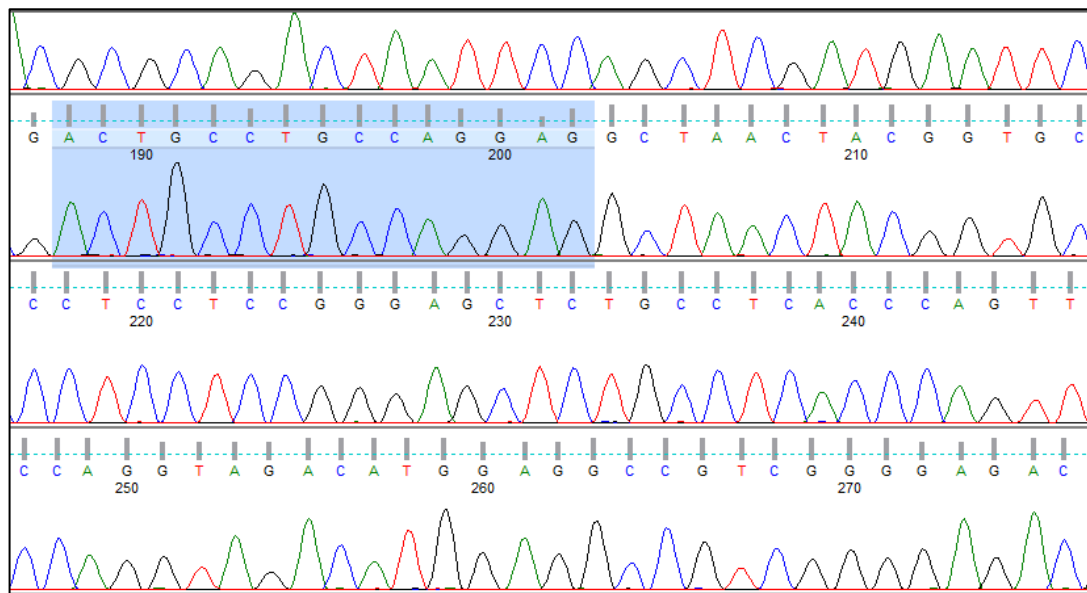
Figure 2.4: Sanger sequencing of the signal peptide cloned into the MBP vector. The highlighted region shows the sequence encoding the BM40 signal peptide. Peaks indicate a single base in the sequence with thymine (red), guanine (black), adenine (green) and cytosine (blue). Sanger sequence result for the signal peptide (query) compared with the ‘actual’ BM40 signal peptide sequence (subject) shows a 100% identity to each other.

2.3.2 RAMP1/2/3-CLR ECD-Histag Cloning Sequence

Following the cloning of the signal peptide sequence, the vector was used to clone the constructs encoding for the RAMP1/2/3-CLR ECDs. As previously stated, bacterial colonies that grew on antibiotic selection plates which could contain the gene insert (RAMP-CLR ECD) were cultured and plasmids extracted/purified. Plasmids were sent for sanger sequencing for confirmation of successfully inserted gene sequences that aligns properly in the vector at both restriction sites. The Sanger sequencing and sequence alignment analysis for the RAMP1-CLR, RAMP2-CLR and RAMP3-CLR ECDs can be found in figures 2.5, 2.6 and 2.7 respectively.

All sequences showed a successfully cloned, in-frame sequence which has 100% sequence identity match in comparison to the 'actual' synthesised gene. This vector could now be purified in larger yields (higher concentrations) and transfected into CHO-K1 cells for transient expression and stable cell line generation.

2.3.2.1 RAMP1-CLR-10xHisTag Sanger sequencing



Score	Expect	Identities	Gaps	Strand
1288 bits(697)	0.0	697/697(100%)	0/697(0%)	Plus/Plus
Query 1	GAATTCGACTGCCTGCCAGGAGGCTAACTACGGTGCCCTCCTCCGGGAGCTCTGCCTCAC	60		
Sbjct 1	GAATTCGACTGCCTGCCAGGAGGCTAACTACGGTGCCCTCCTCCGGGAGCTCTGCCTCAC	60		
Query 61	CCAGTTCAGGTAGACATGGAGGCCGTCGGGGAGACGCTGTGGTGTGACTGGGGCAGGAC	120		
Sbjct 61	CCAGTTCAGGTAGACATGGAGGCCGTCGGGGAGACGCTGTGGTGTGACTGGGGCAGGAC	120		
Query 121	CATCAGGAGCTACAGGGAGCTGGCCGACTGCACCTGGCACATGGCGGAGAAGCTGGGCTG	180		
Sbjct 121	CATCAGGAGCTACAGGGAGCTGGCCGACTGCACCTGGCACATGGCGGAGAAGCTGGGCTG	180		

Figure 2.5: Sanger sequencing of the RAMP1-CLR-10xHisTag into the SP-MBP vector. The Sanger sequencing shows part of the sequence generated from the sequencing analysis. Highlighted in blue is the start of the RAMP1-CLR-10xHisTag sequence. Sanger sequence for the RAMP1-CLR-10xHisTag ECD (query) compared with the 'actual' RAMP1-CLR-10xHisTag ECD sequence (subject) comparison shows a 100% identity to each other.

2.3.2.2 RAMP2-CLR-10xHisTag Sanger Sequence

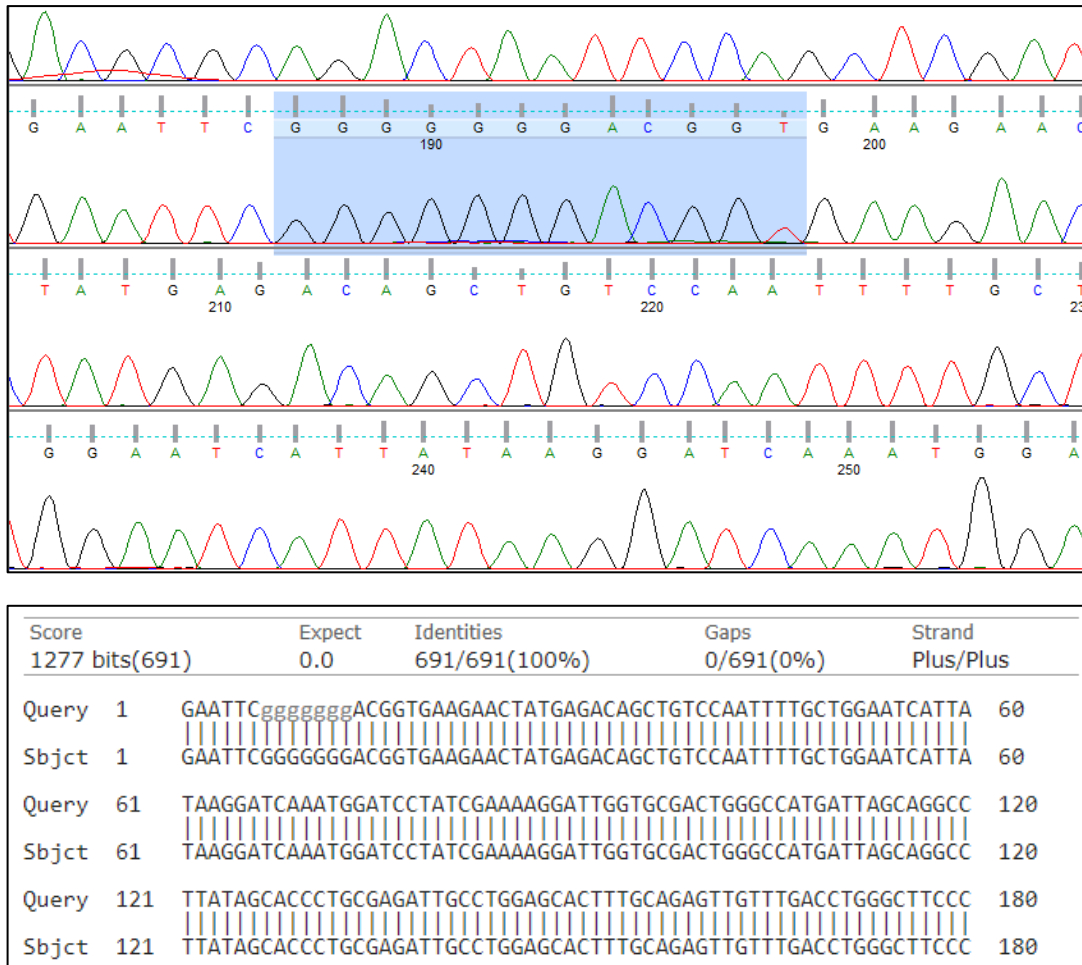


Figure 2.6: Sanger sequencing of the RAMP2-CLR-10xHisTag into the SP-MBP vector. The Sanger sequencing shows part of the sequence generated from the analysis. Highlighted in blue is the start of the RAMP2-CLR-10xHisTag sequence. Sanger sequence for the RAMP2-CLR-10xHisTag ECD (query) compared with the 'actual' RAMP2-CLR-10xHisTag ECD sequence (subject) shows a 100% identity to each other.

2.3.2.3 RAMP3-CLR-10xHisTag Sanger Sequence

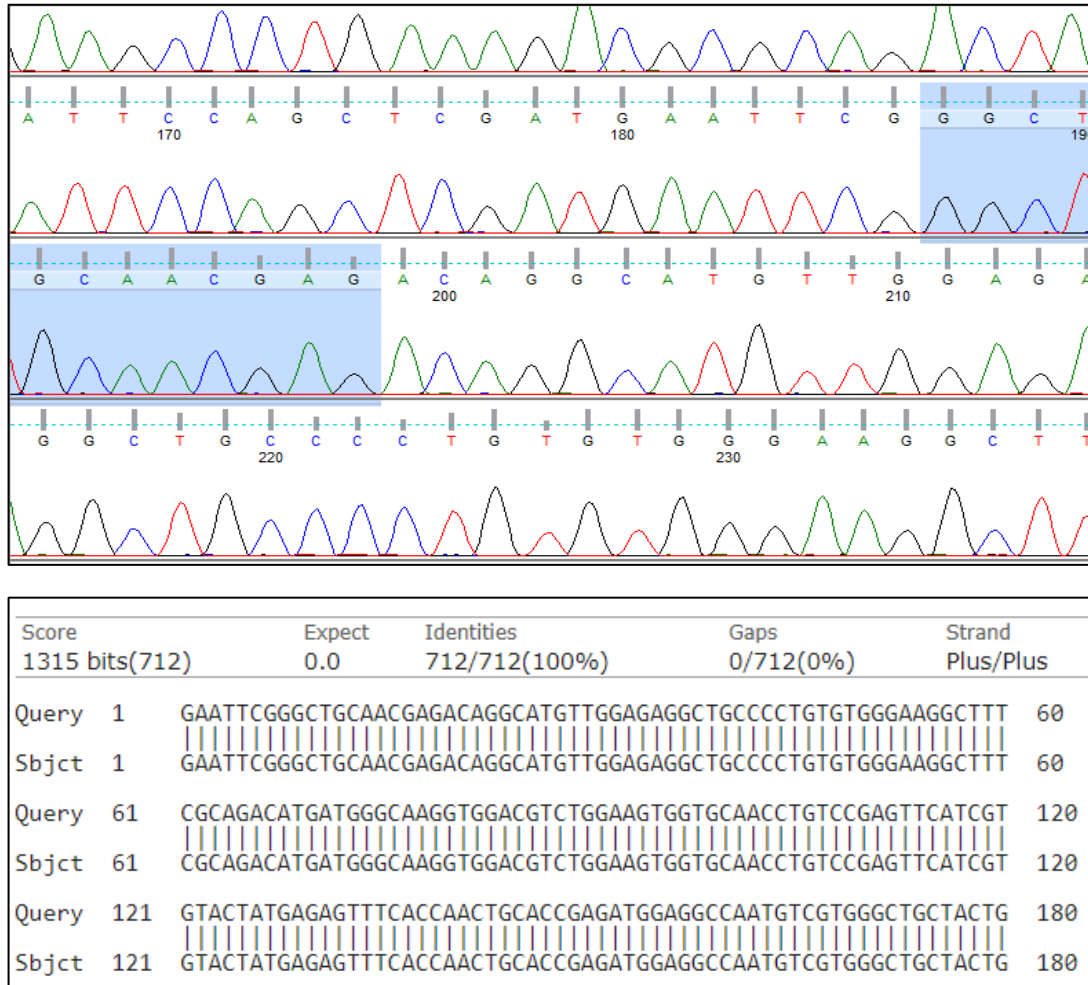


Figure 2.7: Sanger sequencing of the RAMP3-CLR-10xHisTag into the SP-MBP vector. The Sanger sequencing shows part of the sequence generated from the analysis. Highlighted in blue is the start of the RAMP3-CLR-10xHisTag sequence. Sanger sequence for the RAMP3-CLR-10xHisTag ECD (query) compared with the ‘actual’ RAMP3-CLR-10xHisTag ECD sequence (subject) shows a 100% identity to each other.

2.3.3 RAMP1/2/3-CLR Fusion protein expression

The constructs were transfected into CHO-K1 cells and their expression was monitored over 4-5 days post transfection. The expression of the RAMP1, RAMP2 and RAMP3-CLR fusion proteins were monitored using western blotting, initially probing for the MBP tag of the protein (figure 2.8). Media was collected from transfected cells 1-5 days post transfection and from wild type (WT) CHO-K1 cells (no DNA in the transfection material) before being probed with different target antibodies. After visualisation through the chemiluminescence detection, protein sizes were predicted as monomers in all targets. Predicted sizes for all 3 fusion proteins were ~69 kDa, but detected band sizes were ~75kDa, ~80kDa and ~90kDa for RAMP1, RAMP2 and RAMP3-CLR fusion proteins respectively (figure 2.8).

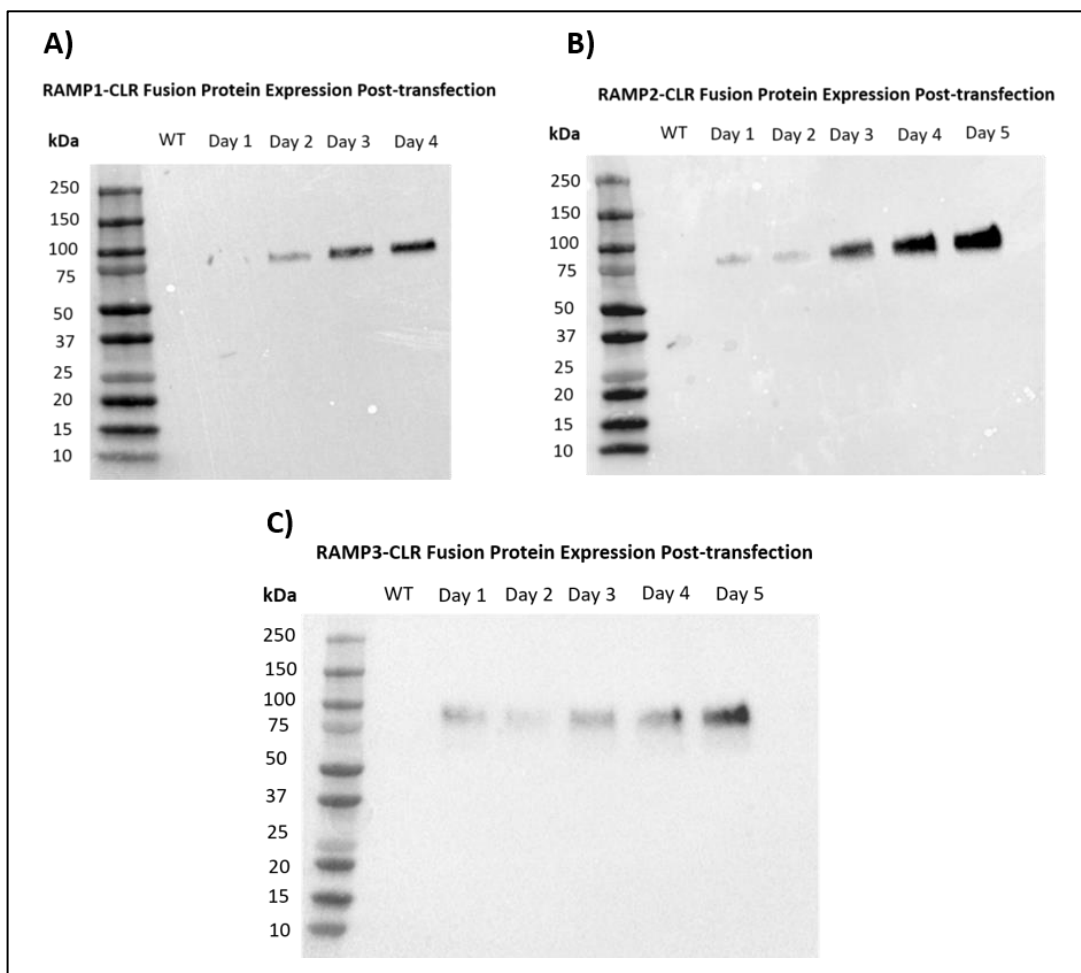


Figure 2.8: Western Blot analysis of the RAMP1/2/3-CLR fusion protein expression in CHO-K1 cells. (A) The expression of the RAMP1-CLR fusion protein 1-4 days post transfection. A single band appears at ~75 kDa. (B) The expression of the RAMP2-CLR fusion protein 1-5 days post transfection. A single band appears at ~80 kDa. (C) The expression of the RAMP3-CLR fusion protein 1-5 days post transfection. A single band appears at ~90 kDa.

2.3.4 Expression of RAMP-CLR Fusion proteins in Cell Lysates

Although the signal peptide was incorporated on to the constructs N-terminus to allow secretion into the cell media, CHO-K1 intracellular proteins should be examined to monitor if any mass fusion protein is being retained in the cell lysates. This would report how efficient the signal peptide is in enabling protein secretion.

Intracellular proteins were extracted and 5 µg total cell protein was used for each sample. The MBP tag of each protein was probed for, with either the MBP only (non-cloned vector containing the gene encoding for the MBP protein only) or the CREB-MBP recombinant protein used as positive controls. There was no detection of any intracellular fusion protein in RAMP1/2/3-CLR transfected CHO cells from days 2-4 (figure 2.9).

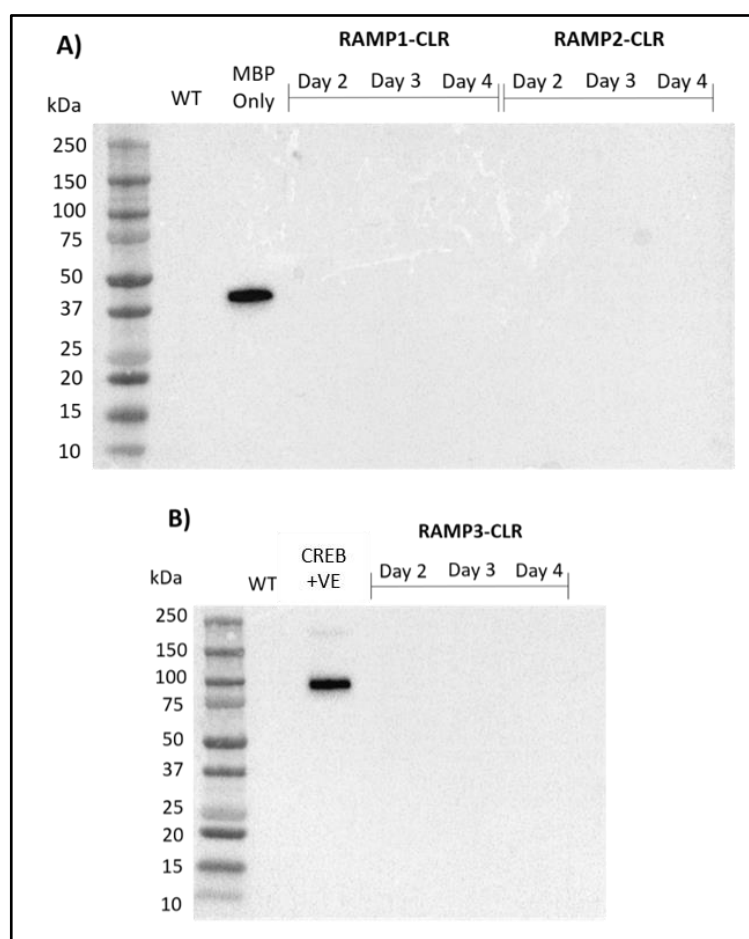


Figure 2.9: Western Blot analysis of the RAMP1/2/3-CLR fusion protein expression in CHO-K1 cell lysates. (A) Shows the no cell lysate expression from days 2-4 for the RAMP1/2-CLR fusion proteins, using the MBP protein only as a control. **(B)** Shows no cell lysate expression from days 2-4 for the RAMP3-CLR fusion protein, using the CREB-MBP recombinant protein as a positive control. WT samples are cell media samples which has no DNA in the transfection material.

2.3.5 Probing for the Fusion Protein Histidine Tag

While probing for the fusion proteins MBP tag would inform us of the protein being expressed in the sample, it does not give a strong indication if there is any presence of the RAMP-CLR-10xHisTag. If an anti-6x HisTag antibody is used, this would inform us of the expression of both the N- and C- terminal tags, and therefore the expression of the RAMP-CLR ECDs.

The RAMP1/2/3-CLR-10xHisTag samples were probed for their histidine tag (figure 2.10). Protein samples from day 4 post-transfection of the RAMP1/2/3-CLR fusion proteins were used to detect any histidine tag in the samples. The western blot shows the detection of a single band in each sample which corresponds to the previous band detected when probing for the MBP tag in the fusion protein (~75, ~80 and ~90 kDa for the RAMP1, RAMP2 and RAMP3-CLR fusion proteins respectively). No protein was detected in the wild type (WT) sample (non-transfected CHO-K1 cells). This shows the expression of both the MBP and histidine tags in the transfected samples, indicating the expression of the protein of interest.

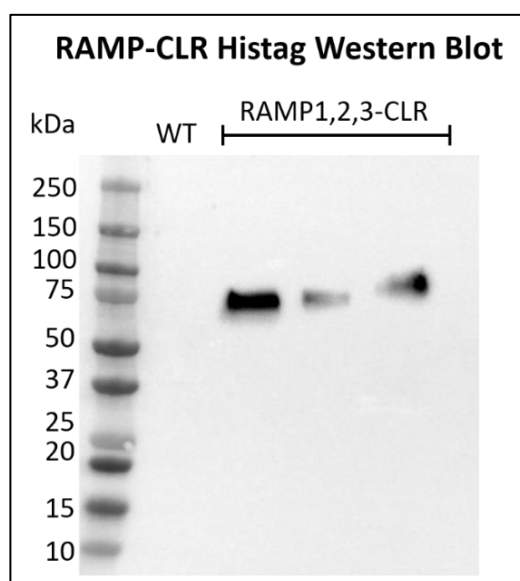


Figure 2.10: Western Blot analysis of the RAMP1/2/3-CLR fusion proteins probing for their 10xhistidine tags. Samples were run against the WT sample (CHO-K1 cells with no DNA in the transfection material) as a negative control. Bands appear at ~75 kDa, ~80 kDa and ~90 kDa for the RAMP1, RAMP2 and RAMP3-CLR fusion proteins respectively.

2.3.6 Deglycosylation of the RAMP1/2/3-CLR fusion proteins

As the RAMP-CLR fusion proteins were successfully being secreted from the transfected CHO-K1 cells, it means they were entering the secretory pathway and therefore undergoing post-translational modification such as the addition of glycans to the protein. The glycans could be removed using PNGase F and the extent of deglycosylation could be analysed.

5 µg of the deglycosylated RAMP-CLR fusion protein samples were applied to western blots for analysis, probing for the MBP tag on the fusion protein (figure 2.11). Analysis shows the deglycosylation process with no enzyme added to the sample (-ve) has no effect on protein size in comparison to the WT samples, which have not undergone any of the deglycosylation process. Samples which have the PNGase F enzyme added (+ve) all show significant reduction in protein size, which is ~5 kDa, ~ 10 kDa and ~25-30 kDa for the RAMP1-CLR, RAMP2-CLR and RAMP3-CLR fusion proteins respectively. This shows the 'actual' (without glycan addition) size of the RAMP-CLR fusion proteins. Protein sizes for all the RAMP-CLR fusion protein are ~69-70 kDa, which matches the predicted sizes for each fusion protein prior to expression in CHO-K1 cells. This shows the change in protein size is due to the addition of glycans during post-translational modifications.

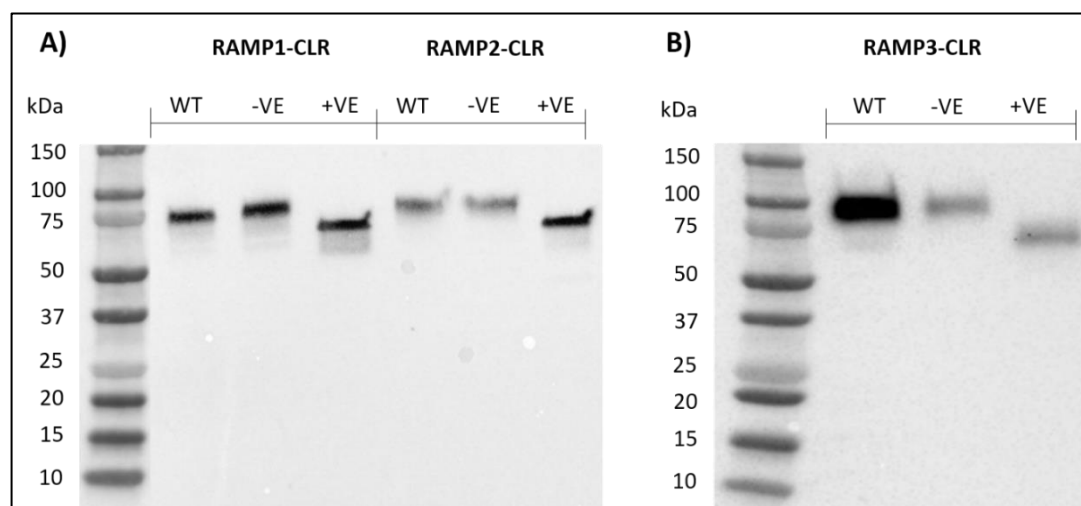


Figure 2.11: Western Blot analysis of the deglycosylated RAMP1/2/3-CLR fusion proteins probing for the proteins MBP tag. (A) Shows the RAMP1-CLR and RAMP2-CLR fusion protein samples with a control undergone no deglycosylation process (WT), undergone the process with no enzyme (-ve) and the sample undergone the process in the presence of PNGase F (+ve). **(B)** Shows the glycosylation patterns for the RAMP3-CLR fusion protein, with sample wells laid out in the same manner as stated.

2.4 Discussion

Here, genes encoding a SP and the RAMP1/2/3-CLR ECDs linked to a 10x histidine tag have been successfully cloned into a vector encoding the MBP protein to generate a N-SP-MBP-RAMP ECD-(Gly-Ser-Glu)₃-CLR ECD -(Gly)₆-10xHistag-C protein construct. Proteins were successfully expressed and secreted into the CHO-K1 cell media and it was determined that the fusion protein was being glycosylated, which gave larger protein sizes than expected (non-deglycosylated).

The mammalian cell type CHO-K1 was used as the host expression system over bacteria. Previous studies have purified and crystallised functional RAMP1 and RAMP2 ECDs tethered to the CLR receptor ECD in bacteria but could not generate a functional RAMP3-CLR ECD fusion protein (Kusano *et al.*, 2008)(Kusano *et al.*, 2012)(ter Haar *et al.*, 2010) (Booe *et al.*, 2015). This suggests that the RAMP3-CLR ECD receptor cannot be functionally expressed in bacteria and requires more advanced folding and modification steps, which only mammalian cells have the proper enzyme systems to do so. RAMP3 contains 4 *N*-linked glycosylation sites in its ECD where in comparison, RAMP1 has 0 and RAMP2 contains 1 potential site, further implying RAMP3 may require these *N*-glycans for proper folding and functionality. A recent study by Roehrkaase *et al.*, 2018 shows a glycosylated MBP-RAMP3-CLR ECD fusion protein expressed in HEK 293T cell media retains affinity for AM2/IMD peptides, implying the additional glycans are responsible for at least the correct protein folding and/or peptide binding (Roehrkaase *et al.*, 2018). This corresponds with previous data generated by Flahaut *et al.*, 2003 who completed direct site mutagenesis on the four suspected glycosylation sites in the RAMP3 ECD and observed substantial inhibition of receptor AM binding, suggesting they are important for protein functionality (Flahaut *et al.*, 2003).

In this instance, CHO-K1 cells were chosen over HEK293 cells due to CHO-K1 cells being generally more robust leading to easier cell culture, with CHO-K1 cells being highly tolerant to changes in temperature, oxygen and pH (Dumont *et al.*, 2015). CHO-K1 cells expressed all 3 RAMP-CLR ECD fusion proteins in the CHO-K1 cell media (figure 2.8) which were glycosylated during the process (figure 2.11). Deglycosylation of the fusion protein (figure 2.11) shows sizes of ~69 kDa which corresponds with previously expressed RAMP-CLR fusion proteins in bacteria (Booe *et al.*, 2015)

2.4.1 Protein Construct Generation and Rationale

The extracellular domains of the RAMP1/2/3-CLR constructs were designed based on previous reports which show the functional expression of RAMP1/2-CLR fusion proteins as well as their ability to be soluble and purified (Kusano *et al.*, 2012)(ter Haar *et al.*, 2010)(Booe *et al.*, 2015). The pSF-CMV-Puro-NH₂-MBP plasmid was used as the basis to express the SP-MBP-RAMP1/2/3 ECD-(glycine-serine-glutamate)₃-CLR ECD-(Glycine)₆-10xHistidine construct in CHO-K1 cells. The (glycine-serine-glutamate)₃ linker between the RAMP and CLR ECDs could be seen as a flexible linker but allows limited protein mobility, and was used in previous RAMP-CLR ECD expression studies, suggesting it has no effect on selectivity or binding affinities of AM or CGRP (Moad and Pioszak, 2013)(Booe *et al.*, 2015). A (glycine)₆ linker was used to tether the CLR ECD and 10xHistag to give flexibility and extend the 10xHistag from the CLR ECD, allowing easier purification by increasing its accessibility (Chen, Zaro and Shen, 2014). Literature reports the use of glycine as a linker in protein purification. Sabourin *et al.*, 2007, were able to report that the flexibility of a (glycine)₈ linker improved *in vitro* function of several epitope-tagged proteins involved in telomere maintenance while being stable against proteolytic enzyme digestion in yeast cells (Sabourin *et al.*, 2007). A (glycine)₆ linker was also expressed in a human serum albumin-natriuretic factor fusion protein by de Bold *et al.*, 2012 and maintains protein bioactivity, suggesting it can be applied without effecting the fusion proteins structural integrity (de Bold *et al.*, 2012)

The expression plasmid used provides a good template as the MBP is positioned at the N-terminal of the RAMP-CLR ECD gene, which has been shown to allow faster protein turnover/folding through the ER and golgi (Reuten *et al.*, 2016). A BM40 signal peptide was cloned at the N-terminus of the MBP which allowed the fusion protein to enter the secretory pathway. This was determined as the fusion protein was detected in cell media and not in the cell lysates (figure 2.8 and 2.9). The secretory pathway enables the protein to be processed through the ER and golgi apparatus, meaning cysteine bonds and post-translational modifications can be made which are crucial for RAMP-CLR protein function, in particular RAMP3 (Flahaut *et al.*, 2003)(Booe *et al.*, 2015)

2.4.2 RAMP-CLR Fusion Protein Expression in CHO cells.

Initially, the fusion protein was transiently transfected into CHO-K1 cells and grown in T175 flask to determine if the fusion proteins were able to be secreted when the media was

changed to FortiCHO media. FortiCHO media is a medium specifically designed to increase yield of recombinant proteins expressed in CHO-K1 cells without the need for FCS supplementation. The removal of FCS would aid purity in downstream applications including protein purification. Western blots were used to measure the presence of the fusion protein expressed by the transfected CHO-K1 cells, targeting specifically the N-terminal MBP tag and the C-terminal Histidine tag (figure 2.8 and 2.10). Antibodies which have been generated to target either the RAMP or CLR protein are limited and data produced by our lab (not published) shows non-specific binding or bands at unexpected molecular weights. It is also made more difficult with, in this instance, only the extracellular domain of the receptor being present. This means those antibodies that target the intracellular or intramembrane domain cannot be used to detect the protein of interest. In this case, the detection of the terminal tags should indicate a fully transcribed protein with the presumption that the RAMP-CLR protein is correctly folded.

The lysates of the proteins were also used to probe for the MBP tag of the fusion protein. As previously stated, this would inform us if any fusion protein is being retained in the cell and would inform us of the efficiency of the tag on secreting the protein. If the majority of the fusion protein was being retained in the cell, a different signal peptide may be required, replacing the BM40 signal peptide. However, in this instance, there was no detectable fusion protein in the lysates of the CHO-K1 cells and therefore the efficiency of the BM40 signal peptide was excellent (figure 2.9). If an increased amount (higher than 5 µg) of cell lysate was used during western blot analysis, some fusion protein may be detected but this would be an insignificant amount being retained in the cell. It would therefore not be worth the time and resources to alter and test different signal peptides in the vector when the BM40 signal peptide appears to be suitable for fusion protein expression. The BM40 signal peptide therefore remained in the vector and was used for future expression studies.

2.4.3 Glycosylation of the RAMP-CLR Fusion Proteins

All fusion proteins were found to be expressed in transfected CHO-K1 cells at slightly higher than expected sizes (figure 2.8). Protein sizes were expected to be ~69 kDa for all the RAMP-CLR fusion proteins but they were predicted to be higher due to posttranslational modifications, including glycosylation. Protein sizes appeared at ~75-80 kDa for the RAMP1/2-CLR and analysis showed bands appeared at ~85-90 kDa for the RAMP3-CLR fusion protein (figure 2.8). The size of the RAMP3-CLR fusion protein was expected to be higher in

comparison to the RAMP1-CLR and RAMP2-CLR fusion proteins, due to the increased number of glycosylation sites on the proteins ECD, but the extent of glycosylation was unknown prior to transfection and expression. The RAMP1-CLR fusion protein is slightly higher than the expected size (~69 kDa) due to glycosylation on the CLR protein, which contains 3 potential sites of glycosylation. This was confirmed when the removal of sugars reduced protein sizes to their predicted molecular weights (figure 2.11). While the CLR may be getting glycosylated in mammalian cells, it appears it would be unlikely to play a key role in correct receptor folding and functionality, with previous RAMP-CLR ECD proteins expressed in bacteria retaining their functionality and holding affinity for their requisite peptides (ter Haar *et al.*, 2010; Booe *et al.*, 2015; Roehrkasse *et al.*, 2018). Mutagenesis studies also do not report any change in ligand potency when altering these glycosylation sites in the CLR receptor (Booe *et al.*, 2015). Analysis of the MBP amino-acid sequence shows it is unlikely to contain any glycosylation sites and therefore was not deemed to be the reason for an increase in protein size.

The RAMP2-CLR fusion protein is therefore expected to be glycosylated on both the CLR and RAMP2 protein. The RAMP2-CLR fusion protein holds 1 potential site of glycosylation and sees a slight increase in protein size in comparison to the RAMP1-CLR fusion protein. This results in the protein being ~5 kDa larger than the RAMP1-CLR fusion protein, but ~15-20 kDa smaller than the RAMP3-CLR fusion protein, which corresponds with the requisite glycosylation sites on each RAMP protein. As mentioned in the glycosylation sites of the CLR receptor, RAMP2 glycosylation does not appear to have a major influence on protein folding and functionality with no reported mutagenesis at this position (RAMP2 Asp130) having an effect on ligand potency/binding (Kusano *et al.*, 2008, 2012; ter Haar *et al.*, 2010; Booe *et al.*, 2018). The RAMP3-CLR fusion protein has a higher amount of glycosylation (~20-25 kDa additional glycans), which was expected due to the increased number of glycosylation sites found on the RAMP3 ECD in comparison to RAMP1 and RAMP2.

2.5 Conclusion

To conclude, the data shows the successful cloning of both the signal peptide and the RAMP1/2/3-CLR-10xHistag proteins at the C-terminus of a vector containing the MBP gene. Constructs were successfully transfected and expressed in CHO-K1 cell media with no detectable amounts of protein being retained in the cell (intracellular lysates). All 3 fusion proteins show the addition of glycans to their ECD structures, with the largest difference seen on the RAMP3-CLR fusion protein.

Chapter 3: Purification of the RAMP3- CLR Fusion Proteins

3.1 Introduction

A large majority of the top selling drugs in the world target GPCRs, including those for respiratory, cardiovascular and neurological diseases. Membrane proteins are attractive as targets because they may not need compounds to penetrate the cells and represent over 40% of drugs (Cooke *et al.*, 2015). Although some drugs are based on crystal structures and rational design, minority of receptor protein structures weren't solved by X-ray crystallography. Only 1% of the 50,000 entries in the Protein Data Bank (PDB) represented membrane proteins in 2008, but structural data on GPCRs has dramatically increased over the past decade with high resolution crystals now available for most major GPCR classes (Carpenter *et al.*, 2008)(Stevens *et al.*, 2012)(Xiang *et al.*, 2016). The generation of this structural data aids structure based drug design (SBDD), which utilises the binding pocket and overall structure of a specific protein to aid drug design against the target receptor (Congreve, Dias and Marshall, 2014)(Shoichet and Kobilka., 2012). Despite these advances, membrane proteins are notoriously difficult to study, particularly because they are hard to express in a conformationally correct form while extracting them from the cell membrane. Addition of detergent to extract proteins from the bi-layer induces protein instability due to irreversible denaturation exposing hydrophobic regions to water, leading to increased protein insolubility (Flecha, 2017). GPCRs are often purified in the presence of ligands to attempt to overcome instability issues, which can be measured by using a thiol-specific fluorochrome which binds to exposed cysteine residues as the protein is unfolded during temperature increases (Errey *et al.*, 2015).

Along with the recent advancements in adding tags when expressing GPCRs in mammalian cells, further developed technology by biotech company Sosei Heptares includes the development of their proprietary STaR technology. This involves mutagenesis of amino acids around the receptor binding pocket to make them more thermostable without compromising their structure or functionality. The stabilisation process is typically based on alanine scanning where every amino acid after each methionine is mutated to an alanine residue, and natural alanine residues are mutated into leucine (Robertson *et al.*, 2011). This enables the easier purification of the full-length GPCR proteins to commence crystallography studies, assessing both the bound/non-bound extracellular, intramembrane and the intracellular segment possibly coupled to the G-protein (Jazayeri *et al.*, 2016)(Rucktooa *et al.*, 2018). Although this method is advantageous for expression of a full-length GPCR, it is not essential when smaller parts of a GPCR are required. In this study, the fusion protein

generated contains only the ECD of the protein. This excludes the highly hydrophobic membrane domain, reducing the chance of protein aggregation and insolubility but at the expense of some possible structural folding differences which may not give the exact 'true' binding composition of the receptors.

In order to utilise recombinant proteins expressed in cells for drug screening and to investigate the binding specifics through crystallography imaging, it is important to generate a pure protein sample which contains only the protein of interest. Samples which are generally impure may give false positive results during biochemical signalling assays due to cross interaction/binding of drugs to other proteins. Other protein impurities may also influence future protein crystallography screening. In order to separate the protein of interest from other proteins in the sample, multiple methods of protein purification can be used.

3.1.1 Affinity Chromatography

Proteins are primarily purified using tags they are co-expressed with and a common technique for this is affinity chromatography. In 1910, Starkenstein was one of the first to apply the technique to bind amylase to insoluble starch, and since then, affinity chromatography has become a widely used purification method. Typically, it relies on a reversible reaction where purification occurs during a biphasic interaction (Urh, Simpson and Zhao, 2009). A molecule is immobilised to a stationary phase (metal ion, ligand, amylose etc.) and the protein sample is passed over the immobilised molecule in a mobile phase. Here, the protein associates to the immobilised molecule on the stationary phase and is separated from the bulk of the proteins in the sample.

3.1.1.1 Immobilised Metal Affinity Chromatography

Specifically, during immobilised metal affinity chromatography (IMAC), a metal ion (Co^{2+} , Ni^{2+} or Cu^{2+}) is immobilised to the stationary phase. This enables the capture of a protein using a small 6-10 histidine amino acid tag on either the C or N-terminus of the recombinant protein. The positioning and size of the tag minimises the likelihood of interference on protein folding, solubility, or function. The tag/metal ion interaction can easily be displaced

with a higher concentration of imidazole, which is structurally identical to the R-group of the histidine amino-acid (figure 3.1).

Many groups have used IMAC as a first enrichment step due to the commercial availability of columns and resin types which offer slightly different characteristics depending on their associated metal ion. These include Ni²⁺-nitrilotriacetate (Ni²⁺-NTA) resins and Talon resins (Co²⁺- cabroxymethylaspartate). For example, Ni²⁺-NTA columns offer a higher affinity to the histidine tag of the fusion protein but therefore a higher amount of background binding, whereas Talon resins offer reduced affinity to the proteins histidine tag but greater likelihood of selectivity (Weiû and Grisshammer, 2002)(Grisshammer, 2009). Recombinant proteins which hold a more accessible histidine tag maybe purified with Talon resins to gain a better sample purity. Using talon resins however could reduce the overall yield gained from the IMAC purification but in turn reduce the need for a multi-step purification method. The balance of the affinity and specificity of the method can minimise protein loss and duration of purification experiments.

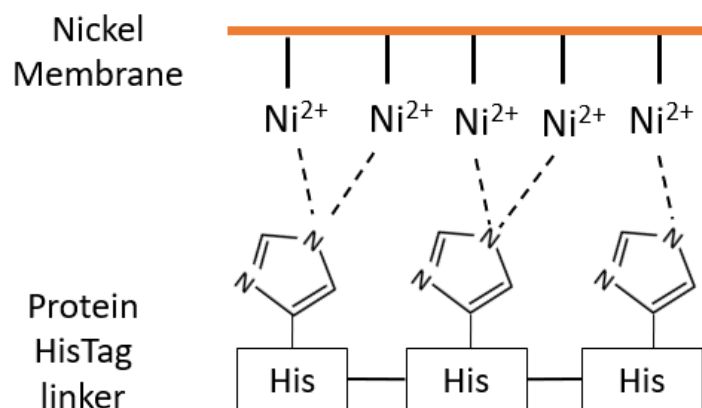


Figure 3.1: The association of a histidine tag to nickel ions during IMAC. The nitrogen group on the imidazole ring forming coordinate covalent bonds (dashed lines) with the immobilised nickel ions.

A typical problem with IMAC is non-specific binding of untagged proteins. Although histidine is a relatively uncommon amino acid, only composing 2% of all protein sequences, if two histidine amino acids are adjacent to each other, they will have affinity to the immobilised metal in the column resin (Schmitt, Hess and Stunnenberg, 1993)(Bornhorst and Falke, 2000). A further issue with this technique is that using metal ions in the purification column can create a charged environment which leads to non-specific ionic interactions of other

proteins to the resin. This can be combated with an increased concentration of NaCl in the binding buffer but high concentrations of NaCl can lead to protein aggregation and precipitation during 'salting out', where the protein competes with the NaCl to hydrogen bond with water in the buffer. As previously stated, a Talon resin packed with cobalt ions could be used as an alternative to reduce the non-specific binding but would reduce the overall protein yield generated from the purification steps whilst not guaranteeing pure protein samples.

3.1.1.2 Purification using the Maltose Binding Protein

Other tags may be used for purification of expressed proteins. The MBP protein is one of the most popular fusion partners for recombinant protein production and purification in bacteria cells. Reports show purities of 70-90% following a single capture step when using the MBP tag expressed in *E. coli* (Lebendiker and Danieli, 2017)(Reuten *et al.*, 2016). As the MBP protein is endogenous to prokaryotes and is not found in eukaryotic cells, it may aid a higher purity when recombinant proteins are expressed in mammalian cells due to less competition from endogenous MBP binding.

Tagging with MBP is a useful method for purification due to its high affinity for amylose. If a column is packed with amylose or dextrin, the protein can be passed through the column and captured. As MBP binds to amylose primarily by hydrogen bonding, it offers the advantage of using buffers that contain high concentrations of salts (up to 1 M NaCl) and therefore can reduce the amount of ionic binding to the column. However as there are no metal ions in the column (unlike IMAC) a charged environment in the column is unlikely and this concentration therefore may not be required (Riggs, 2000). The protein is simply eluted from the column using a higher concentration of maltose (typically 10 mM), which dissociates the protein from the column. The only significant disadvantage of the tag is that it relies on the MBP binding pocket being accessible to bind to the amylose or dextrin in the column. If the protein folds in such a way that the binding pocket is not accessible, it may not be able to bind to the column and therefore will affect protein capture.

Purification utilising the MBP tag after recombinant proteins have been expressed and secreted into mammalian cell media however is generally unreported. Reports which use the MBP tag in mammalian cells focus on the advantage the MBP tag gives in recombinant protein expression and opt for other methods of protein purification (Roehrkasse *et al.*,

2018)(Reuten *et al.*, 2016)(Bokhove *et al.*, 2016). Although reports which utilise the MBP tag for purification after expression in mammalian cells show excellent purity (>85% purity), these have been intracellularly expressed recombinant proteins (Oak, Jansen and Chan, 2019)(Paulsen *et al.*, 2015). It is therefore unclear how effective purifying the protein based on its MBP tag would be when the recombinant protein is secreted into the cell media.

3.1.2 Size Exclusion Chromatography

Size exclusion chromatography (SEC) or gel filtration chromatography, is a powerful and reliable chromatography technique for obtaining information about a proteins size under native conditions. It allows the separation of proteins based on their size and does not rely on buffer compositions, pH, temperature or ionic strength of the protein, although these can influence protein mobility through the column. The protein passing through the column does not interact with the gel matrix and therefore no protein is retained, allowing a 100% protein yield after the purification. The gel is packed with porous particles (beads) which composes the stationary phase of the chromatography system. The beads allow molecules which are smaller than the pores to enter, while those that are too large manoeuvre around the beads. This results in the smaller particles having a longer residence/retention time, while larger molecules have a shorter retention time (figure 3.2). The larger molecules elute first as they avoid the beads and pass quickly through the column, while the smaller molecules elute in the later fractions as they have a larger surface to manoeuvre (figure 3.2). The pore size therefore has a large effect on the protein retention times which should be considered before protein application. In addition to this, the proteins in the sample can be de-salted (buffer exchanged) as the salts are able to enter the pores in the column.

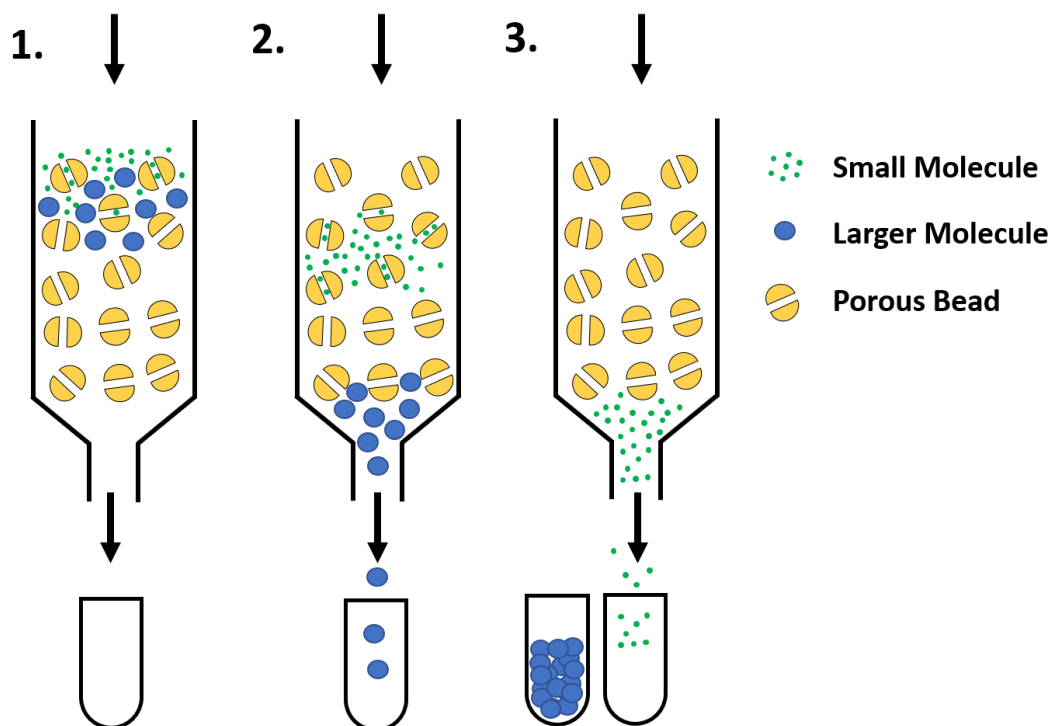


Figure 3.2: Size Exclusion Chromatography process. 1. Shows the application of the protein sample into the column with blue circles representing the larger molecules and green representing smaller molecules. 2. Shows the separation of the proteins the smaller molecules enter the beads. 3. The elution of the different size proteins into different fractions.

While these experimental characteristics may be favourable, the technique relies on separation of protein based on their size. Although columns can be fairly resolute, if proteins are of similar sizes, the purification will not be as efficient. Generally, SEC columns require a 10% difference in molecular mass between proteins in order for the technique to have a good resolution. This therefore can be an issue during second purification steps if non-specific binding is similar to that of the protein of interest. In spite of this, recent reports purifying the RAMP-CLR, CTR, and RAMP-CTR ECD receptors fused to an MBP tag all utilised SEC as a secondary purification step (Lee, Hay and Pioszak, 2016; Roehrkasse *et al.*, 2018). If purification efforts in this study were to produce similar purities after completing an initial purification, it suggests if necessary, SEC could be applied to further purify the protein of interest.

3.1.3 Aim and Objectives

As previously stated in Chapter 2, RAMP-CLR ECD fusion proteins were successfully expressed in CHO-K1 cells with stable cell lines secreting the RAMP-CLR fusion protein. This protein will be expressed in larger quantities and purified based on either its MBP protein, 10x histidine tag or by SEC, to generate a pure protein sample. The number of purifications is ultimately dependant on the protein purity at the end of each experimental step.

Media containing the RAMP-CLR fusion protein from stable cell lines will be collected and concentrated using tangential flow filtration before being applied to column (affinity) chromatography. First attempts will be to purify the RAMP-CLR fusion protein based on its MBP tag and depending on purity and amount captured, the protein will be applied to IMAC as a primary or secondary purification method. Protein capture will be monitored by using western blot analysis and the purity of samples will be assessed by staining SDS-PAGE gels with Coomassie blue protein dye. If the protein needs further purification, it will be applied to SEC. If the yield and protein expression is low from the stable cell lines, the RAMP-CLR fusion proteins may be transiently transfected into CHO-K1 cells using PEI reagent to increase protein expression. The aggregation state of the fusion protein after purification will also be assessed by using non-denaturing, native PAGE.

Hypothesis

- RAMP-CLR fusion proteins can be purified to >95% purities by using affinity and size exclusion chromatography techniques.

Objective

- Purify RAMP-CLR fusion proteins based on their MBP tags and assess protein purity.
- Purify the RAMP-CLR fusion proteins using IMAC and assess protein purity.
- Assess the aggregation state of the purified fusion protein by applying it to native PAGE.

3.2 Methods

3.2.1 Tissue Culture

3.2.1.1 HYPER Flask Culture

All general cell culture techniques for the maintenance and passage of CHO-K1 cells were carried out as previously stated (chapter 2 section 2.2.7). For purification of higher quantities of cell media, CHO-K1 cells were cultured in Corning 1720 cm² High Yield Performance (HYPER) *Flask* culture vessels (HYPERFlask, Corning). Each flask contains the equivalent space as 10 T175 cm² culture vessels, allowing time and incubator space saving while giving increased protein expression. 48×10^6 CHO-K1 cells were added to 560 mL pre-heated full growth serum media (Khans F-12K medium 1X, 10% FBS, 1% 10,000 units of penicillin and 10 mg/mL streptomycin, ThermoFisher Scientific) supplemented with 1 µg/mL puromycin and the media was mixed to enable homogeneity. The media containing cells was poured into the HYPERFlask, holding the flask at a 45° angle to avoid the addition of bubbles into the flask. Once full, the HYPERflask containing cells was placed back in the incubator at 37°C and 5% CO₂. 24 hours later or at ~80% confluency, media was removed, and cells washed gently 4x with 200mL PBS to remove excess media left in the flask. ~560 mL pre-heated CD FortiCHO medium (ThermoFisher Scientific) supplemented with 0.01 mM Sodium Hypoxanthine, 0.016mM Thymidine (HT Supplement, GIBCO) and 1x GlutaMAX (GIBCO) was added to the flask at a 45° angle to avoid addition of air into the flask. The HYPERflask containing fresh medium was placed back into the incubator and cultured. 7 days after the addition of CD FortiCHO medium the cultured cell media was collected for purification.

3.2.1.2 Transient Transfection of CHO-K1 cells.

Although stable cell lines were generated for ease of cell culture, recombinant protein expression levels can often be compromised. Transiently transfected CHO-K1 cells would express a higher amount of the RAMP-CLR fusion protein in the cell media and this could aid protein purification steps and sample purity. Vector DNA encoding the RAMP-CLR fusion protein was transiently transfected into CHO-K1 cells using polyethylenimine (PEI) reagent as previously stated in Chapter 2, Section 2.2.8.1. 5 T175 cm² culture flasks (ThermoFisher Scientific) were transfected and cells were cultured in CD FortiCHO media for 6 days post

transfection. The media was collected and a final volume equated to ~150 mL (30 mL/flask). The media was process in the same way as that collected from the HYPERFlask culture.

3.2.2 Media Sample and RAMP-CLR Fusion Protein Concentration

Prior to sample purification, the media was concentrated to reduce sample volume and aid the ease of purification (reducing time). Cell culture media was applied to a PALL Minimate Tangential Flow Filtration (TFF) System (Pall Corporation) using a 30K molecular weight cut off (MWCO) Minimate TFF capsule. The Minimate TFF system is an efficient method of separation, purification, and concentration of a protein sample. TFF uses crossflow filtration rather than the direct flow or 'dead-end' filtration which places the membrane perpendicular to the feed stream (figure 3.3). Crossflow filtration prevents membrane clogging, which is common in direct flow filtration, with one portion of the sample passing through the membrane (those below the MWCO filtrate) while the remainder of the sample (retentate) is circulated back into the reservoir. A schematic of the TFF system can be found in figure 3.4.

Unless stated, the following steps were all performed at 4°C. A 30,000 MWCO Minimate Tangential Flow Filtration Capsule (Pall Laboratories) was applied to a Minimate Tangential Flow Filtration 115V AC 50/60 Hz (1/pkg) system (Pall Laboratories, Product ID: OAPMP110) System set up was followed as directed in manufactures instructions, with the provided pharmed tubing (3.2 mm) used for sample delivery and sample retention, and 3.2mm (1/8") clear tubing used to direct the filtrate volume to waste. Prior to each run, the retentate tubing was disconnected from the sample reservoir and outlet was directed to filtrate reservoir (waste) with the filtrate tubing. The system was subsequently flushed out with 500 mL H₂O, running the peristaltic pump at 70 rpm. The clamp on the retentate tubing was tightened to achieve a system pressure of ~30 psi for 1 minute to remove any air in the Minimate capsule. Pressure in the system was released by loosening the retentate clamp and the water was allowed to flow through the system until the sample reservoir was run almost empty. This process was repeated twice more with the sample reservoir not being allowed to empty on the last wash run to prevent air entering the system.

Prior to its application to the TFF system, the cultured cell media sample containing the RAMP-CLR fusion protein was spun at 2000 x *g* for 10 minutes to sediment any cell and debris from the culture media. The resulting supernatant was added to the sample reservoir and

the peristaltic pump was run at 70-80 rpm. The retentate clamp was tightened until a system pressure of ~30 psi was achieved, and the sample was left to concentrate for ~2-3 hour, or until the sample volume was reduced to ~10 mL. Once the desired retentate volume was reached, 20 mL of binding buffer used during chromatography purification (see sections 3.2.3.1 and 3.2.3.2) was added to the reservoir for buffer exchange and sample left to concentrate to ~10 mL. This process was repeated 7 times. After the final buffer exchange step was completed, the system was stopped and the clamp on the filtrate tubing tightened to stop liquid flow through this outlet. The clamp on the retentate tubing was released and 30 mL of buffer was added to the reservoir. The buffer was run through the system for 15 minutes at 70 rpm, removing any protein covering the capsule membrane. The retentate tubing was then re-directed into a 50 mL tube and sample collected. A further 10 mL buffer was run through the system and collected giving a final sample volume of 40-50 mL. The concentrated media sample was stored on ice until it was applied to column chromatography.

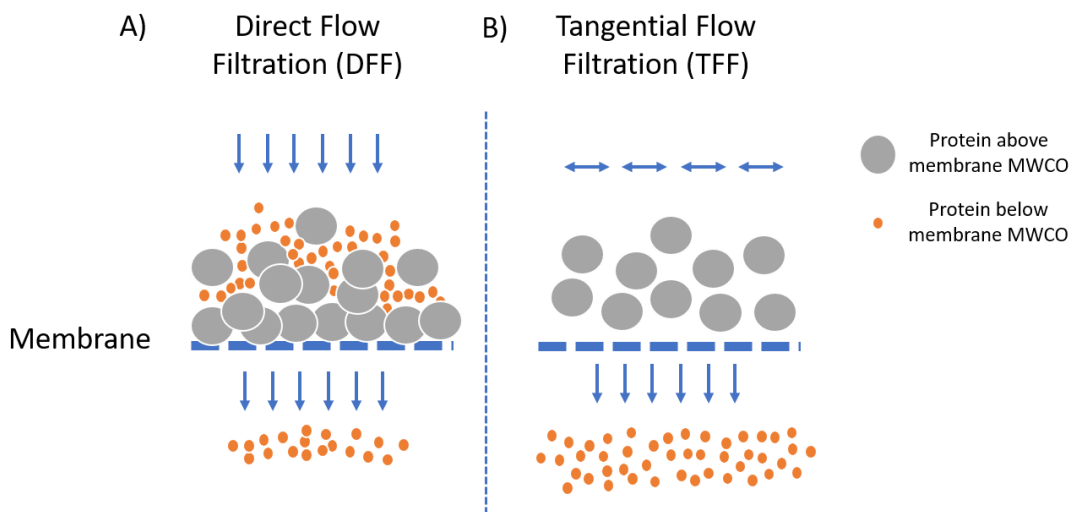


Figure 3.3: Comparison of conventional direct flow filtration and tangential flow filtration (TFF). A) Direct flow filtration shows protein samples flow directly onto the membrane surface leading to clogging and partial size separation. B) The movement of sample across a membrane in a tangential manner shows more effect size separation and prevention of membrane clogging, resulting in a more efficient size separation and protein retention

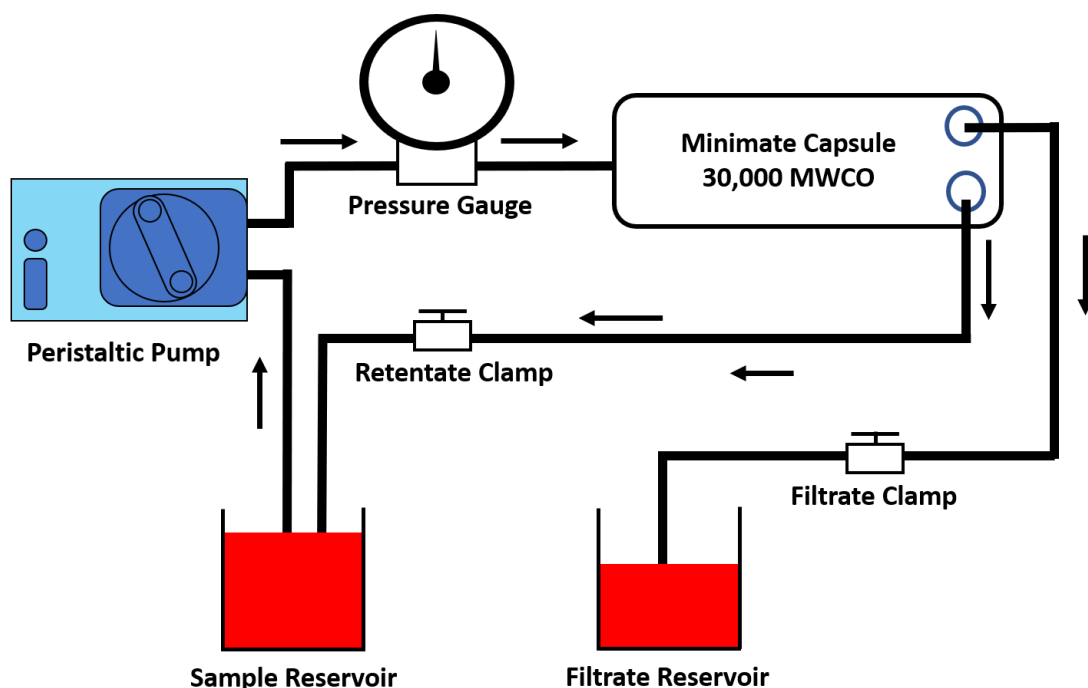


Figure 3.4: Schematic of the TFF system. Arrows show the flow of sample through the system.

After sample concentration, the system was cleaned at room temperature. 200 mL 0.5 M NaOH pre-heated to 40°C was flushed through the system, setting the peristaltic pump at 70 rpm and directing the filtrate and retentate tubing to the filtrate reservoir (waste). The system was stopped when the reservoir was almost emptied. The retentate tubing was redirected back to the reservoir and the filtrate clamp was tightened to stop liquid release at this outlet. 300 mL of pre-heated (40°C) 0.5 M NaOH was added to the reservoir and circulated in the system for ~45 minutes. If the liquid in the sample reservoir appeared discoloured after a short period, the retentate tube was redirected to the waste and the liquid flushed through the system. Fresh pre-heated 0.5 M NaOH was added to the reservoir after redirecting the retentate tubing to the sample reservoir. After washing, the retentate tubing was redirected to the waste and liquid in the system removed. The filtrate clamp was released, and the system was flushed through with 1.5 L of water (room temperature), with both outlet tubing going straight to the filtrate reservoir (waste). The capsule was then ready for use on another concentration cycle. If the capsule was to be left for longer periods without use (>2 days) the water wash step was replaced with 0.05 M NaOH and left at room temperature.

3.2.3 Fusion Protein Purification

To generate a pure protein sample, the media containing the RAMP-CLR fusion protein was applied to column chromatography based on the fusion proteins histidine or MBP tag. The AKTA Start Column Chromatography System (GE Lifesciences) was used to purify the RAMP-CLR fusion proteins. The sample can be applied via a peristaltic pump or injected manually into the system (figure 3.5). As a sample is introduced into the system, it passes through a column of user choice and captures the protein of interest. The non-bound protein which is not captured (flow through), leaves the column and passes through a UV light (280 nm) whose absorbance is monitored by the system. The absorbance of the light is based primarily on tryptophan, tyrosine and partly phenylalanine amino acids due to their aromatic rings ability to absorb UV light. The sample flow through can be collected for protein analysis of any fusion protein not bound to the column, or to be re-run over the column for maximum protein capture in any one run. Upon elution from the column, samples can be eluted into set fractions.

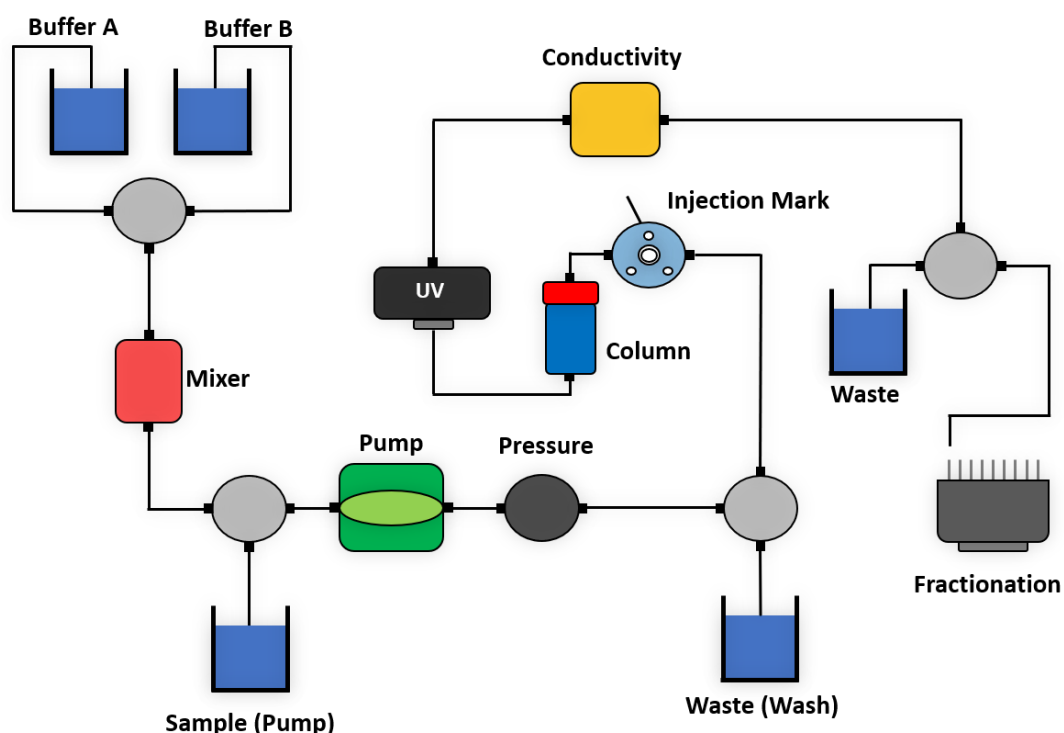


Figure 3.5: Schematic diagram of the AKTA start purification system. Samples are introduced by the pump or injection mark (labelled) which passes through the column and into the UV detector. The eluted samples can be fractionated upon elution.

Prior to each purification, buffer A, buffer B and sample pumps were washed with 10 mL (10

mL/min flow rate) 1 M NaOH followed by 10 mL water. Fraction tubing was washed with 5 mL 1M NaOH followed by 5 mL water at a flow rate of 1 mL/min. This process was also repeated post-purification to clean the system. If the AKTA start system was not going to be used for long periods of time (>5 days) it was stored in 20% ethanol by adding an additional ethanol wash to the previous steps stated.

3.2.3.1 MBP Purification

Unless stated, all purification steps were carried out at 4°C. Prior to the purification, the sample was concentrated using the TFF system as stated in section 3.2.2 and buffer exchanged in binding buffer (50 mM Tris-HCl pH 8, 5% glycerol, 150 mM NaCl). Concentrated samples were applied to the AKTA Start Column Chromatography System (GE Lifesciences) via the pump inlet and run through a 1 mL MBPTrap HP column (GE Lifescience). An MBPTrap column contains a membrane coated in Dextrin Sepharose, which has a high affinity to the MBP protein, allowing protein binding and subsequent purification. Prior to purification, the column was washed with 5 CV (column volumes) of distilled H₂O followed by 15 CV of binding buffer prior to sample purification at a flow rate of 0.5 mL/min. The protein sample was passed through the column at a flow rate of 0.5 mL/min and unbound or weakly bound proteins were washed from the column using 10 CV binding buffer. The protein sample was eluted from the column using the elution buffer (50 mM Tris-HCl pH 8, 5% glycerol, 150 mM NaCl, 10 mM maltose) and fractioned into 1 mL samples. After purification steps were complete, the column was washed with 5 CV 0.5 M NaOH followed by 5 CV water at 1 mL/min. 5 CV of 20% ethanol was added to the column at a flow rate of 1 mL/min and the column was removed from the system, sealed tightly using the screw caps and stored at 4°C until its next use.

Eluted fractions were applied to western blot analysis for the detection of any RAMP-CLR fusion protein eluted from the column. If any proteins were detected, their purity would be tested.

3.2.3.2 IMAC utilising the RAMP-CLR Fusion Protein Histidine Tag

Unless stated, the purification steps were carried out at 4°C. Prior to the histidine tag purification, cell culture media containing the RAMP-CLR fusion protein was concentrated

using the TFF system as stated in section 3.2.2 and buffer exchanged in binding buffer (50 mM Tris-HCl pH 7.4, 10% glycerol and 300 mM NaCl). Post buffer exchange, the media sample containing the RAMP-CLR fusion protein was applied to the AKTA Start Column Chromatography System (GE Lifesciences). A 1 mL or 5 mL HisTrap Excel column (GE Lifescience) was attached to the system. The column is prepacked with Ni Sepharose, consisting of highly cross-linked agarose beads with a chelating ligand, which has been coupled with covalently bound nickel ions. Prior to the application of the sample, the column was washed out with 5 CV of H₂O to remove storage ethanol and equilibrated with 15 CV binding buffer.

The sample and buffers were applied to the column at a flow rate of 0.5 mL/min or 2.5 mL/min for the 1 mL and 5 mL columns respectively. The sample flow through was collected and re-applied over the column for a total of 3 times to ensure maximal binding to the column. The final flow through was collected for analysis. The columns were washed with 20 CV of binding buffer supplemented with 25 mM imidazole to remove the weakly bound proteins to the column. Proteins bound to the column were eluted with elution buffer (50 mM Tris-HCl pH 7.4, 10% glycerol, 300 mM NaCl, 300 mM imidazole) and were pooled into 1 mL fractions. Eluted fractions were applied to western blot analysis probing for the MBP tag to detect any RAMP-CLR fusion protein in the sample, and fractions were also applied to SDS-PAGE with Coomassie blue dye to monitor protein purity. The HiTrap column was washed with 15 CV binding buffer followed by 5 CV H₂O and 5 CV 20% ethanol using a flow rate of 1 mL/min and 5 mL/min for the 1 mL and 5 mL columns respectively. Columns were removed from the system, sealed tightly using the screw caps and stored at 4°C until their next use.

Subsequent fractions containing the RAMP-CLR fusion protein were concentrated using a Pierce Protein Concentrator PES, 10K MWCO, 5-20 mL at 3,700 x *g*, and buffer exchanged 3 times using 15 mL 50 mM Tris-HCl and 300 mM NaCl buffer to a final volume of ~1 mL. The protein was quantified, aliquoted into ~30 µL stocks and stored at -20°C for single use on future functional and drug displacement assays (chapter 4 and 5).

3.2.4 Coomassie Blue SDS-PAGE Staining

Coomassie Blue is a triphenylmethane dye which binds to basic amino acids via a combination of hydrophobic and heteropolar interactions, which allow the visualisation of protein bands in a sample. Protein samples can be separated by SDS-PAGE and stained with Coomassie blue dye, which can indicate how pure a purified protein sample is.

20 μL of protein sample from the protein purification steps were diluted in 25 μL Lamelli buffer (Bio-rad) (62.5 mM Tris-HCl, pH 6.8, 25% glycerol, 2% SDS and 0.01% Bromophenol Blue) and 5 μL 500 mM DTT (ThermoFisher). Samples were heated at 95°C for 10 minutes. 50 μL samples were loaded into a precast 4-20% Mini-PROTEAN TGX protein gel (Bio-Rad) with 10 μL Precision Plus Protein Dual Colour Standards (Bio-Rad). The gels were placed into a Vertical Electrophoresis Cell (Bio-Rad) filled with 1 x Tris/Glycine/SDS buffer (25 mM Tris, 192 mM Glycine, 0.1% SDS, pH 8.3) (Bio-Rad). The samples were run at 150 mV for ~50 minutes. The resulting gel was placed in ~50 mL of QC Collidal Coomassie Blue Stain (BioRad) and allowed to stain overnight on a Stuart shaker at 550 rpm. The gel was placed in H₂O to remove background staining until bands were fully developed. The H₂O was replaced often to ensure full background colour removal. The gel was placed into a transparent plastic film and visualised using the Bio-Rad ChemiDoc XR+ System (Bio-Rad). The purity of the purified protein was determined via densitometry analysis using the ImageLab software (BioRad).

3.2.5 Native PAGE

Protein samples were applied to native gel electrophoresis which would inform us of protein conformation and aggregation state. Prior to the electrophoresis, purified protein samples from the column chromatography (fractions 5-19) were combined in a pre-washed Pierce Protein Concentrator PES, 30K MWCO, 5-20 mL spin column (ThermoFisher Scientific) and spun at 3,700 x *g* in a swing and bucket centrifuge for 15 minutes or until a sample volume of ~500 μL was achieved. The concentrated sample was buffer exchanged 3x with 50 mM Tris-HCl pH 7.4 and 150 mM NaCl, giving a final sample volume of ~1 mL. Concentrated samples were mixed at a 1:1 ratio with native sample buffer (62.5 mM Tris-HCl pH 6.8, 40% glycerol and 0.01% bromophenol blue) (Bio-Rad) to a final sample volume of 20 μL . The entire sample was loaded into a precast 4-20% mini-PROTEAN TGX protein gel (Bio-Rad) in a Vertical Electrophoresis Cell (Bio-Rad) containing 25 mM Tris, 192 mM glycine pH 7.5 buffer. A voltage of 100 mM was applied, and the samples were allowed to run for ~60 minutes or

until samples had adequately migrated down the gel. Gels were stained with Quoitidal Coomassie blue stain (Bio-rad) as previously stated (section 3.2.4).

3.2.6 Western Blot analysis

Purified protein samples were applied to western blots to determine if the fusion protein was being captured by the purification columns. Western blot analysis probing for the MBP tag of the RAMP-CLR fusion protein was carried out as previously stated in chapter 2, section 2.2.13.

3.2.7 Histag ELISA for Protein quantification

Enzyme linked immunosorbent assays (ELISAs) are commonly used plate-based assays which quantify an amount of protein or peptide in a sample. The plate is usually coated in an immobilised antigen or molecule which is specific for the protein of interest. The bound protein is subsequently detected using an antibody linked to an enzyme which is targeted to the protein of interest. In this instance, the surface of the plate was coated in Ni²⁺ ions and blocked with 2% BSA to prevent non-specific binding. The antibody can be detected by using 3,3',5,5'-tetramethylbenzidine (TMB), which is an HRP substrate turning blue during its degradation, with colour intensity having a positive coloration with the amount of HRP in the sample. The reaction can be stopped with sulphuric acid which produces a yellow colour. This can be read at a specific wavelength with increased absorbance indicating a higher protein amount in the well. A schematic of the plate layout can be found in figure 3.6.

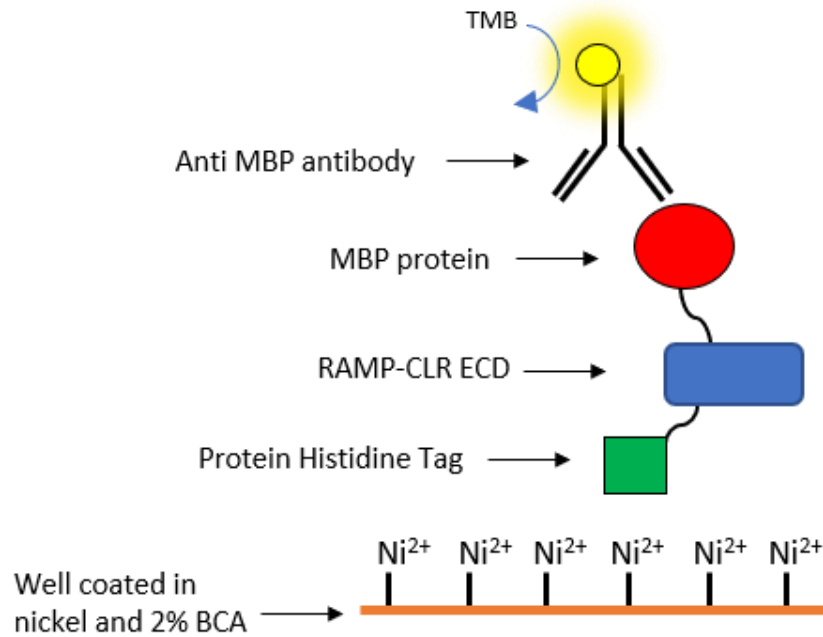


Figure 3.6: Schematic diagram of the Histag ELISA for protein quantification

Purified RAMP-CLR fusion protein samples (fraction 5-19) were combined and concentrated using a Pierce Protein Concentrator PES, 10K MWCO 5-20 mL concentrator (ThermoFisher Scientific) and was buffer exchanged (3 times) in 15 mL 50 mM Tris-HCl pH 7.4 and 150 mM NaCl buffer to remove the imidazole found in the elution step during the IMAC purification. Final collected sample volumes equated to ~1 mL.

Samples were diluted at least 1:1000 in PBS pH 7.4 supplemented with 0.2% tween-20. 100 μL sample was added to a Pierce Nickel Coated Plate, White, 96-well plate (ThermoFisher Scientific) in triplicates along with the standard control protein; CREB1 Recombinant Human Protein N-His.MBP Tag (ThermoFisher Scientific) at specific concentrations (table 3.1). Concentration values of the standard ranged from 0-0.8 $\text{ng}/\mu\text{L}$ (See table 3.1 for dilutions). Samples were subsequently incubated for 1 hour at room temperature on a plate shaker at 570 rpm, followed by 3x wash steps with 200 μL PBS pH 7.4 and 0.2% Tween-20 for 5 minutes with the aid of a plate shaker at 570 rpm. 100 μL of a primary HRP conjugated anti-MBP antibody (Abcam) was diluted 1:4000 (1 μg) and added to each well. Plates were incubated in the dark for 1 hour at room temperature. Sample wells were washed 3x in 200 μL PBS + 0.2% Tween-20 for 5 minutes with the aid of a plate shaker at 570 rpm. 100 μL TMB substrate (ThermoFisher Scientific) was added to each well and the colour was allowed to develop for

~10 minutes. The reaction was stopped using 0.16 M sulphuric acid and colour absorbance was measured at 450 nm using a Perkin Elmer Ensignt Plate (Perkin Elmer). Absorbance values were plotted using GraphPad Prism 7.0 Software (GraphPad) and sample concentrations were calculated by interpolation from the standard curve, with an example shown in figure 3.7.

Table 3.1: Dilution of stock protein for the ELISA standard curves

Tube	Final Protein Concentration (ng/μL)	Dilution buffer (PBS-Tween 20 or Media)	Protein amount
1	0.8	533.83 μL	46.42 μL (10 μg/mL) Stock
2	0.6	120.1 μL	360.25 μL from tube 1
3	0.4	130.16 μL	260.33 μL from tube 2
4	0.2	170.5 μL	170.5 μL from tube 3
5	0.1	121 μL	121 μL from tube 4
6	0.01	198 μL	22 μL from tube 5
7	0	220 μL	0 from tube 6

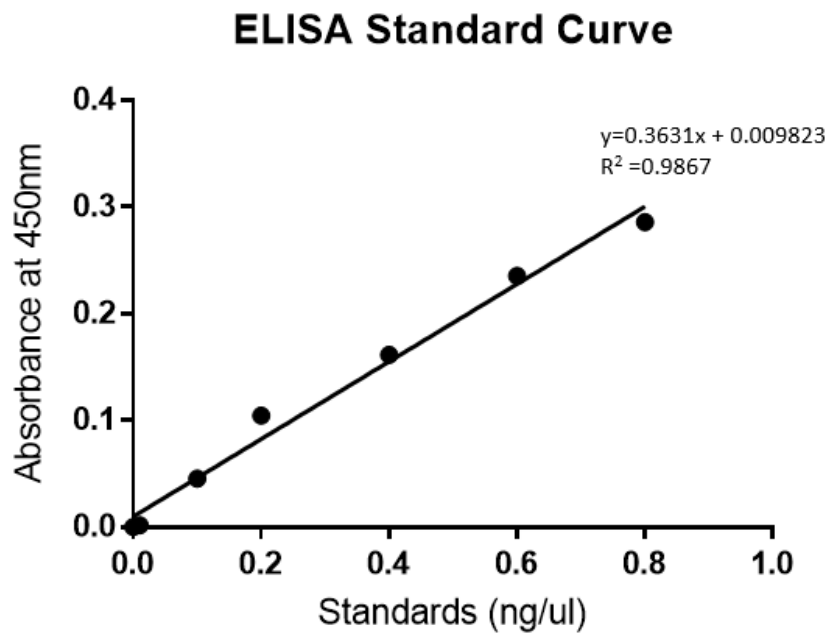


Figure 3.7: Standard curve for the CREB1 Recombinant Human Protein N-His.MBP Tag ELISA.

3.3 Results

3.3.1 Purification Using the MBP Tag

Initially, attempts were made to purify the RAMP3-CLR fusion protein by using the proteins MBP tag. These purifications based on the MBP tag could be compared against future histidine tag purifications, and either be used as a single purification step or be applied as a 2-step approach depending on the protein purity generated.

Despite numerous attempts to capture the RAMP3-CLR fusion protein using this methodology, the protein could not be captured by its MBP tag (figure 3.8). An example of purifying the RAMP3-CLR fusion protein based on its MBP tag is shown in figure 3.8. Western blot analysis shows the protein expression in the pre-purified sample (PP) and no protein eluted from the column in the eluted fractions. All protein (detectable) appears in the flow through following purification, which has not been captured by the column and no protein eluted during column wash.

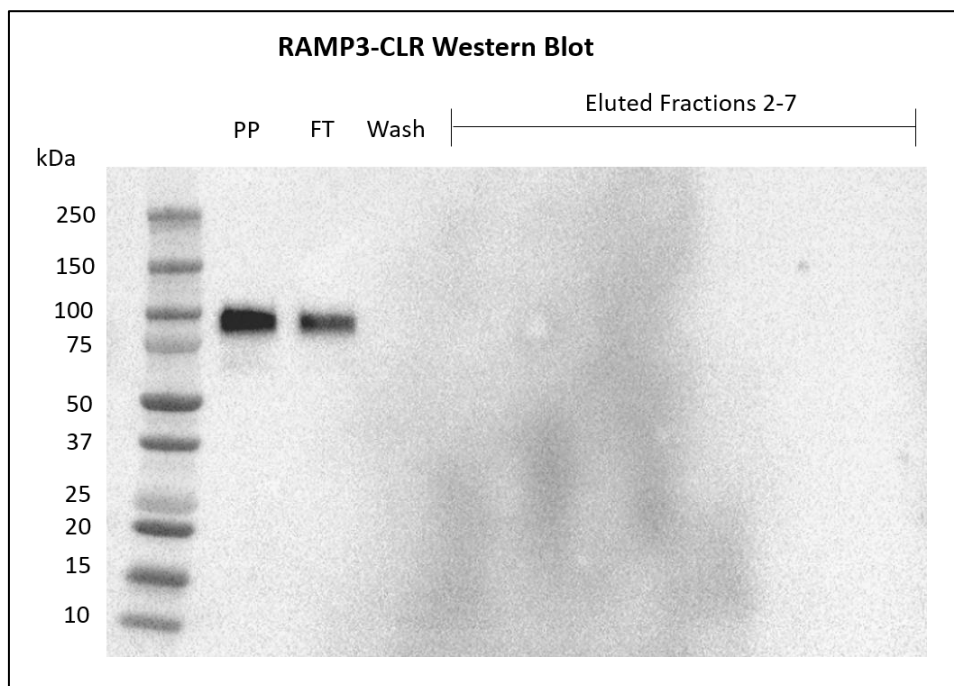


Figure 3.8: Western blot of the RAMP3-CLR fusion protein purification using the MBP tag. Western blot analysis of the RAMP3-CLR fusion protein which has been purified using the MBP tag. The pre-purified (PP) lane shows protein present before the purification. The flow through (FT) shows protein not captured by the amylose column and wash lane shows protein eluted during the column wash steps. The eluted fractions show protein eluted from the column during the elution steps.

3.3.2 Initial RAMP3-CLR Histidine Tag Purification Attempts.

During the initial attempts to purify the fusion protein based on its histidine tag, a 6x histidine tagged protein with no linker from the CLR protein was used, being applied to a 1 mL nickel column. This construct was generated as stated in chapter 2 and the sequence for this can be found in the supplementary section 1. There was no significant difference in protein size in comparison to the 10x histidine tagged protein used during final purification attempts.

The initial IMAC purification attempts were made with the fusion protein secreted into full growth serum medium (supplemented with 10% FCS) with binding buffer containing 1 M NaCl, in comparison to the 300 mM optimised concentration (figure 3.9). The Coomassie blue stain of the pre-purified sample (PP) showed a large bulk of the protein in the media found at ~65 kDa. This protein size was lower than the expected size for the RAMP3-CLR protein (~90 kDa) and therefore likely not to be the protein of interest which was confirmed in the western blot analysis and the fact a small amount of protein of that size was captured by the column.

The eluted/purified fractions show 5 distinct protein bands at ~150 kDa, ~130 kDa, ~90 kDa, ~65 kDa and ~15 kDa in all samples, with the majority of the eluted protein appearing in the later fractions (5-7). The band at ~150 kDa appears the most predominant band in each eluted fraction (figure 3.9). Western blot analysis (figure 3.9) shows the RAMP3-CLR fusion protein corresponds to the band at ~90 kDa, with no indication that the other bands contain the MBP protein, suggesting they are non-specific proteins binding to the column. The entire amount of the RAMP3-CLR fusion protein is not being captured by the column with some being observed in the flow through well (FT). The RAMP3-CLR fusion protein also does not make up the majority of the protein in the eluted fractions, with its purity being estimated at ~10%.

Eluted fractions were combined and quantified, giving a total captured RAMP3-CLR fusion protein amount of ~80 µg. While there were attempts to quantify the amount of RAMP3-CLR fusion protein before it was purified to determine the yield during the purification steps, a large amount of interference was observed in the sample wells during quantification. It was therefore difficult to accurately test the amount of RAMP3-CLR protein gained before purification. Nevertheless, the amount of captured protein and purity was insufficient to be

applied to future assays and structure determination, with a higher protein yield and purity required.

While this purification (figure 3.9) holds a high amount of non-specific binding, the NaCl concentration was required to be increased to 1M during these purification steps in order to reduce the amount of ionic interactions to the column. An example of protein purity before the increase in NaCl concentration can be found in figure 3.10. This resulted in a drastic reduction in background binding, but a high concentration of NaCl was now being used during the purification steps.

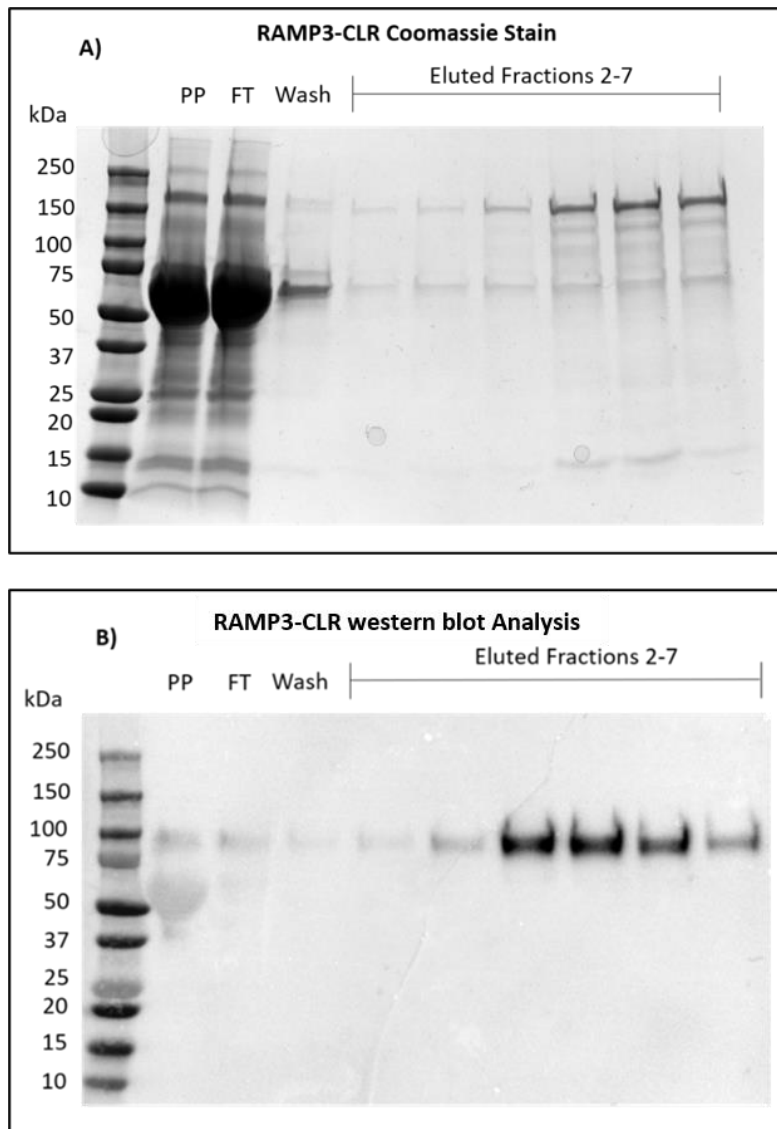


Figure 3.9: RAMP3-CLR fusion protein purity using a 6x histidine tag and FCS supplemented cell growth media. A) The Coomassie blue stain and overall purity of the IMAC in full growth serum media and 1 M NaCl. The pre-purified (PP) lane shows whole protein amount before purification and the flow through (FT) is protein not captured by the column. The ‘wash’ lane is protein removed from the column during wash steps which contain an increased amount of imidazole. Fractions 2-7 show the protein which was eluted from the column. **B)** Shows detection of the MBP tag (western blot) on the RAMP3-CLR fusion protein in the PP, FT, wash step and fractionated samples.

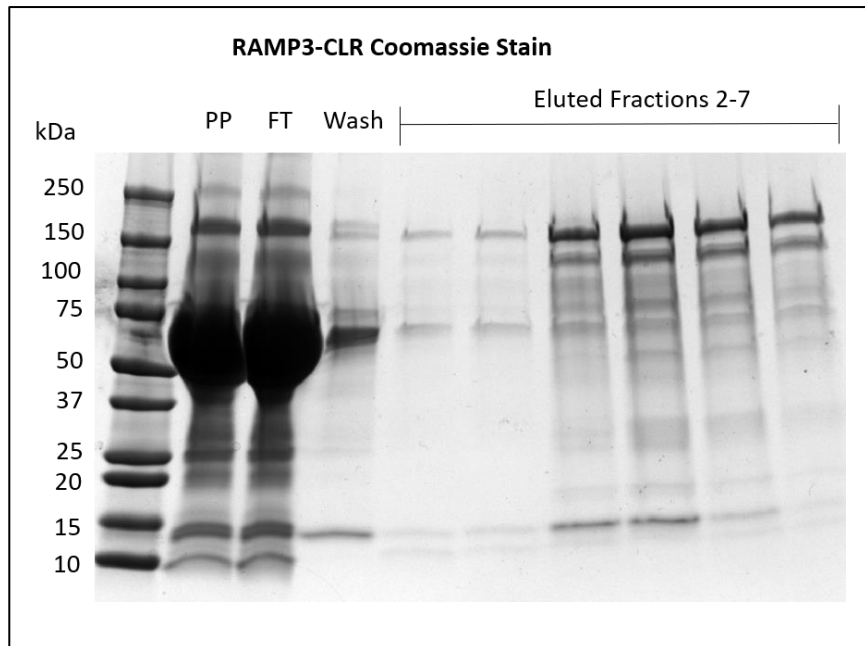


Figure 3.10: RAMP3-CLR fusion protein purity using 300 mM NaCl and FCS supplemented media. The Coomassie blue stain and overall purity of the IMAC in full growth serum media before the increase of NaCl concentration to 1 M. The pre-purified (PP) lane shows whole protein amount before purification and the flow through (FT) is protein not captured by the column. The 'wash' lane is protein removed during the wash steps with increased imidazole concentration. Fractions 2-7 show the protein eluted from the column.

Following the attempts to purify the RAMP3-CLR fusion protein samples after cells were cultured in full growth serum media, culture media was changed to CD FortiCHO chemically defined medium which did not require the addition of FCS. While the removal of the FCS from the media may result in reduced protein expression, it may result in reduced bulk protein in the samples and therefore less competition for the nickel ion sites in the purification columns.

Figure 3.11 shows the purification of the RAMP3-CLR fusion protein after changing culture medium and maintaining the 1M NaCl in the binding buffer. While the majority of the protein remains in the pre-purified well at ~65 kDa, it appears far more reduced than that observed when culturing in the full growth serum medium. The protein found at ~65 kDa in the pre-purified sample could be the FCS remaining in the flask which has not been adequately removed (washed) when changing the media from full growth medium to the chemically defined medium.

The eluted fractions show vastly reduced protein capture in comparison to the full growth serum media (figure 3.11) with 5 faint bands appearing in correspondence with the previous purification attempts (figure 3.9). Western blot analysis (figure 3.11) confirms the capture of the RAMP3-CLR fusion protein in fractions 4-6 (detectable levels) but the majority of the protein remaining in the flow through during purification. Quantified captured amount equated to ~15 µg total RAMP3-CLR fusion protein, which is less than the previous purification using full growth serum media.

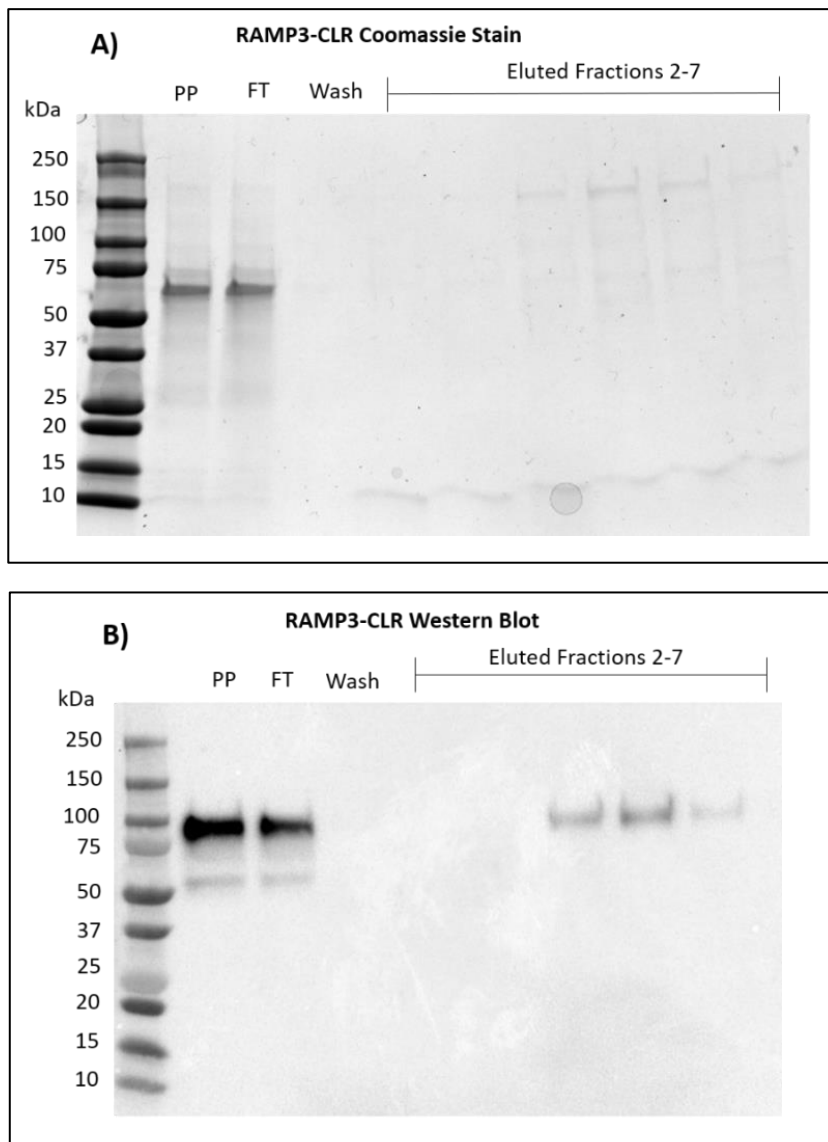


Figure 3.11: Purification of the RAMP3-CLR fusion protein utilising a 6x histidine tag in CD FortiCHO media. A) Shows the Coomassie blue stain and overall purity of the IMAC in full growth serum media and 1 M NaCl. The pre-purified (PP) lane shows whole protein amount before purification and the flow through (FT) is protein not captured by the column. The ‘wash’ lanes shows protein removed from the column during the wash steps containing increased imidazole concentrations. Fractions 2-7 show the protein which was eluted from the column. **B)** Shows detection of the MBP tag of the RAMP3-CLR fusion protein in the PP, FT, wash step and fractionated samples.

3.3.3 Effect of a Longer Histidine Tag on Protein Purification

Previous attempts at purifying the RAMP3-CLR fusion protein which contained a 6x Histidine tag proved to be limited in terms of non-specific binding to the column and low affinity of the RAMP3-CLR fusion protein histidine tag to the nickel ions. This prompted the change to a 10x histidine tag on the fusion protein, which was extended by a glycine linker (6x) from the CLR protein. This construct was generated as previously stated in chapter 2.

The purification and western blot analysis of the purified RAMP1-CLR fusion protein can be found in figure 3.12. With the longer histidine tag, CD FortiCHO media was used to culture the stable RAMP1-CLR fusion protein expressing CHO-K1 cells, a larger chromatography column volume (5 mL) was used in comparison to the previously used 1 mL volume and 300 mM NaCl concentration was used in the binding buffer. The eluted fractions show a much cleaner, purer protein elution from the column although some of the RAMP1-CLR fusion protein was retained in the flow through and wash steps (figure 3.12). The RAMP1-CLR protein was eluted in fractions 7-14 in correspondence to fraction 2-7 from the previous purifications. This was due to the larger bed volume in the column, and therefore protein samples eluting and appearing in the later fractions. RAMP1-CLR fusion proteins appear to be >95% pure with western blots confirming the single band appearing at ~75 kDa to be the RAMP1-CLR fusion protein. After combining fractions, the yield of RAMP1-CLR fusion protein was around 1-1.5 mg per 1 hyperflask of CD FortiCHO media (~560 mL).

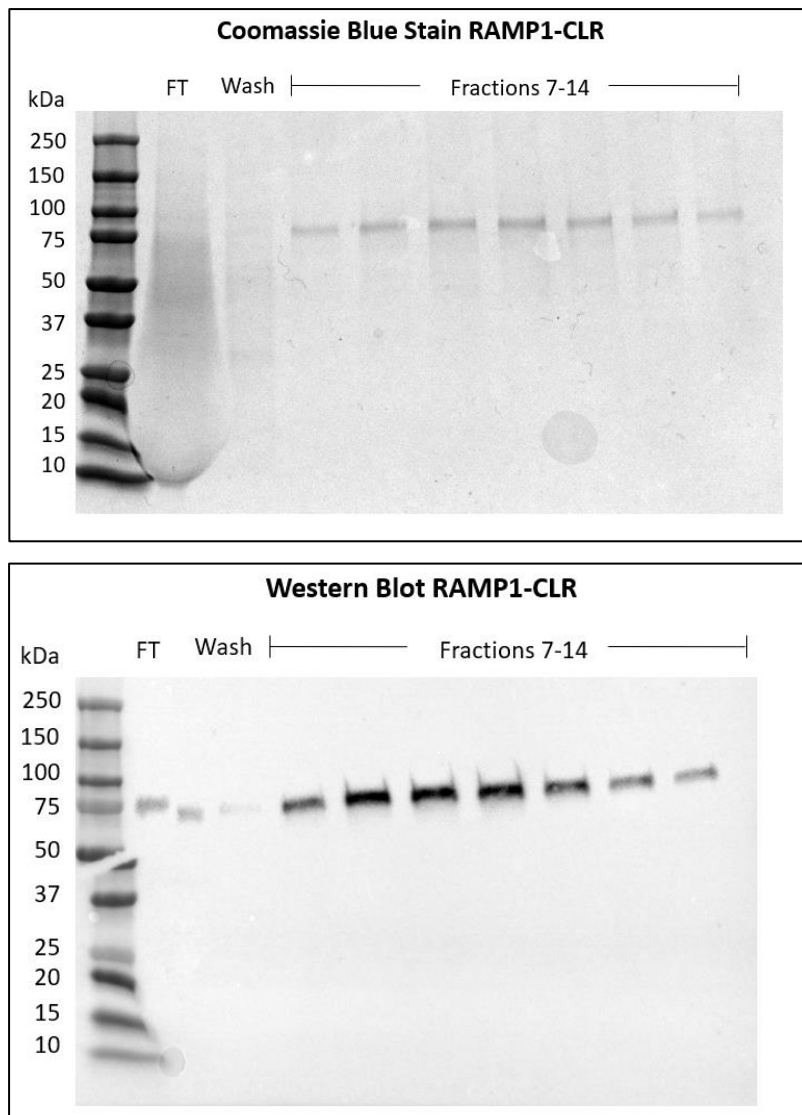


Figure 3.12: Purification of the RAMP1-CLR fusion protein utilising a 10x histidine tag. A) The Coomassie blue stain and overall purity after RAMP1-CLR fusion protein sample application to IMAC. The flow through (FT) lane is protein not captured by the column and the 'wash' lanes show protein removed from the column during the wash steps containing increased imidazole concentrations. Fractions 7-14 show the protein which was eluted from the column. **B)** Western blot analysis shows detection of the MBP tag of the RAMP1-CLR fusion protein in the FT, wash steps and fractionated samples.

3.3.4 RAMP2-CLR Fusion Protein Purification

As seen with the RAMP1-CLR fusion protein, the RAMP2-CLR fusion protein was expressed in CHO-K1 cells and cell media purified via its histidine tag (figures 3.12 and 3.13). When

expressed stably in CHO-K1 cells, although RAMP2-CLR fusion protein was captured during purification steps, expression appeared low and purity was not as high as the RAMP1-CLR fusion protein purification (figure 3.12). The RAMP2-CLR fusion protein in the eluted fractions samples were predicted to be ~60-70% pure (figure 3.13) with no protein being detected in the flow through and wash steps (as seen in the previous RAMP1-CLR fusion protein purification) showing a good overall efficiency in protein capture in the column.

A group of faint bands appeared at ~150 kDa, a single band at ~75-80 kDa and a further faint band at ~50 kDa. The most prominent band appeared at ~75-80 kDa which was detected as the RAMP2-CLR fusion protein in western blot analysis (figure 3.13). While protein samples may be considered pure enough to carry out functional assays, they are not pure enough for structural studies and a second purification step would be required to remove the background binding. The total protein captured was quantified, and the total amount of RAMP2-CLR fusion protein appeared to be ~600 µg from 1 cultured HYPERFlask (~560 mL culture media).

Before purifying larger quantities and optimising secondary purification protocols, the protein was transiently expressed in CHO-K1 cells to increase the amount of protein which could be purified at any one time. The cell culture and purification conditions remained identical to those on the stable CHO-K1 cells expressing the RAMP2-CLR fusion protein, with the exception the vector DNA encoding the RAMP2-CLR fusion protein was transfected into the CHO-K1 cells.

Results in protein purification showed a substantial difference in RAMP2-CLR fusion protein capture and eluted sample purity (figure 3.14). Samples held higher purity, with one single band appearing at ~80 kDa, with minimal or no detectable background binding to the column. This band was confirmed as the RAMP2-CLR fusion protein following western blot analysis, with predicted purity values of >95%. This is a ~30% increase in purity between the transiently transfected CHO-K1 cells and the stable cell line purification. When purifying the transiently transfected RAMP2-CLR fusion protein however, not all the RAMP2-CLR fusion protein was captured by the column indicating column saturation. Combined total amount of quantified protein purified during transiently transfected CHO-K1 cells was ~1.5-2 mg.

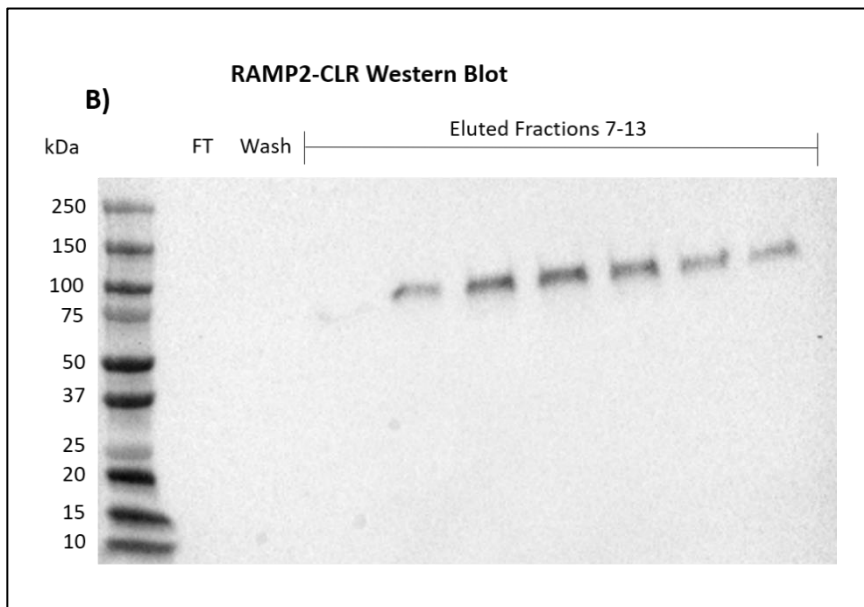
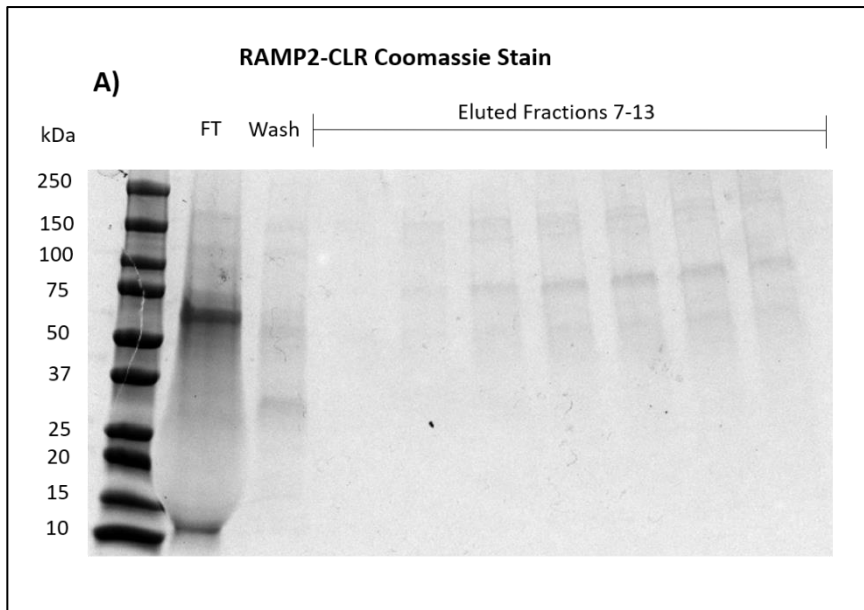


Figure 3.13: Purification of the RAMP2-CLR fusion protein from stable expression in CHO-K1 cells **A)** The Coomassie blue stain and overall purity of the sample after its application to IMAC. The flow through (FT) is protein not captured by the column and the 'wash' lanes show protein removed from the column during the wash steps containing increased imidazole concentrations. Fractions 7-13 show the protein which was eluted from the column. **B)** Western blot analysis shows detection of the MBP tag of the RAMP2-CLR fusion protein in the FT, wash steps and fractionated samples

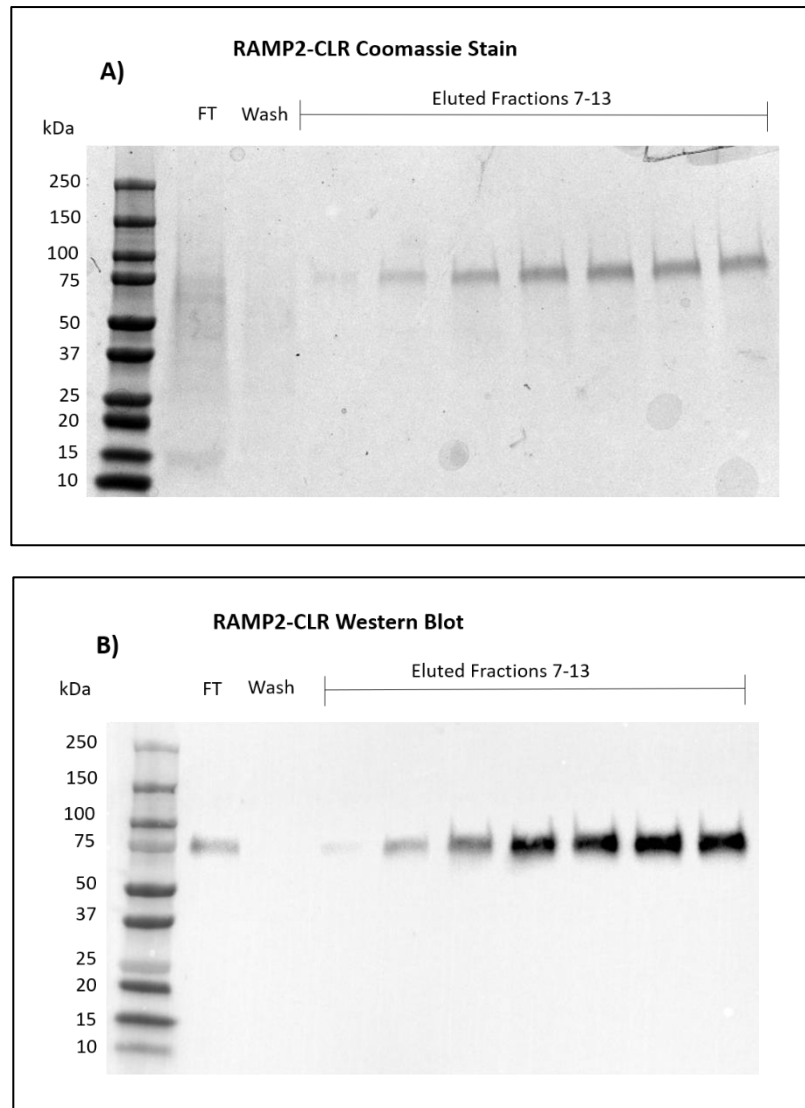


Figure 3.14: Purification of the RAMP2-CLR fusion protein from transiently transfected CHO-K1 cells
A) The Coomassie blue stain and overall purity of the RAMP2-CLR fusion protein sample after its application to IMAC. The flow through (FT) is protein not captured by the column and the ‘wash’ lane shows protein removed from the column during the wash steps containing increased imidazole concentrations. Fractions 7-13 show the protein which was eluted from the column. **B)** Western blot analysis shows detection of the MBP tag of the RAMP2-CLR fusion protein in the FT, wash steps and fractionated samples.

3.3.5 RAMP3-CLR Fusion Protein Purification

Following the previous success of purifying the RAMP1-CLR fusion protein from stably expressing CHO-K1 cells, the RAMP3-CLR fusion was expressed in the same manner. As seen with the RAMP1-CLR fusion protein, the purified RAMP3-CLR fusion protein showed excellent purity based on one histidine tag purification (figure 3.15). A single band appeared

at ~90-100 kDa in the eluted fractions and was predicted to be >95% pure. The western blot analysis detects the RAMP3-CLR fusion protein in the flow through, suggesting all the protein is not being captured and the column is becoming saturated during purification. The combination of the fractions and quantification showed total RAMP3-CLR fusion protein captured equated to ~1.5 mg from 1 HYPERFlask (~560 mL) of stably expressing CHO-K1 cells.

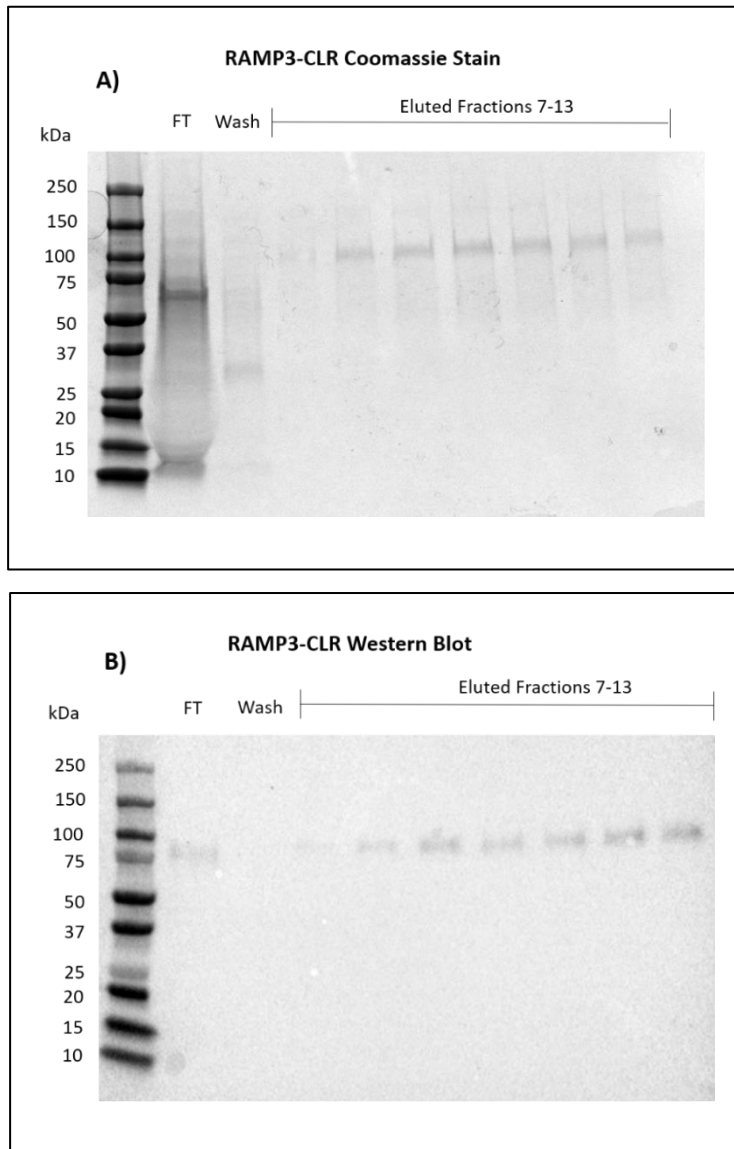


Figure 3.15: Purification of the RAMP3-CLR fusion protein from utilising a 10x histidine tag **A)** Shows the Coomassie blue stain and overall purity of the RAMP3-CLR fusion protein after its application to IMAC. The flow through (FT) is protein not captured by the column and the 'wash' lane show protein removed from the column during the wash steps containing increased imidazole concentrations. Fractions 7-13 show the protein which was eluted from the column. **B)** Shows detection of the MBP tag of the RAMP3-CLR fusion protein in the FT, wash steps and fractionated samples.

3.3.6 Purified RAMP-CLR Fusion Protein Native PAGE Analysis

After the fusion proteins had been purified using column chromatography via their histidine tags, they were applied to native gel electrophoresis to assess the aggregation states of the proteins (figure 3.16). A single band appears in each well containing either the RAMP1-CLR, RAMP2-CLR or RAMP3-CLR fusion proteins indicating the protein is being purified as a single state. The three bands show the RAMP1-CLR fusion protein to have migrated further down the gel, with the RAMP2-CLR fusion protein slightly higher and the RAMP3-CLR fusion protein migrating the least distance.

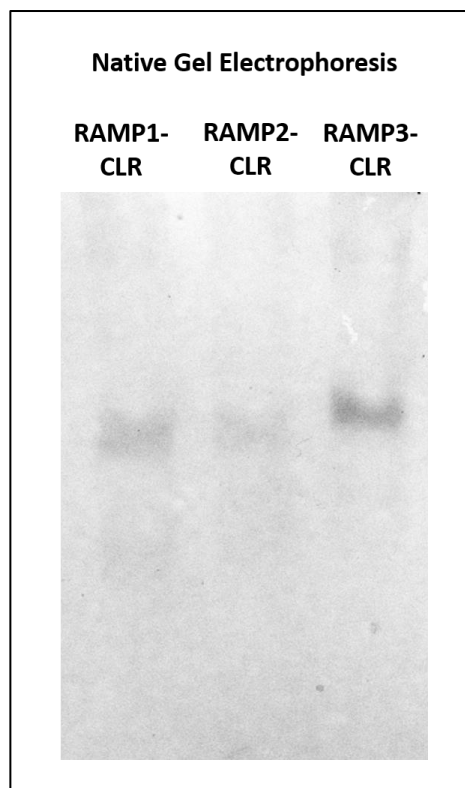


Figure 3.16: Native gel electrophoresis of purified RAMP-CLR fusion proteins. The stained gel after native PAGE shows a single band in all 3 purified protein samples.

3.4 Discussion

Here, it has been demonstrated an efficient and effective 1-step approach for the purification of a RAMP-CLR fusion protein. Purification of the RAMP-CLR fusion proteins via their 10x histidine tag gave protein purities of >95% for all 3 receptor proteins in a single unit state. The methods provided sufficient level of protein capture for future functional assays, fragment screening and structural studies (crystallography screening). In theory, this concept and purification method could be applied to any GPCR ECD linked to a RAMP ECD to give high protein purities.

Initial purification steps used a 6x histidine tag with no linker between itself and the CLR protein which was subsequently changed into a 10x histidine tag with a 6x glycine linker. Different variants of cell culture media were also tested to determine which gave the more desirable characteristics on the purification process. The change of increased histidine tag size and additional glycine linker was the most noticeable change in the purification optimisation steps, with the 10x histidine tag giving vastly increased protein capture and purity in comparison to the 6x histidine tag. While it was thought the longer histidine tag and linker would have a positive effect on protein capture, the purity gained exceeded initial expectations. The longer histidine tag appeared to out-compete the non-specific binding in the column and appeared to almost act as a 'wash step' by competing for nickel sites and removing the background binding to the column.

While theoretically, the 6x histidine tag should have a high affinity to the column, protein capture by the columns in this instance was poor and a large amount of background binding was observed (figures 3.9 and 3.10). The poor association of the histidine tag to the column could partly be due to protein folding and positioning of the fusion protein C-terminus or glycan addition preventing its accessibility. As the tag was not extended from the CLR ECD via a linker, it may have made it difficult to access the nickel in the column, particularly if the MBP protein is folded in such a way where it could interfere with the histidine tag binding. It could also be suggested that the majority of the proteins in the sample (from the FCS in the media) were outcompeting the 6x histidine tag on the RAMP3-CLR fusion protein for nickel ion sites in the column. This needed to be changed in order to increase the amount of fusion protein capture and reduce the non-specific binding to the column.

The change of cell culture medium was crucial to enable proper purification of the protein sample. Initial optimisation steps using the 6x histidine tag saw the requirement of a high

concentration of NaCl (1 M) (figure 3.9) in comparison to 300 mM used in optimised purification steps. While it reduced vastly the amount of non-specific binding in the column from the FCS supplementation in the media, a large amount of NaCl during protein purification may have caused future issues which include protein aggregation when purifying larger quantities.

The change from FCS supplemented media to the CD FortiCHO chemically defined medium to aid purification efforts, captured such minimal amounts of the RAMP3-CLR fusion protein (figure 3.11). It was therefore difficult at that point to determine whether the use of the CD FortiCHO medium was more advantageous than using full growth serum medium in reducing background binding at the cost of reduced protein expression. As it appeared the majority of the RAMP3-CLR fusion protein was not being captured by the column despite little apparent competition from other secreted proteins or those found in the full growth serum media, it suggested that the histidine tag is not readily accessible to the nickel ions in the column and the salt concentration was too high. Therefore, the longer histidine tag and linker was used in order to increase binding to the column while maintaining a lower NaCl concentration in the sample.

3.4.1 RAMP-CLR Column Chromatography

Each RAMP-CLR fusion protein was successfully purified via their 10x histidine tags, reporting a high protein yield and purity in eluted samples (figures 3.12, 3.14 and 3.15). The amount of protein purified in each run varied slightly but the RAMP2-CLR fusion protein gave the higher protein values due to being transiently expressed in CHO-K1 cells. The reason for RAMP2-CLR fusion proteins being expressed only at low levels in stable CHO-K1 cells is unclear, but its purity and capture during purification were significantly increased after transient transfection.

Previous reports by Roehrkassee *et al.*, 2018 showed the purification of RAMP-CLR fusion proteins expressed in mammalian cells with a histidine tag and a MBP protein, can be achieved in a 2-step approach. This was achieved by purifying the proteins by IMAC and further by SEC (Roehrkassee *et al.*, 2018). That report contrasts with this study which generates a purified protein using a 1-step approach based on the proteins histidine tag only. While Roehrkassee *et al.*, 2018 obtained yields of 1-3 mg of purified protein (dependant on the RAMP-CLR fusion protein), the same yield could not be achieved in this study with

protein capture equating to ~1-2 mg of purified protein per run (Roehrkasse *et al.*, 2018). Main differences in protein quantity and yield could be due to transiently expressing the protein in previous studies which generally leads to increased protein expression. Other differences include the use of FCS supplemented media, culturing cells with 4 nM valproic acid (RAMP2-CLR fusion protein) and at 30°C (RAMP3-CLR Fusion protein) to induce mild hypothermia (Roehrkasse *et al.*, 2018). While these conditions could have increased protein expression, they were not tested in this study as the protein capture and yield was sufficient for its application.

While it is important to note that all the expressed protein was not captured during the purification process, in this instance, a sufficient amount was being captured to complete downstream applications. These future studies included functional assays, drug screening and structural studies (crystallography screening). If necessary, the amount of media applied to the purification could be increased along with column bed volume to gain a higher yield. While more of the RAMP-CLR fusion protein could have possibly been captured by increasing the column volume, it could have resulted in a higher amount of non-specific binding to the column. It appeared, particularly during the purifications between the stable and transiently expressed RAMP2-CLR fusion protein, the saturation of the column and using excess RAMP-CLR fusion protein appeared to 'out-compete' the non-specific binding to the column leading to purer protein samples (figure 3.13 and 3.14). This has been previously noted, with the high expression of the GPCR β_2 -adrenergic receptor resulting in >90% sample purity with a single IMAC use due to the protein out-competing the low affinity binding (Hanson *et al.*, 2008). While an increased sample amount could have been used to complement the larger column volume to reduce the non-specific binding it would have increased the overall cost of cell culture methods. In this instance, as the achieved quantity of protein was sufficient for future assays and crystallography screening, increasing bed/sample volume was opted not to be used.

Further purification steps which could have been applied would have included size exclusion chromatography. Although this may have aided the protein purity, it is a time-consuming method of protein separation and avoidance reduced the time of experimental procedures. Although there were no signs of protein aggregation in the native PAGE analysis and all the purified RAMP-CLR fusion proteins appeared generally stable in this study, multiple purification steps and use of protein out of storage can affect its stability and functionality. It is therefore preferred, particularly with crystallography screening, to use the purified protein as soon as possible after the first/last purification steps. However, as only the soluble

ectodomain of the receptor protein was used in this instance, the protein would have been expected to be soluble/stable for future purification steps.

3.4.2 MBP purification

During construct development, the RAMP-CLR fusion proteins were planned to be purified in a 2-step approach. These included IMAC based on the protein's histidine tag and by utilising the MBP tag of the fusion protein which should hold a high affinity for an amylose column. Despite achieving a highly successful capture and protein purity during IMAC, the fusion protein could not be captured by its MBP tag (figure 3.8).

Despite numerous tests and experimental changes which included temperature (4°C- room temperature), buffer/pH differences, flow rates and sample volume, the protein could not be purified based on its MBP tag with minimal amounts being captured. While initially it was suggested glycan addition to the fusion protein makes the binding pocket of the MBP protein inaccessible, this was not the case (data not shown). Glycans were removed from the RAMP-CLR fusion protein in a non-denaturing manner (as described in chapter 4 section 4.2.3) and reapplied to the MBP column. This however had no effect on capturing the MBP protein in the column.

These results contrasts with the previously generated RAMP-CLR fusion proteins expressed in bacteria, which showed successful capture of RAMP-CLR fusion proteins based on their MBP protein and histidine tag (Hill and Pioszak, 2013; Booe *et al.*, 2015). However, more recent reports from the same group, purifying RAMP-CLR fusion proteins after expression in mammalian cells, show purification methods of IMAC and SEC with no reported reason for the change from using the MBP tag (Roehrkasse *et al.*, 2018). It may therefore suggest similar purification characteristic as these attempts, where the RAMP-CLR fusion proteins cannot be captured via their MBP tags when expressed in mammalian cells. This suggests that the RAMP-CLR fusion proteins folds differently in mammalian cells in relation to their protein tags, which could have a subsequent effect on protein functionality. The different positioning of the MBP tag in the RAMP-CLR fusion protein must leave its binding pocket inaccessible to the amylose in the resin, and therefore cannot be captured on the stationary phase. How this could influence the structure of the RAMP or CLR ECD in this instance is unclear at this point, but does not appear to affect protein binding in previous studies, with

reported affinity between the CGRP, AM and AM2/IMD peptides and their requisite receptors (Roehrkasse *et al.*, 2018).

3.4.3 RAMP-CLR Native PAGE Analysis

After purifying the protein using column chromatography, the purified protein samples were combined and assessed using native PAGE. The native PAGE analysis (figure 3.16) shows the purification of a single state protein for the RAMP1-CLR, RAMP2-CLR and RAMP3-CLR fusion proteins. Although from this analysis it is unclear whether the protein exists in a monomeric, dimeric or tetrameric state, the analysis shows the expression and capture of a single state, pure, fusion protein which can be carried forward for functionality testing and crystallography modelling. As native PAGE separates proteins based on charge and size, it can be difficult to interpret the actual size of the protein. Gel migration comparison to SDS-PAGE and to reports by Roehrkasse *et al.*, 2018, who expressed RAMP1/2/3-CLR fusion proteins in a similar manner, suggests the protein is being expressed and captured in a monomeric state (Roehrkasse *et al.*, 2018).

Interestingly, the RAMP2-CLR fusion protein was expressed as a dimer in bacteria and required the mutagenesis of RAMP2 Leu106Arg to prevent this dimerization by disrupting a putative oligomerisation site in the RAMP2 α 3 helices (Kusano *et al.*, 2012)(Booe *et al.*, 2015). The RAMP2 Leu106Arg mutation was applied to recent studies which expressed the RAMP-CLR fusion proteins in mammalian cells but it was not reported whether the mutation was required to express the fusion protein as a monomer in mammalian cells (Roehrkasse *et al.*, 2018). The RAMP2-CLR fusion protein does not appear to be dimerising in this study, suggesting glycosylation of the RAMP-CLR protein is important for monomeric RAMP2-CLR fusion protein expression by preventing the dimerization at this site.

3.5 Conclusion

To conclude, the RAMP1/2/3-CLR fusion proteins were all purified by IMAC to purities of >95% after one chromatography attempt. Attempts were made to purify the protein using its MBP tag, but insufficient or no binding to the column was observed. Native PAGE analysis of the purified protein shows the capture of all 3 RAMP-CLR fusion protein in a single unit state.

Chapter 4: Fusion Protein Functionality and Binding Pocket Occupancy of Novel CGRP and AM₂ Antagonists

4.1 Introduction

The CGRP, AM₁ and AM₂ receptors all have clinical applications where they could be targeted to treat disease (Aggarwal *et al.*, 2013)(Keleg *et al.*, 2007)(Deng *et al.*, 2012)(Zudaire *et al.*, 2006)(Oehler *et al.*, 2001)(Oehler *et al.*, 2003)(Rocchi *et al.*, 2001). Previously, there have been several antagonists developed which show high affinity and selectivity against the CGRP receptor ECD with many being taken forward for clinical trials. (Doods *et al.*, 2000) (Salvatore *et al.*, 2008) (Bell *et al.*, 2010) (Dubowchik, Conway and Xin, 2020). However, there are currently no developed antagonists which show high affinity for the AM₁ receptor and only recent examples of those which target the AM₂ receptor but have yet to be taken to clinical trials (Avgoustou *et al.*, 2020). The structure of the CGRP receptor ECD bound with these antagonists have been developed, outlining key receptor binding pocket residues which aid compound affinity (ter Haar *et al.*, 2010). Comparative analysis of these structures against the ligand free binding pocket of the AM₁ receptor ECD and homology models of the AM₂ receptor ECD, have given an insight into compound selectivity and aided structure-based drug design (SBDD) studies to develop antagonists which are selective for the AM₂ receptor (ter Haar *et al.*, 2010) (Kusano *et al.*, 2012)(Avgoustou *et al.*, 2020).

4.1.1 Olcegepant Binding to the CGRP Receptor

Olcegepant (BIBN4096BS) was developed by Boehringer Ingelheim Pharmaceuticals and was the first drug that could competitively block the effects of the CGRP receptor (figure 4.1). Olcegepant is an extremely potent compound with *in vitro* studies showing an inhibitory constant value (K_i) of 14.4 +/- 6.3 pM, a 150-fold higher affinity than CGRP itself (Doods *et al.*, 2000). Olcegepant retains high selectivity over the closely related amylin, calcitonin and AM₁/AM₂ receptors (Doods *et al.*, 2000). During a multicentre, double-blinded randomised clinical trial (126 patients with migraines) by Olesen *et al.*, 2004, the olcegepant group all showed a significantly higher response in pain-free rate (symptoms of nausea, photophobia and phonophobia) over a period of 24 hours with no adverse side effects observed, in comparison to the placebo group (Olsen *et al.*, 2004)(Rudolf *et al.*, 2005)(Doods *et al.*, 2000). Olcegepant holds slow binding kinetics, with the compound taking around 2 hours to equilibrate on the receptor (association constant K_{ON}), and a K_{OFF} rate (dissociation constant) of 0.0018 min⁻¹ (Schindler and Doods, 2002). Despite the promising results on migraine treatment, the relatively high molecular weight (870 Da) of the compound led to poor oral

bioavailability meaning intravenous administration was the only feasible application, therefore hindering its potential (Doods *et al.*, 2000).

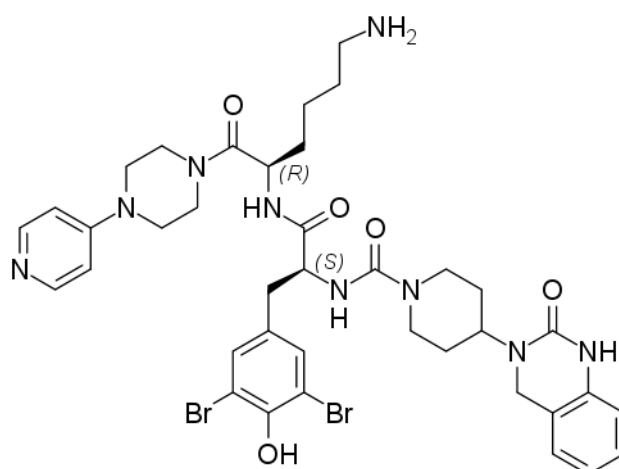


Figure 4.1: Chemical structure of the CGRP antagonist olcegepant

Ter Haar *et al.*, 2010 who reported the ligand free RAMP-CLR ectodomain complex, also report an olcegepant and telcagepant CGRP receptor ECD complex, which revealed key drug-protein interactions. The quinazolinone moiety of olcegepant forms two hydrogen bonds with the NH backbone and carbonyl side chain of CLR Thr122, which enables the compound to extend along the CLR patch (figure 4.2). This results in the stacking of the piperidine ring of olcegepant in an edge-to-face position with the indole group of CLR Trp72 (figure 4.2). This is further aided by the side chain of CLR Trp72 (indole group) undergoing a 70° rotation in comparison to the non-ligand bound complex (ter Haar *et al.*, 2010). A hydrogen bond is also formed here, between the indole NH of CLR Trp72 and the carbonyl group in the backbone of the olcegepant compound. A dibromotyrosyl group extends towards the rear of the binding pocket (toward the RAMP proteins hydrophobic region) where it forms a water-mediated hydrogen bond with CLR Arg38 (α -helices) and RAMP1 Arg67 (α 2) (figure 4.2) (ter Haar *et al.*, 2010). The lysine amino acid terminus of olcegepant extends and forms a salt bridge with the side chain carboxyl group of RAMP1 Asp71, where RAMP1 Trp74 indole group stacks on the dibromotyrosyl aliphatic portion (ter Haar *et al.*, 2010). One final connection between olcegepant and the CGRP receptor is the hydrogen bonds formed with CLR Asp94 from the terminal pyridyl rings, which stack on the aromatic ring of CLR Phe92. The bonds between olcegepant and the CGRP receptor are annotated in figure 4.2.

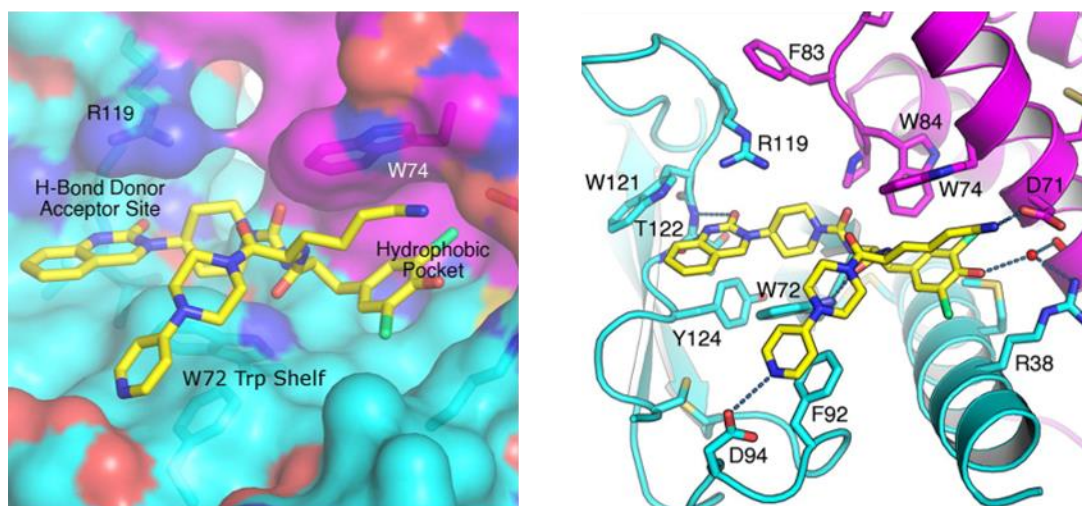


Figure 4.2: Surface and ribbon representation of olcegepant bound to the RAMP1-CLR ECD (PDB: 3N7S). A) Surface representation of how olcegepant fits into the binding pocket of the receptor, with CLR in blue and RAMP1 coloured magenta. B) Cartoon representation of the receptor interaction, with key residues shown in stick figure. CLR is shown in blue and RAMP1 in magenta. Hydrogen bonding between the compound and receptor is highlighted by dashed lines. Image taken from ter Haar *et al.*, 2010 with permission from Elsevier

Olcegepant holds some affinity for the AM_2 receptor (K_B : 407 nM) but is significantly lower than its affinity to the CGRP receptor (Doods *et al.*, 2000) (Hay *et al.*, 2006). As the CLR is likely to remain the same conformation/structure between the CGRP and AM_2 receptors ECD, interactions between CLR and olcegepant are likely to remain consistent between the 2 receptors. This means the loss of affinity must be primarily due to differing interactions toward the RAMP3 protein in the binding pocket. Reduced affinity may be due to the loss of the salt bridge between the olcegepant lysine terminus and RAMP1 Asp71, which is replaced with the polar, uncharged amino acid asparagine in RAMP3. This would result in potential hydrogen bonding only and the loss of a key contact point in olcegepant binding (ter Haar *et al.*, 2010). There is also the change of RAMP1 Trp74 to RAMP3 Glu74, which will result in the loss of the RAMP1 Trp74 indole group stacking on the dibromotyrosyl aliphatic portion of olcegepant. Previous mutagenesis have shown RAMP1 Trp74 is key for antagonist binding (Hay *et al.*, 2006)(Moore *et al.*, 2010). RAMP2 Arg97, equivalent to RAMP1 Ala70 and RAMP3 Thr70, extends further into the binding pocket, altering the size and basic charge at the base of the binding pocket (Archbold *et al.*, 2011). This ultimately changes the structure of the AM_1 receptor binding pocket in comparison to the CGRP and AM_2 receptors and appears to be the primary reason for compound selectivity across all CGRP antagonists.

4.1.2 Telcagepant Association to the CGRP receptor.

The initial success of olcegepant led to the inspiration of other groups to develop different CGRP antagonists that were more suitable for clinical applications, such as those that were orally active. This led to the development of telcagepant (MK-0974) by Merck Research Laboratories and was the first orally developed CGRP antagonist (figure 4.3). Although telcagepant is not as potent as olcegepant, it still showed high affinity and selectivity for the CGRP receptor (table 4.2), with K_i values of $0.78 \text{ nM} \pm 0.05 \text{ nM}$ in SK-N-MC cells and a K_i of $0.77 \pm 0.07 \text{ nM}$ on cells expressing the recombinant CGRP receptor (Salvatore *et al.*, 2008). Despite being less potent than olcegepant, telcagepant showed a much-improved oral bioavailability due to its lower molecular weight, meaning the compound was more clinically applicable (Joshi *et al.*, 2014). These differences are presumably due to the removal of the terminal pyridyl rings and the lysine terminal from olcegepant, leading to less contact areas but a smaller compound. Telcagepant also showed fast association and dissociation kinetics, with a k_{ON} rate of $1.01 \times 10^9 \text{ M}^{-1} \text{ min}^{-1}$ and a k_{OFF} rate of 0.51 min^{-1} (Moore *et al.*, 2009).

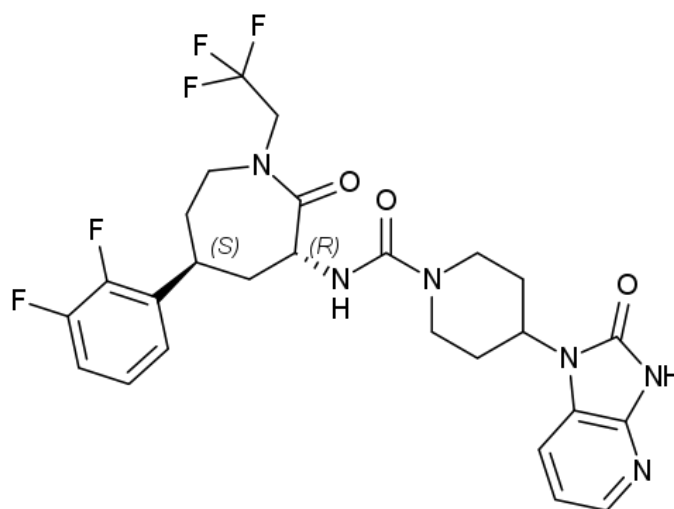


Figure 4.3: The chemical structure of the CGRP antagonist telcagepant

A phase 2 randomised double-blind, parallel dose finding study ranging from 25-600 mg clinical trial of orally administered telcagepant was tested against rizatriptan (10 mg), a 5HT₁-receptor agonist, and placebo to measure the effects of reducing migraine pain (pain-relief, nausea, photophobia and phonophobia). The higher doses of telcagepant were well tolerated, with 300 mg, 400 mg and 600 mg being well comparable in reducing migraine pain with the commercially available rizatriptan (Fan *et al.*, 2008). 3 phase 3 efficacy studies

compared doses of 300 mg telcagepant, 5 mg zolmitriptan (5HT1-receptor agonist) and a placebo. These trials showed that telcagepant and zolmitriptan had similar efficacy levels at all endpoints of pain relief, pain freedom, nausea, absence of photophobia and phonophobia while both being more effective than the placebo at these end points (Ho *et al.*, 2008). Despite these promising results in several stage-II and III studies, there were multiple reports of adverse effects on liver toxicology through aminotransferase activity, which led to the drug being discontinued (Ho *et al.*, 2015)(Ho, Connor and Zhang, 2014)

Ter Haar *et al.*, 2010 were also able to generate the telcagepant bound CGRP receptor model which shows a similar binding topology to olcegepant with some slight differences (figure 4.4). The azabenzimidazolone ring system of telcagepant forms an additional weak hydrogen bond with the backbone of CLR Thr122 in the patch region (H-bond donor acceptor site) in comparison to olcegepant. The edge of the piperidyl moiety of telcagepant stacks on top of the CLR Trp72 indole group, as seen in the binding of olcegepant to the CGRP receptor. The compound extends to the seven-membered caprolactam ring, which forms a weak hydrogen bond with the CLR Trp72 residue via the carbonyl group of telcagepant. A similar interaction is observed in the olcegepant association to CLR Trp72. The telcagepant difluorophenyl group extends further into the binding pocket than the equivalent dibromotyrosyl group of olcegepant. This extension results in telcagepant displacing the water molecule in the binding pocket which was seen to be important for the generation of two hydrogen bonds between olcegepant and the receptor (ter Haar *et al.*, 2010). With this displacement, telcagepant forms primarily hydrophobic interactions with the binding pocket, making contacts with the side chain of CLR Met42, a residue which has previously been implied as important for telcagepant affinity to the CGRP receptor (Miller *et al.*, 2010). The trifluoroethyl group also stacks in an edge-to-face manner with the indole group of RAMP3 Trp74 (ter Haar *et al.*, 2012). A final contact is a hydrophobic interaction arising from the trifluoroethyl region of telcagepant and the side chain of CLR Ile41 (ter Haar *et al.*, 2010). The bonds formed between telcagepant and the CGRP receptor are found in figure 4.4.

Despite olcegepant and telcagepant holding similar topologies to the CGRP receptor, olcegepant is much more potent ($K_D = 45 \text{ pM}$) than telcagepant ($K_D = 1.9 \text{ nM}$) partly due to olcegepant having a much larger molecular weight (870 Da in comparison to 566 Da) and therefore making more contact points to the receptor (Moore *et al.*, 2009)(Schindler and Doods, 2002). These would include olcegepant contacting CLR Asp94 through the extension of its terminal pyridyl rings and the strong salt bridge bond formed between olcegepant lysine extension and RAMP1 Asp71. Despite having a significantly lower MW and making

fewer contact points, telcagepant holds much more constructive contacts with the receptor. These include the additional hydrogen bonds at CLR Thr122 as well as the deeper extension into the binding pocket to form additional hydrophobic interactions with CLR and RAMP side chains. This further includes the displacement of the water molecule which aids binding of olcagepant to the receptor. Ter Haar *et al.*, 2010 describe the displacement of the bound water molecule potentially alters the enthalpic and entropic components of the hydrophobic pocket, which would therefore affect ligand binding (ter Haar *et al.*, 2010). The comparison by overlay of telcagepant and olcagepant to the CGRP receptor is displayed in figure 4.4 and key RAMP/CLR residues for drug binding are summarised in table 4.1.

Telcagepant has minimal binding on the AM₁ receptor (K_i : >100 μ M) and holds little affinity for the AM₂ receptor (K_i : 29 μ M) (Salvatore *et al.*, 2008). These binding affinities are much lower in comparison to olcegepant's potency to the AM₁/AM₂ receptors (table 4.2) (Doods *et al.*, 2000) (Hay *et al.*, 2006). As previously stated with olcegepant, the change of RAMP1 Trp74 to RAMP3 Glu74 will lose the edge to face stacking of the RAMP1 Trp74 indole group with telcagepant difluoroethyl group and subsequent interactions. With olcegepant able to hold more contact points with the AM₂ receptor through its piperazylpiperidine and 6-aminohexyl moieties, it will have a higher affinity than telcagepant, as telcagepant cannot form these extra connections (ter Haar *et al.*, 2010). With a key component for telcagepant binding to the CGRP receptor therefore removed, the additional connections between telcagepant and the AM₂ receptor may not be able to compensate the decreased binding at the Trp74 region and a reduced affinity is observed (ter Haar *et al.*, 2010).

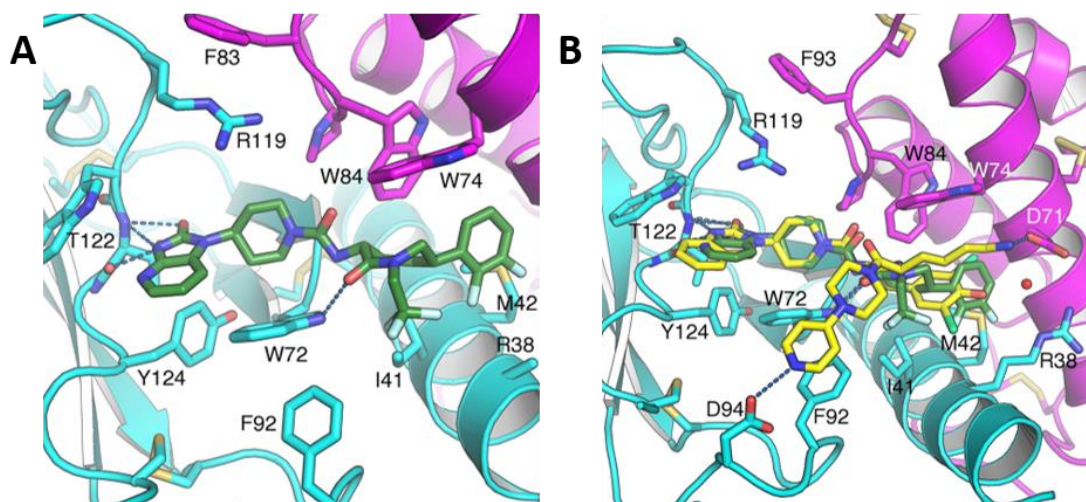


Figure 4.4: Comparison of CGRP receptor binding pocket occupancy by telcegepant and olcegepant
A) Association of telcegepant to the CGRP receptor ECD (PDB: 3N7R). Crystal structure of the CGRP receptor (RAMP1: Magenta, CLR: Light blue) represented in cartoon format with bound telcegepant (green). Key hydrogen bonds between compound and receptor residues are highlighted by dashed lines. **B)** Overlay of olcegepant and telcegepant association to the CGRP receptor ECD. Crystal structure in ribbon format which shows the binding differences between the olcegepant and telcegepant compounds. Noticeable different binding points include olcegepant hydrogen bonding to RAMP1 Asp71 and CLR Asp94. Image taken from ter Haar *et al.*, 2010 with permission from Elsevier

Table 4.1: Key CGRP receptor residues involved in hydrogen bonding to olcegepant and telcegepant

Compound	Key RAMP and CLR residues for antagonist binding
Olcegepant	CLR - residues Thr122, Trp72, Arg67 and Asp94 RAMP1 - residues Arg67, Asp71 and Trp74
Telcegepant	CLR - residues Thr122, Trp72, Met42 and Ile41

4.1.3 Development of MK-3207

During the development of telcegepant, Merck Research Laboratories also discovered the potent CGRP antagonist MK-3207 when investigating the second generation of oral antagonists. Dubiwchik and Xin, 2020, form a comprehensive review on the development of MK-3207, describing the initial formation of a potent ($K_i = 0.04$ nM) tricyclic compound (figure 4.5) which showed poor oral exposures due to its high polar surface and poor

aqueous solubility (Stump *et al.*, 2009). This compound was modified with the goal of improving solubility and led to the formation of MK-3207 (figure 4.5) (Bell *et al.*, 2010).

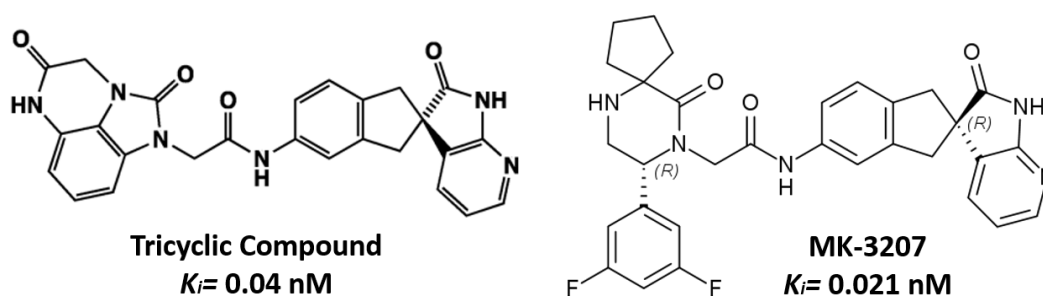


Figure 4.5: The development of CGRP antagonist MK-3207. The tricyclic compound that was used to develop the CGRP antagonist MK-3207

MK-3207 holds a similar structural backbone to telcagepant but does have varying regions. Main substitutions include the addition of an indane group extending from the CLR binding region and the addition of a cyclopentane group replacing the trifluoroethyl group. A final change is the repositioning of the fluorine groups on the benzene ring, from 5,6- to 3,5-difluorophenyl. MK-3207 holds a higher affinity to the CGRP receptor ($K_i = 0.021 \text{ nM}$) than telcagepant (0.78 nM). Presumably, both compounds occupy the same site on the CLR Trp72 and Thr122 due to structural similarities, but extra bonding may occur between MK-3207 and the CGRP receptor. Suspected sites may be weak hydrogen bonding between RAMP1 Trp84 (a key residue for CGRP binding) and the MK-3207 difluorophenyl group and RAMP1 Asp71 and the opposing fluorine group in the MK-3207 difluorophenyl ring (figure 4.6). This is a similar contact as olcegepant, which may give the compound an increased affinity at the AM_2 receptor over the AM_1 receptor, forming contacts at RAMP3 Asn71 in comparison to telcagepant (Salvatore *et al.*, 2010). As with all other CGRP antagonists, RAMP1 Trp74 has also been seen to play a key role in MK-3207 binding to the CGRP receptor (Salvatore *et al.*, 2010) This could be determined to be the main detriment for MK-3207 decreased affinity at the AM_2 receptor, in comparison to the CGRP receptor. The additional binding suggested could compensate the loss of Trp74 in RAMP3 and why the antagonist holds higher affinity to the AM_2 receptor in comparison to telcagepant (table 4.2).

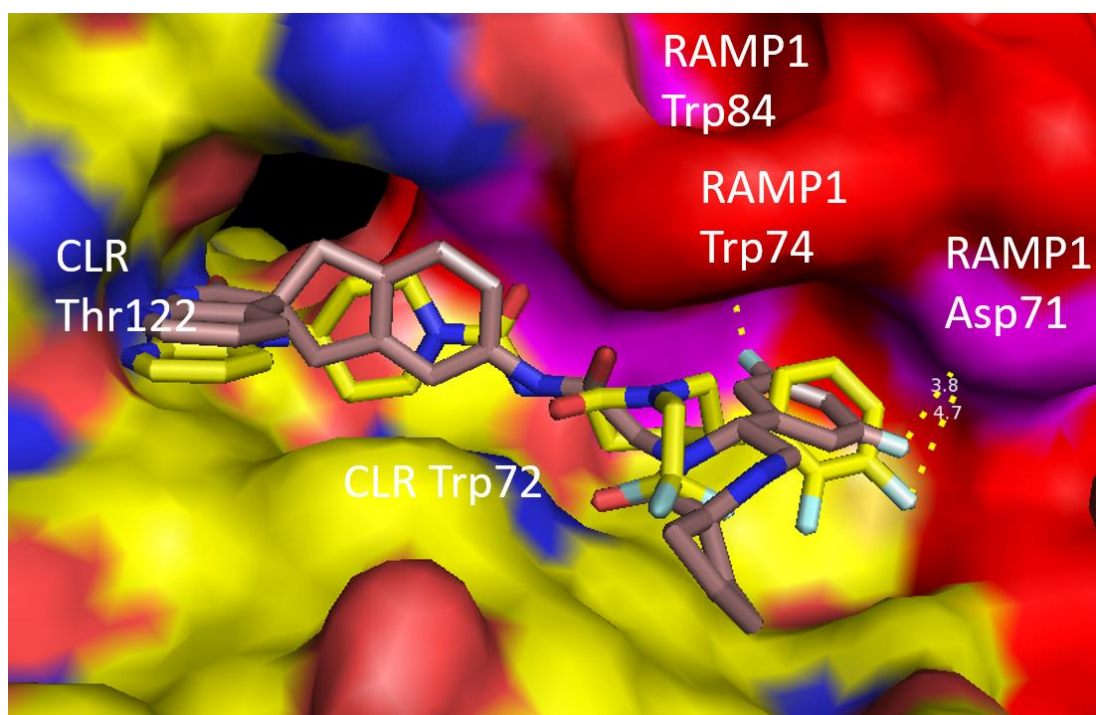


Figure 4.6: Comparison of MK-3207 and telcagepant binding to the CGRP receptor. The crystal model (PDB: 3N7R) of the CGRP receptor with telcagepant (yellow) bound and MK-3207 superimposed (mauve) to predict bond formation with the CLR (yellow) receptor and RAMP1 (red) protein. Magenta colours highlight the potential regions of binding between MK-3207 and the receptor. The distance between the difluorophenyl group of MK-3207/telcagepant to RAMP1 Asp71 is highlighted by the yellow dashed lines and distance given in angstroms (Å). The distance between the difluorophenyl in MK-3207 and RAMP1 Trp84 is further highlighted by yellow dashed lines.

Similar to telcagepant, MK-3207 showed excellent selectivity over the closely related receptor family, AM₁, AM₂ and AMY₃ receptors (> 5,000-fold) but only boasted of moderate selectivity over the AMY₁ receptor (30-fold) (Salvatore *et al.*, 2010). MK-3207 holds a K_{ON} rate of $1.5 \times 10^9 \text{ M}^{-1} \text{ min}^{-1}$ and a K_{OFF} rate of 0.012 min^{-1} for the CGRP receptor (Salvatore *et al.*, 2010). In phase 2 clinical trial testing, doses of 10, 100 and 200 mg MK-3207 all showed significant increase in pain relief in comparison to the placebo, but some patients developed unusual delayed liver enzyme elevations during an extended phase 1 trial which led to MK-3207 being discontinued (Dubowchik, Conway and Xin, 2020) (Hewitt *et al.*, 2011).

4.1.4 The Development of Ubrogепant and Atogepant

Despite telcagepant having limited success due to hepatotoxicity concerns, Merck Research Laboratories believed that the adverse effects were compound related (properties) and not

due to activation of CGRP receptors or non-specific binding in liver cells. Initial thoughts determined this could be caused by the difluorophenyl ring of telcagepant and MK-3207 compounds but the full report of this development which ultimately lead to the discovery of ubrogepant and atogepant structures (figure 4.7) has yet to be published (Dubowchik, Conway and Xin, 2020). Ubrogapant was released as the first commercially available oral CGRP antagonist for treatment of migraine pain. The compounds only differ between their terminal benzene rings, with atogepant changing to a trifluorophenyl group in comparison to the benzene ring of ubrogepant (figure 4.7). The addition of the fluorine groups leads to a slight increase in potency on the CGRP receptor, with ubrogepant holding a K_i of 0.067 nM, while atogepant has a K_i of 0.015 nM. The differing groups in both compounds are the regions which extend towards the RAMP1 protein to make further contact points. The higher potency between ubrogepant and atogepant could be due to atogepants ability to form additional hydrogen bonds with its 2,3,6-trifluorophenyl groups and RAMP1 Asp71 due to their close proximity (figure 4.8).

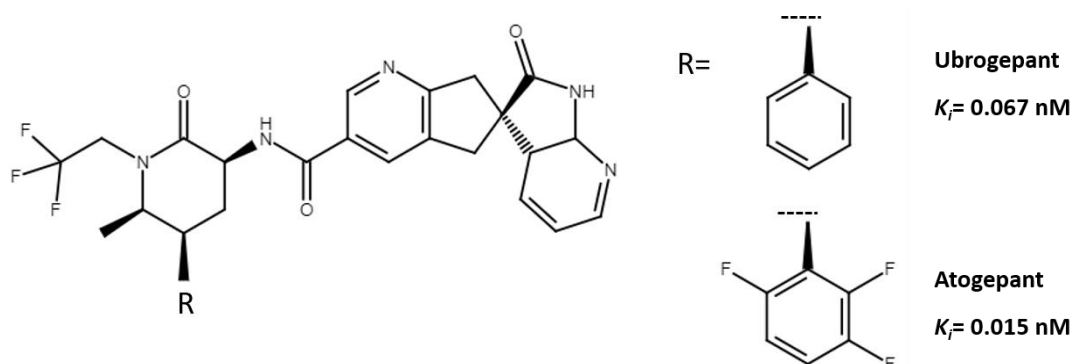


Figure 4.7: Structural comparison of ubrogepant and atogepant

A placebo-controlled, double blind phase 2 dose finding study (n=834) of ubrogepant (1, 10, 25, 50 and 100 mg) showed only the 100 mg dose was effective in pain relief in 2 hours (Voss *et al.*, 2016). In addition to this, several further phase 3 studies showed effective pain freedom at the 100 mg dose with no adverse effects reported, including hepatotoxicity, which was previously an issue with telcagepant and MK-3207 (Dodick *et al.*, 2019)(Lipton *et al.*, 2019) (Trugman *et al.*, 2019). A 12 week phase 2/3 placebo controlled study which patients took either 30 or 60 mg atogepant twice daily, all showed significant reduction in primary efficacy end point (monthly migraines) and showed good tolerance with no reported hepatotoxicity (Dubowchik, Conway and Xin, 2020).

Due to being very structurally similar to telcagepant, it would be safe to predict that contacts would form at similar residues, which may include the CLR Ile41, Met42, Trp72 and Thr122, but what gives ubrogepant and atogepant higher potency than telcagepant is unclear. An example could be an additional hydrogen bond formed between ubrogepant/atogepant and CLR Trp72, with an NH and oxygen group in the compound back bone in close proximity (figure 4.8). Additional contacts may be held between the methyl group of ubrogepant and atogepant extending towards the RAMP1 Trp74 'ceiling' (figure 4.8). It may also be presumed that RAMP1 Trp74 plays a role in compound selectivity and binding as seen with previous CGRP antagonists by forming the top of the hydrophobic pocket and the ligand-protein hydrophobic surface, although a lack of mutagenesis data has not confirmed this (ter Haar *et al.*, 2010) (Mallee *et al.*, 2002).

Despite similarities in structure and functional groups, ubrogepant has a higher affinity for the AM₂ receptor (2059 nM) in comparison to telcagepant (29,000 nM) (Salvatore *et al.*, 2008) (Moore *et al.*, 2020). Superimposing ubrogepant and telcagepant on the CGRP receptor model (figure 4.8) shows the repositioning of the trifluoromethyl group. Whether this repositioning enables stronger interactions between the trifluoromethyl group in ubrogepant and the CLR protein has not yet been determined but could compensate the loss of the RAMP1 Trp74 side chain on RAMP3, leading to a higher affinity on the AM2 receptor. Ultimately, until the crystal structure is found or mutagenesis of these amino acids effect on compound binding is completed, these interactions remain unknown. The binding affinities of olcegepant, telcagepant, MK-3207 and ubrogepant and summarised in table 4.2.

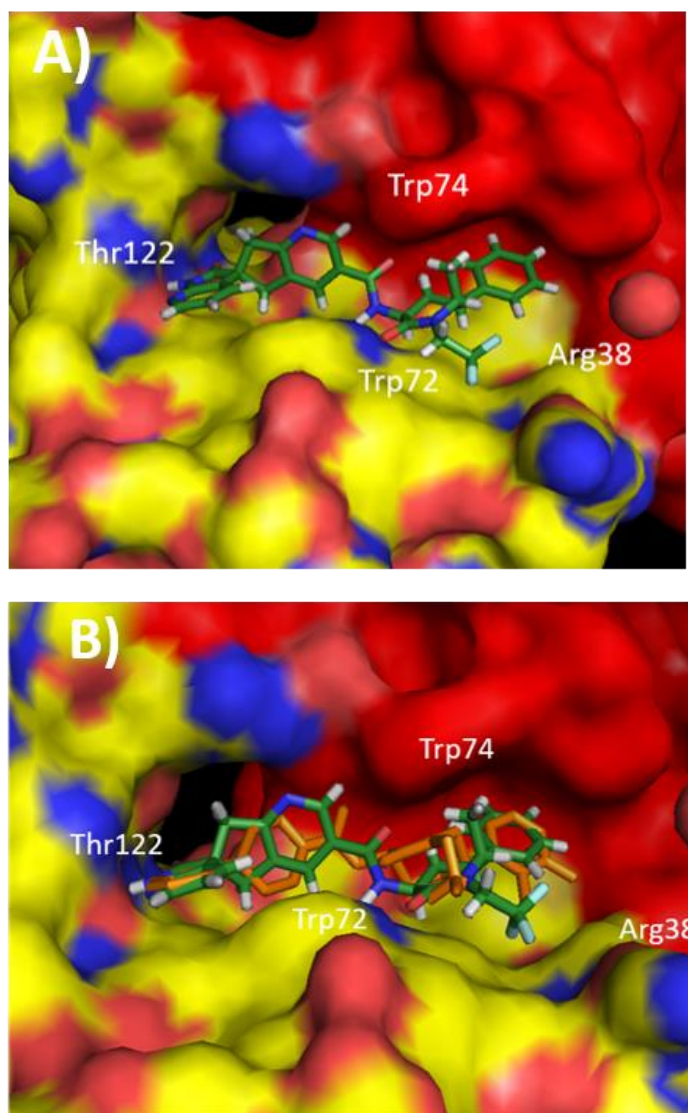


Figure 4.8: Comparison of ubrogepant and telcagepant binding pocket occupancy in the CGRP receptor. **A)** Surface representation of the CGRP receptor with ubrogepant (green) superimposed in the binding pocket. The crystal structure of the CGRP receptor (PDB: 3N7R) with the RAMP1 protein (red) and CLR receptor (yellow). These colours are consistent across the figure. **B)** Surface representation of the CGRP receptor comparing binding pocket occupancy of telcagepant (orange) and ubrogepant. Surface view comparison may give an insight into differing connections between the two compounds and the receptor.

Antagonist	Receptor (human)	Affinity (nM)	K_{ON} Rate	K_{OFF} Rate	Reference
Olcegepant	CGRP	0.01 (K_B)	2 hours	0.0018 min ⁻¹	(Doods <i>et al.</i> , 2000) (Schindler and Doods, 2002)
	AM ₁	>10,000 (K_B)	-	-	(Hay <i>et al.</i> , 2003)
	AM ₂	407 (K_B)	-	-	(Hay <i>et al.</i> , 2006)
	CTR	>10,000 (K_B)	-	-	
	AMY ₁	36 (K_B)	-	-	
	AMY ₃	≤10,000 (K_B)	-	-	
Telcagepant	CGRP	0.8 (K_i)	1.01 x 10 ⁹ M ⁻¹ min ⁻¹	0.51 min ⁻¹	(Salvatore <i>et al.</i> , 2008)
	AM ₁	>100,000 (K_i)	-	-	(Stump <i>et al.</i> , 2009)
	AM ₂	29,000 (K_i)	-	-	
	CTR	>100,000 (K_i)	-	-	
	AMY ₁	190 (K_i)	-	-	(Stump <i>et al.</i> , 2009)
	AMY ₃	>100,000 (K_i)	-	-	
MK-3207	CGRP	0.02 (K_i)	1.5 x 10 ⁹ M ⁻¹ min ⁻¹	0.012 min ⁻¹	(Salvatore <i>et al.</i> , 2010)
	AM ₁	16,500 (K_i)	-	-	
	AM ₂	156 (K_i)	-	-	
	CTR	1900 (K_i)	-	-	
	AMY ₁	0.8 (K_i)	-	-	
	AMY ₃	128 (K_i)	-	-	
Ubgrogepant	CGRP	0.08 (K_i)	-	-	(Moore <i>et al.</i> , 2020)
	AM ₁	>20,000 (K_i)	-	-	
	AM ₂	2059 (K_i)	-	-	

Table 4.2: Pharmacological Properties of CGRP antagonists

4.1.5 Aim and Objectives

Using antagonists which have previously been generated to target the CGRP receptor and utilising SBDD, small molecule antagonist have been developed by our lab which selectively target both the CGRP and AM₂ receptors over the closely related receptor family members (AM₁, AMY and CTR receptors)(Avgoustou *et al.*, 2020). Despite knowing these compounds can antagonise these receptors and prevent cAMP release, how they occupy the receptor binding pocket is still unknown. By exploiting an HTRF (homologous time resolved fluorescence) response using the purified RAMP-CLR fusion protein (outlined in chapter 3), these developed compounds can be tested against each other along with previously developed CGRP antagonists which may give an insight into their receptor site occupancy.

Hypothesis:

- Purified RAMP1-3-CLR ECD fusion proteins will show high specificity and selectivity for CGRP and AM₂ antagonists.
- Deglycosylation of the RAMP1-3-CLR ECDs will not affect the affinity of CGRP and AM₂ receptor antagonists against the fusion proteins.

Aims and Objectives

- Optimise the HTRF response and determine the binding constants of a BODIPY-FL labelled SHF-1036 compound (SHF-1257).
- Investigate fusion protein functionality by displacing the SHF-1257 compound with the calcitonin family of peptides.
- Determine fusion protein selectivity by displacing SHF-1257 with other novel compounds that occupy similar binding sites.
- Displace SHF-1257 with commercially available CGRP antagonists (olcegepant and telcegepant) to give an insight into binding pocket occupancy
- Determine whether glycan addition to the fusion proteins has a role in compound binding.

4.2 Methods

4.2.1 Homogenous Time Resolved Fluorescence (HTRF) Assay: Saturation binding

To investigate the functionality of the purified fusion proteins, samples were applied to an HTRF assay. The assay relies on the binding of both an acceptor and donor molecule to the fusion protein and the energy transmission to the donor from the acceptor, which fluoresces at a specific wavelength upon its activation. In this instance, the donor is an anti-HisTag Terbium cryptate antibody (CisBio) and the acceptor is the novel AM₂ antagonist SHF-1036 tagged with a BODIPY-FL tag (SHF-1257). The binding affinities and structures of both SHF-1036 and SHF-1257 are summarised in table 4.3. The antibody is excited and emits light at a specific wavelength and is responsible for transferring the energy, when in close proximity, to the BODIPY-FL labelled drug (SHF-1257) (figure 4.9). This in turn emits a light at a specific wavelength which is directly proportional to the amount of drug bound to the receptor. Terbium Cryptate has an absorbance peak of ~330 nm which results in 4 distinct emission peaks being observed at 490 nm, 545 nm, 590 nm and 620 nm (figure 4.10). The BODIPY-FL label has a spectral overlap with the terbium between the wavelengths 420 nm and 530 nm (figure 4.10). If energy is transferred from the terbium cryptate to the BODIPY-FL tag, the BODIPY-FL tag will emit light at ~520 nm (figure 4.9).

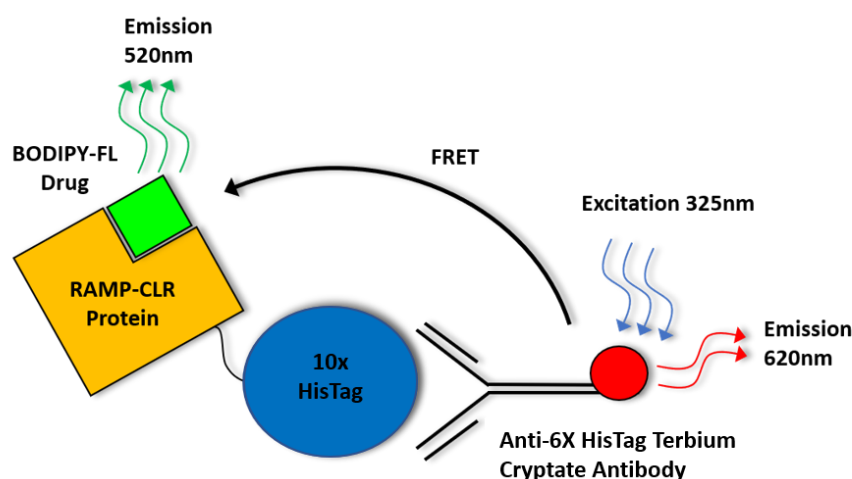


Figure 4.9: Schematic of the HTRF assay. The histidine tag on the fusion protein is targeted by a terbium cryptate antibody which is excited at 325nm. The antibody gives off its own emission of 620nm which was used to normalise the response within each well/reaction. When in close proximity, FRET (fluorescence energy transfer) will occur from the antibody after excitation, which will in turn excite any BODIPY-FL labelled drug which is in close proximity to the antibody ie compound which has bound to the fusion protein. The BODIPY tag will emit a light at ~520 nm which will be directly proportional to the amount of drug bound to the fusion protein.

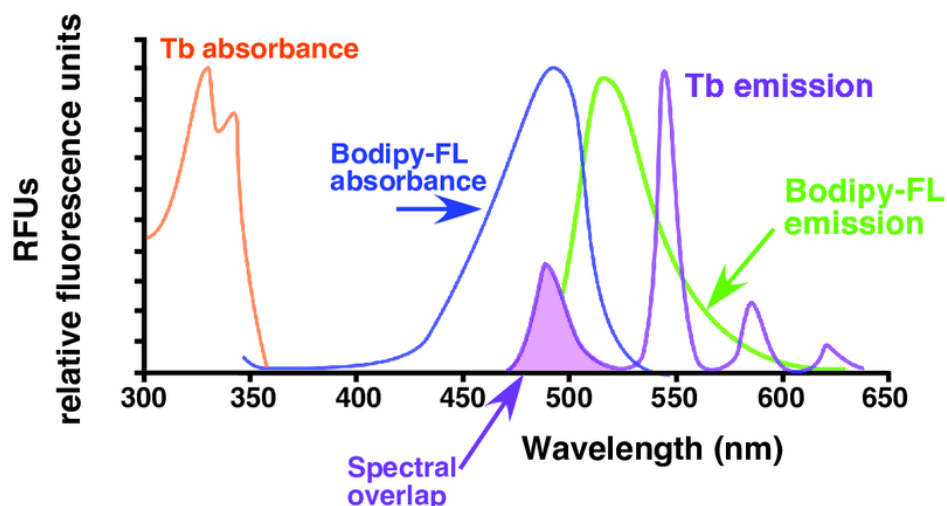
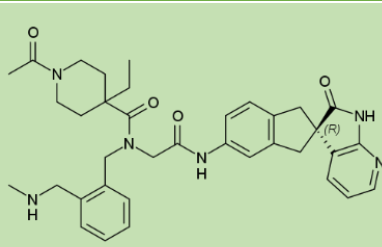
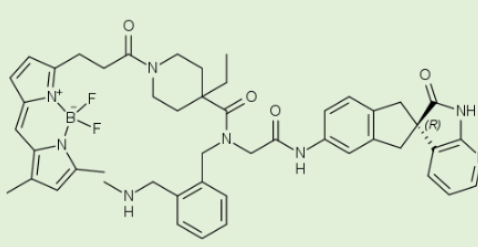


Figure 4.10: Spectral view of terbium cryptate and BODIPY-FL. Terbium cryptate (Tb) absorbance is highlighted in red, having a peak absorbance of ~ 330 nm. The BODIPY-FL absorbance is shown in blue and the spectral overlap of Tb emission and BODIPY absorbance is highlighted in the area coloured purple. BODIPY-FL emission is highlighted in green. Image taken from Nanazashvili *et al.*, 2018 under the terms of the Creative Commons Attribution-NonCommercial-No Derivatives license.

Table 4.3: Structures and binding affinities of SHF-1036 and SHF-1257. Compounds were synthesised by Olivier Zirimwabagabo and Professor Joe Harrity (Department of Chemistry and biological engineering, University of Sheffield) and IC₅₀ values were determined using a cAMP assay by Dr Paris Avgoustou (Department of Oncology and Metabolism, University of Sheffield).

Compound	Compound Structure	IC ₅₀ value (CGRP receptor)	IC ₅₀ value (AM ₂ receptor)
SHF-1036		2.44 nM [pIC ₅₀ 8.61] (M)	1.42 nM [pIC ₅₀ 8.85] (M)
SHF-1257		16.9 nM [pIC ₅₀ 7.77] (M)	32.4 nM [pIC ₅₀ 7.49] (M)

Prior to the experiment, fusion protein samples were purified as previously stated in chapter 3, section 3.2.3.2. Samples and buffers were allowed to cool to room temperature before use. Prior to use, all samples and reagents were diluted into the assay buffer, which contained 25 mM HEPES pH 7.4, 150 mM NaCl, 0.3% Triton-X (Sigma Aldrich) and 0.2% fatty acid free BSA Perkin Elmer). The Described experiments were carried out in reduced lighting (dark) to prevent photobleaching of the SHF-1257 BODIPY-FL tag.

10 μ L of 200 nM RAMP-CLR fusion protein was added to a ProxiPlate 384-Well Plus (PerkinElmer) followed by 5 μ L SHF-1257 BODIPY-FL compound at an in well (final) concentration of 0.01 nM to 50 nM. Novel compounds were synthesised in house at the Department of Chemistry and Biological Engineering, University of Sheffield, by Dr Jean-Olivier Zirimwabagabo and Professor Joe Harrity. All previous cAMP studies which showed the described affinity values (IC₅₀ values) of the compounds to the full-length receptor were carried out by Dr Paris Avgoustou (Department of Oncology and Metabolism, The University of Sheffield). A MAb anti-HisTag Terbium Cryptate donor (Cisbio) (stock: 8.8 ng/ μ L) was diluted 1:100 and 5 μ L was added to each well, giving a final working volume of 20 μ L/well, and plates were incubated at room temperature overnight in the dark to allow the antibody to bind and drug to equilibrate. Each sample was carried out in triplicates and well composition can be seen in figure 4.11. The sample plate was placed in EnSight Plate Reader (PerkinElmer) and was read using the parameters laid out in table 4.4.

To measure background signal, 5 μ L of 400 nM RAMP-CLR fusion protein was added to the well followed by 5 μ L anti-HisTag Terbium cryptate antibody (1:100 dilution prior to addition). 5 μ L of SHF-1257 at an in well concentration ranging from 0.01 nM to 50 nM was added to the plate. 5 μ L unlabelled SHF-1257 compound (SHF-1036) was added to each well in excess, at a final in well concentration of 5 μ M. This would displace any SHF-1257 binding to the receptor and give the background signal of the drug in the well which can be used to generate the 'actual' binding signal of SHF-1257 and the receptor.

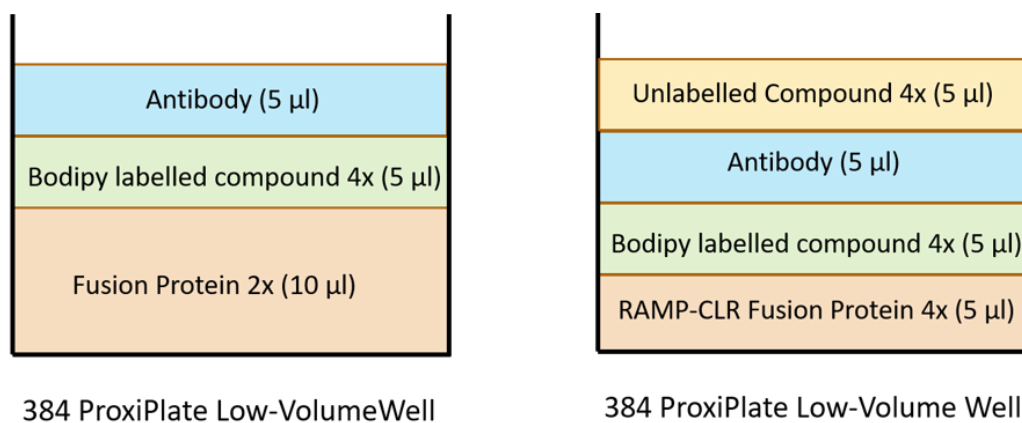


Figure 4:11: Well layout for the HTRF assay. Layouts for the saturation binding curve (left) and the displacement assay (right) are outlined.

Table 4.4: Parameters used during HTRF plate reading

Time-Resolved Fluorescence (RFUs)	
Integration Delay	50µs
Integration	300µs
	Top Read
Wavelengths	
Excitation 1	340 nm
Emission 1	514 nm
Excitation 2	340 nm
Emission 2	620 nm
Sensitivity	
Flashes	200

4.2.2 HTRF Assay: Drug Displacement Response

To investigate which compounds have preferable binding to the CGRP and AM₂ receptors, novel small molecules were competed against SHF-1257 utilising the HTRF assay as previously mentioned (section 4.2.1). The 75% binding maximum (drug concentration) of SHF-1257 for each fusion protein receptor was estimated from the saturation binding curves

and used to carry out subsequent drug displacement studies. As previously stated (section 4.2.1), experiments were carried out under reduced lighting. 5 μ L of 400 nM RAMP-CLR fusion protein was added to a 384-well ProxiPlate (PerkinElmer) in triplicates followed by 5 μ L of SHF-1257 at concentrations of 20 nM and 80 nM for the RAMP1-CLR and RAMP3-CLR fusion protein respectively (final SHF-1257 in-well concentration: 5 nM and 20 nM for RAMP1-CLR and RAMP3-CLR reactions respectively). 5 μ L of novel small molecule antagonist (unlabelled) or human peptides α CGRP (Sigma), AM₁₋₅₂ (Bachem), CT and rat AMY (AnaSpec) were added to the plate in a dose response manner, ranging from an in-plate concentration of 100 μ M to 10 pM. During peptide dilutions, siliconized 1.7 mL/ 0.65 mL microcentrifuge tubes (Sigma) and Ultra fine 0.1-10 μ L (VWR) or Bevel Point 1-200 μ L pipette tips (VWR) were used. MAb anti-HisTag Terbium Cryptate (Cisbio) (stock 8.8 ng/ μ L) was diluted 1:100 and 5 μ L was added to the plate. The final working (in-well) volume equated to 20 μ L and well layout can be seen in figure 4.11). The solution was allowed to equilibrate overnight at room temperature and the HTRF signal was measured as previously stated (section 4.2.1).

4.2.3 Native Protein Deglycosylation

In order to assess the importance of the added glycans on peptide binding, the sugars needed to be removed before being applied to the HTRF assay (section 4.2.1). 20 μ g of purified RAMP1/3-CLR fusion proteins were added to 5 μ L GlycoProfile Buffer 2 (Sigma Aldrich) and 2 μ L PNGase F enzyme (New England Biolabs). DNase and RNase free water was added to the reaction to make up a final volume of 50 μ L. Samples were placed in a ProFlex PCR System Thermocycler (ThermoFisher Scientific) for 10 hours at 37°C and stored at 4°C until use. The efficiency of protein deglycosylation was assessed using the western blot analysis as stated in chapter 2, section 2.2.12, 2.2.13, and 2.2.14.

4.2.4 Hisbead purification

Prior to the assessment of glycan addition on the fusion protein functionality, impurities and PNGase F enzyme needed to be removed from the sample. Protein samples were applied to Dynabeads HisTag isolation and Pulldown Kit (ThermoFisher Scientific). Dynabeads are magnetic beads coated in cobalt ions (Co²⁺) and their purification is based on immobilised metal affinity chromatography, where the 10x histidine tag on the fusion protein has a high

affinity for the cobalt ions on the beads. This allows the isolated proteins to be removed from other impurities by applying the sample to a magnet for separation. The isolated protein can then be eluted from the beads with the addition of imidazole in the wash buffers.

Prior to the isolation, protein samples were diluted to 700 μL (20 $\mu\text{g}/700 \mu\text{L}$) in the binding buffer (25 mM HEPES pH 7.4, 150 mM NaCl, 0.3% triton-X100 and 0.5% BSA). Dynabead stocks were vortexed thoroughly prior to use to ensure proper suspension and 50 μL (2 mg) was added to a 1.5ml Eppendorf tube (Sigma Aldrich) and applied to a magnet for 2 minutes or until the beads had separated from the solution. The supernatant was discarded, and beads were washed 2 times for 3 minutes in 500 μL binding buffer with the aid of a stuart roller, rotating at 14 rpm. Supernatant was discarded each time during bead separation. After the wash steps were complete, the entire protein sample (700 μL) was added to the magnetic beads and allowed to mix on a stuart rotator for 15 minutes at 14 rpm. The sample was applied to the magnet for 2 minutes and the supernatant discarded. Each sample containing magnetic beads and isolated protein was washed 4x for 3 minutes by adding 500 μL of the binding buffer and placing on a stuart rotor at 14 rpm. The samples were applied to the magnet and supernatant discarded between each wash. The protein was eluted from the magnet by adding 300 μL elution buffer (binding buffer containing 300 mM imidazole). The samples were incubated at room temperature on a stuart rotator for 10 minutes and applied to the magnet for 2 minutes to separate the beads from the eluted protein. The supernatant was collected and applied to a Pierce Protein Concentrator PES, 10 MWCO, 0.5 mL (ThermoFisher) spin column. Columns were spun at 13,000 x *g* and buffer exchanged 3x in 400 μL 25 mM HEPES pH 7.4 and 150 mM NaCl buffer. Protein samples were diluted to 50 μL prior to their collection from the spin column. Protein samples were quantified using the HisTag ELISA as stated in chapter 3, section 3.2.7.

4.2.3.1 Deglycosylation assessment on protein functionality

During assessment of glycan removal on protein functionality, 5 μL of SHF-1257 at concentrations of 20 nM and 80 nM for the deglycosylated RAMP1-CLR and RAMP3-CLR fusion protein reactions respectively were added to the 384-well proxiplate plus plate (Perkin Elmer) followed by 5 μL of 400 nM deglycosylated RAMP1-CLR or RAMP3-CLR fusion proteins. Non-deglycosylated protein (WT) was also run and used as the 100% maximum signal. This was followed by 5 μL of the reaction buffer (25 mM HEPES pH 7.4 and 150 mM

NaCl, 0.3% Triton-X and 0.2% fatty acid free BSA) or 100 nM of SHF-770 or SHF-771 as displacement controls. MAb anti-HisTag Terbium Cryptate (Cisbio) (Stock 8.8 ng/ μ L) was diluted 1:100 in the reaction buffer and 5 μ L added to each well. The final working (in-well) volume equated to 20 μ L. The solution was allowed to equilibrate overnight at room temperature. The HTRF signal was measured as previously stated (table 4.4).

4.2.5 Data analysis

Data generated from the HTRF responses were plotted into graphs using the GraphPad Prism Version 7.0 (GraphPad software, Inc). Values were first normalised by the fret signal measured (520 nm) divided by the antibody emission signal (620 nm) in each individual well. During saturation binding curves using SHF-1257, a one-site specific binding curve was applied to the data sets after removal of the background binding values (wells with excess unlabelled compound). This generated the maximum signal (Bmax) and the Kd values of SHF-1257 association to the receptors, allowing the comparison of SHF-1257 affinity for each receptor. The data was further normalised to 100% binding (Bmax value) and 0% binding (no protein in the sample) and plotted with a one-site specific binding curve to gain percentage values of drug occupancy on the receptor at specific concentrations. From this, 75% of the Bmax for drug association to each receptor was used for subsequent HTRF displacement studies.

During non-labelled compound displacement, data was normalised as mention to the FRET signal measured against the antibody emission signal. The data sets were further normalised to the 100% maximum SHF-1257 binding (addition of SHF-1257 with no unlabelled antagonist) and 0% SHF-1257 binding (excess unlabelled antagonist). Data was plotted in GraphPad Prism with a non-linear 4 parametric curve. During ligand displacement assays, due to the dose response curve being incomplete (complete inhibition of SHF-1257 binding not achieved), a constraint of equal to 0 was applied to bottom of the curve to gain more accurate IC50 values. Graphs were plotted with standard deviation (SD) bars unless stated in the figure legend.

4.3 Results

4.3.1 Spectral Scan of HTRF Response

As previously stated in the material and methods section, to generate a HTRF response, a terbium cryptate antibody against the receptor proteins histidine tag was used as an electron donor and a BODIPY-FL labelled novel AM₂ antagonist (SHF-1257) designed to target the CGRP and AM₂ receptors was used as an electron acceptor. This would generate an emission of light at a specific wavelength (~520 nm) which would directly correlate to the amount of drug bound to the receptor.

First off, the peak emissions of the antibody and the FRET signal were investigated by measuring a spectral scan of the wells. Drug concentrations ranging from 0.1 nM – 10 μM were left overnight to equilibrate and the spectral scan of each dose was measured in comparison to a no FRET control (no SHF-1257 drug). The spectral scan of the anti-HisTag terbium conjugated antibody is displayed in figure 4.12. The graph shows the 4 energy emission peaks of the terbium cryptate antibody which are located at ~487 nm, 547 nm, 585 nm and 620 nm, with the peak spectral overlap of terbium cryptate and BODIPY-FL being between ~490 nm and 530 nm. This correlates with the reported emission spectrum of terbium cryptate (figure 4.10). The addition of SHF-1257 shows the increase of another peak across an emission range of 500-520nm (figure 4.12). This corresponds to the BODIPY-FL emission spectra, which has a reported peak emission of ~520nm, having a direct correlation with increased drug concentration increasing the signal produced at ~520 nm. The emission data also shows an indirect correlation, as the concentration of SHF-1257 is increased (subsequent 520 nm emission peak), there is a decreased emission peak at the 490 nm, which corresponds to a terbium cryptate emission peak but also the spectral overlap between the acceptor and donor (figure 4.12). This implies the BODIPY-FL tag is absorbing the light at this specific wavelength

Peak emissions from the FRET response appear to be slight dose dependant, with lower SHF-1257 concentrations (0.1 nM – 100 nM) showing peak emission of ~514 nm while the emission of the higher concentrations (1 μM-10 μM) shifting to values ranging from 515-525 nm, with the spectra more scattered and disorientated at the higher concentration (10 μM), although the spectra still follows the same trend. This may indicate the HTRF response could have a BODIPY-FL concentration limit of around 100 nM before background signalling begins to be indistinguishable from the real signal generated by SHF-1257 binding to the receptor.

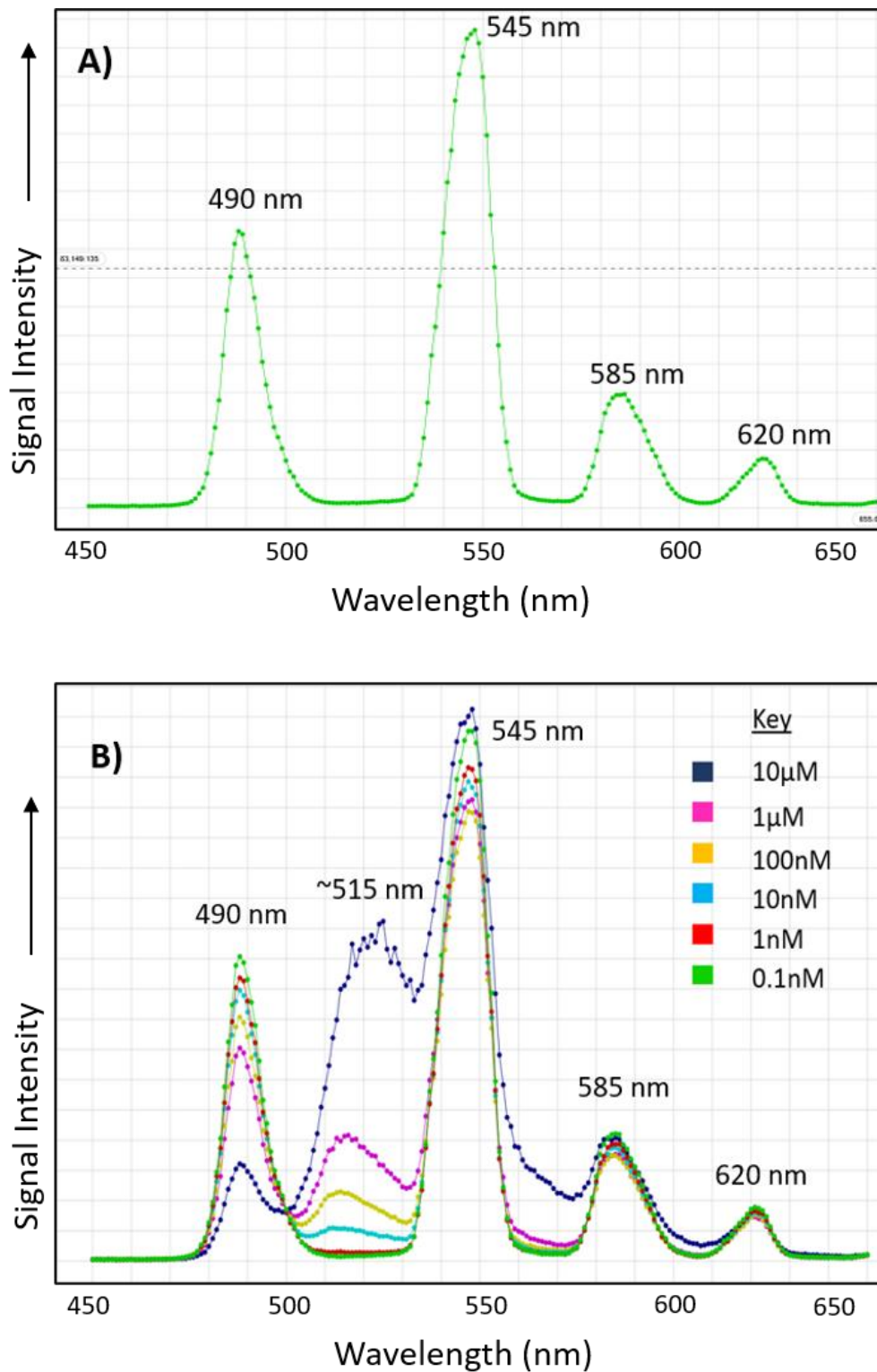


Figure 4.12: Spectral scan of the HTRF assay **A)** The spectral scan of the terbium cryptate antibody without SHF-1257. The standard emission peaks after excitation at 340 nm are annotated which include those at 490 nm, 545 nm, 585 nm and 620 nm. **B)** The spectral scan of terbium cryptate antibody with SHF-1257. The emission peaks are shown after the excitation of the well at 340 nm. There is an additional peak in comparison to -ve SHF-1257 compound, indicating a FRET response from the antibody to the BODIPY-FL tag.

4.3.2 Saturation Binding Curves of the RAMP1/3-CLR Fusion Proteins

After spectral data confirmed that the terbium cryptate antibody was able to transfer energy efficiently to the BODIPY-FL tag on SHF-1257, the compound could now be applied to investigate SHF-1257 K_d values on the purified RAMP-CLR fusion protein. Figure 4.13 shows the saturation binding curve of the SHF-1257 compound on both the RAMP1-CLR fusion protein and the RAMP3-CLR fusion protein. SHF-1257 shows an increased affinity for RAMP3-CLR (K_d: 2.85 nM) over the RAMP1-CLR fusion protein (K_d 11.2 nM) (table 4.5). Previous preliminary cAMP assays show IC₅₀ values for SHF-1257 on the full length receptor are 32.4 nM and 16.9 nM for the AM₂ and CGRP receptors respectively (data not shown). The HTRF data generated shows similar levels of inhibition in terms of concentration, but cAMP data shows the CGRP receptor holds a higher inhibitory value for SHF-1257 in comparison to the AM₂ receptor. The binding and curve values from the saturation binding are summarised in table 4.5.

The 75% maximum binding for the compounds on each receptor was calculated by interpolating the concentration from the saturation binding curve, giving values of ~20 nM for the RAMP1-CLR fusion protein and ~5 nM for the RAMP3-CLR fusion protein.

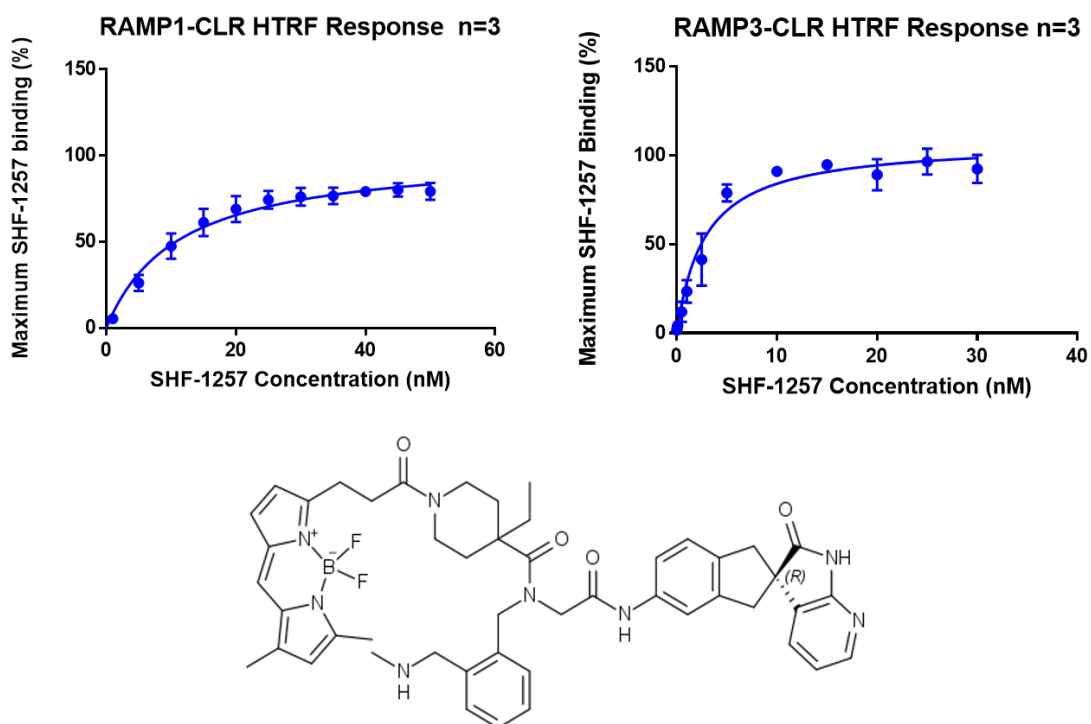


Figure 4.13: HTRF saturation binding curves of SHF-1257 on the RAMP1/3-CLR fusion proteins. SHF-1257 concentration is plotted against the maximum binding to the receptor, with a one-site K_i curve fitted to the plot. The structure of SHF-1257 used for binding saturation is shown.

Table 4.5: Values of SHF-1257 saturation binding to the RAMP1/3-CLR fusion proteins

Onsite Specific Binding: Curve Binding	Fusion Protein	
	RAMP1-CLR	RAMP3-CLR
Dissociation Constant (Kd)	11.2 nM	2.85 nM
Std Error: Kd	+/- 1.097	+/- 0.2942
95% profile likelihood (Kd)	9.236 to 13.37	2.331 to 4.412
R square	0.9439	0.964
75% Maximum	~20 nM	~5 nM
Bmax (Top)	101.8	107.7

A range of SHF-1257 concentrations were tested in order to gain the saturation curve of SHF-1257 on the receptor, ranging from 0.01 nM – 10 μ M. SHF-1257 concentrations above 200 nM were avoided due to the non-specific binding masking the real signal generated by drug binding. As preliminary cAMP inhibition data using SHF-1257 on AM₁ receptors showed affinity at only high concentrations (IC₅₀ 36.5 μ M), it subsequently meant we could not apply the SHF-1257 for binding responses to the RAMP2-CLR fusion protein. Therefore, the binding and displacement assays were focused on the RAMP1-CLR and RAMP3-CLR fusion proteins which showed SHF-1257 affinity below 200 nM.

4.3.3 Displacement of SHF-1257 Using the Calcitonin Family of Peptides

Peptides from the calcitonin family were used to assess the displacement of SHF-1257 and their affinity to the fusion proteins. Concentrations were applied in a dose response manner ranging from 100 μ M-1 nM. On the RAMP1-CLR fusion protein, IC₅₀ values equated to ~12.49 μ M and ~8.19 μ M for the CGRP and AM peptides respectively with CT, AMY and IMD having little or no displacement of SHF-1257 (figure 4.14).

Ligand IC₅₀ values on the RAMP3-CLR fusion protein were ~13.56 μ M, ~1.89 μ M, ~20.81 μ M and ~5.091 μ M for CGRP, AM, AMY and IMD respectively with the calcitonin peptide showing no displacement of SHF-1257 (figure 4.14). The IC₅₀ values of the calcitonin family of

peptides on displacing SHF-1257 from the RAMP1/3-CLR fusion proteins are summarised in table 4.6.

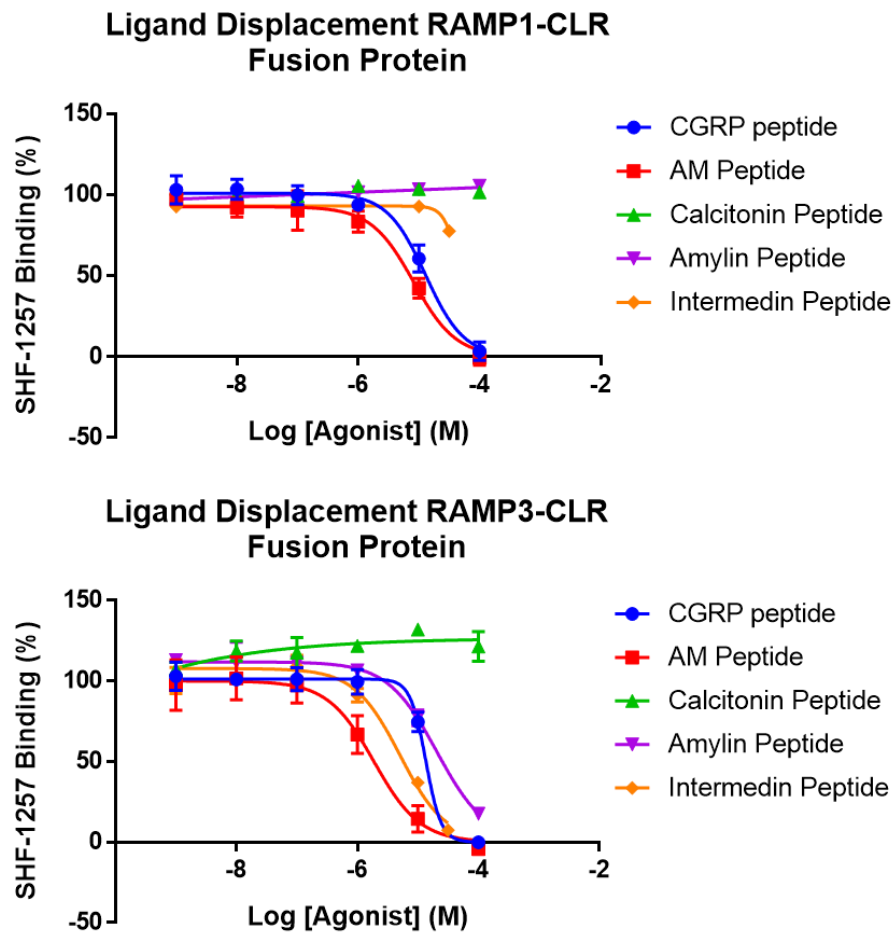


Figure 4.14: Ligand displacement curves. Dose response curves of CGRP, AM, CT, AMY and IMD displacing SHF-1257 from the RAMP1-CLR and RAMP3-CLR fusion proteins.

Table 4.6: Parameters of ligand displacement curves on the RAMP1-CLR and RAMP3-CLR fusion proteins

Fusion Protein	Parameters				
	IC50	Std Error LogIC50	R Square	Best Fit Top	Best Fit Bottom
CGRP					
RAMP1-CLR	~12.49 μ M	+/- 0.04028	0.9698	101.1	=0
RAMP3-CLR	~13.56 μ M	+/- 0.3705	0.9404	99.86	=0
AM					
RAMP1-CLR	~8.19 μ M	+/- 0.005007	0.9607	92.64	=0
RAMP3-CLR	~1.89 μ M	+/- 0.07684	0.9404	99.86	=0
CT					
RAMP1-CLR	-	-	-	-	-
RAMP3-CLR	-	-	-	-	-
AMY					
RAMP1-CLR	-	-	-	-	-
RAMP3-CLR	~20.81 μ M	+/- 0.05462	0.9812	111.8	=0
IMD					
RAMP1-CLR	49.4 μ M	+/- 0.2254	0.9123	93.03	=0
RAMP3-CLR	~5.091 μ M	+/- 0.06476	0.9753	107.8	=0

4.3.4 Displacement of SHF-1257 Using Novel Compounds.

During the drug development process, multiple compounds were tested against the CGRP and AM₂ receptors, holding different potencies for the receptor. As the purified fusion protein only contains the ECD of the receptor, it's important to validate its functionality in comparison to the full-length receptor to investigate whether it still holds the same level of specificity and affinity for novel compounds. Initial HTRF results on the saturation curves are implying that the protein is retaining its structure, but further investigation needs to be completed. The competition assays generated by the HTRF assay will also be more comparable with the previously generated cAMP data, with the exception of compound-compound competition, rather than compound-ligand inhibition. The structure of all the tested small molecules can be found in figure 4.16.

Firstly, the non-labelled SHF-1257 compound (SHF-1036) was tested to see if it retained its binding affinity for the receptor. As the compounds are designed to compete for a similar binding site on the receptor, their potencies should remain consistent if the fusion protein is properly folded. When testing cAMP inhibition on the full-length receptor, SHF-1036 holds an IC₅₀ value of 2.44 nM for the CGRP receptor and 1.42 nM for the AM₂ receptor. Figure 4.15 shows the SHF-1036 inhibition curve on both the RAMP1-CLR and RAMP3-CLR fusion proteins. The IC₅₀ values generated by the curve show values of 4.441 nM and 0.5664 nM for the RAMP1-CLR and RAMP3-CLR fusion proteins respectively. This corresponds very closely with cAMP inhibitory data, showing SHF-1036 holds a higher affinity for the AM₂ receptor over the CGRP receptor while having similar IC₅₀ values. With the inhibitory concentrations between the ECD and the full-length receptor being closely related, it is implying that the ECD structure of the receptor is being properly folded and is functional. The IC₅₀ and curve best fit values are summarised in table 4.7.

RAMP1-CLR and RAMP3-CLR 1036 Displacement

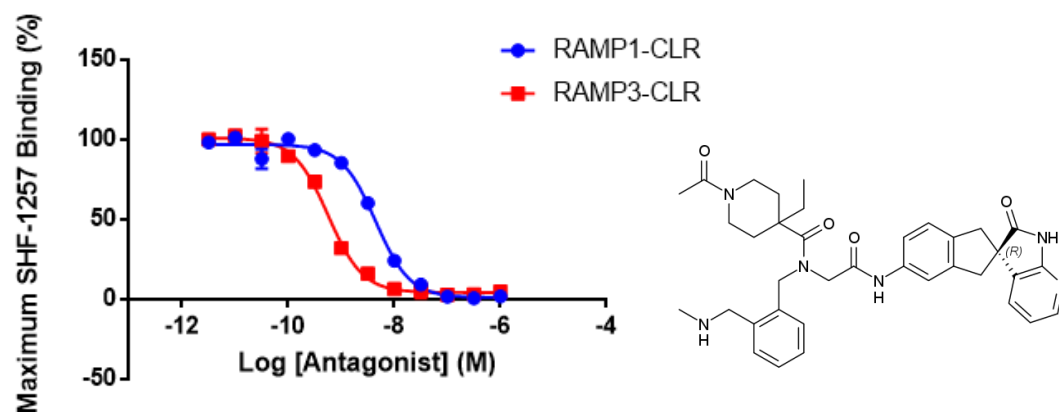


Figure 4.15: Dose response curve of SHF-1036 displacing SHF-1257 on the RAMP1/3-CLR fusion protein receptors. SHF-1036 concentration (x-axis) is plotted against SHF-1257 binding % (y-axis). SHF-1257 concentration remains constant through each dose of SHF-1036. The structure of SHF-1036 is also shown.

Table 4.7: Values of SHF-1036 displacing SHF-1257 on the RAMP1/3-CLR fusion proteins

IC50 Dose Response Curve: SHF-1036	Fusion Protein	
	RAMP1-CLR	RAMP3-CLR
IC50 Value	4.441 nM (Log-8.35)	0.5664 nM (Log-9.24)
Std Error: IC50	+/- 0.03584	+/- 0.02944
R square	0.9924	0.9949
Best fit: Top	96.76	100.8
Best Fit: Bottom	0.997	4.128

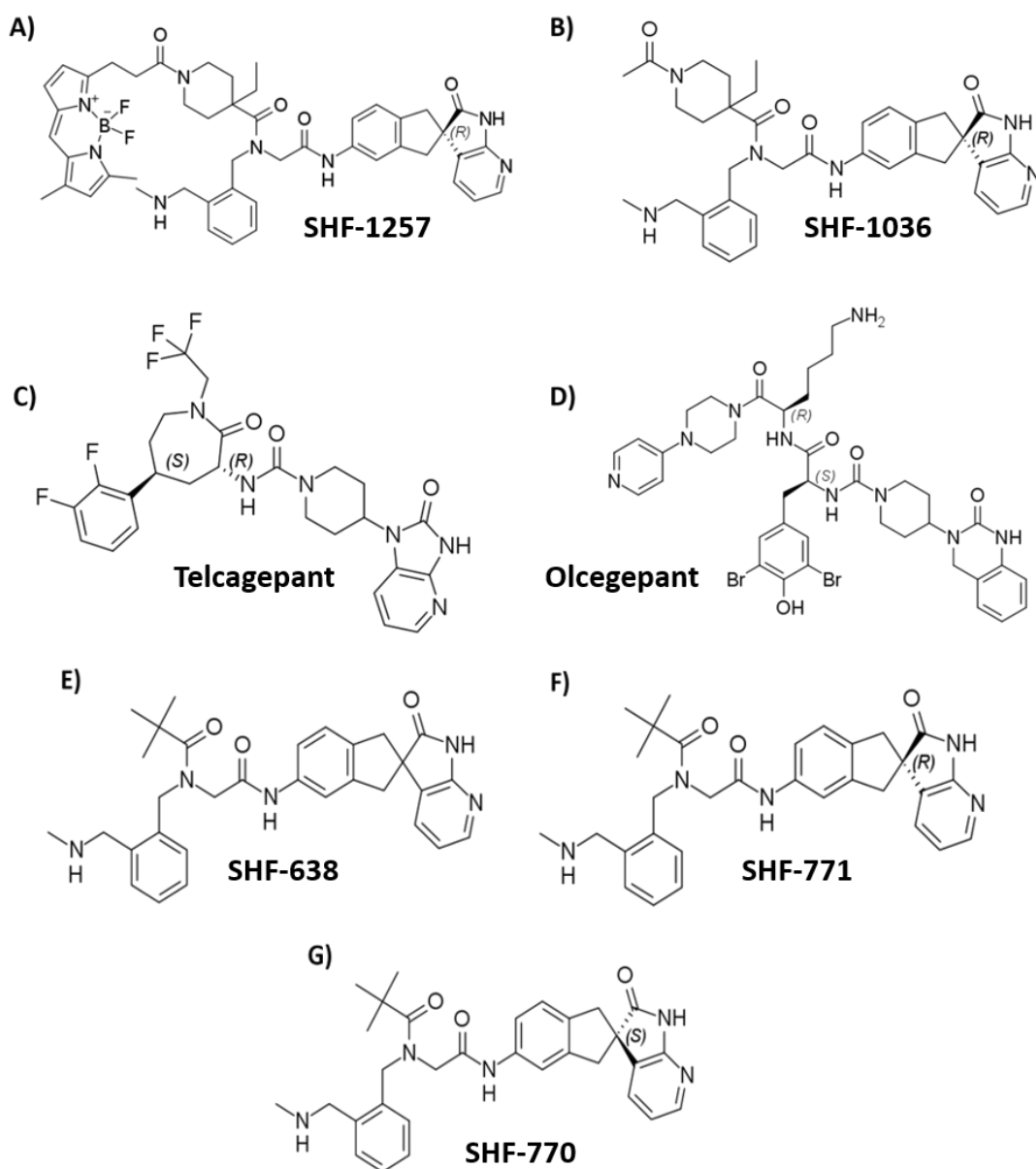


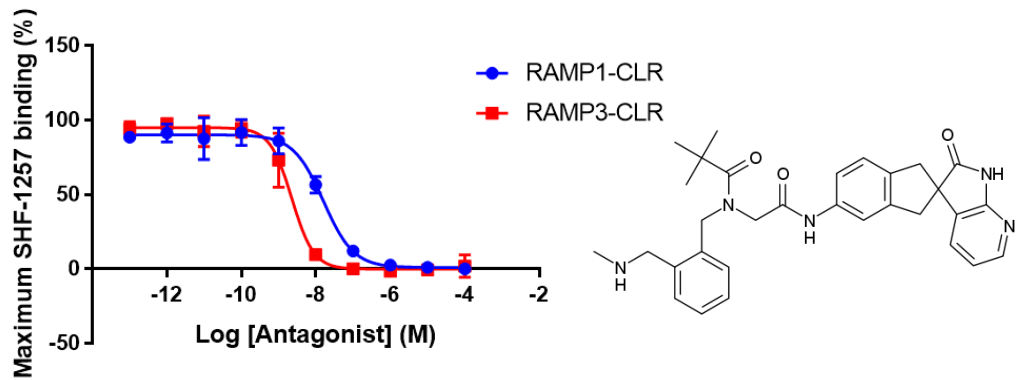
Figure 4.16: Structures of CGRP and AM₂ receptor antagonists. Structural comparison of commercially available CGRP antagonists olcegepant and telcagepant (C and D) with novel SHF small molecules which have potency for both the CGRP and AM₂ receptors. All compounds were applied to the HTRF assay.

To further assess the proper functionality and specificity of the purified protein receptors, another high and low affinity novel small molecule were used to see if the proteins retain their proper binding affinities. Compounds chosen were SHF-638, SHF-770 and SHF-771. SHF-770 and SHF-771 are two racemates of one another, while SHF-638 is a mixture of these compounds. On the full length receptors, interestingly, the two different racemic models give very different IC₅₀ values from one another, with SHF-771 holding an IC₅₀ value of 5.08

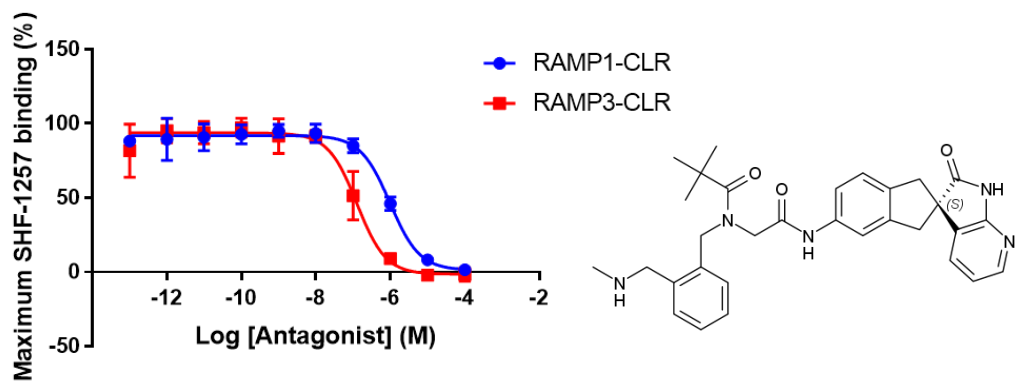
nM and 1.31 nM for the CGRP and AM₂ receptors respectively, while SHF-770 holds IC₅₀ values of 199 nM for the CGRP receptor, and 239 nM for the AM₂ receptor. Despite SHF-638 being a mix of these two compounds, it holds IC₅₀ values of 5.86 nM and 1.33 nM for the CGRP and AM₂ receptors respectively, which are almost identical to the values generated by SHF-771.

When applied to the HTRF assay, these values remain consistent with preliminary cAMP data. Figure 4.17 show the dose response displacement curve of SHF-638, 770 and 771 on both the RAMP1-CLR and RAMP3-CLR fusion proteins. The IC₅₀ values for SHF-1257 displacement on the RAMP1-CLR fusion protein were 16.34 nM, 100.7 nM and 9.027 nM for SHF-638, 770 and 771 respectively. The IC₅₀ values for SHF-1257 displacement on the RAMP3-CLR fusion protein were 2.387 nM, 132 nM and 1.535 nM for SHF-638, 770 and 771 respectively. Further information about the IC₅₀ curves can be found in tables 4.8 and 4.9. These run adjacent with the cAMP data, with SHF-638 and 771 holding similar IC₅₀ values with slight selectivity over the AM₂ receptor. The IC₅₀ values for SHF-770 shows increased binding for receptor proteins when using the HTRF assay in comparison to the cAMP assay, but still show lower affinities than SHF-638 and SHF-771. This is presumably due to experimental differences between the two assays, but the same trend is followed. The association of the SHF-638 series of compounds not only shows a functional purified receptor fusion protein, but they retain their capability and specificity of binding high and low affinity compounds.

RAMP1-CLR and RAMP3-CLR SHF-638 Displacement



RAMP1-CLR and RAMP3-CLR SHF-770 Displacement



RAMP1-CLR and RAMP3-CLR SHF-771 Displacement

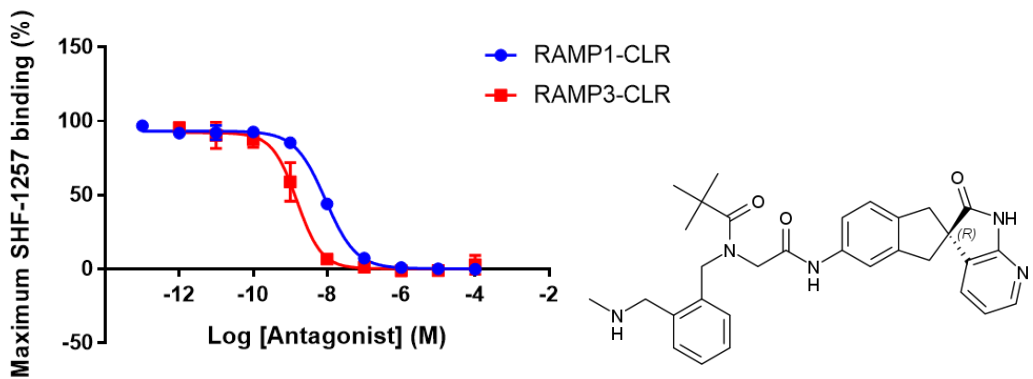


Figure 4.17: Dose response curves of SHF-638, SHF-770 and SHF-771 on the RAMP1 and RAMP3-CLR fusion proteins. The binding potencies of each novel compound as well as the structural differences between each novel small molecule is presented.

Table 4.8: IC50 values for SHF-638, SHF-770 and SHF-771 on the RAMP1-CLR fusion protein

IC50 Dose Response Curve: RAMP1-CLR Fusion Protein	Compound		
	SHF-638	SHF-770	SHF-771
IC50 Value	16.34 nM	100.7 nM	9.027 nM
Std Error: LogIC50	+/- 0.06484	+/- 0.06833	+/- 0.02817
R square	0.9788	0.9708	0.9957
Best fit: Top	90.39	92.22	93.54
Best Fit: Bottom	0.4775	0.24	-0.006383

Table 4.9: IC50 values for SHF-638, SHF-770 and SHF-771 on the RAMP3-CLR fusion protein

IC50 Dose Response Curve: RAMP3-CLR Fusion Protein	Compound		
	SHF-638	SHF-770	SHF-771
IC50 Value	2.387 nM	132 nM	1.535 nM
Std Error: LogIC50	+/- 0.07854	+/-0.08478	0.06231
R square	0.9713	0.9609	0.9802
Best fit: Top	96.55	94.01	93.65
Best Fit: Bottom	-1.787	-2.495	-0.7535

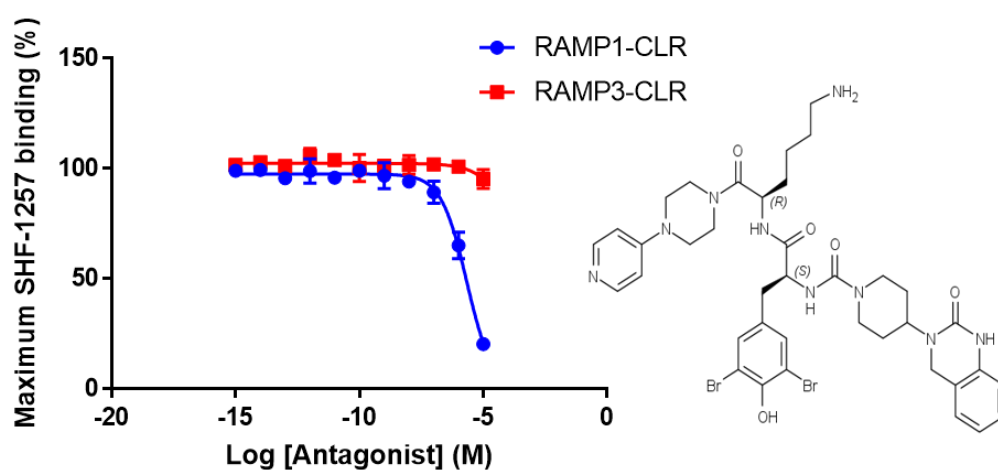
4.3.5 SHF-1257 Displacement Using Previously Developed CGRP Antagonists

In an effort to predict the specific residues novel SHF- compounds are/are not associating with on both the CLR receptor and RAMP proteins, the SHF-1257 compound can be displaced with previously developed CGRP antagonists of known binding capabilities. Previously reported data by ter Haar *et al.*, 2010 shows the crystal structures of both olcegepant and telcagepant bound to the CGRP receptor as stated in sections 4.1.1 and 4.1.2. This identified the key amino acids which are responsible for their association to the CGRP receptor. If these same compounds are applied to the HTRF assay, predictions can be made after displacement data and structural comparison into the specific residues, particularly on the CLR receptor and the RAMP1 protein, that novel compounds are associating with.

Olcegepant was applied to the HTRF assay in a dose response manner (figure 4.18), with concentrations ranging from 100 μ M - 0.001 pM on the RAMP1/3-CLR fusion proteins. Data

generated shows a decrease in SHF-1257 binding (increased displacement) in olcegepant concentrations higher than 100 nM on the RAMP1-CLR fusion protein. Although the curve does not reach 100% inhibition (0% SHF-1257 binding), the predicted IC₅₀ of olcegepant on displacing SHF-1257 on the RAMP1-CLR fusion protein is 1.74 μM. Further to this, there appears to be a minimal amount of displacement on the RAMP3-CLR fusion protein, with the possibility of slight SHF-1257 displacement at the highest concentration of olcegepant, 10 μM. Higher concentrations of olcegepant beyond 100 μM may not give an accurate representation of drug displacement, with higher concentrations of DMSO from drug storage likely to have an increasing role in the assay binding kinetics.

RAMP1/3-CLR fusion protein olcegepant displacement



RAMP1/3-CLR fusion protein Telcagepant Displacement

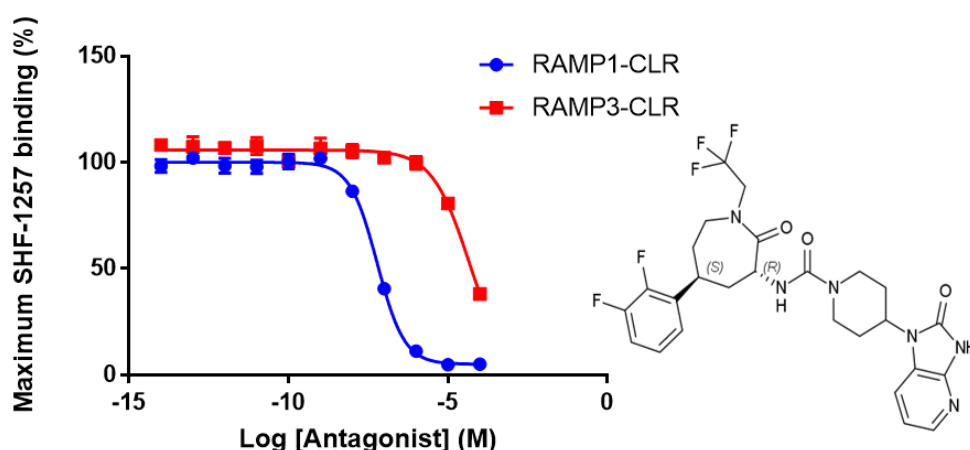


Figure 4.18: Olcegepant and Telcagepant HTRF displacement curves. The dose response of olcegepant and telcagepant displacing SHF-1257 from the RAMP1-CLR and RAMP3-CLR fusion proteins.

Telcagepant was also applied to the HTRF assay to displace the SHF-1257 compound (figure 4.18) in a dose response manner, with concentrations ranging from 100 μ M to 0.001 pM. Data shows a very different response to the displacement curves generated by the olcegepant dose response. Telcagepant displacement on the RAMP1-CLR fusion protein shows both maximum and minimum inhibition, with an IC₅₀ value of 60.71 nM. Telcagepant also shows a slight response on the RAMP3-CLR fusion protein, which is similar to olcegepant on the RAMP1-CLR fusion protein. While 100% displacement of SHF-1257 is not achieved, a predicted IC₅₀ value is \sim 22.27 μ M. While these IC₅₀ values are showing increased displacement of SHF-1257 in comparison to olcegepant on both the RAMP1-CLR and RAMP3-CLR fusion proteins, telcagepant is not as potent as previously generated cAMP data on inhibiting peptide binding on the full-length protein receptor. Telcagepant shows IC₅₀ values of 1.99 nM and 2.49 μ M on the CGRP and AM₂ receptors respectively on peptide inhibition. This implies they are occupying similar residues within the receptor binding pocket, but different contact points will be seen between the novel compounds and telcagepant.

4.3.6 Effect on Glycan Addition on Receptor Binding Capabilities

Receptor proteins are extensively modified during protein production, which includes the addition of glycans to the protein structure. Glycans can be added for numerous reasons, including proper protein folding and increased protein stability. With both fusion proteins being used for future crystallography studies, glycan addition may result in receptor surfaces being structurally different, and therefore effect the alignment of proteins during the crystal formation process. The removal of these sugars may therefore be necessary, but the effect on functionality after glycan removal needs to be investigated to ensure the structure and binding pocket of the receptor is not compromised. Both RAMP1/3-CLR proteins were not incubated at 37°C with no enzyme (WT), were incubated at 37°C without the presence of enzyme (-ve) or were incubated at 37°C with enzyme (+ve).

Both the RAMP1-CLR and RAMP3-CLR fusion proteins were deglycosylated and applied to the HTRF assay, with SHF-770 and SHF-771 being applied as a displacement control at a single dose of 100 nM. 100 nM SHF-770 has minimal displacement on both glycosylated RAMP1-CLR and RAMP3-CLR fusion proteins. 100 nM SHF-771 however, displaces SHF-1257 by \sim 55.96% and \sim 93.31% on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively (figure 4.17). These doses can therefore be used as a determinant of the receptor retaining proper functionality and affinity for compounds. When glycans were removed, it has no

effect on the capability of the compounds (SHF-1257, SHF-1036, SHF-770 and SHF-771) to bind to either the RAMP1-CLR or the RAMP3-CLR fusion proteins (figures 4.19 and 4.20). Despite undergoing native deglycosylation overnight at 37°C, this had no effect on the deglycosylated (+ve) and non-deglycosylated (-ve) versions of the protein in comparison to the wild type (WT) protein, which had not been incubated at 37°C (figures 4.19 and 4.20). All protein samples retained binding of the SHF-1257 compound.

Additional to this, SHF-1036, SHF-770 and SHF-771 retained their affinity over the receptor. Excess addition of SHF-1036 saw the complete inhibition of SHF-1257 binding on both fusion protein receptors (figures 4.19 and 4.20). SHF-771 reported displacement values of 54.26% and 91.12% on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively, while SHF-770 has no effect on SHF-1257 displacement on both fusion proteins (figures 4.19 and 4.20). All displacement values and column statistics can be found in table 4.10 and 4.11.

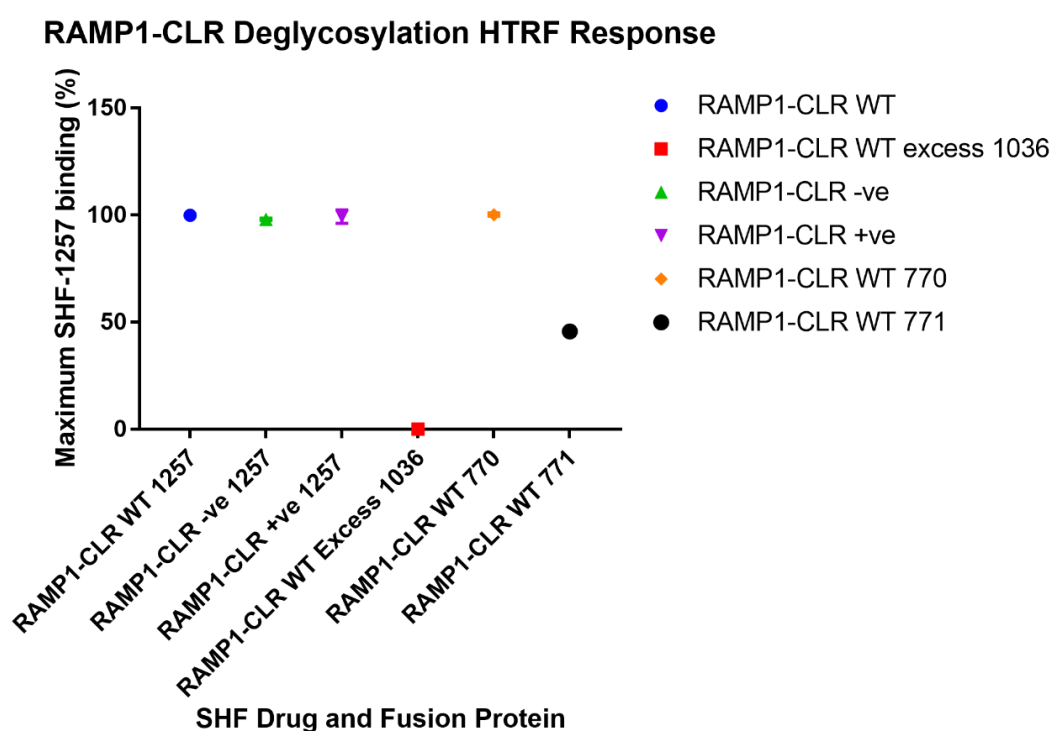


Figure 4.19: HTRF response post deglycosylation of the RAMP1-CLR fusion protein. Fusion protein with novel antagonist plotted against maximum SHF-1257 binding to the receptor. WT is protein not incubated at 37°C. -ve sample is protein incubated overnight at 37°C without enzyme (PNGaseF). +ve sample is protein incubated at 37°C in the presence of enzyme (PNGase F).

Column Statistics	Fusion Protein and Competing Novel Small molecule					
	RAMP1- CLR WT	RAMP1- CLR WT SHF-1036	RAMP1- CLR -ve enzyme	RAMP1- CLR +ve enzyme	RAMP1- CLR SHF- 770	RAMP1- CLR SHF- 771
SHF-1257 Binding (%)	100	0.0206	98.05	99.31	100.2	45.74
Std error mean +/- (SEM)	1.539	0.02722	0.3576	1.796	0.4917	0.6772

Table 4.10: SHF-1257 association values to glycosylated and deglycosylated RAMP1-CLR fusion

RAMP3-CLR Deglycosylation HTRF Response

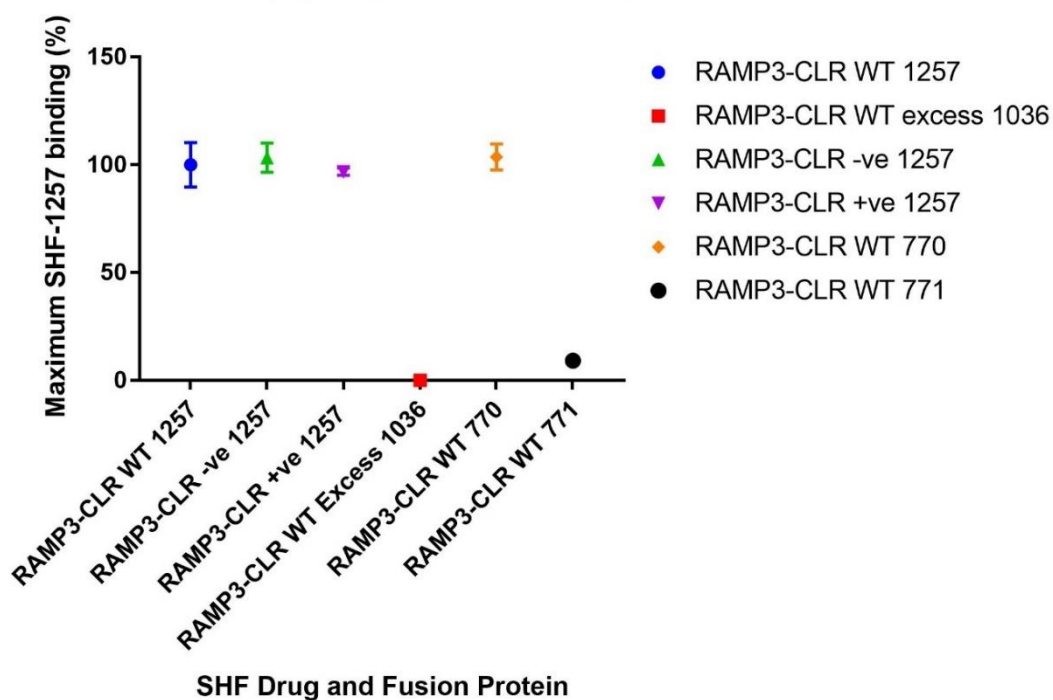


Figure 4.20: HTRF response post deglycosylation of the RAMP3-CLR fusion protein. Fusion protein with novel antagonist plotted against maximum SHF-1257 binding to the receptor. WT is protein not incubated at 37°C. -ve sample is protein incubated overnight at 37°C without enzyme (PNGaseF). +ve sample is protein incubated at 37°C in the presence of enzyme (PNGase F).

Table 4.11: SHF-1257 association values to the glycosylated and deglycosylated RAMP3-CLR fusion protein

Column Statistics	Fusion Protein and Competing Novel Small molecule					
	RAMP3-CLR WT	RAMP3-CLR WT SHF-1036	RAMP3-CLR -ve enzyme	RAMP3-CLR +ve enzyme	RAMP3-CLR SHF-770	RAMP3-CLR SHF-771
SHF-1257 Binding (%)	100	0.01659	103.3	96.85	103.7	9.186
Std error mean +/- (SEM)	5.97	0.3847	3.916	0.9923	3.466	0.5682

4.4 Discussion

Here, it has been shown that purified RAMP-CLR fusion proteins retain affinity for their endogenous peptides and hold high selectivity and specificity for the novel SHF series of antagonists. Results show antagonists may hold contacts with different residues in the receptor binding pocket in comparison with previously developed CGRP antagonists due to a vast reduction in their potencies when displacing novel antagonist SHF-1257. Results state glycan addition to the receptor has no direct role in antagonist binding to the receptor ECD, with no observed differences in SHF-1257 affinity between the glycosylated and deglycosylated fusion proteins, stating they are required for correct receptor folding only.

SHF-1036 is currently the lead compound being used to potentially target the CGRP and AM₂ receptors to treat migraine pain and reduce pancreatic cancer progression. While it holds a high affinity for the receptor, when applying the compound to an HTRF assay, displacement cannot be measured without the addition of a tag that fluoresces when bound to the RAMP-CLR fusion protein. In turn, the compound SHF-1257 was generated to measure the association of novel compounds to the receptors and potentially determine suspected binding sites in the receptor.

The SHF-1257 compound is structurally similar to the novel lead compound SHF-1036 but has the addition of a BODIPY-FL labelled fluorescent dye/tag (*4,4-difluoro-4-bora-3a,4a-diaza-s-indacene*) at the terminus of the compound (see table 4.3). While it is structurally similar in every aspect to SHF-1036 with the exception of the BODIPY-FL tag, it results in a loss of potency to the CGRP and AM₂ receptors in comparison to its unlabelled counterpart. The label was added to the region of the compound which will extend further away from the binding pocket, and comparisons can be made between the pyridyl rings of olcegepant which extends further out of the pocket to contact CLR Asp94 (see figure 4.16 for structural comparison) (ter Haar *et al.*, 2010). While the addition of this tag has evidently affected compound association to the receptors, this would keep loss of potency to a minimum. During preliminary studies investigating the inhibition of cAMP release on full length receptors, the addition of the label results in a loss of potency of the compound, from Ki 2.44 nM [pIC₅₀ (M) 8.61] to 16.9 nM [pIC₅₀ (M) 7.77] on the CGRP receptor and from 1.42 nM [pIC₅₀ (M) 8.85] to 32.4 nM [pIC₅₀ (M) 7.49] on the AM₂ receptor. This corresponds directly to the HTRF data presented on the purified fusion protein (figures 4.13 and 4.15) which shows an affinity reduction from 4.578 nM (IC₅₀) to 11.2 nM (K_d) and 0.549 nM (IC₅₀) to 2.82 nM (K_d) on the CGRP and AM₂ receptors respectively.

The HTRF assay may have lower numerical values for drug association to the receptor in comparison to the cAMP assays, as the drugs are competing for one receptor protein in each reaction and therefore reduces possible non-specific binding between other receptor proteins and cell membrane. Inhibition values also may vary slightly due to the cAMP assay competing with the receptor peptides (CGRP or AM) which will occupy additional amino acids not contacted by the compound. In the HTRF assay compounds are competing with other compounds which have been designed to target almost identical aspects of the binding pocket and therefore may provide more direct competition.

Despite this, the results remain consistent between both assays, with the BODIPY-FL tag on SHF-1257 reducing the affinity of the antagonist. While its addition was designed to have minimal interference with receptor binding, the tag is having an effect on receptor-compound interaction and therefore is affecting its association to the binding pocket. The BODIPY-FL label extends the drug structure beyond the length of its higher potency counterpart (SHF-1036) (figure 4.16) which could result in the compound not packing as efficiently against the receptor binding pocket. It is most likely disrupting the contacts formed between the secondary amine group which extends from the benzene ring, and sits in close proximity to the BODIPY-FL tag in SHF-1257 (figure 4.16). This functional group is likely to hydrogen bond with RAMP1 Asp71/ RAMP3 Asn71 (as contacted in olcegepant binding) and therefore if SHF-1257 is not packing in the binding pocket as efficiently, this bond may be weakened or lost (ter Haar *et al.*, 2010).

Despite the loss in affinity between SHF-1036 and SHF-1257, the affinity of the compounds against the purified fusion protein ECD are extremely similar to their full-length receptor. This is implying that despite only the extracellular domain of the receptors being expressed, they remain fully functional and additional tags (MBP and Histag) are having no effect on protein structure and binding capabilities. This corresponds with the previous literature which shows the expression of RAMP-CLR ECDs with histidine and MBP tags in both bacterial and mammalian cell lines having no effect on receptor binding (Booe *et al.*, 2015) (Roehrkasse *et al.*, 2018). This suggests a fully functional binding pocket which can be used to investigate and compare drug association to the receptor to gain a further insight into the binding properties of the novel antagonists.

4.4.1 Peptide dissociation of SHF-1257

Ligands from the calcitonin family of peptides displaced the SHF-1257 compound on the RAMP1-CLR and RAMP3-CLR fusion proteins at much lower affinities than their full-length receptor counterparts. This was expected due to the expression of the receptors ECD only and has been observed in previous studies which express the ECD of the receptors in a similar manner (Booe *et al.*, 2015)(Liang *et al.*, 2017)(Roehrkasse *et al.*, 2018). This enables the C-terminus of the peptide only contacting the receptor, resulting in a loss of affinity as the N-terminus cannot make contacts with the ECL as seen in the full length receptor crystal models (Liang *et al.*, 2019)(Liang *et al.*, 2020). The ligands appear to retain their rank order for the RAMP3-CLR fusion protein receptor (figure 4.14), with AM and IMD holding the highest displacement affinity (IC₅₀) on the receptor. CGRP hold a slightly higher affinity to the receptor than AMY, and CT appears to hold no affinity for the RAMP3-CLR fusion protein ECD which matches the rank order of the peptides. It should be noted that the peptide concentrations do not capture the bottom of the displacement curve for the RAMP3-CLR ECD protein response and therefore the accuracy of these IC₅₀ values maybe difficult to interpret.

The data generated for the peptide displacement on the RAMP1-CLR fusion protein, could suggest it is not retaining its selectivity for peptides (figure 4.14), with AM holding a slightly higher IC₅₀ value (8.19 μM) than CGRP (12.49 μM). As the fusion protein still retains high accuracy and selectivity for both high and low affinity novel compounds (SHF-770 and SHF-771, figure 4.17) it may suggest that the antagonist (SHF-1257) is blocking key contact points for AM association to the receptor over those implicated for CGRP association. This would result in the AM peptide having a higher displacement affinity as it competes directly with key residues rather than just blocking the binding pocket. Examples could include CLR Trp72 which plays a larger role in AM association to the receptors than CGRP (Booe *et al.*, 2015). With the interface of the SHF series of compounds designed to target the CLR Trp72 residue, it implies it is competing directly with AM Tyr52. This would contribute to AM holding a higher IC₅₀ value when displacing SHF-1257 on the RAMP1-CLR fusion protein than CGRP. This could also explain why the SHF series of compounds have a reduced potency on the CGRP receptor than the AM₂ receptor as they compete less directly with residues involved in ligand binding. This theory could be applied to the study by Roehrkasse *et al.*, 2018 who purified RAMP-CLR fusion proteins in HEK293 cells and displaced fluorescein isothiocyanate (FITC) labelled AM from each protein with the calcitonin family of peptides to determine

receptor functionality. IMD peptide was seen to hold a higher displacement affinity (IC₅₀) to the RAMP1-CLR fusion protein than the CGRP peptide as it competed more directly with the FITC-AM peptide (Roehrkasse *et al.*, 2018).

4.4.2 Olcegepant and Telcagepant Association to the RAMP1-CLR Fusion Protein

Interestingly, telcagepant showed a much higher affinity and displacement of SHF-1257 on the RAMP1-CLR fusion protein in comparison to olcegepant, despite olcegepant having a much higher potency on the full length CGRP and AM₂ receptor (figure 4.18). Previous cAMP data shows olcegepant inhibits peptide binding to the full-length receptor at concentrations of 0.096 nM and 299 nM on the CGRP and AM₂ receptors respectively, which are highly different from the SHF-1257 displacement. This implies they are occupying different sites and residues in the receptor binding pocket for peptide inhibition. SHF-1257 bonding to the fusion protein receptors is more likely to be consistent with that of telcagepant, such as holding a higher number of contacts with CLR Thr122 (ter Haar *et al.*, 2010). SHF-1257 is structurally similar to telcagepant, having the azabenzimidazolone ring that could form the extra contacts at CLR Thr122 in comparison to olcegepant (ter Haar *et al.*, 2010). This could partially explain why telcagepant shows increased SHF-1257 displacement on the RAMP1-CLR fusion protein than olcegepant, due to being more competitive at CLR Thr122 residue. This would therefore result in telcagepant having a higher affinity for the fusion protein.

It further implies similar activity and contacts to that of telcagepant within the CGRP receptor binding pocket. As seen in the models generated by ter Haar *et al.*, 2010, olcegepant forms two water mediated hydrogen bonds with CLR Arg38 and RAMP1 Arg67 in the binding pocket, where in contrast, telcagepant displaces this water molecule potentially altering ligand binding further through enthalpic and entropic changes (ter Haar *et al.*, 2010). With olcegepant having a much lower potency than expected on the RAMP1-CLR fusion protein receptor, it could imply that the SHF-1257 is displacing the water molecule in a similar manner. Presuming the compounds hold a similar site of occupancy, the benzene ring and secondary amine group in SHF-1257 could be extending into the binding pocket, similar to the difluorophenyl ring of telcagepant (ter Haar *et al.*, 2010). This would displace the water molecule in the binding pocket and subsequently affect olcegepant's association to the receptor. If there is no water molecule in the binding pocket, olcegepant cannot make the hydrogen bonds to the appropriate CLR and RAMP residues. To coincide with this, the

SHF-1257 structure is unlikely to make the same contact at CLR Asp94 due to the lack of ringed extensions towards this residue, which is observed from the terminal pyridyl ring found in olcegepant (ter Haar *et al.*, 2010). This implies the water molecule is being displaced and therefore is only competing with olcegepant at CLR Thr122, contacts at the Trp shelf in CLR (CLR Trp72) and the salt bridge at RAMP1 Asp71. This could mean SHF-1257 has the same mechanistic properties to telcegepant in relation to its association to the CGRP receptor.

As telcegepant does not retain its IC₅₀ values on the RAMP1-CLR fusion protein and there is a significant decrease in potency (IC₅₀ values 1.99 nM on the full-length receptor cAMP assays to 60.71 nM on the RAMP1-CLR fusion protein in the HTRF assay), it could be presumed SHF-1257 is not forming the same bonds to the receptor. Despite presumably forming similar bonds at CLR Trp72 and Thr122, it suggests different bonds are being formed with the interface and RAMP region in the receptor (ter Haar *et al.*, 2010). The SHF-1257 compound holds the extension of the piperidine ring (with a methyl group extending from the base of the ring) into the methyl and carbonyl group terminal, whereas telcegepant holds the smaller trifluoromethyl group at the equivalent position (figure 4.16). It is likely that the methyl group of SHF-1257 extends towards the roof of the hydrophobic pocket and forms contacts with RAMP1 Trp74, similar to that of the predicted ubrogepant binding (figure 4.8). The piperidine ring would then extend along the CLR protein towards the CLR loop 4 and make contacts with the CLR receptor, somewhat similar to the pyridine ring of olcegepant but have a different region of occupancy (ter Haar *et al.*, 2010). An additional contact is likely to be made between the secondary amine group in SHF-1257 and RAMP1 Asp71, which is not made during telcegepant binding to the CGRP receptor (ter Haar *et al.*, 2010). These examples could lead to the overall increased affinity of SHF-1257 to the RAMP1-CLR fusion protein in comparison to telcegepant and would explain the loss of potency (IC₅₀ value) of the compound during displacement assays.

4.4.3 Olcegepant and Telcegepant Association to the RAMP3-CLR Fusion Protein

The olcegepant and telcegepant data presented on the displacement of SHF-1257 from the RAMP3-CLR fusion protein (figure 4.18) proposes further that the AM₂ receptor holds a binding pocket which retains certain similarities to the AM₁ receptor. Both olcegepant and telcegepant have a lower affinity at the RAMP3-CLR fusion protein over the RAMP1-CLR

fusion protein when displacing SHF-1257, but coinciding with the RAMP1-CLR fusion protein, telcagepant holds a higher affinity over the RAMP3-CLR fusion protein in comparison to olcagepant. Watkins *et al.*, 2014 describe that RAMP2 Glu101 is a residue which sits deeper into the binding pocket in comparison to its equivalent counterpart in RAMP1, Trp74, which in turn could explain the loss of affinity from the antagonists between the receptors (Watkins *et al.*, 2014). The equivalent residue in RAMP3 is Glu74, which could lead to the formation of a deeper binding pocket as seen in RAMP2 (Booe *et al.*, 2015). RAMP1 Trp74 has been previously linked to be a vital residue in CGRP antagonist binding with mutagenesis leading to significant loss in telcagepant and olcagepant binding, with a loss of the overall hydrophobic surface in the receptor (Hay *et al.*, 2006) (Moore *et al.*, 2010) (ter Haar *et al.*, 2010) (Kusano *et al.*, 2008). If the binding pocket is deeper, and with the loss of the large ligand-protein surface, the extension of the CGRP antagonists are unlikely to extend far enough into the binding pocket and form further contacts with the receptor. The loss of contacts could include CLR Met42 and Ile41 or RAMP1 Asp72, Arg67 (water mediated) and CLR Arg38 (water mediated) for telcagepant and olcagepant respectively (ter Haar *et al.*, 2010). While this explains the loss of telcagepant affinity for the receptor, it fails to propose why the SHF-1257 and other antagonists in the SHF series retain affinity for both receptors.

Although SHF-1257 may rely partly on the Trp74 residue in RAMP1, the secondary amine group which extends from the benzene ring of RAMP/CLR end, is likely to extend further into the binding pocket. This is likely to hydrogen bond with the side chain of RAMP3 Asn71, similar to that of RAMP1 Asp71, and would therefore retain binding at the AM₂ receptor (see figure 4.22 for example of SHF-638 superimposition on the receptor). Telcagepant would not be able to produce similar connections and while it could be presumed that olcagepant lysine terminal would extend to make the contact, the change of residue from the aspartic acid to asparagine would result in the loss of the salt bridge (ter Haar *et al.*, 2010). This would result in olcagepant forming a weaker hydrogen bond due to the change from the charged amino acid to the uncharged amino acid side chain of asparagine, if it is still in close proximity to the residue (ter Haar *et al.*, 2010). This would also partially explain the further loss of displacement from olcagepant on the RAMP3-CLR fusion protein, in comparison to the RAMP1-CLR fusion protein as the contacts become more competitive.

The data further reiterates that SHF-1257 is likely not to be associating with CLR Asp94 on either the RAMP1 or RAMP3-CLR fusion proteins. With telcagepant retaining a higher affinity than olcagepant on the RAMP3-CLR fusion protein, it could be suggested that this is through connections with the CLR protein, which is likely to retain a similar conformation between

the three ligand-free receptors. It is likely the same interactions are formed with the CLR receptor, with the 3 main interactions being between telcagepant and CLR Thr122 and those involved with CLR Trp72 (ter Haar *et al.*, 2010). This suggests why telcagepant holds higher affinity to the RAMP3-CLR fusion protein than olcegepant, by being more competitive with the interactions at CLR Thr122 and SHF-1257 not competing with olcegepant at CLR Asp94. Olcegepant has a higher potency than telcagepant when competing on the full length AM₂ receptor, with K_i values of 299 nM and 2.4 μM respectively, but it further implicates SHF-1257 holds a much similar topology to telcagepant rather than olcegepant, particular at the CLR end of the compound (ter Haar *et al.*, 2010) (Hay *et al.*, 2006) (Doods *et al.*, 2000).

4.4.5 Comparison of SHF-638 Enantiomers and SHF-1036 Displacement

SHF-1036 and the SHF-638 structures are very analogous and would presumably hold a similar topology in the receptor binding pocket. As the body of the SHF-638 and SHF-1036/1257 compounds remain identical (figure 4.21), interactions could form between CLR Thr122 via the azaindole group, CLR Trp72 with the carbonyl groups in their compound backbone and RAMP1 Asp71 (RAMP3 Asn71) through the secondary amine from the terminal benzene ring of all compounds (figure 4.22). The only distinct difference between the two compounds is the piperidine ring which extends to a terminal carbonyl and methyl group in SHF-1036, where the equivalent in SHF-638 is a trimethyl group (figure 4.21). The extension of the piperidine ring to the carbonyl and methyl group terminal means SHF-1036 is likely to extend further across the CLR protein and make further contacts, possibly between CLR Phe92. The superimposition of SHF-770/771 on the CGRP receptor (figure 4.22) shows no obvious interactions with the receptor from its trimethyl group. The extra contacts with SHF-1036 would therefore give the compound a higher affinity/potency to the receptor.

The racemates of the SHF-638 series however prompts interesting results. Like many compounds on the market, SHF-638 was generated as a racemate consisting of two enantiomers (figure 4.21). These two enantiomers were also generated as pure versions to compare the binding affinities/differences to the CGRP and AM₂ receptors. Data presented shows the “rectus” or right (**R**) has a significantly higher potency on both the CGRP (9.027 nM) and AM₂ (1.535 nM) receptors in comparison to the “sinister” or left (**S**) enantiomer, SHF-770 (CGRP: 100.7 nM AM₂: 132 nM) (figure 4.17) . Despite being structurally identical,

the binding potencies and affinity for the receptor is clearly being dictated by the positioning of the functional groups in the binding pocket and how they fit into the receptor.

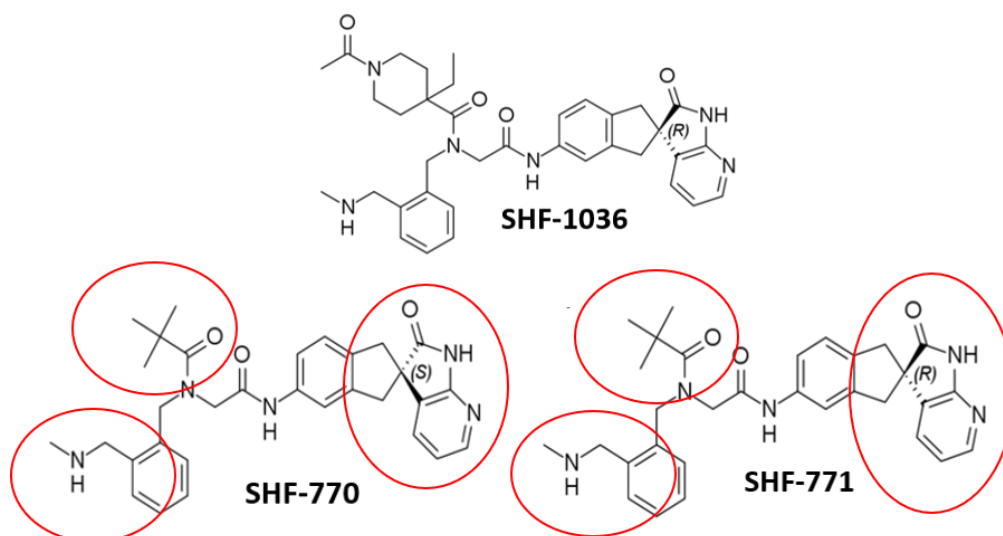


Figure 4.21: Structural comparison of SHF-1036, SHF-770 and SHF-771. Annotated (red) are the regions on SHF-770 and SHF-771 are regions which are responsible for compound binding to the receptor and are suspected sites for binding differences between themselves and SHF-1036.

If both enantiomers are superimposed into the RAMP1-CLR ECD crystal structure (PDB: 3N7R), distinct positional differences are observed (figure 4.22). The contacts at the CLR Thr122 are presumably the same but as the compound extends toward the CLR Trp72 residue, the second carbonyl group of the compound is subsequently repositioned. CLR Trp72 has been seen to be key for peptide binding and therefore dictate how effective antagonists are on blocking receptor signalling if they can occupy this region (Watkins *et al.*, 2014) (Booe *et al.*, 2015).

When investigating the distance between the carbonyl groups of SHF-770 and SHF-771 to the nitrogen in the tryptophan side chain (indole group on CLR Trp72), the carbonyl group of SHF-771 is approximately 2.8 Å away from the nitrogen group, whereas SHF-770 is further away, at 5.2 Å (figure 4.22). This would likely result in no hydrogen bond forming between SHF-770 and CLR Trp72 and could explain the reduction in potency for the receptor. It could also state the importance of hydrogen bond formation at this position, which should be considered in future drug design.

Further to this, the change in enantiomer leads to the slight re positioning of the secondary amine at the RAMP/CLR end of the compound (figure. 4.22). The secondary amine is predicted to extend and associate with RAMP1 Asp71 (RAMP3 Asn71) as previously

described when discussing olcegepant and telcagepant displacing SHF-1257 from the receptor (section 4.4.2 and 4.4.3). As this group is ubiquitous across all the recently generated SHF series, it is presumed this contact is made by all these compounds. While its predicted repositioning enables SHF-770 to be slightly closer to the RAMP1 Asp71 (2.8 Å rather than 3.0 Å in SHF-771) it subsequently positions itself further from the nitrogen group in RAMP1 Trp74 (5.0 Å). While it is not known whether the compound would form hydrogen bonding with RAMP1 Trp74, SHF-771 is 4.0 Å away from the nitrogen atom, and therefore depending upon the proper position of the compound, a weak hydrogen bond may be feasible, which would give SHF-771 further increased affinity for the receptor (figure 4.22).

SHF-770 does however appear to be more rotated and form a face-to-face stacking with the indole group on RAMP1 Trp74 (figure 4.22). This would allow hydrophobic interactions to form between the benzene ring of SHF-770 and Trp74, partly explaining why there would not be a complete loss off affinity to the receptor when the contact at CLR Trp72 is lost.

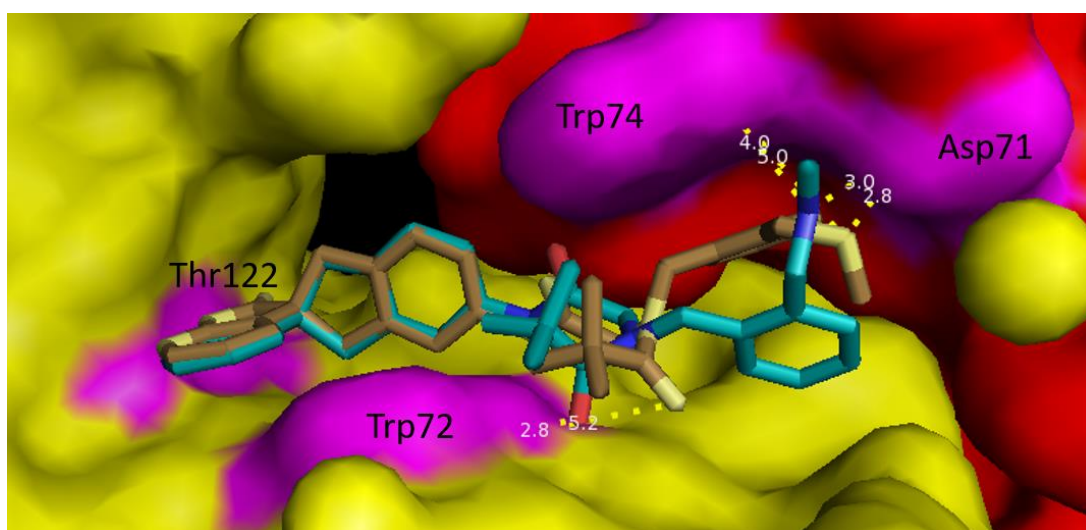


Figure 4.22: Superimposition of the SHF-770 and SHF-771 onto the CGRP receptor. The CLR receptor is coloured yellow and RAMP1 red. Residues believed to be important for SHF-770 (brown) and SHF-771 (blue) binding are coloured magenta and labelled. Distance of functional groups to key labelled residues are highlighted by dashed yellow lines and distance given in angstroms (Å).

4.4.6 Effect of Glycans on Antagonist Binding

Previous literature has reported the crystal structures of CGRP bound to the RAMP1-CLR ECD and AM bound to the RAMP2-CLR ECD, but there is yet to be any definitive structure

reported of the RAMP3-CLR ECD (ter Haar *et al.*, 2010; Booe *et al.*, 2015). Booe *et al.*, 2015 obtained the bacterial produced RAMP1/2-CLR crystals but reported that the generation of a functional RAMP3-CLR ECD receptor was not possible in bacteria. While they later reported that a RAMP3-CLR fusion protein could be expressed in mammalian cells, they did not report whether the glycan addition was solely important for protein folding, or if they were required for receptor binding capabilities (Roehrkasse *et al.*, 2018). Early literature has stated the importance of glycans to be added to RAMP3 to enable proper protein expression, but not if the sugars are required for ligand binding (Flahaut *et al.*, 2003). On both the RAMP1-CLR and RAMP3-CLR, the glycans were removed, purified, and reapplied to the HTRF assay. This subsequently showed no differences in binding of the SHF-1257 compound with the WT protein, in comparison to the receptor protein with the glycans removed (figure 4.19 and 4.20). This implies the glycans hold a primary role in proper protein folding and expression, with no effect on the structural integrity or directly assisting compounds associating to the receptor after the protein has been folded.

Although glycans should be added to every expressed RAMP3-CLR fusion protein, the sugars are unlikely to be added in a uniformed manner, and the extent of glycosylation is likely to differ between each fusion protein. This would lead to some structural variances on the surface of each protein and could interfere with screening attempts by preventing the unilateral lattice packing of proteins for crystal formation. As the glycans are not having a role in peptide binding, this means if necessary, glycans can be removed during crystallography screening to aid crystal formation while maintaining a functional receptor binding pocket.

Conclusion

To conclude, the data shows functionally expressed and purified RAMP-CLR fusion proteins which hold affinity for their endogenous peptides. Novel SHF antagonists appear to be holding similar binding topologies as telcagepant in comparison to other CGRP antagonists of known binding properties (olcegepant) and their receptor occupancy has been speculated. The removal of glycans from the fusion proteins does not affect antagonist affinity to the receptor, therefore implying glycan addition is critical for receptor expression only and has no role in compound association to the receptor.

Chapter 5: Fragment Screening and Compound Development Against the RAMP-CLR receptors

5.1 Introduction

Novel compounds can be generated through a design technique known as structure-based drug design (SBDD). When the receptor structure has been solved by X-ray crystallography, nuclear magnetic resonance spectroscopy (NMR) or cryogenic electron microscopy (cryo-EM) techniques, the binding pocket of the receptor can be assessed to aid drug design (Booe *et al.*, 2015) (Wasko *et al.*, 2015) (Liang *et al.*, 2019) (Liang *et al.*, 2020). If the crystal model has not been generated, a homology model can be created from closely related family receptors to predict the structure of the binding pocket and aid the design of novel compounds (Booe *et al.*, 2015). To be able to generate a homology model for SBDD however, a sequence similarity of >25% is required between the two proteins and is still limited to phylogenetic similarity (Rost and Sander, 1996) (Muhammed and Aki-Yalcin, 2019). Further from this, old drugs which have targeted a similar sub-type of receptor in comparison to the target receptor, are seen as the best basis for the discovery of a new drug (Wermuth, 2006)(Besnard *et al.*, 2013).

A variant of SBDD could be described as fragment-based drug design (FBDD). Fragments can be designed to target a specific region of the receptor binding pocket and these fragments can effectively serve as starting points for the development of lead candidate profiles (Wasko *et al.*, 2015). The advantage of using fragments as a starting point is that the process places special emphasis on creating viable, synthesizable molecules with relatively low cost of production (Wasko *et al.*, 2015). Different fragments that hold distinguishable variants which aid compound affinity to the receptor can be combined to generate a higher affinity fragment. Once different fragments have been identified which target different regions of the binding pocket, they can be linked together for the development of a full-length receptor antagonist and increase overall compound affinity to the receptor (Erlanson, 2012) (Mondal *et al.*, 2016) (Kirsch *et al.*, 2019).

5.1.1 Structural Comparison of Previous CGRP Antagonists

Analysing the structure of previous CGRP antagonists on their association to the CGRP receptor, they are predominantly made of three distinct regions/fragments: the CLR interacting region, the RAMP-CLR interface and the RAMP/CLR interacting region (figure 5.1)(Archbold *et al.*, 2010). The CLR binding region is defined by the region that associates to the CLR receptor only, which in CGRP receptor antagonists, is the quinazolinone structure

forming hydrogen bonds with CLR Thr122 (figure 5.1) (ter Haar *et al.*, 2010). While these interactions are speculative in more recent compounds (MK-3207, ubrogepant and atogepant), they are almost identical structure wise with telcagepant and olcegepant at the CLR binding region, which are compounds known to associate to CLR Thr122 (ter Haar *et al.*, 2010) (see figure 5.1 for model of antagonist binding). As the drug extends along the binding pocket, it meets the RAMP-CLR interface which is the region holding contacts with both the CLR receptor and RAMP protein (figure 5.1). This region could be described as the semi-variable region, as while each structure is designed to target the CLR Trp72 residue (telcagepant, olcegepant, ubrogepant and atogepant) their structures vary and may form contacts with additional residues in the binding pocket (see figure 5.1). An example being the interactions formed between olcegepant and RAMP1 Asp71, which is not seen in telcagepant binding (figure 5.1) (ter Haar *et al.*, 2010). The terminus of the compounds then extends to the variable RAMP/CLR binding region, where the compounds interact with different residues from either the RAMP protein or the CLR receptor, which is entirely dependent on the compound structure. Examples of the binding regions of olcegepant and telcagepant can be seen in figure 5.1

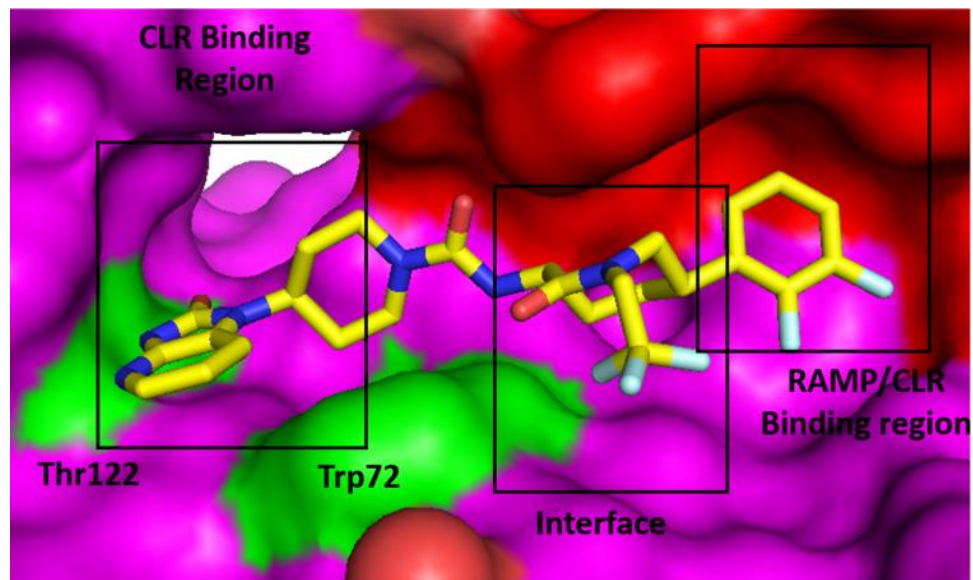
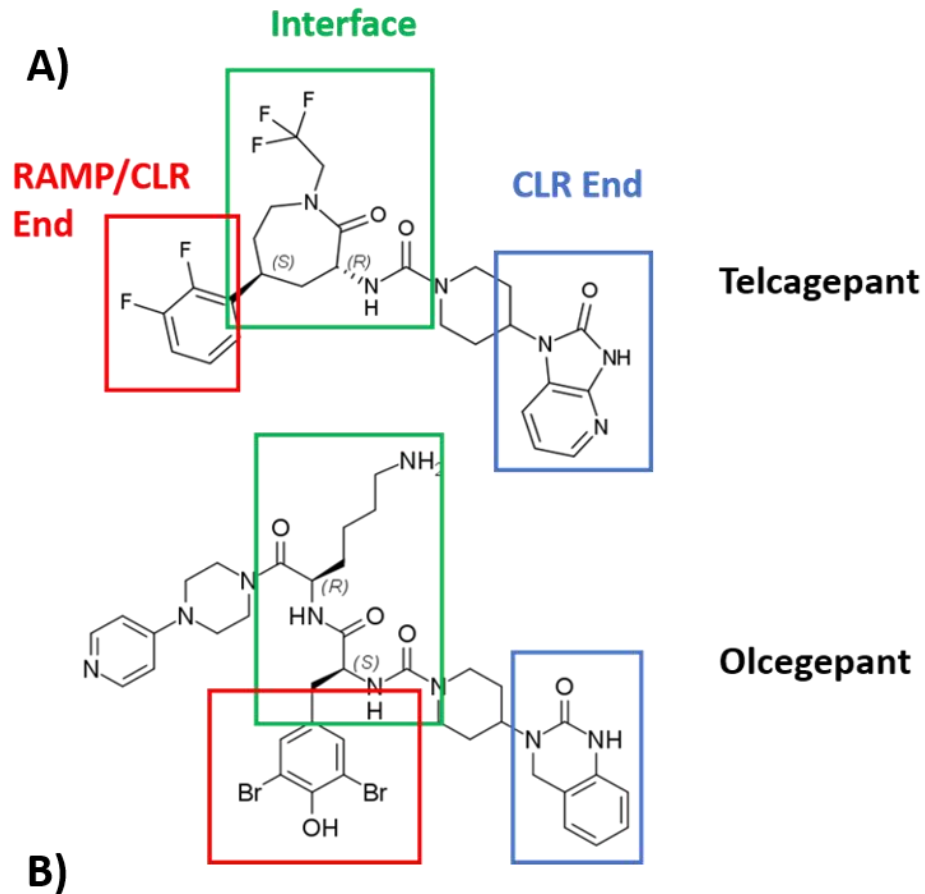


Figure 5.1: The identification of the CLR bind region, the interface and the CLR/RAMP binding region. (A) CGRP antagonists telcagepant (top) and olcegepant (bottom) with their 3 'binding domains' when associating to the CGRP receptor. The CLR end is highlighted in blue, the interface in green, and the RAMP/CLR end in red. Image adapted from Archbold *et al.*, 2011. **(B)** Outlines the positioning of the CLR binding region, interface and RAMP/CLR binding region of telcagepant (yellow) on the CGRP receptor (PDB: 3N7R). The CLR receptor is coloured magenta and RAMP1 protein coloured red. CLR Thr122 and Trp72, which are key for antagonists binding are annotated and coloured green.

5.2 Aim and Objectives

As previously shown in chapter 4, a HTRF binding assay has been developed and the purified RAMP1/3-CLR fusion proteins were proven functional, holding accurate affinities for novel CGRP and AM₂ receptor antagonists. By combining the developed HTRF binding assay and the SBDD technique implemented by the previously developed CGRP antagonists, fragments have been designed to target either the RAMP protein or CLR receptor. Fragments which hold the highest affinity for the receptor will be applied to cAMP assays to assess how well they antagonise the receptor. These may then be linked together to generate a full-length antagonist to determine if they block peptide binding and activation of the receptors. Identified fragments which hold preferential binding to the CGRP or AM₂ receptor could then be applied to future drug design when generating compounds to target other GPCR receptors that associate to RAMP proteins.

Hypothesis:

- High affinity RAMP and CLR fragments can be identified by displacing novel SHF-1257 from RAMP-CLR fusion proteins
- Fragments will not inhibit a cAMP response from CGRP, AM₁ and AM₂ receptors as effectively as full-length compounds.

Objectives:

- Investigate the binding affinities of novel fragments designed to target either the RAMP or CLR protein and identify those that hold a high affinity to the receptor.
- Select the two highest affinity fragments and determine if they inhibit cAMP release in cells overexpressing the CGRP, AM₁ and AM₂ receptors.

5.2 Materials and Methods

5.2.1 HTRF Binding Assay: Single Dose Response of RAMP/CLR Fragments

As outlined in chapter 4, the HTRF assay was applied for fragment displacement of the SHF-1257 compound on both the RAMP1-CLR and RAMP3-CLR fusion proteins. During screening, a single dose of a fragment (10 μ M) designed to target either the RAMP or CLR protein within the receptor was used to displace SHF-1257. Therefore, a decrease in HTRF signal shows increased binding of the fragment to the fusion protein.

Prior to the experiment, the fusion proteins were purified as stated in chapter 3 section 3.2.3.2 but purified fractions were pooled and buffer exchanged in 25 mM HEPES pH 7.4, 150 mM NaCl, 0.2% Triton-X and 0.3% fatty-acid free BSA (Perkin Elmer) (working buffer). 5 μ L of 400 nM RAMP1-CLR or RAMP3-CLR fusion protein was added to each well of a 384-Well Low Volume ProxiPlate (PerkinElmer) followed by 5 μ L SHF-1257 at concentrations of 20 nM or 80 nM for the RAMP3-CLR and RAMP1-CLR fusion proteins respectively. 5 μ L of 40 μ M unlabelled fragment designed to target either the RAMP or CLR protein in the receptor was added to the reactions. Fragments were synthesised in house at the Department of Chemistry and Biological Engineering, University of Sheffield, by Olivier Zirimwabagabo and Professor Joe Harrity. MAb anti-HisTag terbium Cryptate (Cisbio) (stock: 8.8 ng/ μ L) was diluted 1:100 in working buffer and 5 μ L was added to each well. Final well volumes equated to 20 μ L. Control wells used contained excess (10 μ M) SHF-1036 (in the place of the fragment compound) to measure 0% SHF-1257 binding or no unlabelled fragment/compound to measure 100% binding signal maximum. Reactions were left overnight to equilibrate at room temperature and the HTRF signal was measured as previously stated in chapter 4 section 4.2.1. All reactions were plated in triplicates and 3 experimental repeats were run.

5.2.2 cAMP Assay: Peptide Stimulation on Overexpressing Cells

Identified fragments which hold a high affinity for displacing the SHF-1257 compound from the RAMP-CLR fusion proteins were used in a competition assay, measuring cAMP levels in cells overexpressing the CGRP, AM₁ and AM₂ receptors. cAMP is a secondary messenger which is activated when the CGRP, AM₁ and AM₂ receptors are activated and its loss of production would therefore indicate successful blocking of the receptor. To measure a cAMP response in these cell lines, the LANCE cAMP 384 well Kit (PerkinElmer) was used which

utilises time-resolved fluorescence resonance energy transfer (TR-FRET). The kit relies on the increase of cAMP through adenylate cyclase activity from the activation of GPCRs. The kit uses competition between a europium/streptavidin-labelled cAMP (EU/SA-cAMP) tracer complex and endogenous cAMP released from cells to compete for binding sites on cAMP specific antibodies conjugated with Alexa Fluor 647. When EU-cAMP binds to the anti-cAMP antibody, a light pulse at 340 nm excites the EU/SA-cAMP molecule which can transfer to the Alexa Fluor on the antibody due to their close proximity (figure 5.2). This in-turn emits light at 665 nm. When endogenously released cAMP binds to the anti-cAMP antibody, FRET does not occur and therefore no emission of light is observed at 665 nm (figure 5.2). This means a low signal is due to high ligand stimulation.

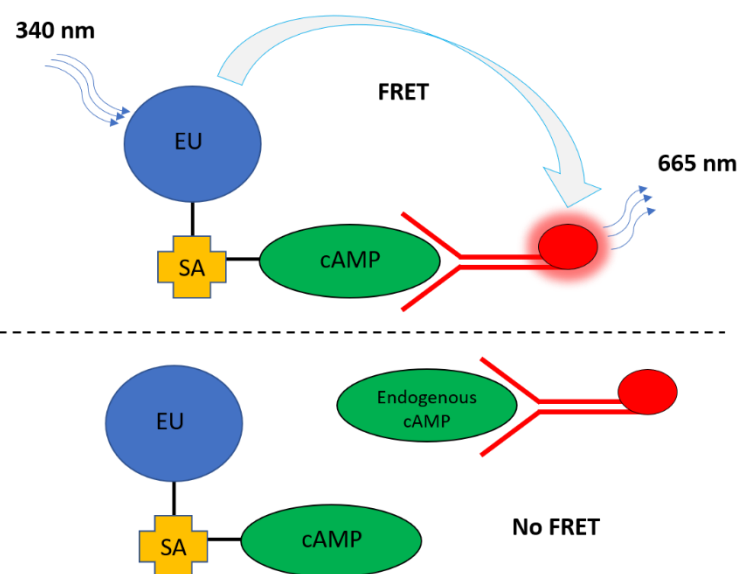


Figure 5.2: Schematic of the cAMP assay used to test CGRP, AM₁ and AM₂ receptor inhibition. The diagram shows how the production of endogenous cAMP from cells leads to the loss of FRET signal which therefore results in a low signal emission. If cAMP is not released from cells, the anti-cAMP antibody binds to the EU/SA/cAMP complex allowing a FRET response to occur

Prior to the start of the assay, a detection mix was made. 1.5 μ L Eu/SA (europium/streptavidin) was added to 25.5 μ L detection buffer and 5 μ L biotin labelled cAMP was added to 25 μ L detection buffer. 25 μ L of the diluted Eu/SA and biotin-cAMP solutions were mixed with 2.545 mL detection buffer at least 30 minutes prior to its addition to the plate to allow the Eu/SA-Biotin cAMP complex to form. The solutions were made up in the dark and the working solution covered in foil until use.

Serial dilutions of the ligand or forskolin for each assay were made in stimulation buffer, which contained 14 mL 1X Hanks balanced salt solution (-CaCl₂, -MgCl₂) (ThermoFisher Scientific), 75 μ L 1 M HEPES pH 7.4, 30 μ L 250 mM phosphodiesterase inhibitor, 3-isobutyl-1-methylxanthine (IBMX) (Sigma) and 100 μ L 7.5% solution BSA stabiliser (Perkin Elmer) (see table 5.1 for dilutions). Siliconised pipette tips were used during the preparation and dilution of ligands to prevent dilution cross over. Blanks were made with stimulation buffer only to allow the measurement of background signal in the reaction (0% stimulation). 4 mM forskolin was used as a positive/maximum stimulation during the cAMP assays. Forskolin is a membrane permeable labdane diterpene and is a direct stimulator of adenylate cyclase. This would produce the maximum possible stimulation of cAMP in cells and therefore can act as a positive control and a 100% cAMP maximum. The ligands/forskolin dilutions were added to a 384 Well OptiPlate (Perkin Elmer) (6 μ L/well) and the plate was spun at 1000 x g for 1 minute to sediment the solution.

Table 5.1: Ligand dilutions for cAMP stimulation

Tube	Final Concentration	Stimulation buffer (μ L)	Peptide amount (μ L)
1	2.00E-06	193	6.6 (60uM stock)
2	2.00E-07	180	20 from tube 1
3	2.00E-08	180	20 from tube 2
4	2.00E-09	180	20 from tube 3
5	2.00E-10	180	20 from tube 4
6	2.00E-11	180	20 from tube 5
7	2.00E-12	180	20 from tube 6
8	2.00E-13	180	20 from tube 7

1 mL aliquots of frozen CGRP, AM₁ and AM₂ overexpressing cells were thawed at 37°C in a water bath for ~30 seconds or until the liquid was almost fully thawed. Overexpressing cells (O/E) were 1321N1 cell lines (human brain astrocytoma cells) were purchased from DiscoverX and had been stably transfected with the CLR receptor and the RAMP1-3 proteins. The cell containing media was transferred into a 15 mL falcon tube and 1 mL of PBS was added to the cells and spun at 1,000 x g for 4 minutes. The supernatant was discarded, and cells were resuspended in fresh PBS and spun again at 1,000 x g for 4 minutes. The

supernatant was discarded, and cells were resuspended in 1 mL pre-heated stimulation buffer. Cells were then counted as previously stated (Chapter 1, section 2.2.7.3) and diluted to 2,500 cells/6 μ L.

An anti-cAMP antibody (from the PerkinElmer Kit) was added to the cells at a 1:100 dilution. Cells containing the antibody were then added to the plate (6 μ L/well), and the plate was spun at 1000 x *g* for 1 minute to sediment the liquid. The plate was briefly vortexed and incubated in the dark at 30°C for 20 minutes to allow the solutions to react (ligands to stimulate cAMP release). Following the incubation, the detection mix containing the Eu/SA-biotin cAMP complex was added to each well (12 μ L/well) and the plate was spun at 1000 x *g* for 1 minute to sediment the solution. The plate was briefly vortexed and incubated in the dark for 1 hour. After incubation, the plate was read using an EnSight multimode plate reader (PerkinElmer) (see 5.2 for conditions).

Table 5.2: Plate reader settings for the cAMP assay

Parameter	Instrument Settings
Excitation Filter	Lamp: 111 (UV2 320)
Emission	1) 615 2) 665
Delay Time	70 μ s
Number of Flashes	100
Window Time	100 μ s
Total Time	170 μ s

5.2.3 cAMP Assay: Competition Assay between fragments and peptides

The EC50 values generated from the previous ligand stimulation experiments (section 5.2.2) were used during cAMP competition assays with fragment antagonists. The detection mix, cell handling/amounts, antibody dilution and CGRP and AM ligands were diluted as previously stated in section 5.2.2. Novel fragment antagonists were diluted in the stimulation buffer (buffer components stated in section 5.2.2) to a final in-well concentration ranging from 1 mM to 10 nM. See table 5.3 for dilution set up. 3 μ L of the fragment antagonists were added to a 384-well OptiPlate (Perkin Elmer) and plates were spun a 1,000 x *g* for 1 minute to sediment the liquid. 6 μ L of cells overexpressing the CGRP, AM₁ or AM₂ receptors with anti-cAMP antibody were added to each well (2,500 cells/6 μ L)

and plates were spun at 1,000 x *g* for 1 minute. Plates were incubated in the dark and allowed to equilibrate for 30 minutes at room temperature.

3 μ L of 2.35 nM CGRP, 1.072 nM AM or 3.38 nM AM, were added to the wells containing the CGRP, AM₁ and AM₂ overexpressing cells respectively. Plates were spun at 1,000 x *g* for 1 minute and further incubated at 30°C in the dark for 20 minutes. 12 μ L of the pre-made detection mix was added to each well and the reaction was left in the dark for 1 hour at room temperature. After incubation, plates were read using an Ensign multimode plate reader as previously stated (table 5.2).

Table 5. 3: Fragment Dilutions from cAMP assay

Tube	Final Concentration	Stimulation buffer (μ L)	Fragment amount (μ L)
1	4.00E-03	57	3 (20 mM stock)
2	4.00E-04	90	10 from tube 1
3	4.00E-05	90	10 from tube 2
4	4.00E-06	90	10 from tube 3
5	4.00E-07	90	10 from tube 4
6	4.00E-08	90	10 from tube 5

5.2.3 Data analysis

5.2.3.1 Fragment Displacement Data

Data generated from the HTRF responses were plotted using GraphPad Prism Version 7.0 (GraphPad software, Inc). Values were first normalised by the fret signal measured (520 nm) divided by the antibody emission signal (620 nm) in each individual well. Single dose displacement responses (RAMP and CLR fragments) were further normalised to the 100% value of SHF-1257 binding calculated without the addition of unlabelled fragment antagonist, and 0% SHF-1257 binding calculated through the addition of excess SHF-1036 compound (unlabelled SHF-1257). The data was plotted as groups. Data bars represented show mean and standard deviation (SD) unless otherwise stated. To calculate significant difference between groups, an unpaired parametric T-test was used, along with an F-test to calculate possible significant difference in SD variation. Groups which showed significant

variance in SD (non-equal SD, $p < 0.05$), a Welch's T-test was applied to compare possible significant displacement differences between groups rather than the unpaired T-test. Comparisons which generated p values of <0.05 were considered significant.

5.2.3.2 cAMP Data

Data generated from the cAMP assays were plotted using GraphPad Prism Version 7.0 (GraphPad software, Inc). During peptide stimulation, data was normalised to a 100% cAMP response value (well containing excess forskolin) and 0% cAMP stimulation (well containing no ligand- buffer only). A non-linear 4 parametric curve was used to generate the EC₅₀ values of peptide stimulation.

During competition assays, data was normalised to the 100% receptor inhibition value (no ligand in the well- buffer only) and a 0% receptor inhibition value (no antagonist in the well, ligand only). Normalised data was plotted into GraphPad Prism Version 7.0 (GraphPad Software, Inc). A non-linear 4 parametric curve with 'top' constraint between 0 and 100 was applied to the data sets to calculate the estimated IC₅₀ values of the compounds.

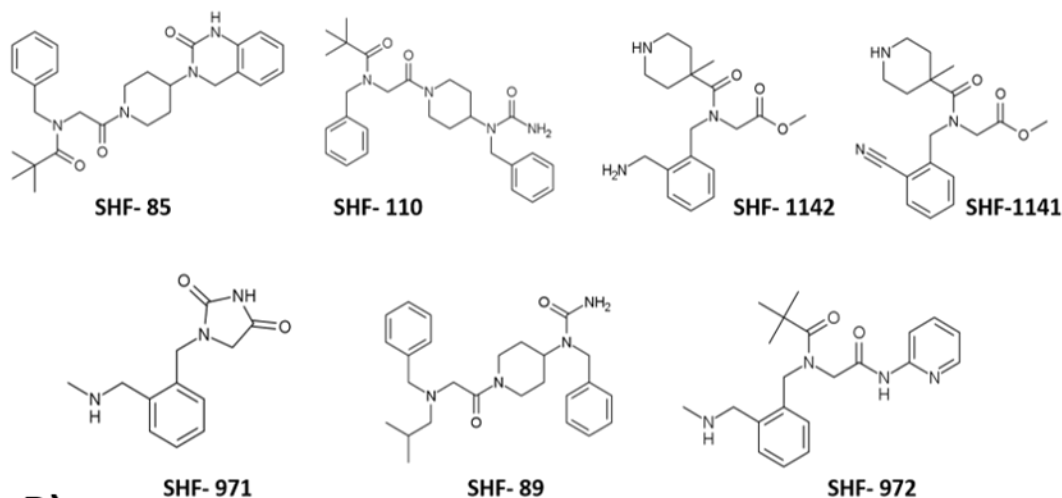
5.3 Results

5.3.1 RAMP Fragment Single Dose Displacement

With the protein functionality confirmed (chapter 4 section 4.3.2), fragments designed to target the RAMP protein in the receptor (RAMP fragments) were tested to determine if any structures held selectivity over the RAMP1 or RAMP3 proteins and whether any particular fragment held high affinity at either receptor. RAMP fragments were designed based on the previously developed CGRP antagonists and RAMP3 homology models created by mutagenesis of the CGRP receptor ECD (see chapter 1 figure 1.16 for model outline). A more in-depth outline of RAMP fragment design and receptor targeting is discussed in section 5.4.1 of this chapter.

The fragments were added as a single 10 μ M dose to monitor the displacement of the SHF-1257 compound. 3 experimental repeats were carried out with 3 repeat wells in each test (9 repeats total). For ease of view, fragment structures and plots were separated into two figures (figure 5.3 and figure 5.4).

A)



B)

RAMP Fragment Point Displacement HTRF Assay n=3

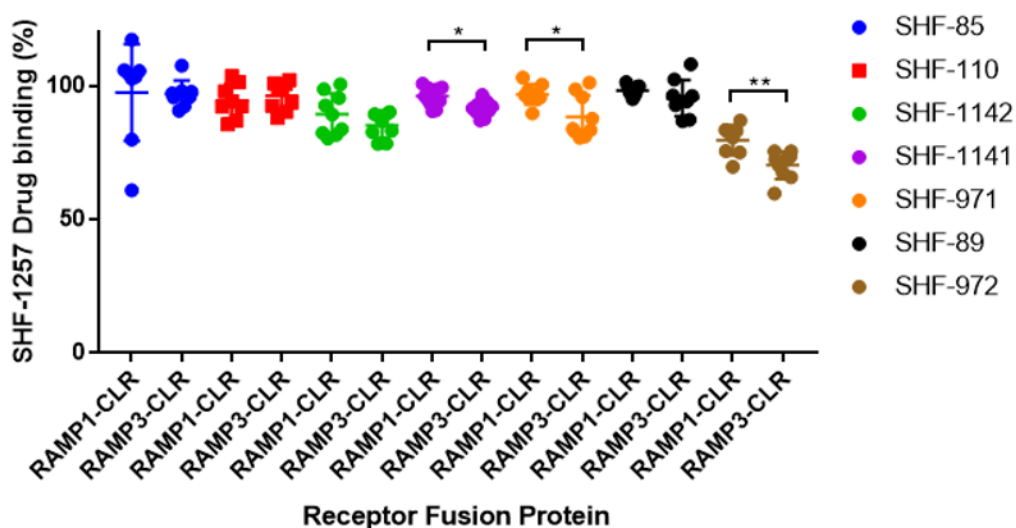
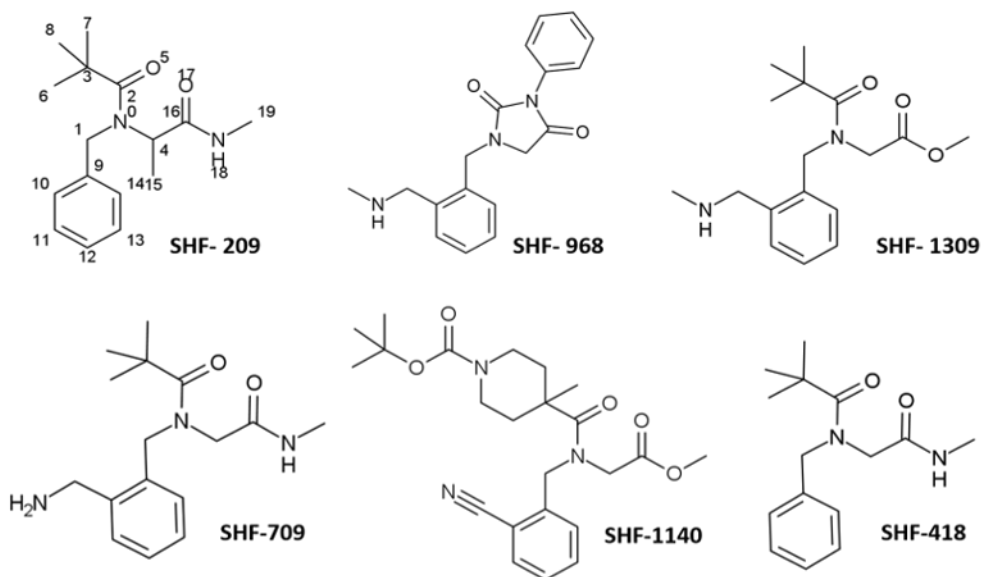


Figure 5.3: Single dose (10 μ M) displacement of SHF-1257 from RAMP1 and RAMP3-CLR fusion proteins using RAMP targeting fragments. (A) Structures of the RAMP fragments used during SHF-1257 displacement. (B) Percentage binding of SHF-1257 after the addition of a single dose of RAMP fragment. Fragment displacements on the two fusion proteins were plotted against each other and statistical differences were reported using an unpaired T-test. A F-test was also carried out to determine distribution of values around the mean. If significantly different ($p < 0.05$) a welches T-test was carried out to compare differences between data sets. Significant binding differences between groups ($p < 0.05$) are signified by stars (* ≤ 0.05 , ** ≤ 0.01 , * ≤ 0.001 , **** ≤ 0.0001)**

A)



B)

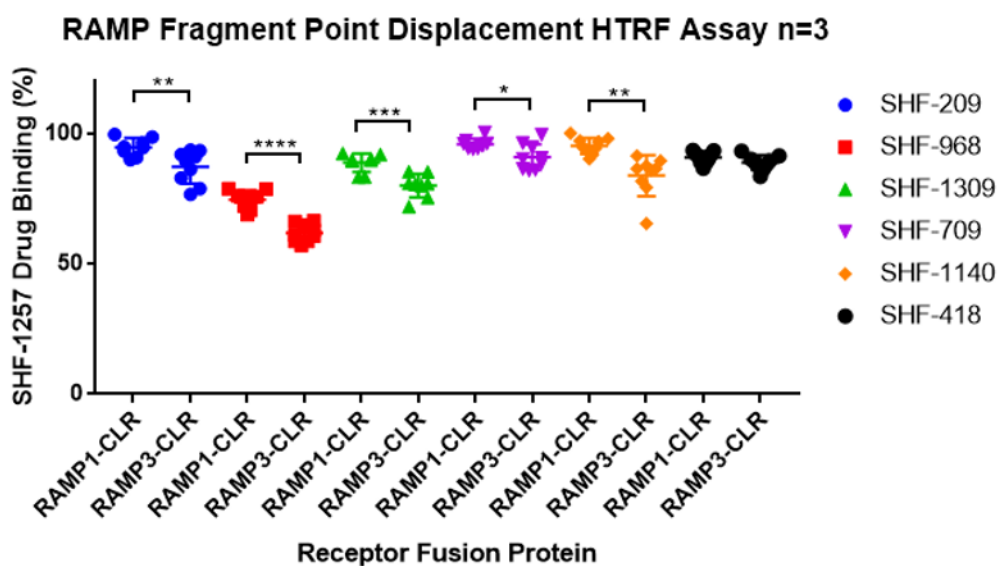


Figure 5.4: Single dose (10 μ M) displacement of SHF-1257 from RAMP1 and RAMP3-CLR fusion proteins using RAMP targeting fragments. (A) Structures of the RAMP fragments used during SHF-1257 displacement. (B) Percentage binding of SHF-1257 after the addition of a single dose of RAMP fragment. Fragment displacements on the two fusion proteins were plotted against each other and statistical differences were reported using an unpaired T-test. A F-test was also carried out to determine distribution of values around the mean. If significantly different ($p < 0.05$) a welches T-test was carried out to compare differences between data sets. Significant binding differences between groups ($p < 0.05$) are signified by stars (* ≤ 0.05 , ** ≤ 0.01 , * ≤ 0.001 , **** ≤ 0.0001)**

Across the range of fragments tested, displacement differences occurred when testing fragments between the 2 fusion proteins. SHF-968 shows the highest significant displacement difference between the 2 fusion proteins ($p < 0.0001$) with a mean binding difference of 12.85%, favouring the RAMP3-CLR fusion protein. Displacement of SHF-1257 by SHF-968 equates to $\sim 25.31\%$ on the RAMP1-CLR fusion protein and $\sim 38.16\%$ on the RAMP3-CLR fusion protein. SHF-968 offers the highest mean displacement of SHF-1257 across both fusion protein receptors (table 5.4). A Fragment which is structurally similar to SHF-968 includes SHF-971 (figure 5.5). The fragments only differ with SHF-938 holding an aniline ring (benzene linked to a nitrogen group) on the terminus of the fragment which enables significantly higher selectivity and binding affinity to the receptors ($P < 0.0001$). SHF-971 displaces the SHF-1257 antagonist by only 3.16% on the RAMP1-CLR fusion protein and 11.55% on the RAMP3-CLR fusion protein, a large decrease from SHF-968 which has 25.31% and 38.16% displacement on the requisite fusion proteins.

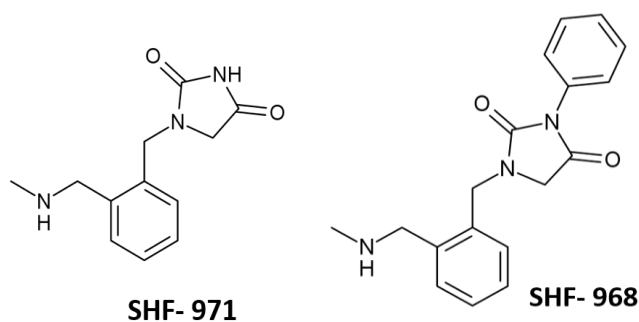


Figure 5. 5 Structural comparison of RAMP fragments SHF-971 and SHF-968

After structural comparison of the fragments, many hold similar structures with slightly varying functional groups. These subsequently enable differing affinities and selectivity over the receptors. Fragments which could be grouped based on their structure are SHF-418, SHF-209, SHF-709, SHF-1309 and SHF-927 (figure 5.6). SHF-418 is the simplest form of the fragments mentioned, displacing the SHF-1257 compound from the RAMP1-CLR and RAMP3-CLR fusion proteins by 9.23% and 11.15% respectively but offers no significant selectivity between the two receptor subtypes ($p = 0.1358$). SHF-209 has the addition of a methyl group on the compound backbone but otherwise remains structurally identical to SHF-418 (figure 5.6). SHF-209 offers increased selectivity for the RAMP3-CLR fusion protein over the RAMP1-CLR fusion protein ($p = 0.0087$) with displacements of 5.24% and 12.87% on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively. SHF-418 offers significantly

higher displacement of SHF-1257 on the RAMP1-CLR fusion protein in comparison to SHF-209 ($p = 0.0129$) but there was no significant difference in displacement on the RAMP3-CLR fusion protein between SHF-209 and SHF-418 ($p = 0.4703$).

SHF-709 is also structurally similar to SHF-418, holding the addition of a primary amine group from the terminal benzene ring, but does not have the additional methyl group found on the backbone of the SHF-209 structure (figure 5.6). SHF-709 holds significant selectivity for the RAMP3-CLR fusion protein over the RAMP1-CLR fusion protein ($p = 0.0167$), with mean displacements of 4.03% and 9.09% on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively. SHF-709 also has a significantly lower affinity on the RAMP1-CLR fusion protein in comparison to SHF-418 ($p = 0.0002$) but no significant difference on the RAMP3-CLR fusion protein between the 2 compounds ($p = 0.2907$). As with SHF-209, this implies that SHF-709 is increasing selectivity for the RAMP3-CLR fusion protein over the RAMP1-CLR fusion by reducing the fragments affinity for the RAMP1-CLR fusion protein.

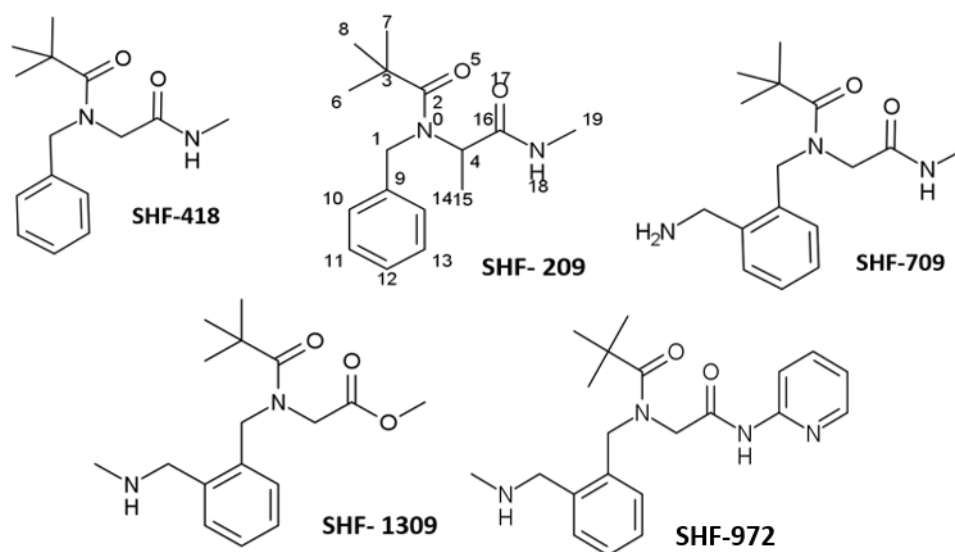


Figure 5.6: Structural comparison of RAMP fragments SHF-418, SHF-209, SHF-709, SHF-1309, SHF-972

SHF-1309 is a derivative of the SHF-418 group, and is structurally similar to SHF-709 but contains an extended methyl group on the primary amine group, changing it to a secondary amine on the fragment terminal, and the change of the nitrogen atom to oxygen group (generating an ester bond) on the opposite terminal end (figure 5.6). This leads to an

increased selectivity for the RAMP3-CLR fusion protein over the RAMP1-CLR fusion protein ($p = 0.0007$) with an 8.81% displacement difference between the two fusion proteins. Along with increased selectivity, the overall binding affinity for the two fusion proteins was increased, with a mean displacement of 11.19% on the RAMP1-CLR fusion protein and 20% on the RAMP3-CLR fusion protein, the highest mean displacement yet from the SHF-418 derivatives.

The SHF-972 fragment could be described as a structural development from SHF-1309 due to compound similarities but for SHF-972 holding a pyridyl group on the fragment terminus in place of the methyl group and linked by a nitrile group (in comparison to SHF-1309) (figure 5.6). These changes on SHF-972 lead to a significant increase in binding affinity to both fusion proteins in comparison to SHF-1309, ($p = 0.0017$ on the RAMP1-CLR receptor)($p = 0.001$ on the RAMP3-CLR receptor), with a mean displacement of 20.33% on the RAMP1-CLR fusion protein and 29.68% on the RAMP3-CLR fusion protein. This is an increase of 9.14% on the RAMP1-CLR fusion protein and 9.68% on the RAMP3-CLR fusion protein in comparison to SHF-1309. Despite the increase in binding affinity, the selectivity of SHF-972 between the two fusion proteins is decreased but still holds preferential affinity for the RAMP3-CLR fusion protein ($p = 0.0028$) in comparison to SHF-1309. SHF-972 is the most developed derivative of the SHF-418 structures and boasts of an increased displacement of 11.1% on the RAMP1-CLR fusion protein and 18.5% on the RAMP3-CLR fusion protein (in comparison to SHF-418), increasing the overall receptor binding affinities for both fusion protein receptors.

Three other fragments, SHF-1140, SHF-1141 and SHF-1142, also share similar structural confirmations (figure 5.7). SHF-1140 is another fragment which holds selectivity for the RAMP3-CLR fusion protein over the RAMP1-CLR fusion protein ($p = 0.0009$) with mean displacement values of 4.75% and 16.2% on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively. While the mean displacement on the RAMP-CLR fusion proteins by SHF-1140 is not as significant as other RAMP fragments such as SHF-968 ($p < 0.0001$ for both fusion proteins), it offers structural differences which may enable selectivity.

SHF-1140 is structurally similar to SHF-1141, with the exception of the removal of the ester bond leading to a trimethyl group on SHF-1140 (figure 5.7). The loss of this group shows a decrease in overall selectivity of SHF-1141 on the RAMP3-CLR and RAMP1-CLR fusion proteins ($p = 0.011$), but still remains selective for the RAMP3-CLR fusion protein. It also results in the loss of affinity for the RAMP3-CLR fusion protein in comparison to SHF-1140. Mean displacement by SHF-1141 is 3.744% and 8.48% on the RAMP1-CLR and RAMP3-CLR

fusion proteins respectively, resulting in an average loss of 7.72% displacement on the RAMP3-CLR fusion protein in comparison to SHF-1140. This indicates the additional groups in SHF-1140 potentially increases the fragment binding affinity to the RAMP3-CLR fusion protein.

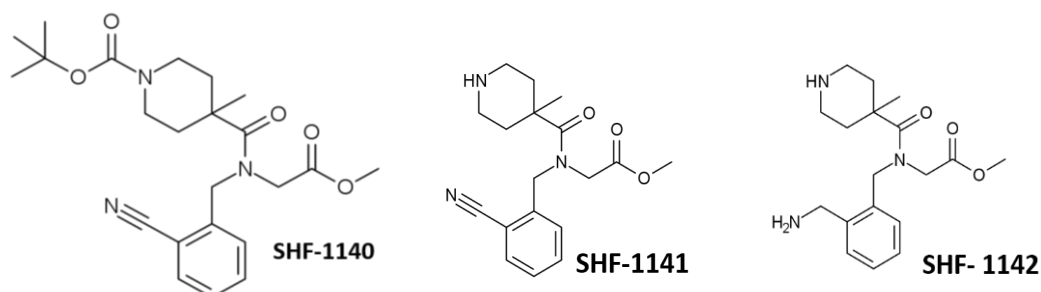


Figure 5.7: Structural comparison of RAMP fragments SHF-1140, SHF-1141 and SHF-1142

Further from this, SHF-1142 remains structurally identical to SHF-1141, with the exception of the change of the nitrile group on the terminus of SHF-1141 structure to a primary amine group in SHF-1142 (figure 5.7). The change resulted in there being no significant difference in SHF-1142 selectivity between the RAMP1-CLR and RAMP3-CLR fusion proteins ($p = 0.1813$), but did provide significantly higher affinity to both the RAMP1-CLR ($p = 0.0352$) and the RAMP3-CLR ($p = 0.0041$) fusion proteins in comparison to SHF-1141. The mean displacement values obtained were 10.5% and 14.74% for the RAMP1-CLR and RAMP3-CLR fusion proteins respectively, ~6-7% higher affinity for both fusion proteins. This change in nitrile bond results in a higher affinity to the fusion protein receptors, but less selectivity between the two RAMP proteins. All average displacement values and measurements from the RAMP fragment displacement can be found in table 5.4.

Table 5.4: Pharmacological profiles of RAMP fragments on displacing SHF-1257

Fragment used to displace SHF-1257	Fusion Protein: Average Displacement Values				
	RAMP1-CLR		RAMP3-CLR		p Value
	% Displacement	SEM	% Displacement	SEM	
SHF-85	2.416	+/- 6.398	3.031	+/- 1.783	0.9275
SHF-110	5.645	+/- 2.270	4.637	+/- 1.734	0.4872
SHF-1142	10.501	+/- 2.590	14.739	+/- 1.576	0.1813
SHF-1141	3.886	+/- 1.248	8.485	+/- 1	0.011
SHF-971	3.058	+/- 1.253	11.553	+/- 2.672	0.0117
SHF-89	1.67	+/- 0.685	4.546	+/- 2.242	0.2376
SHF-972	20.33	+/-2.002	29.676	+/- 1.718	0.0028
SHF-209	5.241	+/- 1.217	12.866	+/- 2.118	0.0087
SHF-968	25.312	+/- 1.140	38.023	+/- 1.163	<0.0001
SHF-1309	11.104	+/- 1.251	20	+/- 1.596	0.0007
SHF-709	4.031	+/- 0.752	9.089	+/- 1.635	0.0167
SHF-1140	4.754	+/- 1.048	12.106	+/- 2.606	0.0009
SHF-418	9.229	+/- 0.781	11.152	+/- 0.943	0.1358

5.3.2 CLR Fragments Single Point Displacement

Fragments have been designed which should determine the optimal structure which holds the highest affinity for the CLR protein. As noted, when testing the RAMP fragments, a single 10 μ M dose was used to monitor the displacement of the SHF-1257 compound. A boidpy labelled CLR fragment based on the high affinity SHF-1036 lead compound was generated to investigate the displacement of fragments (rather than the full length SHF-1257 compound) in an attempt to gain a higher degree of accuracy in results generated. However, attempted optimisation of labelled fragment binding saw high levels of background signalling due to the concentration of the fragment required (> 200 nM) leading to inconsistent results being generated. This meant a labelled fragment could not be used in the assay due to its low affinity to the fusion protein receptors and SHF-1257 was used during displacement testing. 3 experimental repeats were carried out with 3 repeat wells in each test (9 repeats total).

For ease of view, fragment structures and plots were separated into two figures (figure 5.8 and figure 5.9).

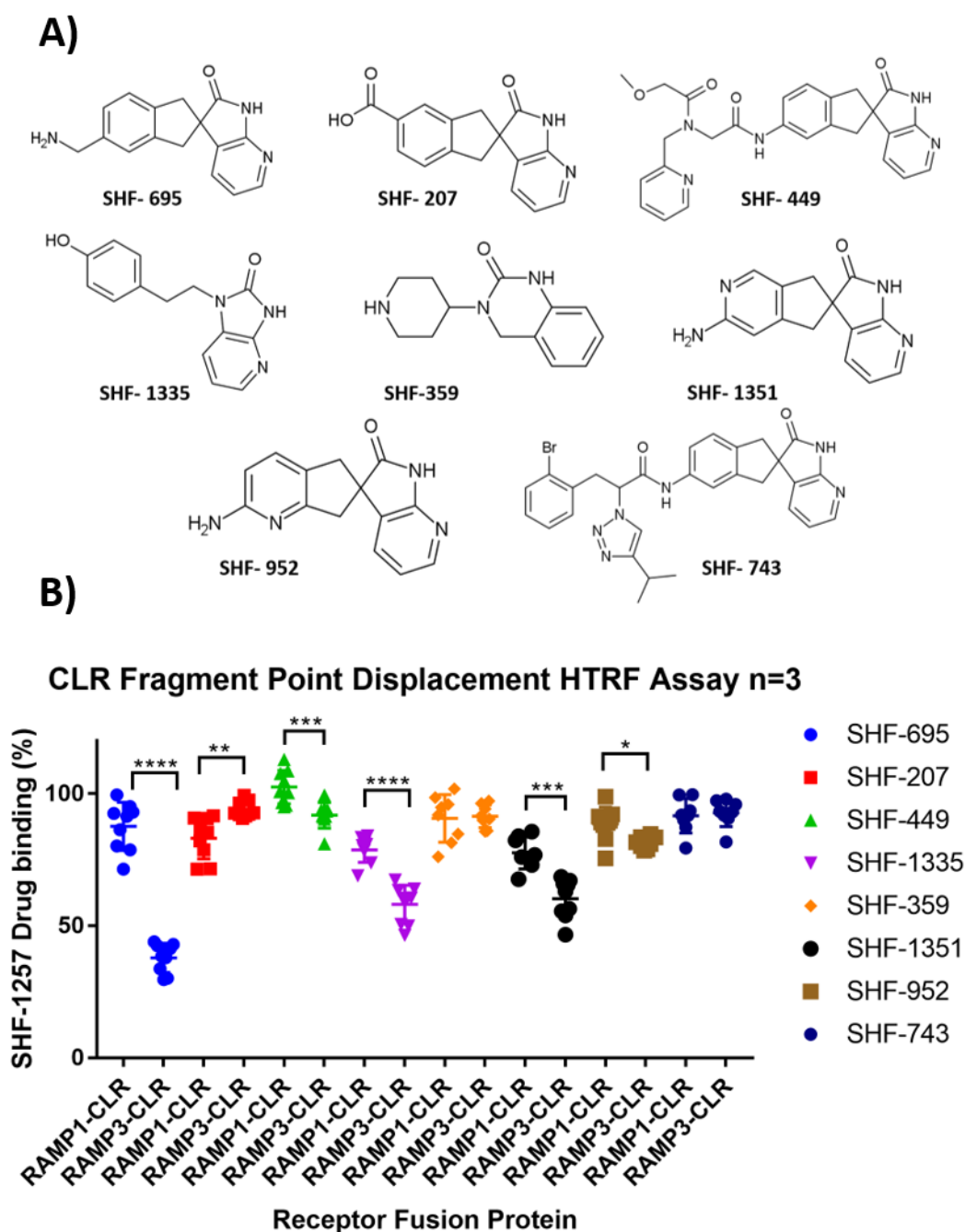
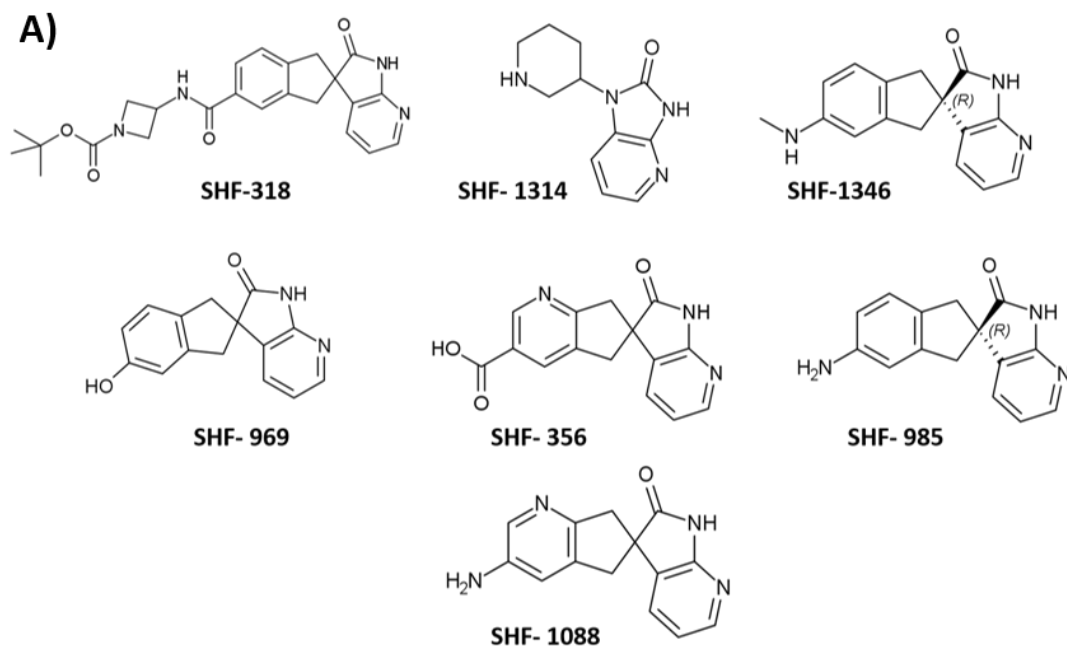


Figure 5.8: Single dose (10 μ M) displacement of SHF-1257 from RAMP1 and RAMP3-CLR fusion proteins using CLR targeting fragments. (A) Structures of the CLR fragments used during SHF-1257 displacement. (B) Percentage binding of SHF-1257 after the addition of a single dose of CLR fragment. Fragment displacements on the two fusion proteins were plotted against each other and statistical differences were reported using an unpaired T-test. A F-test was also carried out to determine distribution of values around the mean. If significantly different ($p < 0.05$) a welches T-test was carried out to compare differences between data sets. Significant binding differences between groups ($p < 0.05$) are signified by stars (* ≤ 0.05 , ** ≤ 0.01 , * ≤ 0.001 , **** ≤ 0.0001)**



B)

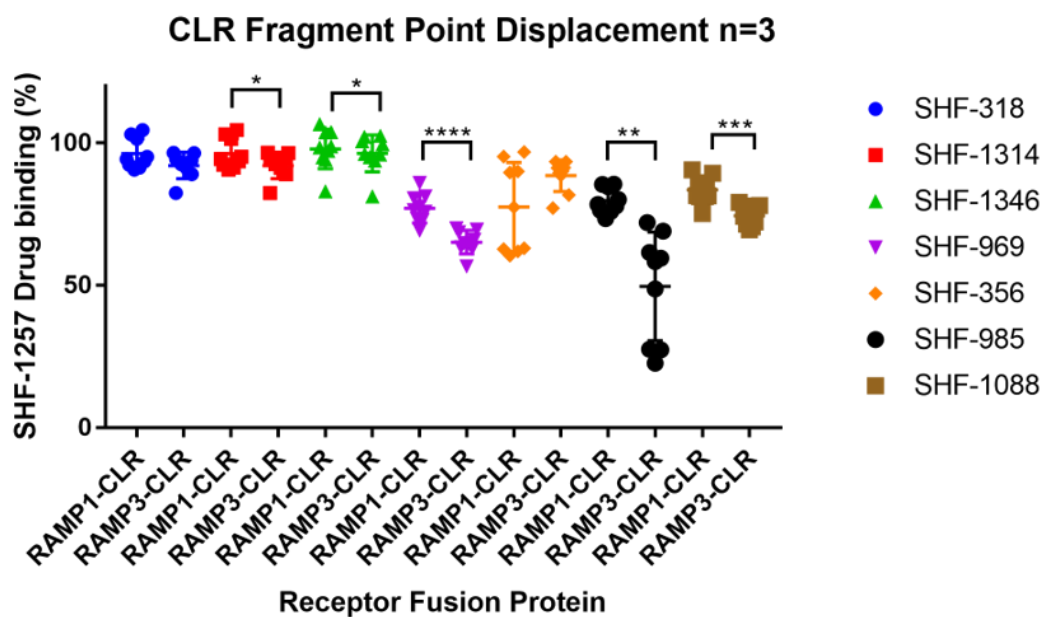


Figure 5.9: Single dose (10 μ M) displacement of SHF-1257 from RAMP1 and RAMP3-CLR fusion proteins using CLR fragments. (A) Structures of the CLR fragments used during SHF-1257 displacement. (B) Percentage binding of SHF-1257 after the addition of a single dose of CLR fragment. Fragment displacements on the two fusion proteins were plotted against each other and statistical differences were reported using an unpaired T-test. A F-test was also carried out to determine distribution of values around the mean. If significantly different ($p < 0.05$) a welches T-test was carried out to compare differences between data sets. Significant binding differences between groups ($p < 0.05$) are signified by stars (* ≤ 0.05 , ** ≤ 0.01 , * ≤ 0.001 , **** ≤ 0.0001)**

Comparison of the CLR fragments shows structural similarities but vastly different displacement percentages on the RAMP1 and RAMP3-CLR fusion proteins. A similar group of CLR fragments could be reported in SHF-1351, SHF-952, SHF-1088, SHF-985, SHF-1346 and SHF-695 due to their azaindole group connected to the indane group and a branched amine group (figure 5.10). The simplest fragment could be seen as SHF-1088. This fragment leads to a mean displacement of 16.46% on the RAMP1-CLR fusion protein and 25.5% on the RAMP3-CLR fusion protein. Statistically, there is also an increase in selectivity for the RAMP3-CLR fusion protein over the RAMP1-CLR fusion protein ($p = 0.0003$).

SHF-1351 is structurally identical to SHF-1088, with the exception of the nitrogen atom replacing carbon-5 in the indane group rather than carbon-6 (figure 5.10). This results in an average displacement value of 22.37% on the RAMP1-CLR fusion protein and 39.71% on the RAMP3-CLR fusion protein, with the fragment showing selectivity for the RAMP3-CLR fusion protein ($p < 0.0001$). In comparison to SHF-1088, SHF-1351 shows significantly higher displacement values on the RAMP1-CLR fusion protein ($p = 0.0391$) and on the RAMP3-CLR fusion protein ($p < 0.0001$), with mean displacement values of 5.91% and 14.21% higher respectively on displacing SHF-1257. The data suggests that the replacement of the nitrogen group leads to increased potency to the RAMP1/3-CLR fusion proteins.

Another structurally similar fragment is SHF-952. It is structurally the same as SHF-1351 and SHF-1088 but repositions the nitrogen group to carbon-3 instead of carbon 6 or 5 (figure 5.10). SHF-952 displaces the SHF-1257 fragment by 11.47% on the RAMP1-CLR fusion protein and 18.51% on the RAMP3-CLR fusion protein, which shows slight selectivity for the RAMP3-CLR fusion protein ($p = 0.0104$). When compared with its higher affinity counterpart SHF-1351, it shows significantly reduced displacement on the RAMP1-CLR fusion protein ($p = 0.0027$) (decrease of 10.9%) and on the RAMP3-CLR fusion protein ($p < 0.0001$) (decrease of 21.2%). Additional to this, if you compare SHF-952 displacement to SHF-1088, there is no significant change in potency to the RAMP1-CLR fusion protein but there is significantly less binding to the RAMP3-CLR fusion protein ($p = 0.0003$) (decrease of 6.99%).

SHF-695 is structurally similar but contains no nitrogen atoms in the indane structure and extends the amine group by an extra hydrocarbon bond in comparison to SHF-1351, SHF-952 and SHF-985 (figure 5.10). This results in displacement values of 12.38% and 62.15% on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively. SHF-695 is highly selective for the RAMP3-CLR fusion protein in comparison to the RAMP1-CLR fusion protein ($p < 0.0001$).

and displacement on the RAMP3-CLR fusion protein is the largest mean value of any CLR fragment on either fusion protein (figures 5.8 and 5.9).

SHF-985 and SHF-1346 are also structurally similar and are the right (R) enantiomers of their compounds, offering interesting comparisons (figure 5.10). SHF-985 holds the same structure as the previously described CLR fragments, with the indane and azaindole groups and a primary amine group from the indane, similar to that of SHF-1088 but with no nitrogen group in the indane moiety (figure 5.10). SHF-1346 holds this same structure, but a methyl group is extended from the terminal amine, to give a secondary amine group (figure 5.10). SHF-985 shows mean displacement values of 20.24% on the RAMP1-CLR fusion protein and 50.34% on the RAMP3-CLR fusion protein but values are spread in relation to RAMP3-CLR fusion protein displacement. Binding also shows significant selectivity for the RAMP3-CLR fusion protein over the RAMP1-CLR fusion protein ($p = 0.0013$). The change of this primary amine group on SHF-1346 results in a loss in affinity to both receptors. The mean displacement of SHF-1257 on the RAMP1-CLR fusion protein and RAMP3-CLR fusion protein is 2.1% and 7.73% respectively. This is an average decrease of 18.14% on the RAMP1-CLR fusion protein and 42.61% on the RAMP3-CLR fusion protein, stating the position of this primary amine group is key for increasing fragment affinity on the receptor.

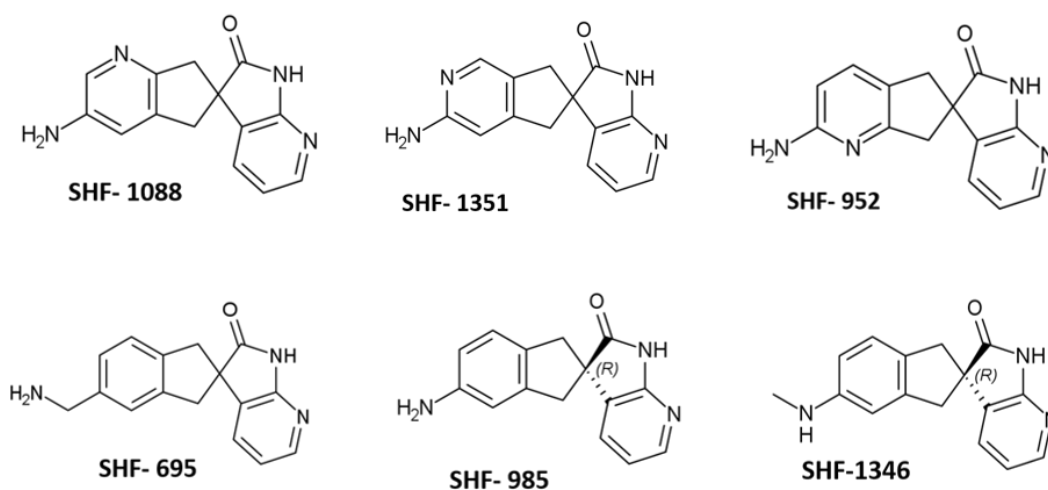


Figure 5.10: Structural comparison of CLR fragments SHF-1088, SHF-1351, SHF-952, SHF-695, SHF-985, SHF-1346

Another similar group of fragments includes SHF-969, SHF-356 and SHF-207 (figure 5.11). While these fragments are similar to those previously described, the branched primary amine group is changed between hydroxyl- or carbonyl- group variants (figure 5.10 and

5.11). SHF-969 could be described as the simplest form, with the azaindole group attached to the indane group and the extension of a hydroxyl group in place of the primary amine shown in previous fragments. SHF-969 showed a mean SHF-1257 displacement of 22.93% and 34.83% on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively, with SHF-969 being significantly selective for the RAMP3-CLR fusion protein ($p < 0.0001$)

SHF-207 holds a similar structure to SHF-969, holding the same indane group connected to the azaindole group but holds a carboxyl group in comparison to the hydroxyl group in the SHF-207 (figure 5.11). This group is also positioned on the benzene carbon-5 in the indole group in comparison to the carbon-4 in SHF-969. Changes result in SHF-207 displacing SHF-1257 by 16.81% and 6.07% on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively. This is a significant decrease ($p < 0.0001$) in SHF-1257 displacement on the RAMP3-CLR fusion protein in comparison to SHF-969, with a mean displacement difference of 28.76%. There is no significant difference in displacement between the SHF-207 and SHF-969 on the RAMP1-CLR fusion protein. Interestingly, SHF-207 holds a higher selectivity for the RAMP1-CLR fusion protein over the RAMP3-CLR fusion protein ($p = 0.0031$) due to the reduction of displacement on the RAMP3-CLR fusion protein. SHF-207 is the only fragment of all CLR and RAMP fragments which holds significant selectivity for the RAMP1-CLR fusion protein.

SHF-356 holds the same azaindole group connected to the indane structure and holds a carboxyl group in the equivalent hydroxyl group of SHF-969 (figure 5.11). It also contains a nitrogen atom in place of the carbon-6 of the benzene ring of the indole group which can be compared to that of SHF-1088 (figure 5.10). SHF-356 displaces SHF-1257 by an average of 22.49% and 12.48% on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively. Despite there being an average of 10% difference in displacement values between the two fusion protein receptors, due to the large deviation of values around the mean on SHF-356 on the RAMP1-CLR fusion protein, it resulted in there being no significant difference in selectivity between the RAMP1/3-CLR fusion proteins ($p = 0.0637$). A comparison between SHF-356 and SHF-1088 which hold identical structures with the only difference being between the primary amine and carboxyl groups, show SHF-1088 has significantly more displacement on the RAMP3-CLR fusion protein ($p < 0.0001$) in comparison to SHF-356 and no significant difference on the RAMP1-CLR fusion protein. This shows that the addition of the amine group on the benzene ring in SHF-1088 is having a greater effect on RAMP3-CLR fusion protein fragment affinity in comparison to the hydroxyl group on SHF-356.

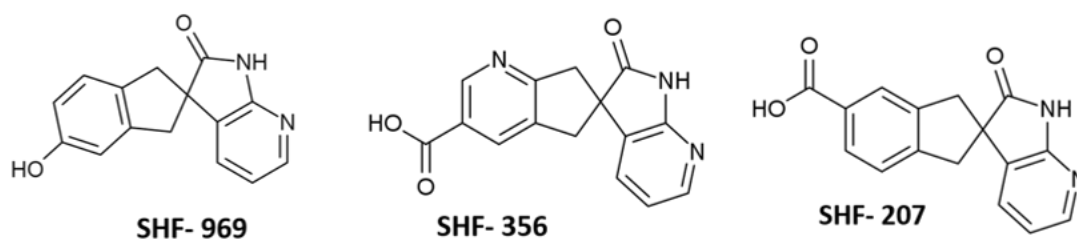


Figure 5. 11: Structural comparison of CLR fragments SHF-969, SHF-356 and SHF-207

Other comparable fragments include SHF-395, SHF-1314 and SHF-1335 (figure 5.12). These are comparable due to the change of the indole group into the piperidine ring, which is similar to the CLR end of the early developed CGRP antagonists, olcegepant and telcegepant (figure 5.1). There is also the slight change in the addition of an amine group which links the piperidine ring to the indole group (quinazoline group in SHF-359) (figure 5.12). SHF-359 displaces SHF-1257 by an average of 9.32% and 8.56% on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively with no significant selectivity for either fusion protein receptor. SHF-1314 changes back to the azaindole group but maintains the nitrogen group linking to the piperidine ring (figure 5.12). There is also a repositioning of the nitrogen in the pyridyl ring, replacing the carbon-3 instead of the carbon-4 atom in SHF-359 (figure 5.12). This leads to a displacement of -3.8% on the RAMP1-CLR fusion protein and 3.63% on the RAMP3-CLR fusion protein, with SHF-1314 being slightly selective for the RAMP3-CLR fusion protein over the RAMP1-CLR fusion protein ($p = 0.042$). Despite this, the overall displacement of SHF-1257 on both receptors by SHF-1314 is minimal in comparison to other fragments.

Another fragment comparable to SHF-1314 is SHF-1335 (figure 5.12). The SHF-1335 fragment holds the azaindole group with the added nitrogen, but the extended ringed structured changes. The group extends further away from the body of the fragment by an extra hydrocarbon bond followed by the addition of a phenol group terminus, in comparison to the pyridyl group of SHF-1314. These changes lead to average displacement values of 21.21% and 41.9% on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively. The fragment is also significantly selective for the RAMP3-CLR fusion protein over the RAMP1-CLR fusion protein ($p < 0.0001$). These displacement values are also significantly higher than the previously tested fragments which hold similar structures (SHF-359 and SHF-1314) on both fusion proteins ($p < 0.0001$) with displacements differences of >10% and >30% between the fragments on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively. All CLR fragment displacement data is summarised in table 5.5.

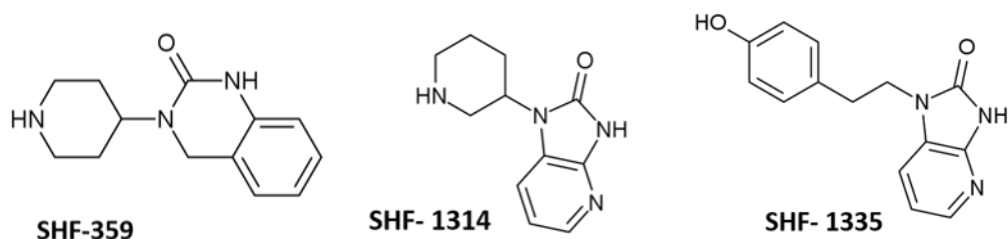


Figure 5.12: Structural comparison of CLR fragments SHF-359, SHF-1314 and SHF-1335

Table 5.5: Pharmacological profiles of CLR fragments on displacing SHF-1257

Fragment used to displace SHF-1257	Fusion Protein: Average Displacement Values				
	RAMP1-CLR		RAMP3-CLR		p Value
	% Displacement	SEM	% Displacement	SEM	
SHF-695	12.384	+/- 3.029	62.147	+/- 1.831	< 0.0001
SHF-207	16.807	+/- 2.618	6.071	+/- 1.016	p= 0.0031
SHF-449	-2.535	+/- 2.120	8.106	+/- 1.612	p= 0.0010
SHF-1335	21.208	+/- 1.577	41.805	+/- 2.432	p< 0.0001
SHF-395	9.319	+/- 3.176	8.565	+/- 1.529	p= 0.8337
SHF-1351	22.371	+/- 2.133	39.713	+/- 2.476	p= 0.0001
SHF-952	11.471	+/- 2.141	18.506	+/- 0.808	p= 0.0104
SHF-743	8.335	+/- 2.307	7.249	+/- 1.833	p= 0.7174
SHF-318	3.718	+/- 1.140	38.023	+/- 1.163	p= 0.1041
SHF-1314	11.104	+/- 1.766	7.922	+/- 1.638	p= 0.0420
SHF-1346	2.102	+/- 2.326	3.633	+/- 2.159	p= 0.0471
SHF-969	22.972	+/- 1.679	34.831	+/- 1.388	p< 0.0001
SHF-356	22.403	+/- 5.213	11.479	+/- 1.841	p= 0.0744
SHF-985	20.242	+/- 1.502	50.339	+/- 6.356	p= 0.0013
SHF-1088	16.458	+/- 1.581	25.501	+/- 1.215	p= 0.0003

5.3.3 Fragment Inhibition on Receptor Overexpressing Cell lines

The 2 highest affinity RAMP and CLR fragments from the displacement experiments (SHF-695 and SHF-968) were applied to a cAMP assay to measure their abilities to inhibit a cAMP response on a full length CGRP, AM₁ and AM₂ receptors. The fragments were applied either singularly or together, to assess how effectively they antagonise the receptor binding site. SHF-85 was used which acted as a low affinity control due to its low displacement value on the single point displacement assay.

Initially, ligands were applied to the O/E cell lines to generate EC₅₀ values, which would be incorporated into future competition assays. The CGRP peptide stimulation on the CGRP receptor O/E cells produced an EC₅₀ value of 0.588 nM and AM peptide stimulation on AM₁ and AM₂ O/E cells produced EC₅₀ values of 0.269 nM and 0.846 nM respectively (figure 5.13). EC₅₀ values and parameter analysis is summarised in table 5.6. These values were incorporated into the competition assay with the RAMP/CLR fragments.

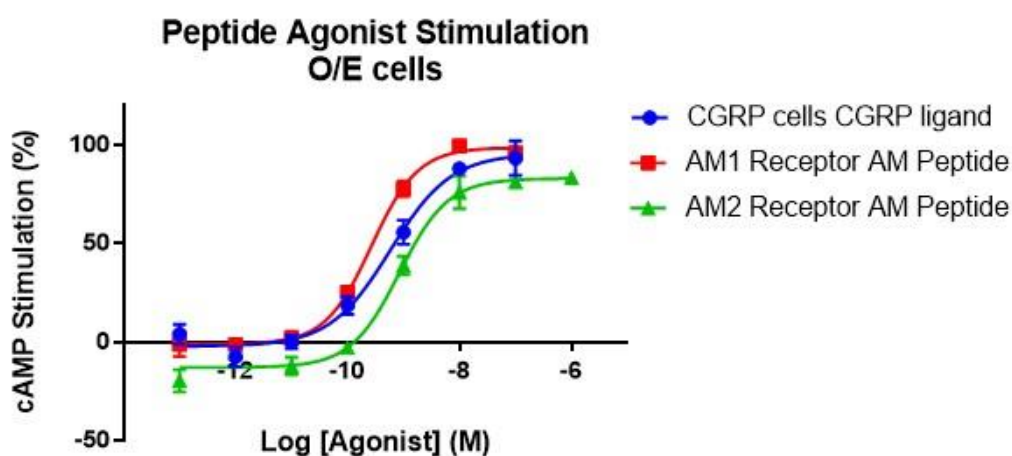


Figure 5.13: Agonist stimulation of 1231N1 O/E cell lines. The potency of the CGRP and AM on their receptors and their ability to stimulate a cAMP response is outlined.

Table 5.6: Best fit values of ligand stimulation of CGRP, AM₁ and AM₂ O/E cells

EC50 Dose Response Curve: Peptides	Receptor O/E cells		
	CGRP cells	AM ₁ Cells	AM ₂ cells
EC50 Value	0.588 nM (Log-9.231)	0.269 nM (Log-9.57)	0.846 (Log-9.072)
Std Error: LogEC50	+/- 0.09342	+/- 0.04202	+/- 0.07661
R square	0.9805	0.9948	0.9876
Best fit: Top	95.49	98.58	82.88
Best Fit: Bottom	-2.291	-1.276	-13

Fragment responses on the CGRP receptor showed IC₅₀ values of 56.59 μ M, 86.92 μ M and 76.02 μ M for SHF-695, SHF-968 and SHF-695/968 combined wells respectively (figure 5.14). The SHF-85 response on the CGRP receptor produced ambiguous results in the response curve. Fragment responses on the AM₁ receptor showed ambiguous results in the response curves for SHF-695, SHF-695/968 combined wells and SHF-85 (figure 5.14). The response curve for the SHF-968 was not converged (figure 5.14). Fragment responses on the AM₂ receptor O/E cells showed IC₅₀ values of 55.35 μ M and 216.3 μ M for SHF-695 and SHF-695/968 combined wells respectively (figure 5.14). Both the SHF-968 and SHF-85 responses on the AM₂ receptor produced ambiguous curves. The IC₅₀ values and parameters of the response curves are outlined in table 5.7.

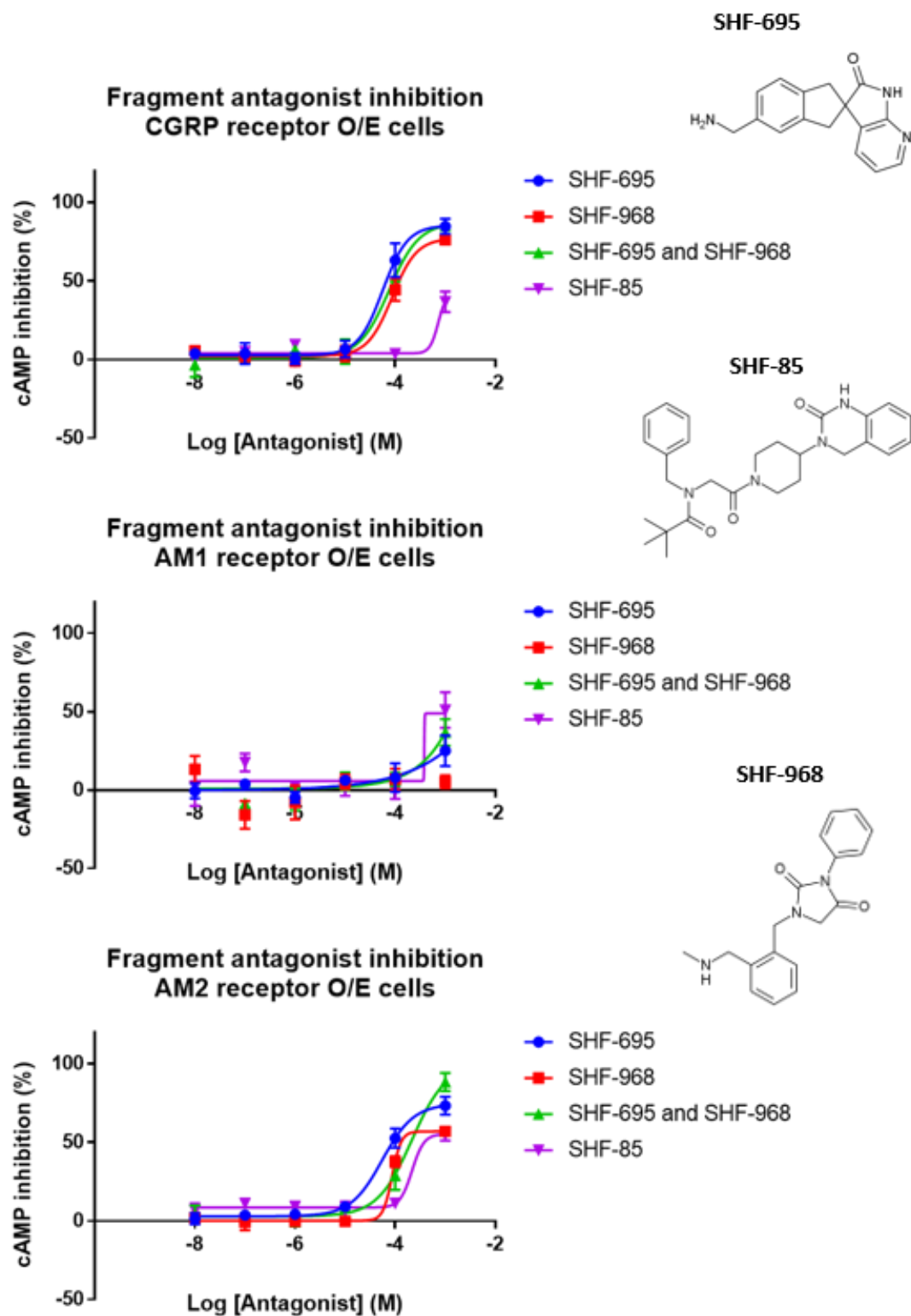


Figure 5.14: Activity of small molecule fragment antagonists against the CGRP, AM₁ and AM₂ receptors. Fragment antagonists were tested for their ability to block peptide interactions with their receptors to inhibit a cAMP response. The structures of the fragments used during the assay are outlined.

Table 5.7: Best fit values from fragment inhibition of the CGRP, AM₁ and AM₂ receptors.

Receptor	Best Fit Values			
	IC ₅₀	LogIC ₅₀ Std Error	Top	Bottom
SHF-695				
CGRP	56.59 μM (Log-4.247)	+/- 0.1444	85.240	2.729
AM₁	-	-	-	-
AM₂	55.35 μM (Log-4.257)	+/- 0.1008	74.440	2.901
SHF-968				
CGRP	86.92 μM (Log-4.061)	+/- 0.06363	77.090	2.100
AM₁	-	-	-	-
AM₂	-	-	-	-
SHF-968 and SHF-695				
CGRP	76.02 μM (Log-4.119)	+/- 0.08882	86.92	1.176
AM₁	-	-	-	-
AM₂	216.3 μM	+/- 0.6048	~100 (hit constraint)	2.800
SHF-85				
CGRP	-	-	-	-
AM₁	-	-	-	-
AM₂	-	-	-	-

5.4 Discussion

The main aim of this chapter was to identify high affinity fragments designed to target the CLR region or RAMP binding region of the RAMP1-CLR and RAMP3-CLR fusion proteins. Here, high affinity fragments against the CLR binding region and RAMP binding region have been identified, which show selectivity for the RAMP3-CLR fusion protein over the RAMP1-CLR fusion protein. Interestingly, there appeared to be more displacement variance when testing fragments which occupied the CLR binding region of the receptor, despite the CLR protein sequence being identical between the 2 RAMP-CLR fusion proteins. The fragments were tested on their ability to antagonise the full-length receptor. When tested either singularly or combined, inhibition showed IC₅₀ values in the μM range. Key differences and functional groups from each fragment which was allowing higher affinity binding will now be assessed and combined to design full-length compounds which could hold high potencies for the CGRP and AM₂ receptors.

5.4.1 RAMP Fragment Association to the RAMP1/3-CLR Fusion Proteins

The results from the RAMP fragment analysis shows the progression in generating a fragment which displaces the SHF-1257 compound at the RAMP end of the receptor. It is important to note that while we have identified fragments which show selectivity over the RAMP3-CLR fusion protein, these fragments were displacing SHF-1257, a novel CGRP and AM₂ antagonist with a BODIPY-FL tag. The addition of the tag resulted in the loss of binding potency to the receptor in comparison to the non-labelled compound (SHF-1036, chapter 4, table 4.3) and therefore may not fit the receptor binding pocket as efficiently, particularly at the RAMP/CLR end where the compound has had the addition of the tag. Displacement may therefore not give true values of where the non-tagged SHF-1036 compound is associated to the receptor. On a final note, fragments are competing to displace the SHF-1257 compound and not with the endogenous receptor peptide. Fragments may therefore be competing with residues which may have no effect on the peptide binding and CGRP/AM₂ receptor activation. They may therefore bind with a higher affinity but not antagonise the receptor and stop signal transduction.

Nevertheless, we have shown the design of fragments which hold high affinity to the receptor. While 10 μM remains a high concentration for drug displacement, fragments were attempting to displace the full-length compound while only targeting one region of the

receptor. It would therefore require a higher concentration to displace SHF-1257 in comparison to conventional full-length antagonists, as seen in the SHF-638 and SHF-1036 compound displacements (Chapter 4, section 4.3.4). As the RAMP/CLR end of the compound is expected to hold less contacts to the RAMP-CLR fusion proteins, there was expected to be less overall displacement of the compound on the fusion protein, but a higher amount of variance between the fusion proteins and each individual fragments due to different residue occupancy.

RAMP fragments were designed to target regions of both the RAMP3 protein and the CLR receptor, which would form the interface of the compound (figure 5.15). Fragments were attempting to target RAMP3 residues Asp71 and Glu74, and CLR residues Arg38, Thr37 and Phe92 (figure 5.15). Comparisons could be made to that of olcegepant which extends its pyridyl rings to form contacts with CLR Phe92 and Asp94 (ter Haar *et al.*, 2010) (figure 5.15). Figure 5.15 outlines the region of the receptor which was being targeted by the RAMP fragment.

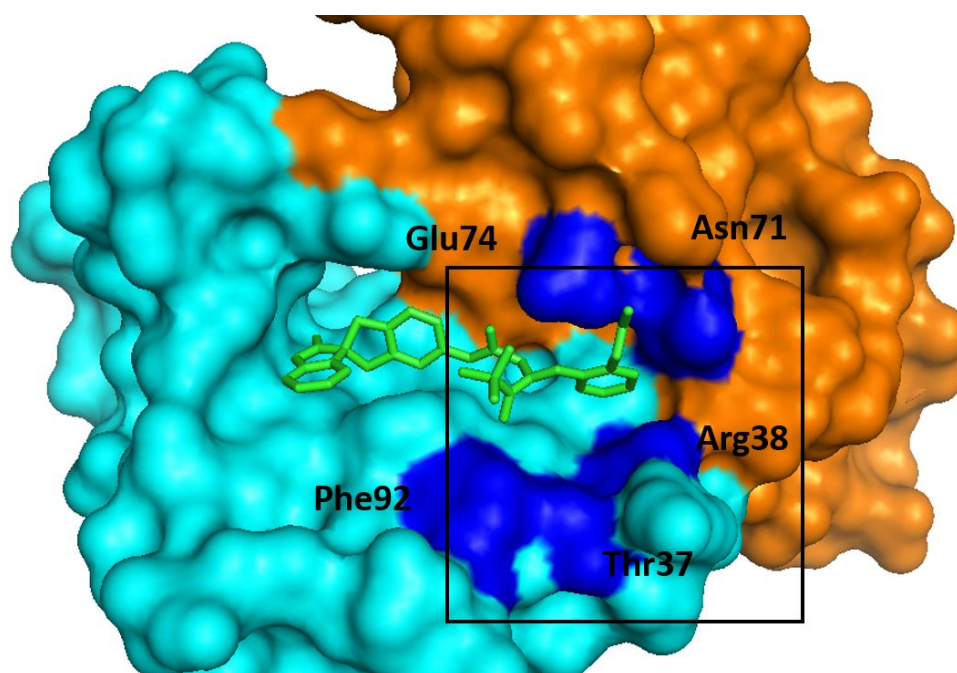


Figure 5.15: Region of RAMP fragment targeting during FBDD. The homology model of the RAMP3 (orange) and CLR (cyan) ECD superimposed with SHF-638 (green). Residues which were being targeted by the RAMP fragment during drug design are coloured blue and overall region outlined with the black box.

Fragments tested varied in characteristics which could potentially be utilised depending on which traits are desired from the generated compound. SHF-968 showed the highest displacement value on both fusion proteins (25.31% and 38.02% on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively) and provided the largest significant increase in selectivity between the RAMP1 and RAMP3-CLR fusion proteins (~12.7% difference, $p < 0.0001$) (figure 5.4). The fragment is distinguishable due to its hydantoin centre, while most other fragments hold 2 inward facing, carbonyl groups in the equivalent position. This could be consistently used across the fragments to gain a higher affinity to both protein receptors. The hydantoin centre does not appear to form any vital bonds with the RAMP-CLR fusion protein (SHF-971) due to low SHF-1257 displacement and therefore could be changeable to allow easier design for fragment/compound development.

Fragments could possibly hold a higher affinity if there was the addition of the benzene ring to extend the fragment terminus connected by an amine linker (towards the CLR binding region), which was seen to increase the affinity when added to SHF-972 (figure 5.6) (20.33% and 29.67% displacement on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively). This could be incorporated on the hydantoin centre of SHF-968 at the same position of the double bonded oxygen group (figure 5.16). However, it may be seen that the addition of an oxyl group in place of the amine bond could increase the selectivity of the compound for the RAMP3-CLR fusion protein (comparison between SHF-709 and SHF-1309), although it is unclear at this point why this would enable selectivity (figure 5.16).

The addition of the benzene group with amine linker (combination of SHF-968 and SHF-972) may cause stability issues in the compound however, with the oxygen group (which would change to a hydroxyl group upon the addition of the benzene group, figure 5.16) from the imidazoline centre in close proximity, it may lead to compound decomposition (figure 5.16 outlines the compound decomposition issues). If the now hydroxyl group from the imidazoline centre, was changed to a hydrocarbon group (methyl), it would partly solve potential decomposition issues. Additional to this, the nitrogen group at the top of the imidazoline structure would only be separated by a single bonded hydrocarbon residue from the added amine bond and would therefore form instability issues (figure 5.16). The nitrogen atom was therefore removed and changed to a hydrocarbon group to prevent further issues (figure 5.16). This now changes the centre of the fragment to a pyrrolidine centre (figure 5.16). The importance of the removed oxygen and nitrogen group on compound binding/potency however is unclear but did not show vast displacement before the addition of the benzene ring on either fusion protein (SHF-971: 3.05% and 11.553% displacement on

the RAMP1-CLR and RAMP3-CLR fusion protein respectively in comparison to SHF-968 with the benzene ring: 25.31% and 38.02% displacement on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively) (figure 5.3 and 5.4).

Keeping the pyrrolidine group at the centre of the compounds would maintain the positioning of the benzene ring extension (see SHF-968), however it may reposition the amine linker which extends the newly added benzene group toward the CLR binding region (figure 5.16). Previously developed novel antagonists SHF-1036 and the SHF-638 group, as well as the CGRP antagonist MK-3207 all have the alternate inward facing carbonyl (ketone) groups. As these carbonyl groups are poised near the CLR Trp72, an important residue for peptides associating to the CGRP and AM₁ receptor, it may be important to keep this motility group to block signal transduction rather than increasing the compounds affinity to the receptor (Booe *et al.*, 2015). It's important to note this benzene group is now extending towards the CLR end, and therefore the secondary amine linker and benzene group may be edited during CLR fragment analysis and compound development (figure 5.16).

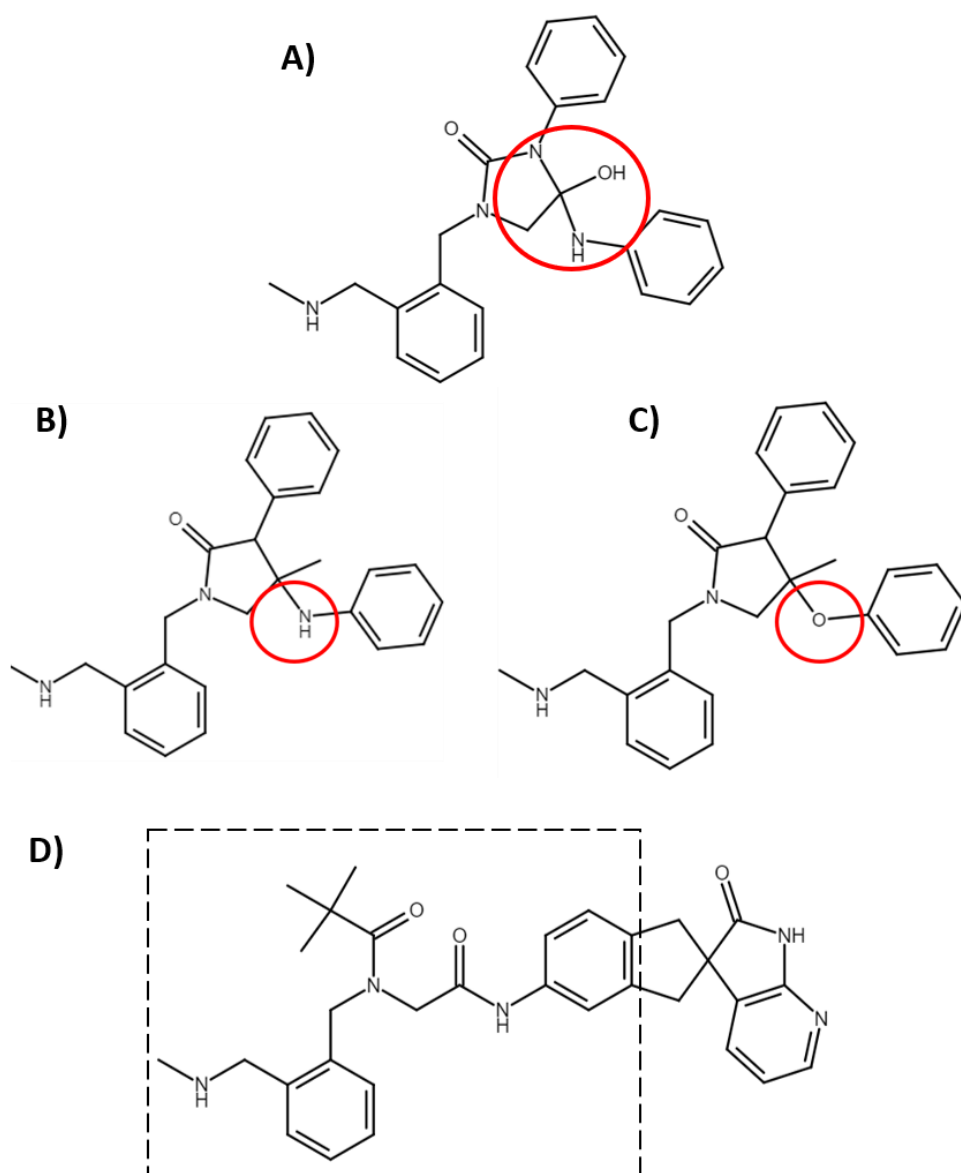


Figure 5.16: Development of the RAMP fragment. (A) Highlights the regions of the compound which is prone to decomposition after combining RAMP fragment SHF-968 and SHF-972. (B and C) Highlighted is the change in linker that may give the fragment a higher selectivity for the RAMP3-CLR fusion protein. (D) The box around compound SHF-638 shows the region currently generating the RAMP fragment and how the added benzene ring is extending towards the CLR end of the compound

From the HTRF displacement data, the addition of the carbamate bond leading to the trimethyl group between SHF-1140 and SHF-1141 shows increased binding affinity for the RAMP3-CLR fusion protein (8.48% binding increase) and minimal difference for the RAMP1-CLR fusion protein (3.77% binding increase). A similar group is found in the developed SHF-1036 antagonist and is the only region which differs from SHF-638, which must subsequently allow increased potency against the CGRP and AM₂ receptors (figure 4.15 and 4.17). This

region and the differences between SHF-1036 and SHF-638 is outlined in figure 5.17. SHF-1140 however, only significantly increases binding affinity for the RAMP3-CLR fusion protein despite displacing the same SHF-1257 antagonist from each fusion protein receptor. This must therefore imply structural differences between SHF-1140 and the equivalent group in SHF-1036 are important for association to the CGRP receptor. See figure 5.17 for the outline in structural differences. The noticeable difference includes the removal of the trimethyl group and the carbamate bond in SHF-1140 which is replaced with a methyl and carbonyl group in SHF-1036. There is also an extra hydrocarbon extension to the methyl group at the start of the chain in SHF-1036, before the pyridyl group is formed (figure 5.17). This extension of the methyl group before the pyridyl ring (figure 5.17) may form contacts with the Trp74 'ceiling' of the hydrophobic pocket of the CGRP receptor, which would give increased affinity to the receptor as seen in previous CGRP antagonists (ter Haar *et al.*, 2010) (Miller *et al.*, 2010)(Qi *et al.*, 2011). The RAMP1 Trp74 changes to a glutamic acid residue in RAMP2 and RAMP3, resulting in the loss of this interaction (Miller *et al.*, 2010). A comparison may be made with the methyl group of ubrogepant and atogepant, which appears to extend towards and form contacts with RAMP1 Trp74 (see chapter 4, section 4.1.4 for more detail). Depending on the required fragment characteristic, the methyl group may be extended to give increased CGRP receptor binding or the methyl group could be shortened and the addition of the carbamate bond and trimethyl group to give selectivity for the AM₂ receptor (figure 5.17). A combination of these may give increased affinity for both receptor proteins. Figure 5.17 outlines these changes between fragments.

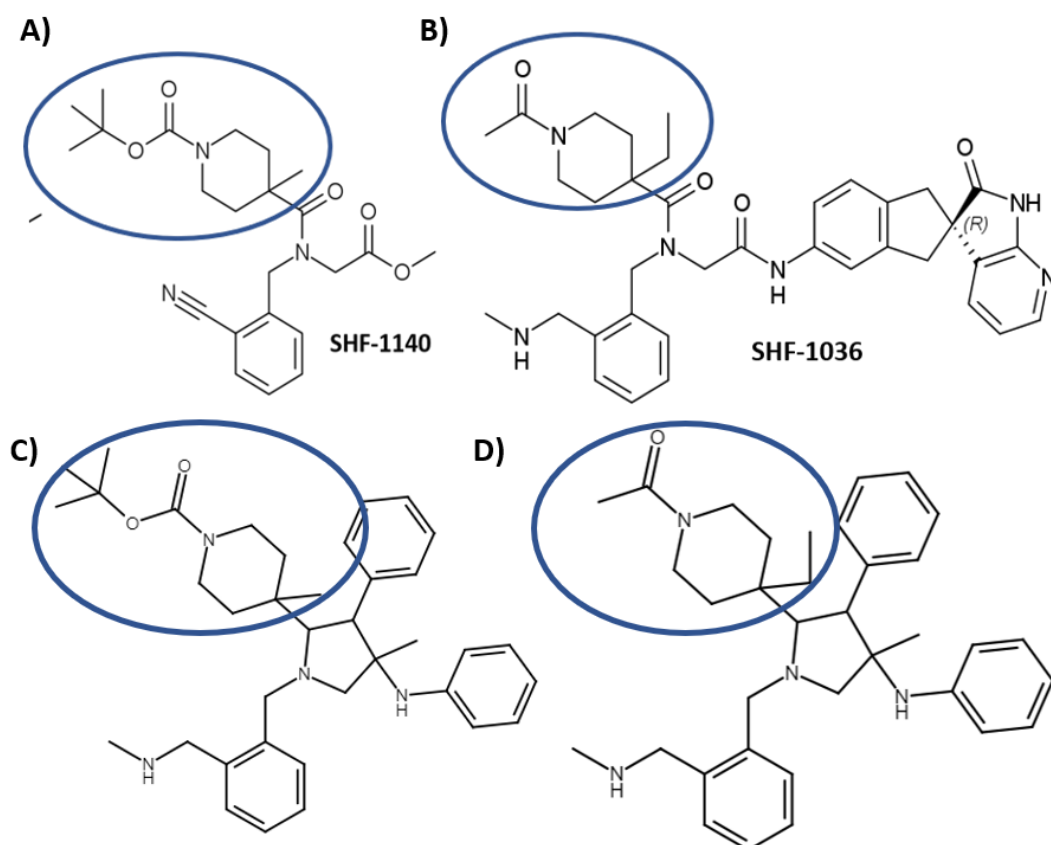


Figure 5.17: Development of the RAMP fragment and possible changes for receptor affinity. (A and B) Outlines structural similarities and differences when comparing the group in question between the SHF-1140 fragment and SHF-1036 antagonist. **C)** The changes in the RAMP fragment that could favour RAMP3-CLR fusion protein selectivity. **D)** Changes in the motility group from (C) that could favour both RAMP1-CLR and RAMP3-CLR fusion protein affinity.

A final addition to the fragment may be to include the methyl group on the pyrrolidine centre. The adaptation of SHF-418 to SHF-209 allowed an additional methyl group in the centre of the structure which showed increased selectivity for the RAMP3-CLR fusion protein receptor by decreasing its affinity for the RAMP1-CLR fusion protein. Although SHF-209 has no hydantoin/pyrrolidine group at its centre, a methyl group could be added in an equivalent position (figure 5.18). This may increase the selectivity for the RAMP3-CLR fusion protein and should not be added to a CGRP and AM₂ selective antagonist. It is unclear at this point as to why the addition of the methyl group is decreasing the affinity to the RAMP1-CLR fusion protein. Depending on binding pocket occupancy, the methyl group positioning could be interfering with interactions formed between the fragments benzene ring and RAMP1 Trp74 (figure 5.18). Similar interactions are observed in the benzene rings of previously developed

CGRP antagonist (olcegepant and telcegepant) and RAMP1 Trp74 which were important for antagonist binding (Hay *et al.*, 2006) (Moore *et al.*, 2010) (Miller *et al.*, 2010) (ter Haar *et al.*, 2010).

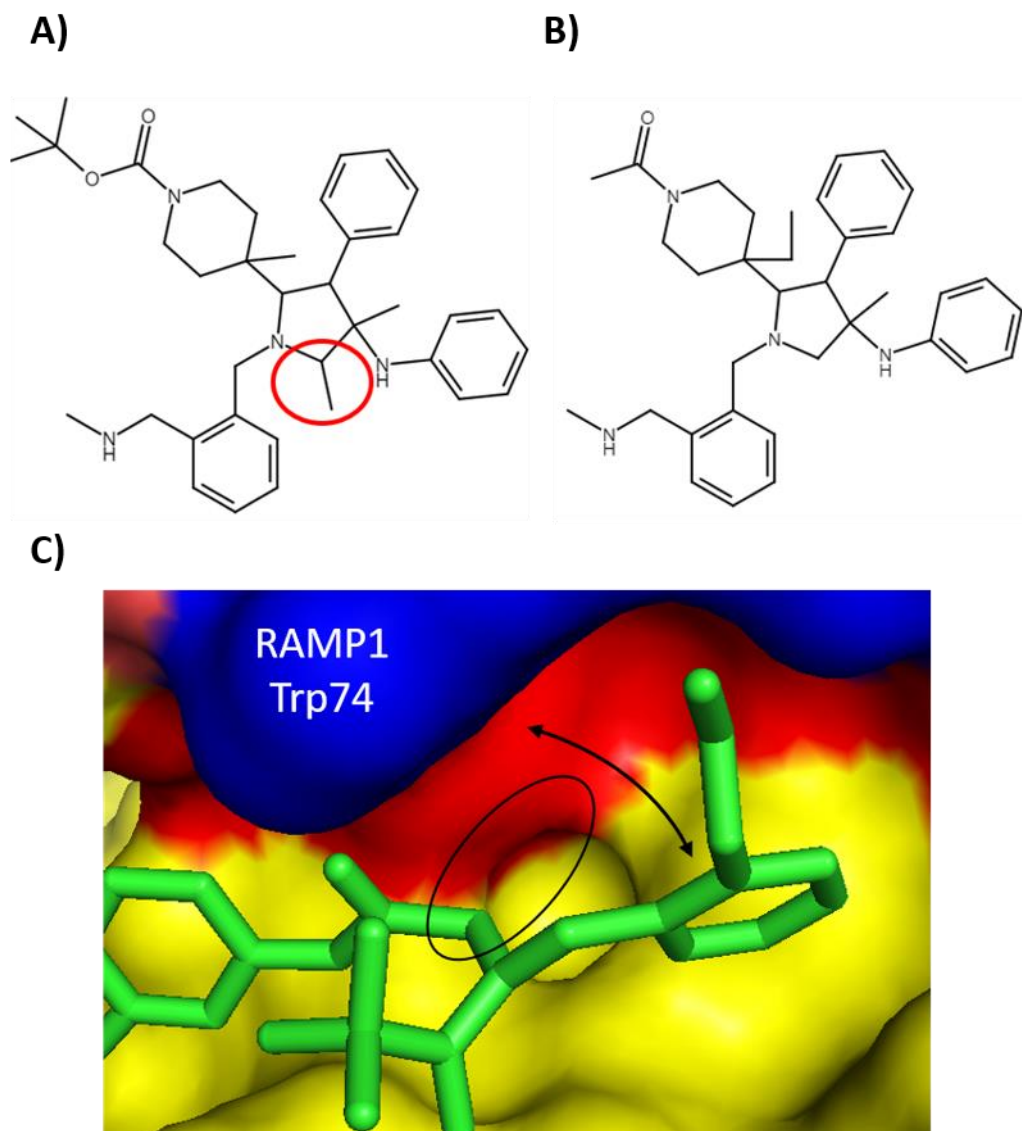


Figure 5.18: The final adaptation of the RAMP fragment. (A and B) The addition of the methyl group to the pyrroline centre (left) which may increase fragment selectivity for the RAMP3-CLR fusion protein (A). This was not included on the CGRP and AM₂ selective antagonist (B) as it could reduce binding to the RAMP1-CLR fusion protein. **(C)** Shows the binding pocket with the RAMP/CLR binding region of SHF-638 (green) in the CGRP receptor. RAMP1 is coloured red and CLR yellow. The RAMP1 Trp74 residue is coloured blue. The black arrow indicates possible hydrophobic interactions between the benzene ring of SHF-638 and the indole group of RAMP1 Trp74. The black ring indicates the positioning of the added methyl group which may interfere with this interaction and give selectivity for the RAMP3-CLR fusion protein.

The suggested fragment should be synthesised and reapplied to the HTRF assay to determine if there is an increased displacement of SHF-1257 from both the RAMP1/3-CLR fusion proteins. They should also be applied to the cAMP assay to determine whether they block peptide signalling through the receptors. Although these fragments are large and could significantly inhibit receptor activation, it may be expected that without the CLR fragment coupling, they would not be as potent as previously generated compounds. The CLR binding region of the compounds occupies peptide binding residues on the CLR receptor (Thr122 and Trp72) which are vital for peptide stimulation of the CGRP receptor (ter Haar *et al.*, 2010) (Booe *et al.*, 2015) (Watkins *et al.*, 2014). Until they have been synthesised and tested however, the possibility of them being developed into compounds in their own right (without the CLR binding region) is unknown.

5.4.2 CLR Fragment Displacement on RAMP-CLR Fusion Proteins

As stated with the RAMP fragment displacement, while we have identified CLR fragments which show lower and higher affinities to the receptor, they are displacing the SHF-1257 antagonist which contains a boidpy-FL tag. Despite this tag being placed towards the compounds interface and away from the binding pocket, it has still affected compound potency to the receptor (see chapter 4, section 4.2.1). While it remains unclear which specific residues and interactions may have been affected, it may have had an effect on how the compound is associating to the CLR protein in the receptor. Fragments which show a higher displacement percentage therefore may not have the same effect on the unlabelled version of the compound.

As with the rationale on designing the RAMP fragments, the previously generated CGRP antagonists (olcegepant, telcagepant and MK-3207) were used to design fragments which could best fit the CLR binding region of the receptor (see figure 5.19). Primarily, CLR Thr122 and Trp72 residues were focused to form interactions with the fragment as their mutagenesis saw a significant decrease in a cAMP response, implying they are important for peptide binding and/or receptor activation (Booe *et al.*, 2015). The majority of fragments consisted of the azabenzimidazoline structure linked to an indane group, as seen with the MK-3207 compound structure. Figure 5.19 outlines the CLR residues and the site occupancy which was being targeted during CLR fragment design.

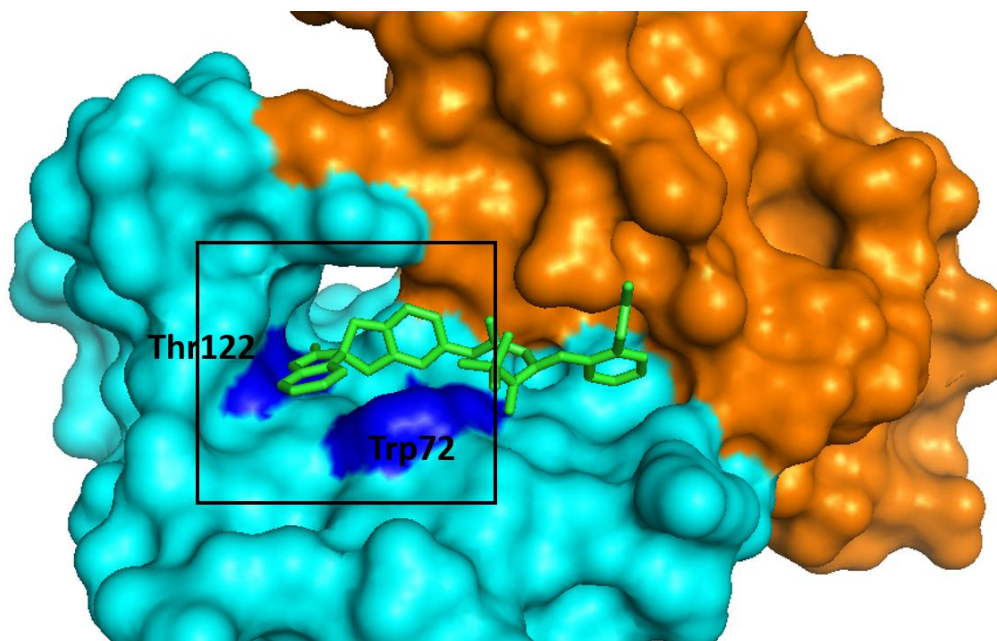


Figure 5.19: Region of CLR fragment targeting during FBDD. The homology model of the RAMP3 (orange) and CLR (cyan)ECD superimposed with SHF-638 (green). Residues which were being targeted by the CLR fragment during drug design are coloured blue and labelled, with the overall region outlined with the black box.

The results of the CLR fragment analysis show interesting responses in terms of small structural changes leading to huge differences on SHF-1257 displacement, as well as compounds showing a much larger increase in RAMP3-CLR selectivity (figures 5.8 and 5.9). While compounds were expected to have larger differences in SHF-1257 displacement due to the likely hood of disrupting a higher number of contacts in comparison to the RAMP fragments, it was expected there would be similar levels of displacement for the same compound on the 2 different fusion proteins. This was due to the CLR protein remaining genetically identical between each fusion protein, and therefore CLR fragments were expected to associate to the receptor in the same manner. Data analysis however showed this was not the case. There were much larger differences in SHF-1257 displacement using the same fragments on the two different fusion protein receptors, with up to almost 50% difference (SHF-695) in displacement between the 2 fusion proteins. This is therefore implying differences in the CLR receptor conformation between the 2 fusion proteins or differences in the way the fragments are interacting with the receptors. It implies that RAMPs could be allosterically modulating the CLR receptors into a conformation where they are able to associate to their appropriate peptides. An example of RAMPs altering the

conformation could be those observed in the CGRP receptor, where the CGRP peptide binding appears to induce a 'clamp like' movement on the CLR loops 3 and 4 and the rotation of RAMP1 Phe83 in comparison to the non-ligand bound version (ter Haar *et al.*, 2010; Booe *et al.*, 2015). This could explain why CGRP peptide association to RAMP1 Trp84 is so important, that the association to this residue is the main factor which induces these conformational changes within the receptor.

In some contrast to this, as the SHF- series of compounds are tailored to bind to RAMP3 in the RAMP binding region (see figure 5.15) and not RAMP1, it could have some subsequent effects on how the compound packs at the CLR binding region. These changes could be RAMP1 Trp74 to RAMP3 Glu74 and RAMP1 Asp71 to RAMP3 Asn71 as they interact with the amine group of the SHF- series of compounds (Kusano *et al.*, 2008). An example model is found with SHF-638 in figure 5.15. This could lead to weaker interactions formed between the SHF-1257 compound and CLR Trp72 and Thr122 in the CLR binding region which are more readily displaced in the RAMP1-CLR fusion protein. Until the crystal images of the CGRP and AM₂ receptor bound with novel SHF- antagonists are resolved however, this remains speculation. In order to properly observe whether RAMPs are allosterically modulating the CLR receptor, displacing a labelled peptide over a labelled novel antagonist may have been more desirable, as seen in previous RAMP-CLR fusion protein functionality studies (Booe *et al.*, 2015) (Rohrkasse *et al.*, 2018).

5.4.3 CLR Fragment development

The fragments tested showed differing results on SHF-1257 displacement, which could be utilised depending on what characteristics may be desired from the fragment. SHF-695 could be utilised to generate a compound which could have increased selectivity for the AM₂ receptor, boasting of 62.147% SHF-1257 displacement on the RAMP3-CLR fusion protein while only maintaining 12.384% displacement on the RAMP1-CLR fusion protein. An alternative starting fragment could be the use of SHF-985, which holds 50.339% displacement on the RAMP3-CLR fusion protein while having a modest, 20.242% displacement of SHF-1257 on the RAMP1-CLR fusion protein. This is one of the highest mean values for displacement on the RAMP1-CLR fusion protein (table 5.5). The structures do not have much variance, with SHF-695 primary amine group being extended from the body of the fragment by an additional hydrocarbon bond and SHF-985 being the right hand (R)

racemic version of its compound. The left hand (S) was not available for testing and therefore the exact importance of this is unknown. Comparison of the fragments against developed compounds (SHF-1036) show that it is predicted the amine group is located primarily near the side chain of CLR Trp72, a key residue for previous CGRP antagonists and peptide binding to the receptor (ter Haar *et al.*, 2010; Booe *et al.*, 2015). It may therefore support the theory of different CLR conformations, with the repositioning of the amine group being required to favour fragment association to the RAMP1-CLR or RAMP3-CLR fusion protein due to the RAMP protein altering the receptor conformation.

The next development of the CLR fragment could be the addition of the nitrogen group in the benzene ring of the indole group (figure 5.20). SHF-1351, SHF-1088 and SHF-952 all showed the repositioning of the nitrogen group in the fragment had a large effect on fragment association to the fusion protein receptors, with SHF-1351 showing overall the highest increase in displacement on both the RAMP1-CLR and RAMP3-CLR fusion proteins (22.371% and 39.713% displacement on the RAMP1-CLR and RAMP3-CLR fusion protein respectively).

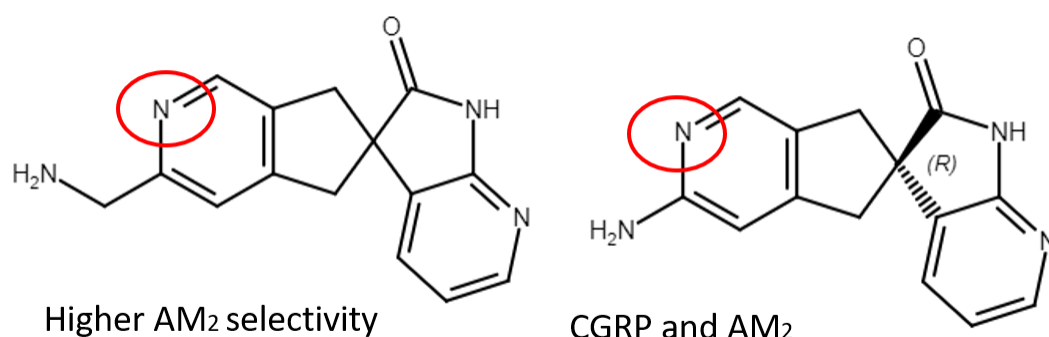


Figure 5.20: Development of the CLR fragment. Differences in 2 fragments that may enable AM₂ selectivity and fragments which may have high affinity over the CGRP and AM₂ receptor. The nitrogen group highlighted (red) was added to both fragments do increase receptor affinity.

An additional fragment which saw an increase in potency to the receptors was SHF-1335. SHF-1335 appears to mimic the CLR end found in early CGRP antagonists, olcegepant and telcegepant, and could be a better fit for the CGRP receptor in comparison to more recently developed antagonists (Avgoustou *et al.*, 2020). The compound extends from the azaindane group with 2 hydrocarbon bonds leading to a phenol group. While this fragment displaces

SHF-1257 by 21.2% and 41.8% on the RAMP1-CLR and RAMP3-CLR fusion proteins respectively, it is structurally different to the fragments which saw the addition of a nitrogen group on the benzene ring (SHF-1351, SHF-1088 and SHF-952). When incorporating this nitrogen group, it could therefore take more testing to determine its optimal position, but comparison and structural overlay of SHF-1335 and SHF-695, shows prime position could be the carbon-6 of the phenol ring in SHF-1335 (figure 5.21). This comparison also shows the hydroxyl group of SHF-1335 and primary amine group of SHF-695 occupy a similar space and therefore presumably contact the same residue(s) in the receptor. The nitrogen from SHF-1351 testing has been incorporated into the SHF-1335 structure to generate a fragment to consider for full compound generation (figure 5.21)

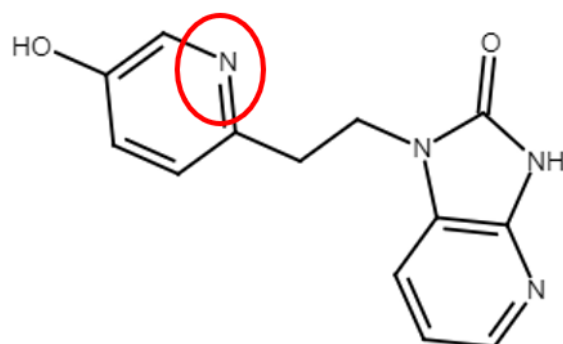


Figure 5.21: Development of the CLR fragment using SHF-1335. An alternative for the CLR fragment which similar resembles early CGRP antagonists (olcegepant and telcagepant). Nitrogen group (red) was included after superimposition on to the SHF-1351 which may aid fragment binding to the receptor.

5.4.4 Combining RAMP and CLR Fragments for Full Length Compounds

Combining the structures from both the CLR and RAMP fragments could generate a full-length compound which is able to antagonise the AM_2 and/or the CGRP receptor. Using the previously generated compounds, SHF-1036 and the SHF-638 group, the fragments were combined to adopt a similar conformation in an attempt to gain the most suitable structure. However, it is important to note that due to the current COVID-19 pandemic and access restrictions, these compounds have yet to be synthesised, and their potencies and selectivity on the CGRP, AM_1 and AM_2 receptors has yet to be assessed.

5.4.4.1 Compounds Selective for the AM₂ receptor

Compounds which may show increased selectivity for the AM₂ receptor over the CGRP and AM₁ receptor have been generated using the data previously laid out in sections 5.41 and 5.43. To generate AM₂ selective compounds, the RAMP/CLR fragments which could hold selectivity for the AM₂ were combined (figure. 5.22) The benzene ring which extended towards the CLR end of the compound in the RAMP fragments (figure 5.16) is presumably the benzene ring found on the indole group of the CLR fragments. If an extra benzene group was added, it would result in a long compound which may not effectively fit in the receptor binding pocket. This was therefore incorporated as the CLR end of the fragment.

Linking the fragments together could have been done in multiple ways. The CLR fragments showed that the extension of its primary amine group from the benzene ring by an extra hydrocarbon chain resulted in a high increase in RAMP3-CLR fusion protein affinity and selectivity (SHF-695, figure 5.8). This is something which should be considered during the compound design. Although limited preliminary data may suggest an extended linker could aid drug binding, the extra extension may cause implications in drug size/length. In order to maintain the correct positioning of this amine group, while maintaining the structure of the pyrroline centre, it creates a longer chain of hydrocarbon bonds which may affect the compound packaging in the binding pocket (figure 5.22). The shortening of this chain therefore may be necessary, but this shortens the amine group extension, which could lead to a loss of potency and selectivity for the AM₂ receptor due to the repositioning of this group (figure 5.22). In comparison to other developed fragments, the positioning of the amine group is near the CLR Trp72 residue which may be contacted by other elements in the compound, such as the nitrogen atom in the benzene ring (CLR end).

While the shortening of the chain would enable the pyrroline group to remain in the centre, it would still mean the hydroxyl group extended from the pyrroline centre could not be included due to compound decomposition issues, as explained (section 5.4.1). However, if the longer linker is maintained, this hydroxyl group could be reincorporated into the structure, as the secondary amine group in the linker is placed further away (5.22). As stated, this group (hydroxyl group) is in a similar position as carbonyl groups in full-length novel antagonists which have been linked with forming key hydrogen bonds to CLR Trp72, and therefore may be important for compound binding (ter Haar *et al.*, 2010).

While the new compound design includes a pyrroline centre, it replaces the ketone groups generated from previously developed compounds. As these groups were previously at the centre of the compounds (SHF-1036, SHF-638 series and MK-3207), they could instigate the positioning of the branched groups. It may therefore be important to keep this centre to ensure the correct location and positioning of these functional groups (figure 5.22). From this, additional compounds have been designed that would hold a similar structural backbone to those found in previous compounds (SHF-1036 and SHF-638), with both the longer and shorter extension of the amine group (figure 5.22). While this may aid functional group positioning, it may remove the bond formed at CLR Trp72 and the ketone group of SHF-1036/638 as the benzene ring replaces it (figure 5.22). Whether the benzene group which replaces it has a similar effect, is yet to be determined. The nitrogen and carbonyl groups located in this region may be interchangeable to give the best binding affinity to the receptor, but until the compound is synthesised and tested, the importance of their positioning remains unknown.



Figure 5.22: Developed compounds for targeted AM_2 selectivity after combing RAMP and CLR fragments. (A) Shows the developed compound SHF-638 for structural comparison. Compounds B) and D) contain the original structure of the antagonist backbone (in SHF-638 and SHF-1036) in an attempt to maintain functional group positioning. Compound C) maintains the pyrroline centre from RAMP fragment testing. B) and C) maintain to longer linkers from the CLR end to the interface while compound D) holds the shorter linker as seen in SHF-1036 and SHF-638.

5.4.4.2 Compounds Selective for the CGRP and AM₂ receptors

Along with the generation of selective AM₂ antagonists, compounds which could have affinity for both the CGRP and AM₂ receptors were generated to have a dual target effect using the same rationale. As described, primary differences included the exclusion of the carbamate bond and trimethyl group on the interface region of the compound, as well as the shortening of the amine group on the CLR end (sections 5.4.1 and 5.4.3). The shorter branching of the amine group (CLR binding region) enabled the easier linking of the two fragments, allowing a linker that resembles those on previously generated SHF antagonists. There is also the extension of the methyl group by an extra hydrocarbon bond prior to the development of the pyridyl ring on the interface region (figure 5.23).

The trimethyl terminal being removed may provide a higher potency on the CGRP receptor, with reduction in selectivity for the AM₂ receptor (figure 5.23). This region of the compound is somewhat comparable with the pyridine rings of olcegepant, which branches to make contacts with CLR Asp94 in its loop 4 (ter Haar *et al.*, 2010). This could mean RAMP3 is altering the conformation and positioning of the CLR loop 4 which makes it more accessible for the trimethyl group to form interactions with CLR in the AM₂ receptor. As seen with the design of the AM₂ selective compounds, there has been an additional compound generated which holds a similar structural centre to that of previously generated AM₂ and CGRP receptor antagonists to correctly position the functional groups (figure 5.23). The designed compounds can be found in figure 5.23.

As previously stated, the shortening of the linker to the CLR end on the newly designed compound would require the removal of the alcohol group, which is extended from the pyrroline centre (RAMP fragment) (figure 5.16). This may disrupt key bonding between the compound and the receptor and therefore a lower potency maybe expected with this compound. If a lower potency is observed, it would inform us of key groups in the compound without the generation of a crystal image. Additional to this, the methyl group extending from the pyrroline group (bottom) has been removed which lowered affinity for the RAMP1-CLR fusion protein.

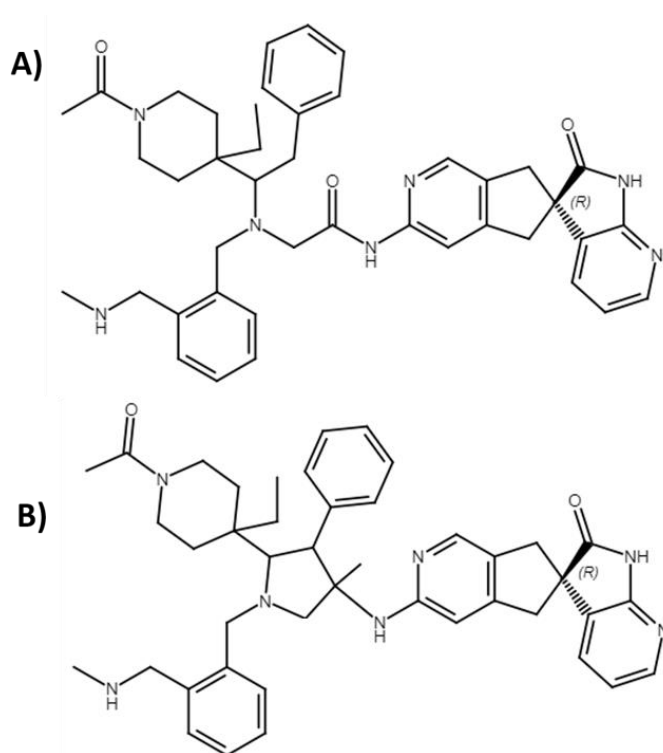


Figure 5.23: Developed compounds for targeted CGRP and AM₂ receptors after combing RAMP and CLR fragments. Compound A) holds the original structural centre as found in previous SHF-1036 and SHF-638 family of antagonists while compound B) maintains the pyrroline centre as seen in the RAMP fragment testing.

5.4.5 Compounds Designed with SHF-1335 Fragment

SHF-1335 was previously edited to include the extra nitrogen group in the benzene ring (CLR binding region) to potentially increase compound potency for the CGRP and AM₂ receptor (figure 5.21). This fragment was then incorporated with the designed RAMP fragments which could hold potency over the CGRP and AM₂ receptor, or be selective for the AM₂ receptor.

The pyrroline centre was maintained in the compound due to the ability to maintain the positioning of functional groups. When attempting to design a compound with SHF-1335 and the previously used ketone groups in SHF-1036 and SHF-638 structures, it would likely disrupt the positioning of 1 or more functional groups in the CLR binding region, interface or RAMP/CLR binding region of the compound. This could therefore affect the potency of the compound to the receptor. This is due to the alcohol group from the benzene ring in SHF-1335 positioned differently to the amine groups found in CLR fragments SHF-695 and

SHF-985. It was therefore difficult to incorporate the fragment into the 2-ketone centre previously used. For this reason, compounds designed with the pyrroline centre only were used to hold structural conformation of the compound.

As seen in the previously designed fragments (figure 5.22), the shortening of the compounds would result in the removal of the hydroxyl group from the pyrroline centre. In an attempt to keep this functional group which has been suggested to be potentially important for hydrogen bonding to the CLR receptor, an amide bond was incorporated into the structure (figure 5.24) (ter Haar *et al.*, 2010). With this, the alcohol on the phenol group of SHF-1335 was replaced with a nitrogen group (figure 5.24). While an ester bond could have been incorporated into the compound, which would have kept the hydroxyl group from the benzene ring and formed an ester bond to the pyrroline ring, esters are more prone to hydrolysis particularly in screening assays and therefore would have affected compound structure. An amine group would lead to a more stable structure. How the change from the hydroxyl group to an amine group on SHF-1335 effects compound potency however is unknown. While the introduction of this amide bond would maintain this carbonyl group (ketone) it would further increase the linker between the RAMP/CLR region and interface from the CLR binding region. It may therefore disrupt proper compound packaging into the receptor and decrease compound potency. Until the compound is synthesised however, how disrupted the association of the compound to the receptor remains unknown. Designed compounds which contain the SHF-1335 CLR fragment and the incorporated amide bond are found in figure 5.24.

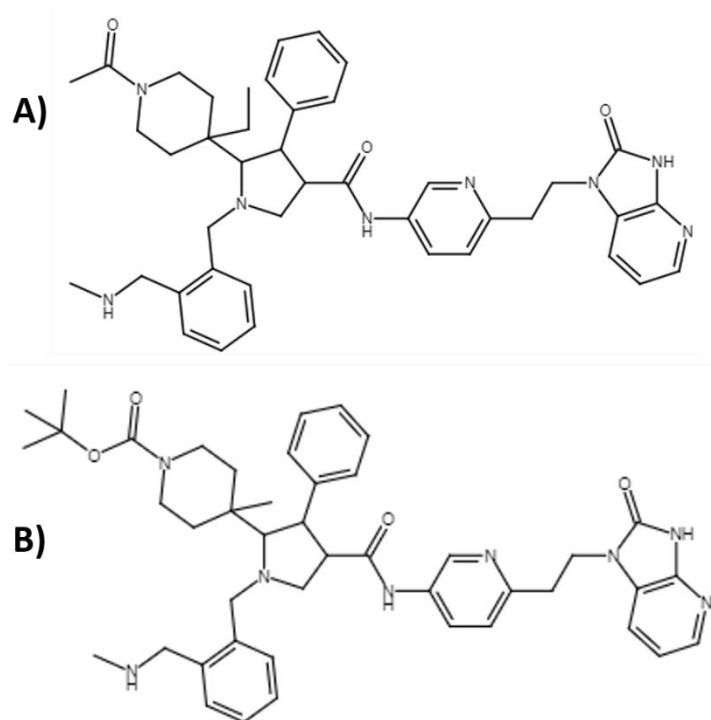


Figure 5.24: Developed compounds for targeted for CGRP or AM₂ selectivity using the CLR fragment SHF-1335. Compound A) has been designed to have selectivity for the CGRP and AM₂ receptor and compound B) has been designed to selectively target the AM₂ receptor.

5.4.6 Effect on cAMP Inhibition

In this study, the fragments were applied to the cAMP assay without linking the two regions together. Ideally upon testing how effective compounds are in inhibiting a cAMP response, the fragments need to be linked together in order to increase overall affinity of the compound to the receptor binding pocket. The linking of the fragments forms the interface of the compound.

Nonetheless, the fragments were assessed on the CGRP, AM₁ and AM₂ receptors competing with their higher affinity ligands (CGRP or AM). Interestingly, there were no major differences observed between the CLR and RAMP fragments on the same receptor, with fragments showing similar displacement values across both the RAMP and CLR regions (figure 5.14). CGRP and AM peptides which occupy the receptor do not hold many contact points with the RAMP region of the receptor, with peptide interactions mainly being found between the CLR receptor in the binding pocket. Fragments which have been designed to target the RAMP region are unlikely to disrupt these key contact points in the binding pocket

but should provide overall increased affinity and selectivity to the receptor for the full-length antagonist. However, there is some inhibition which matched that observed in the CLR fragment inhibition, which is likely to compete more directly with key contact points in the receptor binding pocket. Although RAMP fragment SHF-968 is likely to form contact points at CLR Trp72, it should have a reduced role in receptor inhibition on the CGRP receptor due to CLR Trp72 having a smaller role in peptide binding, in comparison to the AM₁ and AM₂ receptors (Booe *et al.*, 2015). It could suggest that the RAMP fragment is having an indirect role in preventing receptor activation allosterically, or preventing the shift in the RAMP1 protein which is observed in its ligand free and ligand bound crystal models (ter Haar *et al.*, 2010)(Booe *et al.*, 2015). It could imply that the RAMP region of the compound not only increases the overall affinity and receptor selectivity of the compound but also could have an indirect role in structure modulation. This is however speculative, and until the crystal model of the receptor is generated, this remains unknown.

Overall, the fragment data states the importance of the interface region/linker in the antagonist development. Full length antagonists which hold the RAMP, CLR and interface region in both the SHF- series and previously developed antagonists hold reported nM IC₅₀ values against the receptor, with the combined SHF- fragments (SHF-695/968) holding μ M to mM values. This states the linker/interface is important for increasing overall compound affinity to the receptor and could aid the correct positioning of functional groups within the binding pocket. Until the compounds have been synthesised together with the connecting interface/linker, the true affinity of these compounds remains unknown.

Conclusion

To conclude, fragments designed to target the RAMP/CLR region or CLR binding region of the receptors have been tested and high affinity fragments which displace novel SHF-1257 from both the CGRP and AM₂ receptors have been identified. While structural information from the higher affinity RAMP/CLR region and CLR binding region of the fragments has been gathered, manipulated and linked together to generate a full-length compound, their potency on the receptor is yet to be determined. Preliminary cAMP inhibition studies show the importance of the interface and linking of fragments to gain a high affinity compound. CLR fragment data analysis is also supportive of the theory RAMPs alter the confirmation of the CLR receptor and act as an allosteric modulator for receptor activation.

Chapter 6: Crystallography Screening of the RAMP-CLR ECD

6.1 Introduction

Protein crystallography is the process of forming macromolecular protein crystals which allow visualisation and identification of protein structure. The process involves 5 main steps which include obtaining large amounts of the protein of interest, purifying the target protein, crystallisation of the protein, data collection and finally structural determination of the protein crystal (Bijelic, Rompel and Ppo, 2018) (figure 6.1). Once the protein is crystallised, an X-ray light is applied to the protein crystal. X-rays are used due to their ability to produce highly resolute images as their exhibiting wavelengths fall in the range of 0.1-100 Å (Bijelic, Rompel and Ppo, 2018). This allows the visualisation of proteins in atomic detail since these wavelengths fall into interatomic distances e.g a disulfide bond (C-C) is ~ 1.5 Å (Rupp, 2009). As the X-ray light is passed through the protein crystal, the photons interact with the electrons in the protein and are scattered (figure 6.2). The scattered X-ray light hits a detector and with each segment (unit cell) of the crystal retaining highly similar structural motifs, this scattered X-ray diffraction data can be used to generate a 3D electron density model of the protein of interest through computation algorithms and analysis. The analysis output shows the electron density map for an average unit cell of the protein crystal where the residues are automatically added by the computer software (if possible to be determined) or manually. The crystallography process is summarised in figure 6.1.

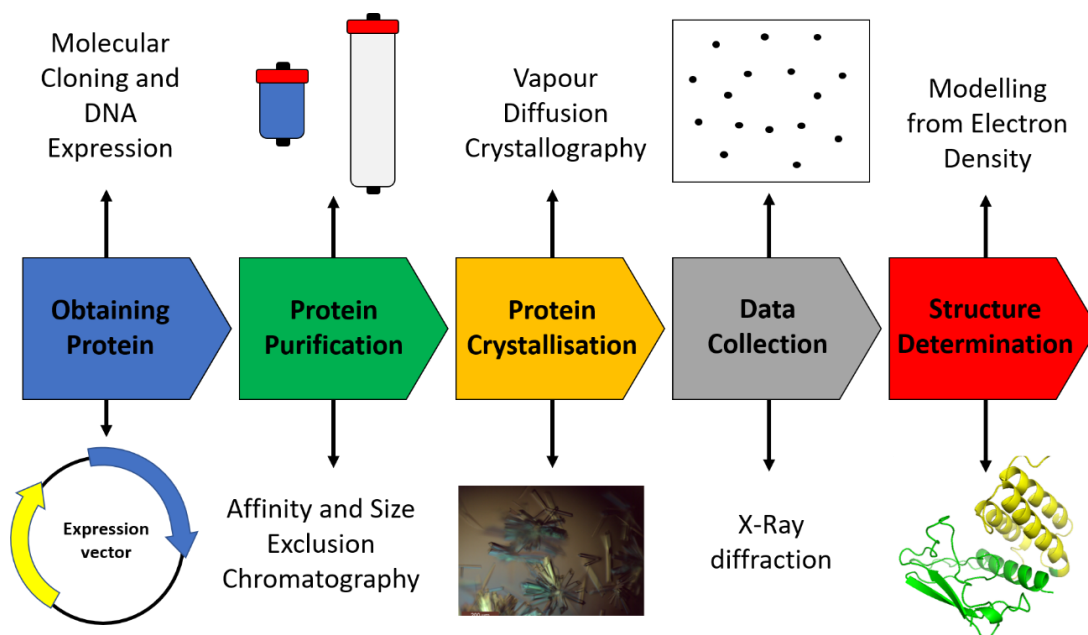


Figure 6.1: Summary of the crystallography process. This include sample preparation (expression and purification) to protein crystallisation and image detection.

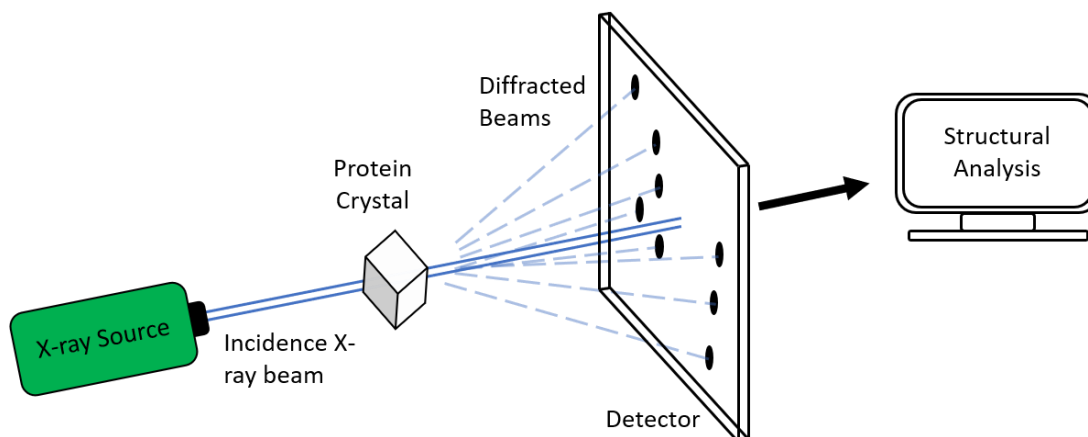


Figure 6.2: X-ray diffraction of a protein crystal. An X-ray hits the protein crystal where it diffracts and hits a detector. The strength of detection depends upon protein alignment. The data generates an electron density map, where it is sent for computation analysis for structure determination.

6.1.1 Formation of Protein Crystals

In order to visualise protein structure through X-ray diffraction crystallography, proteins need to be crystallised. If an X-ray light is applied to a sample containing a high concentration of non-crystallised protein, the X-rays scatter randomly in an unordered manner, leading to no information gathered about protein structure. If however, the protein in the sample is aligned in an ordered array where the proteins are in the same orientation (a crystal) and an X-ray light is applied, the X-ray would scatter in a more ordered fashion (figure. 6.2). Formation of protein crystals increases signal output due to 'fixing' the protein in a specific structure. As the diffraction of X-rays relies exclusively on electrons in the protein crystal and therefore positioning of the bonds in the protein, if these are in a fixed uniformed manner, the signal output at a specific point is much larger (McPherson and Gavira, 2014).

The formation of protein crystals involves bringing the macromolecule to supersaturation, which is defined by the molecule being at a higher concentration than solubility at a given temperature. This can be achieved through a number of setups, which include hinging/sitting drop vapour diffusion, liquid-liquid free interface diffusion and micro dialysis (Thomas and Rice, 1989) (Salemme, 1972) (Bijelic, Rompel and Ppo, 2018). If a protein sample is equilibrated with salts and a precipitation reagent under specific conditions, the water content in the sample will be insufficient to maintain the hydration of proteins

(McPherson and Gavira, 2014). The sample will enter supersaturation, where the protein will start to associate to each other and precipitate out of the solution (Rupp, 2015). This will occur in a uniformed manner if the conditions and procedure is done correctly.

At the region of supersaturation, the process can be divided into two further regions which include the metastable and the nucleation regions (Ducruix and Giege, 1992). These regions are outlined in the phase diagram in figure 6.3. As the conditions become supersaturated, where under specific conditions, protein nucleation occurs as the protein aggregates into a uniformed crystalline structure, which transpires in the nucleation zone of supersaturation. During initial nucleation, multiple clusters of proteins can gather to form spherical aggregates, with free proteins/molecules in the solution continually associating to each protein nucleus (Li *et al.*, 2008). On the other hand, proteins are often dissociating from the nuclei and dissolving at the same time, and is only when the radius of the ordered aggregates exceeds the critical value (solution becomes the optimal level of supersaturation with correct pH and temperature) the gathering of the protein macromolecules exceeds those dissociating from the protein nuclei (McPherson, 1999) (Li *et al.*, 2008). From the nuclei, the crystal is able to grow as more protein is aggregated to its centre. In some contrast to this, the metastable region is a region of a lower degree of supersaturation and is able to develop crystals at a much slower rate but no nucleation of protein occurs (McPherson and Gavira, 2014). It therefore may be favourable, under ideal conditions, for the protein sample to form one nuclei just into the nucleation phase before returning the metastable region. The return into the metastable region would allow slow growth of the crystal which is more likely to be of higher order and avoid defective formation until equilibrium is reached (Asherie, 2004).

If conditions are not perfect, such as if the protein or precipitate solvent concentrations are too high, the protein reaches the precipitation zone (figure 6.3). In this zone, protein sample aggregates in a disordered manner, leading to large insoluble protein aggregates whose structure cannot be determined through X-ray diffraction (Bijelic, Rompel and Ppo, 2018). In contrast to this, if the protein and precipitation concentrations are too low, supersaturation will not be achieved, and crystals will not form in the solution. It is therefore important to use the correct concentration of protein and precipitation solvents for an adequate level of supersaturation (Bijelic, Rompel and Ppo, 2018).

Once the protein has nucleated and crystals have developed, growth will eventually cease as the protein hits its terminal size. The most obvious reasoning for this is the decrease in the protein concentration (solute) to the point where equilibrium has been reached

between the solid and solution phases, but other factors include protein purity and impurities in the crystallisation solution (Li *et al.*, 2008). If a larger crystal is required for imaging (to gain a higher amount of diffraction data) a slight increase in protein concentration is more likely to increase crystal size without reaching the precipitation zone, but may leave the protein in the nucleation region which can give less order in the crystal structures.

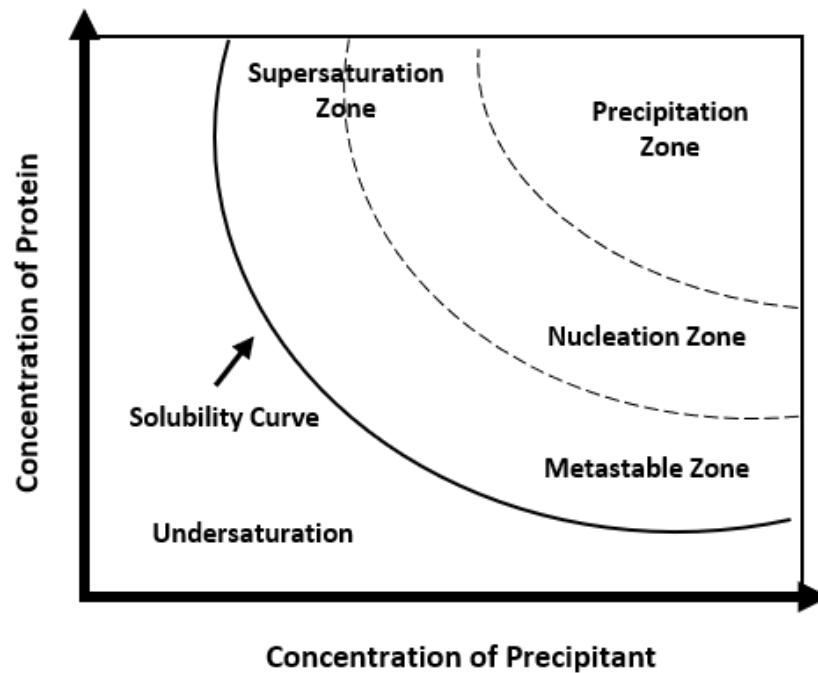


Figure 6.3: Phase diagram of protein supersaturation. Outlined are the different phases which enable the protein (solute) to eluted out of solution in a more ordered structure (nucleation/metastable zone) or as a disordered protein aggregate (precipitation zone)

6.1.2.1 Outcome during Condition Optimisation

As many different conditions are screened during the crystallography process, there are typical outcomes which can be observed. These can include amorphous precipitation which produces a brown matter with no distinct shape or structure when the precipitant and protein concentrations are too high, or on occasions, samples can differentiate into different phases when the solute (protein) is mixed with a high precipitant concentration (Dessau and Modis, 2011). In successful cases, protein crystals may be observed in the weeks or months following the screening setup but these conditions are most likely required to be optimised to generate a protein crystal of high enough quality for X-ray diffraction. Crystals appear in

a variety of different shapes and structures, and appearance is not usually linked to how well crystals diffract. Typical screening outcomes can be seen in figure 6.4 (Dessau and Modis, 2011) (Dessau, Daniel and Hirsch, 2006).

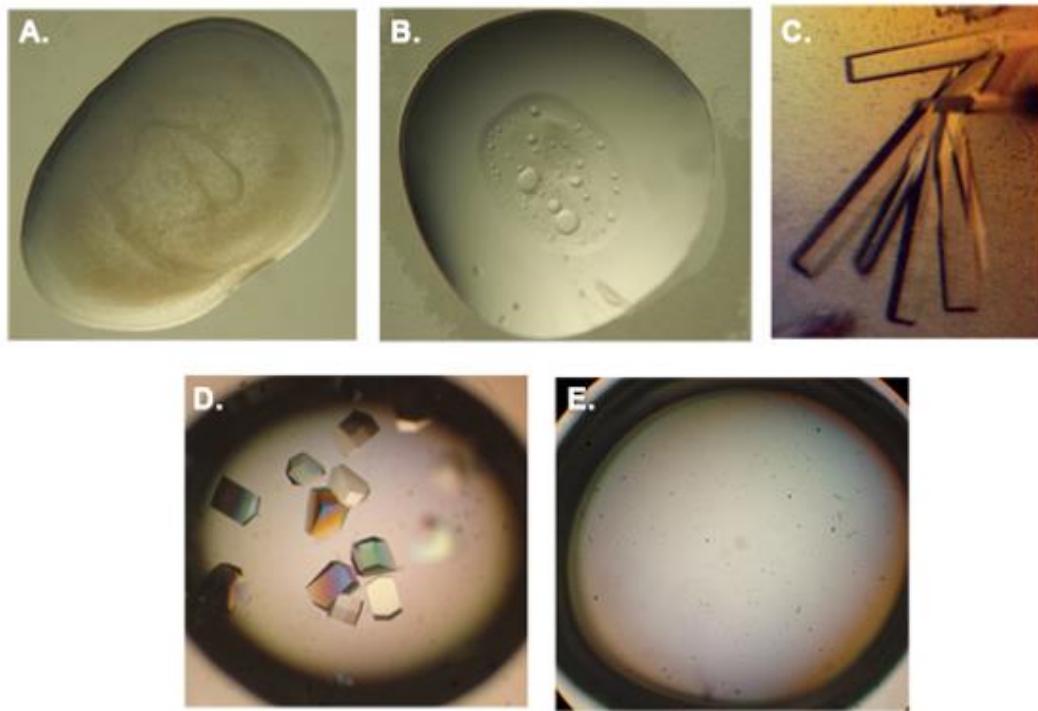


Figure 6.4: Typical outcomes from crystallography experiment. (A) Amorphous protein precipitation when protein or precipitate concentration is too high. Produces a disorientated brown matter. **(B)** Typical phase separation between protein and precipitant usually when the precipitate volume is too high. **(C)** Rod shaped crystals of COP9 signalosome subunit 7 (AtCNS7) (Dessau Daniel and Hirsch, 2006). **(D)** Crystals of lysozyme which are structurally different in comparison to the rod-like crystals. **(E)** An example of an undersaturated drop which usually remains clear during the screening process. Image taken from Dessau and Modis, 2011 with permission from Journal of Visualised Experiments (JoVE)

6.1.2 Imaging the RAMP-CLR ECD structures

6.1.2.1 Multiwavelength Anomalous Dispersion

Early images of the ligand-free and antagonist bound (olcegepant and telcagepant) RAMP1-CLR and RAMP2-CLR receptors were solved using X-ray diffraction crystallography with multiwavelength anomalous dispersion (MAD) (Kusano *et al.*, 2008) (ter Haar *et al.*, 2010)

(Kusano *et al.*, 2012). During image development of an unknown structure, information is required about the wavelength (angle), the amplitude and the phase relation of the wavelengths (Smyth and Martin, 2000). Information from the X-ray wavelength and amplitude can be determined from the experimental setup and spot intensity on the detector, but information cannot be gathered about the phase of the diffracted wavelengths. If wavelengths are in phase, they form constructive phasing where the amplitude is increased but if they are out of phase, they interfere with each other and amplitude is reduced and cannot be refined (figure 6.5) (destructive interference) (Smyth and Martin, 2000). MAD incorporates selenomethionine into the proteins amino-acid structure in place of methionine and allows researchers to take advantage of its anomalous signal at different wavelengths due to the high electron density of metal ions (Barton *et al.*, 2006). This can ultimately resolve the phase issue and refine the electron density map of the protein. This technique however requires complete (or near complete) methionine substitution, which often does not occur, and methionines must be ordered within the macromolecular crystal and not interfere with protein conformation (Barton *et al.*, 2009).

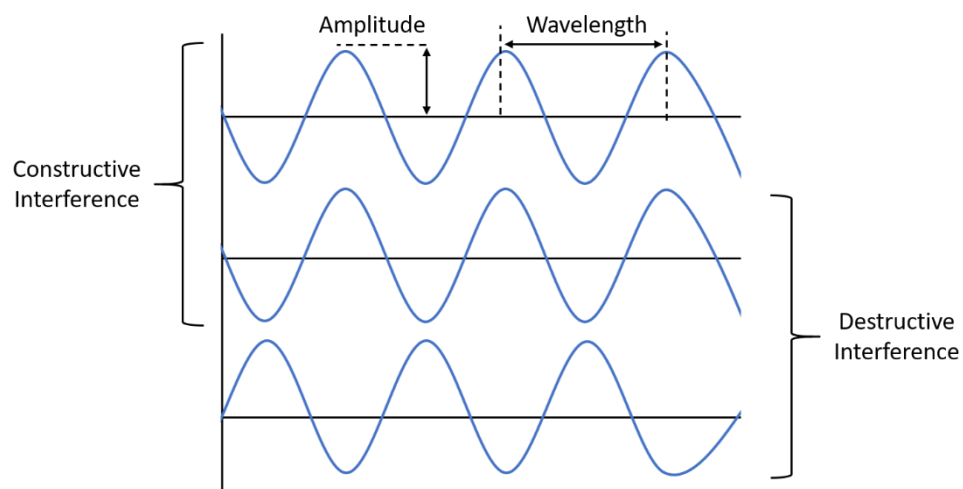


Figure 6.5: In phase and out of phase diffraction waves. The in-phase diffraction waves (constructive interference) and out of phase diffraction waves (destructive interference) are outlined. The amplitude and wavelength unit, which is required for determining electron density is also outlined.

6.1.2.2 Molecular Replacement of Ligand Bound Receptors

The RAMP1-CLR and RAMP2-CLR ligand bound crystals structures generated by Booe *et al.*, 2015 were able to image the protein crystals through molecular replacement rather than MAD. Molecular replacement uses a known molecular model to solve the structure of an unknown molecule as a solution to overcome the phase problem (Evans and McCoy, 2008).

The known model can be used to assume the positioning/approximates of the unknown structure, with the different possible orientations being tested to generate the best predicted diffraction in comparison to the known structure (Evan and McCoy, 2008). Once orientation is optimised, the diffracted phases from the known structure as utilised for the unknown structure and an initial map is calculated (Evan and McCoy, 2008). The technique however heavily relies on a model of high sequence identity ($\geq 25\%$) and structure, although high sequence identity does not always lead to high tertiary structural similarities (Scapin, 2013). In this instance, when the ligand bound RAMP1-CLR and RAMP2-CLR ECD models underwent molecular replacement with the ligand free RAMP1-CLR (PDB: 3N7S) and RAMP2-CLR ECD (PDB: 3AQF) models, they generated good $R_{\text{work}}/R_{\text{free}}$ values of 0.200/0.243 and 0.157/0.200 for the RAMP1-CLR and RAMP2-CLR ECD respectively (Booe *et al.*, 2015). The molecular replacement R values of the ligand bound RAMP1-CLR ECD were higher in comparison to that of the RAMP2-CLR ECD due to the shift in the α -helices of RAMP1 in the ligand bound structure, whereas there was little shifting between the ligand bound/free RAMP2-CLR ECD (ter Haar *et al.*, 2010) (Kusano *et al.*, 2012) (Booe *et al.*, 2015).

6.1.2.3 Cryo-EM

More recently, for the full length CGRP, AM_1 and AM_2 receptors, cryogenic electron microscopy (cryo-EM) has been utilised for the structural imaging of the proteins (Liang *et al.*, 2019) (Liang *et al.*, 2020). Cryo-EM does not rely on the protein to be crystallised before imaging, meaning it can be utilised for very large proteins or those that are proving difficult to crystallise (Murata and Wolf, 2018). The process involves the rapid freezing of protein in ethane chilled by liquid nitrogen, to form a thin layer of vitreous ice (amorphous state) to reduce radiation damage during imaging (Fujiyoshi, 2013) (Murata and Wolf, 2018). This avoids ice crystal formation which can affect macromolecular structure and interfere with imaging during electron microscopy (Thompson *et al.*, 2016). Electrons pass through the vitreous ice containing the protein and scatter, hitting the detector and forming 2D images on electron micrograph (Renaud *et al.*, 2018). As proteins are not fixed in a specific orientation, the 2D images are grouped and averaged to amplify the real signal (Renaud *et al.*, 2018). The relationship between groups is obtained *in silico* to generate a 3D image of the macromolecular molecule (Renaud *et al.*, 2018).

Originally, cryo-EM was limited to macromolecular structures which held a molecular weight of >500 kDa molecular due to poor development of defocused images in smaller macromolecules, but this has recently been vastly decreased to ~65 kDa due to technology advancements (Thompson *et al.*, 2016) (Khoshouei *et al.*, 2017). This may give cryo-EM advantages over X-ray crystallography as it can image proteins in their native states, including all post-translational modifications, which can hinder the crystal development during X-ray diffraction (Bai *et al.*, 2015) (Renaud *et al.*, 2018). This could be directly applicable to the RAMP3-CLR ECD, which requires post-translational modifications for proper protein folding for protein functionality (Flahaut *et al.*, 2003)(Booe *et al.*, 2015) (Roehrkasse *et al.*, 2018). Cryo-EM could therefore be opted for image development if crystals of the RAMP3-CLR ECD prove to be difficult to develop.

6.1.5 Aims and Objectives

The ECDs of the RAMP-CLR proteins have been successfully expressed and purified from CHO-K1 culture media as seen in Chapters 2 and 3. The purified protein will now be concentrated to give high concentrations of fusion protein and will be screened under different conditions, including with and without compounds, in attempts to gain protein crystals. The protein samples will be applied to the sitting drop vapor diffusion method of protein crystallography. Previous published data has all showed the progression and development of the structural representation of RAMP-CLR receptor families but as of yet, no work has shown a well-defined and accurate N-terminus (extracellular domain) structure of the AM₂ receptor. With the use of novel compounds which target the N-terminus of the AM₂ receptor, it is hoped this would resolve this issue.

Hypothesis: The RAMP1/2/3-CLR ECD fusion proteins can be crystallised with and without novel CGRP and AM₂ selective antagonists.

Objective 1: Screen the purified RAMP1/2/3-CLR ECD fusion proteins with different salts, precipitation reagents/concentrations and pH to generate protein crystals with and without novel antagonists.

Objective 2: Screen deglycosylated RAMP1/2/3-CLR ECD fusion proteins under different conditions to generate protein crystals.

6.2 Methods

6.2.1 Protein Concentration and Buffer Exchange

Unless stated the following steps were all carried out on ice or at 4°C. The RAMP-CLR fusion protein samples were expressed and purified as stated in chapter 3. Purified fractions were pooled and concentrated using a Pierce Protein Concentrator PES, 30K MWCO, 5-20 mL (ThermoFisher) by centrifugation at 3700 x *g* until a subsequent volume of ~1 mL was achieved (~15 minutes). The samples were buffer exchanged 3x in 15 mL 50 mM Tris-HCl pH 7.4 and 150 mM NaCl giving a resulting protein volume of ~500 µL after each buffer exchange step. The entire resulting RAMP-CLR fusion protein sample was transferred into a Pierce Protein Concentrator PES, 10K MWCO, 0.5 mL (Thermofisher) and spun at 11,000 x *g* until a final volume of ~50 µL was achieved. The sample was further buffer exchanged 2 times with 500 µL 50 mM Tris-HCl pH 7.4 and 150 mM NaCl and a final sample volume of ~50 µL was achieved. The entire protein sample was collected from the spin column and transferred into a 0.5 mL tube and stored on ice until further use. During RAMP-CLR fusion protein deglycosylation screening, glycans were removed from the protein under native conditions as previously stated in chapter 4, section 4.2.3. Deglycosylated protein samples were subsequently reapplied to the purification process as stated followed by the concentration and buffer exchange steps.

6.2.2 Pre-incubation with CGRP and AM₂ novel antagonist

During the crystal screening process, the RAMP-CLR fusion proteins were incubated either with or without novel CGRP and AM₂ receptor antagonists. Antagonists and ligands can be crystallised with the protein, which can inform researchers of their binding pocket occupancy and therefore aid future SBDD.

Prior to the pre-incubation of the RAMP-CLR fusion protein with the novel receptor antagonists, the concentration of the protein was determined by using the histidine tag ELISA as previously described in chapter 3, section 3.2.7. Novel CGRP and AM₂ antagonist SHF-1036 which can associate to the RAMP1-CLR and RAMP3-CLR fusion proteins (see chapter 4, section 4.3.4 for data) was incubated with the RAMP-CLR fusion proteins, on ice, at a 1:1.13 (protein: SHF-1036) molar ratio. Samples were incubated for 1 hour prior to their

plating for crystallography screening. Protein samples which were not pre-incubated with the ligand were left on ice until they were screened.

6.2.3 Protein Crystallography Screening

In attempts to generate RAMP-CLR fusion protein crystals, the hanging drop vapor diffusion method was utilised. To reduce workload and buffer preparation, pre-made crystallisation kits were used to screen multiple conditions in any one attempt. The kits used during the screening were ProPlex™, JCSG-*plus* HT-96/FX-96 and PACT *premier*™ HT-96/FX-96 (Molecular Dimensions) Crystallography trial kits provide 96 different incubation conditions per kit. Each kit uses different incubation conditions, covering medium to high MW of polyethylene glycol (PEG precipitant) including lower PEG concentrations, a range of pH values from 4-8.5 and effect of different anions and cations in the screening process. 50 µL of each of the different crystallisation solutions was added separately into the reservoir of an MCR 2 Well Crystallisation plate (Hampton Research) (figure 6.6).

The Mosquito Crystal liquid handler with humidity chamber (sptlabtech) was used to aid the crystallography process. A schematic of the system can be found in figure 6.6. The Mosquito system is a fast, accurate and highly precise method of liquid dispensing for crystallography screening purposes which is able to pipette nanoliters of liquid and have 0 cross-contamination. Prior to the addition of the plate and protein samples in the chamber, the humidifier was turned on and the chamber humidity was allowed to reach >75% before starting the dispensing process. The plate containing different crystallisation solutions was placed in the plate holder on the mosquito system and the sample block which holds the protein reservoirs (5 µL max volume) was mounted (figure. 6.6) 4.5 µL of the RAMP-CLR fusion proteins was added to the sample reservoir and humidity chamber closed. The humidity was allowed to reach >75% before the operating system was started. The mosquito system was programmed to dispense 100 nL of protein sample into the droplet well before the addition of 100 nL of the crystallisation solution from the reservoir to each droplet in the plate.

Once dispensing was complete, the plate was sealed using Crystal Clear Sealing Tape (Hampton Research) and was transferred into a 17°C incubator and crystals allowed to form. Wells were check twice a week using a light microscope with any positive wells recorded.

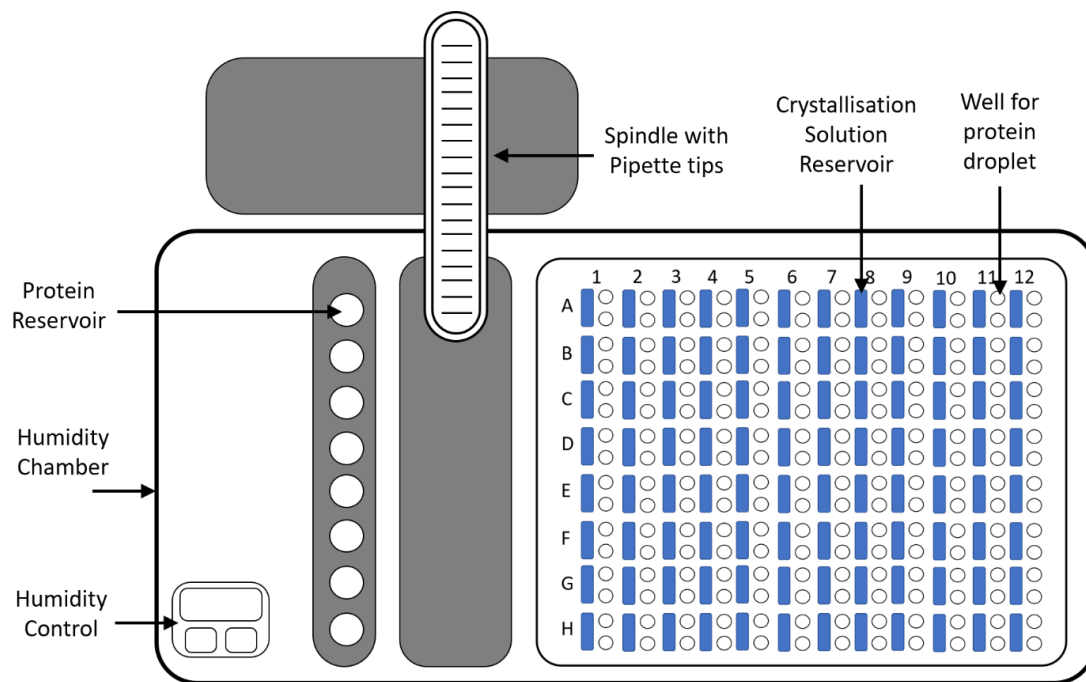


Figure 6.6: Schematic of the Mosquito Liquid Dispensing system. Outlined is the system used for liquid handling to aid protein layout and setup for crystallography screening.

6.3 Results

6.3.1 Crystal Screening

Screening tests were focused on the RAMP1-CLR and RAMP3-CLR fusion proteins with and without the addition of novel compounds. Screening attempts were made on the AM₁ receptor based on previous models and conditions. As the ligand bound and ligand free models had already been generated for the AM₁ receptor and the fact the novel small molecule antagonists hold little affinity for the AM₁ receptor, efforts were primarily focused on the RAMP1-CLR and RAMP3-CLR fusion proteins to determine binding pocket occupancy of the novel compounds. Although the glycosylated RAMP2-CLR fusion protein was trialled using previously established conditions (19% PEG3350, 0.1 M Tris-HCl pH 8.3, 225 nM sodium acetate, 20% ethylene glycol, 30 mg/mL protein concentration), no crystals were observed in these trial runs.

Both the RAMP1-CLR and RAMP3-CLR fusion proteins were screened using protein concentrations ranging from 10-20 mg/mL. No protein crystals or protein precipitate were observed during the RAMP1-CLR fusion protein screens at the range of concentrations given. The RAMP3-CLR fusion protein however showed multiple protein crystal developments. One well in the PACT-premier screen (D11), which equated to 0.2 M calcium chloride dihydrate, 0.1 M Tris pH 8 and 20% w/v PEG6000. This used a protein concentration of 20 mg/mL, the highest protein concentration tested and was supplemented with novel CGRP and AM₂ antagonist, SHF-1036 (figure 6.7). Small crystals appeared approximately 7-8 week after the initial screening conditions were set up and crystals were allowed to be developed for several weeks.

Additional crystals appeared during the deglycosylated RAMP3-CLR fusion protein screening (figure 6.8). A single crystal appeared in the A1 well of the JCSG *plus* HT-96/FX-96 screen, which equated to 0.2 M lithium sulfate, 0.1 M sodium acetate pH 4.5, 50% PEG 400 and 20 mg/mL protein concentration (figure 6.8). This screen was supplemented with novel CGRP and AM₂ antagonist, SHF-1036. Due to the current COVID-19 pandemic and restricted access to university facilities, it is unclear how long it took for the crystal to form in the well.

A final crystal appeared during the RAMP3-CLR fusion protein screening (figure 6.9). A single crystal appeared in the F10 well of the PACT *premier* screen, which equated to 0.02 M sodium/potassium phosphate, 0.1 M BisTris propane pH 6.5, 20% PEG 3350 and 15 mg/mL protein concentration. This screen was not supplemented with any novel CGRP or AM₂

antagonist. Crystals appeared approximately 10 weeks after initial screening was set up and crystals were allowed to develop. All developed crystals and their crystalline solutions are outlined in table. 6.1.

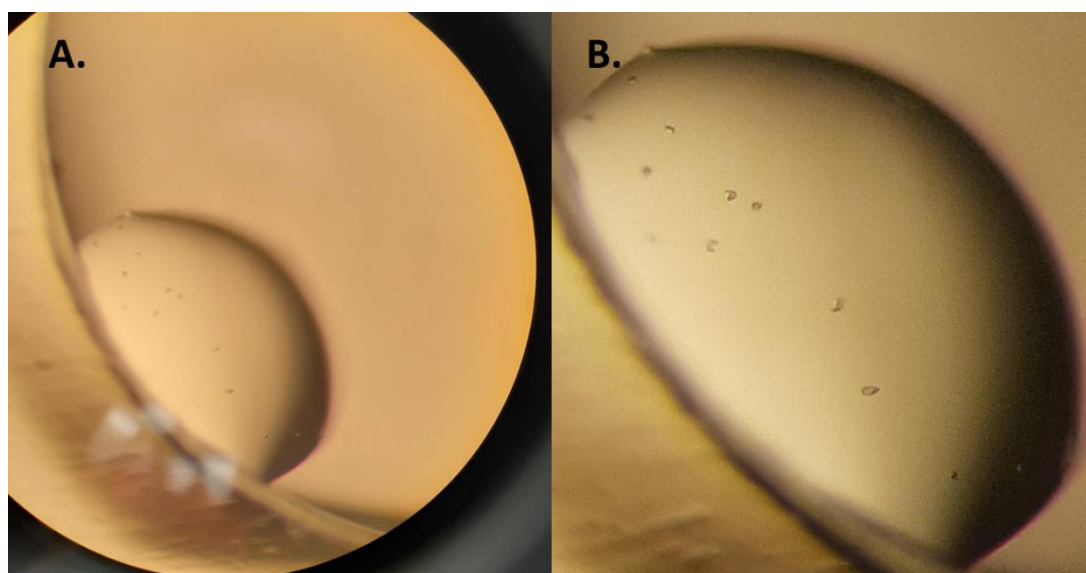


Figure 6.7: RAMP3-CLR fusion protein crystal with SHF-1036 novel antagonist. Image taken of well D11 of the PACT-premier screening kit. Image was taken 2 months after the initial crystal set up. Both images **A)** and **B)** show the same droplet at different magnifications.

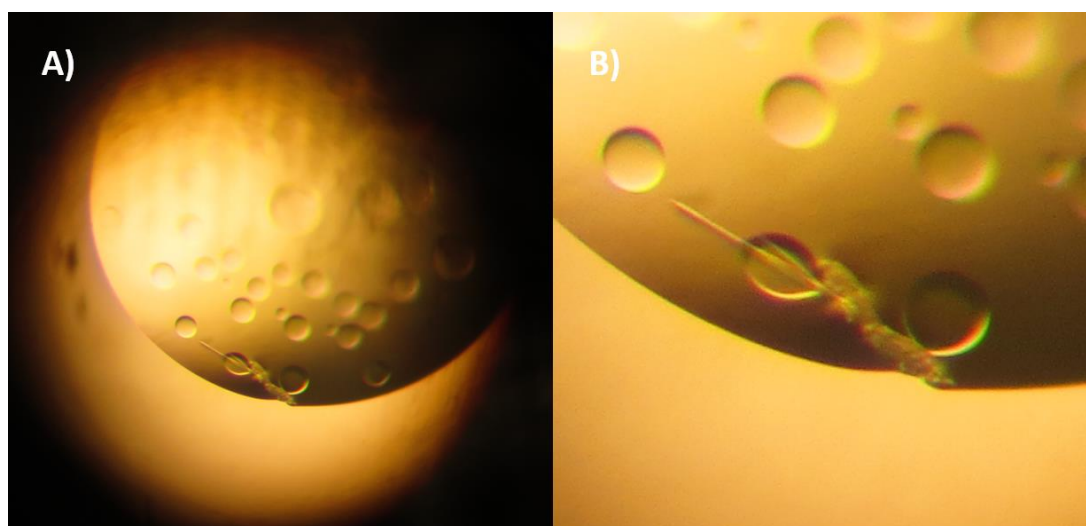


Figure 6.8: Deglycosylated RAMP3-CLR fusion protein crystal with SHF-1036 novel antagonist. Image taken of well A1 in the JCSG plus HT-96/FX-96 screen ~4.5 months after the initial crystal set up. Both images **A)** and **B)** show the same droplet at different magnifications.

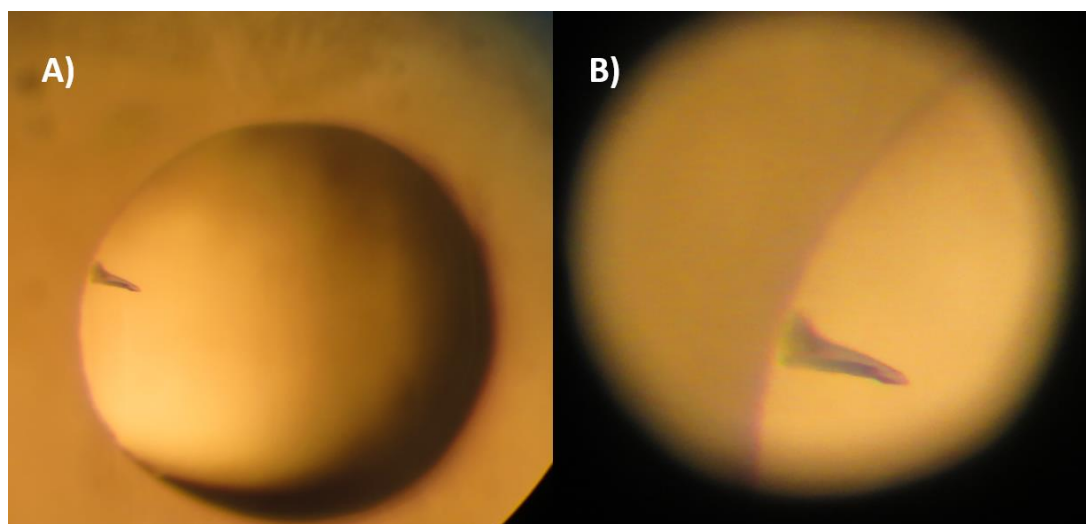


Figure 6.9: Glycosylated RAMP3-CLR fusion protein crystal without addition of novel antagonist. Image taken of well F10 of the PACT-premier HT-96/FX-96 screen. Image was taken ~5 months after the initial crystal set up. Both images **A)** and **B)** show the same droplet at different magnifications.

Table 6.1: Screening conditions which resulted in protein crystals

Fusion Protein State		Screening Kit and well	Crystalline solution in reservoir
Glycosylated RAMP3-CLR fusion protein	ECD	PACT <i>Premier</i> HT-96/FX-96 (Molecular Dimensions)	0.2 M calcium chloride dihydrate, 0.1 M Tris pH 8, 20% w/v PEG6000, 20 mg/mL fusion protein and SHF-1036 antagonist
Deglycosylated RAMP3-CLR fusion protein	ECD	JCSG <i>Plus</i> HT-96/FX-96 (Molecular Dimensions)	0.2 M lithium sulphate, 0.1 M sodium acetate pH 4.5, 50% PEG400, 20 mg/mL fusion protein and SHF-1036 antagonist
Glycosylated RAMP3-CLR fusion protein	ECD	PACT <i>Premier</i> HT-96/FX-96 (Molecular Dimensions)	0.002 M sodium/potassium phosphate, 0.1 M BisTris propane, pH 6.5, 20% PEG3350, 15 mg/mL fusion protein.

6.4 Discussion and Future Work

Here, the purified RAMP3-CLR ECD fusion protein has been crystallised in its glycosylated state with SHF-1036, its deglycosylated state with SHF-1036 and its glycosylated state without SHF-1036. Although it is likely that the crystals need further optimising in order to generate the structural images, once this is complete, the data would outline any conformation differences induced by SHF-1036 binding, its key residues of occupancy, and whether the glycans majorly modify the active conformation of the RAMP3-CLR ECD. It is necessary at this point to mention the current COVID-19 pandemic has hindered the progression of crystal development. Crystals from the initial screening experiments were to be taken for X-ray diffraction analysis to determine if they were if macromolecular content (or salt crystals) but this was not possible due to access and operational restrictions.

6.4.1 RAMP2-CLR Crystal Screening

The RAMP2-CLR fusion protein was screened with the previously published conditions (19% PEG3350, 0.1 M Tris-HCl pH 8.3, 225 nM sodium acetate, 20% ethylene glycol) which generated the ligand bound crystal structure of the RAMP2-CLR ECD (Booe *et al.*, 2015). In this instance however, these conditions were unable to produce RAMP2-CLR fusion protein crystals. This is most likely due to the expression system of choice, with previous crystals generated from bacterially expressed proteins (Booe *et al.*, 2015; Roehrkasse *et al.*, 2018). Bacterially expressed proteins lack glycan addition and therefore would affect the solubility of the protein. Although the RAMP2-CLR fusion protein did not require glycan addition to become a functional protein, it is likely that they affect the protein surface conformation. As it was previously determined that glycans were not necessary for protein functionality (chapter 4, section 4.3.6), future protein crystal screening experiments should include the removal these additional glycans and re applied under the stated conditions.

6.4.2 RAMP1-CLR and RAMP3-CLR Fusion Protein Crystals

The data shows the RAMP3-CLR fusion protein could be crystallised under 3 conditions (table 6.1). Although buffers and salts differ between each crystallised condition, the protein conformation and/or properties within each sample is different. This include a glycosylated

and deglycosylated version of the RAMP3-CLR fusion protein with SHF-1036, which means the hydrophobicity between each protein is likely to have changed. A further crystal generated which does not contain the SHF-1036 antagonists could change the conformation of the RAMP3-CLR fusion protein, as seen with the ligand bound, ligand-free and antagonist bound CGRP receptor (ter Haar *et al.*, 2010)(Booe *et al.*, 2015). This could affect how the protein aligns in the crystal and therefore may require different salts and precipitants for its crystallisation.

The crystals generated will be sent to the national Diamond Light Source synchrotron facility (Oxford) for X-ray diffraction and data collection. This can then be used to assess the binding pocket of the RAMP3-CLR fusion protein and compare it against the previously generated RAMP1-CLR and RAMP2-CLR ECD fusion proteins. As the crystals generated hold the novel CGRP and AM₂ selective antagonists, SHF-1036, binding occupancy of these antagonists can be assessed. The structural image of the binding pocket would also aid future SBDD studies, on the CLR receptor family and other receptors which may interact with the RAMP3 protein.

Although crystals have been developed, it does not guarantee that the crystals are of macromolecular content or the protein of interest. Crystals could be comprised of small impurities in the protein sample or salts from the buffers, that crystallise under certain conditions. If crystals are of macromolecular content and contain the RAMP3-CLR fusion protein bound with SHF-1036, it does not ensure a successful X-ray diffraction pattern and may not give a resolute image. As previously stated in section 6.1.2.3, glycosylated proteins could be difficult to crystallise due to the non-uniformed manner in which glycans are added to the protein which could result in different surface shape and structure between each fusion protein (Baker *et al.*, 1994) (Renaud *et al.*, 2018). It has however been reported that glycan addition to proteins has little difference on a proteins ability to crystallise than previously thought (Mesters and Hilgenfeld, 2007). Nonetheless, if it is found the crystals generated and not macromolecular, and the protein remains difficult to crystallise in later attempts, cryo-EM could be chosen to image the protein due to recent success of its application in full length CGRP, AM₁ and AM₂ receptors. (Liang *et al.*, 2019) (Liang *et al.*, 2020)

It also does not indicate how ordered the crystal structure is, although assessing the time taken in this instance for the crystal to form, it could suggest the protein was held in low metastable phase which leads to a long time required for crystal development. In order to potentially gain better diffraction data, a larger crystal maybe required. The larger the

crystal, in general, a higher amount of protein is encapsulated and therefore would lead to better diffraction output. In this instance the 0.2 M calcium chloride, 0.1 M Tris and 20% w/v PEG produced crystals in the plate. For larger crystal generation, these conditions should be optimised further, which involves slightly changing the parameters which already exist in the reaction.

Initial experiments may increase protein concentration within the sample. Previous RAMP-CLR ECD fusion proteins which have been tagged with the MBP protein and histidine tag were all crystallised after concentrating the proteins to 30 mg/mL (Booe *et al.*, 2015; Roehrkasse *et al.*, 2018). As the construct being used is the soluble N-terminus of the receptor and has a largely hydrophilic MBP tag, it may be required to increase the protein concentration within the sample in order to gain larger crystals at a faster rate. In relation to the RAMP1-CLR fusion protein, it is the probable explanation to there being no crystals observed during the screening process. There was no obvious amorphous protein precipitate or phase separation within the screening attempts indicating the protein concentrations were not high enough during the crystal screens and should be increased in future screening tests. Too high concentration could however have an effect on how ordered the protein is in the structure and therefore should be monitored carefully. It would also be important to note that the previously generated RAMP-CLR fusion protein crystals were all bacterially expressed protein and therefore will not have glycan addition to the receptor, which will influence protein solubility.

In this instance, it therefore maybe more ideal to alter the pH and increase the salt concentration in the sample to generate larger protein crystals in the RAMP3-CLR fusion protein. When examining the conditions used in the screening kits, different pH values were applied when using calcium carbonate and 20% w/v PEG 6000, which ranged from pH 5-8. pH 8 was the only pH value where crystal growth was observed which was the very upper end of the pH values tested. A sensible and viable option for future screening optimisation studies would therefore be to alter the pH around this value, possibly ranging from pH 7-9 to encourage crystal growth. Additional to altering the pH value, the salt concentration could also be altered. In this instance, increasing the calcium chloride concentration could have a significant role in crystal development, by decreasing the time of crystal formation and making the fusion protein less soluble in the buffer water diffused from the droplet.

6.5 Conclusion

To conclude, RAMP3-CLR fusion protein with novel CGRP and AM₂ antagonist, SHF-1036, has been crystallised using a protein concentration of 20 mg/mL and a crystalline solution of 0.2 M calcium chloride, 0.1 M Tris and 20% w/v PEG6000 and are ready for X-ray diffraction testing. If necessary, to potentially increase macromolecular crystal size and improve diffraction data generated the protein concentration should be initially increased before making amendments to the pH of the solution and altering the calcium chloride concentration. Glycans should also be removed from the RAMP1-CLR and RAMP2-CLR fusion proteins and reapplied to the crystal screening process under the previously published conditions in attempts to generate protein crystals.

Chapter 7: General Discussion

7.0 General Discussion

The aim of this study was to express and purify RAMP-CLR ECD fusion proteins in mammalian cells and test their functionality by displacing novel compounds from the receptor. Fragments designed to target the CLR or RAMP region of the receptor could then be applied to gain high affinity hits. This would aid future drug development when targeting RAMP-CLR receptors or other GPCRs which associate with RAMP proteins. This could also determine potential binding pocket occupancy of novel compounds and how compounds occupying different regions of the receptor could lead to a higher binding affinity. The purified protein could then be utilised for crystallography modelling to determine the structure of the receptor extracellular domains.

Previous reports have developed both the ligand free and ligand/compound bound extracellular domain structures of the CGRP and AM₁ receptors, with more recent studies showing the full length receptor complex of the CGRP, AM₁ and AM₂ receptors (ter Haar *et al.*, 2010; Kusano *et al.*, 2012; Booe *et al.*, 2015; Liang *et al.*, 2019; Liang *et al.*, 2020). These structures show specific amino acids residues in the ECD of the receptors enable ligands to selectively bind to their requisite targets (chapter 1, section 1.15) (Kusano *et al.*, 2008, 2012; ter Haar *et al.*, 2010; Booe *et al.*, 2015). The development of these structures has enabled SBDD for novel compounds to selectively target the CGRP and AM₂ receptors (Avgoustou *et al.*, 2020). Despite these advancements there has been no definitive structure of the AM₂ receptor extracellular domain and therefore it has been difficult to determine how novel compounds, which target the receptors N-terminus, occupy the binding pocket (Liang *et al.*, 2019, 2020). Without the structure of the N-terminus of the AM₂ receptor, it is also difficult to predict the binding pocket differences between the CGRP, AM₁ and AM₂ receptors which aid ligand binding and selectivity. This has therefore hindered the development of these compounds. This could have been partly due to previous studies expressing and purifying proteins in bacteria, hindering the structural development.

This study has successfully expressed and purified functional extracellular domains of the CGRP, AM₁ and AM₂ receptors which retain high affinity and selectivity for novel receptor antagonists (chapters 2 and 3). Novel small molecule antagonists were applied against the receptor and their binding pocket occupancies compared against previous CGRP antagonists, with novel compounds showed similarities in binding pocket occupancy to that of telcagepant, particularly at the CLR region of the novel small molecule (chapter 4). Fragments have been tested which are designed to target the RAMP protein within the

receptor, and displacement data outlines fragments which are selective for the RAMP1 or RAMP3 proteins (chapter 5, figure 5.3 and 5.4). These fragments could be applied to future SBDD studies when targeting other GPCRs which associate to RAMP proteins to generate high affinity and selective antagonists. Additional to this, CLR fragments were designed to target the CLR protein within the receptor binding pocket which generated interesting results (chapter 5, figure 5.8 and 5.9). Data indicates RAMPs may modulate the GPCR conformation within the binding pocket which could enable ligand selectivity between the receptors (chapter 5, figure 5.8 and 5.9).

Further to this, the AM₂ ECD could have been crystallised in its glycosylated (with and without novel antagonist) and deglycosylated (with antagonist) form, by applying the protein to the sitting drop vapor diffusion method of protein crystallography (chapter 6, figure 6.7, 6.8 and 6.9). Crystals are ready to be harvested and imaged by X-ray diffraction to determine if they are of macromolecular content and possibly generate a protein image. If the structure of the RAMP3-CLR ECD was generated, it could aid future SBDD when targeting the ECD and other receptors which associate to the RAMP3 protein. It could also determine how the novel compounds are occupying the binding pocket which could aid drug design to generate a higher affinity compound. However, due to the current COVID-19 pandemic and access restrictions, this work has been limited and the crystals have yet to be applied to X-ray diffraction.

7.1 The RAMP-CLR Fusion Protein can be Purified in a 1-Step Approach.

Using the RAMP-CLR extracellular domains tagged to the MBP protein and 10x Histidine tag, a functional and soluble protein was able to be expressed in CHO-K1 cells and purified in large amounts through IMAC. A longer histidine tag with a glycine linker and removal of FCS from the cell culture medium proved to be vital for increased protein purity and yield (chapter 3, sections 3.3.2 and 3.3.3).

In contrast to the previously published data which purified RAMP-CLR fusion proteins, only 1 purification was required during this study (Booe *et al.*, 2015; Roherasse *et al.*, 2018). IMAC utilising the proteins histidine tag in a one-step approach was sufficient to remove non-specific proteins in the sample, leading to samples of >95% purity for all RAMP-CLR fusion proteins (figure 3.12, 3.14 and 3.15). Although the MBP tag could not be incorporated into the purification attempts, it remained tethered to the fusion protein as the tag could have

provided more favourable expression characteristics (discussed in chapter 2, section 2.1.4). Previous studies had to incorporate a dual purification method, using either IMAC and the MBP tag (bacterial expression) or IMAC and SEC (mammalian expression) (Booe *et al.*, 2015; Roehrkasse *et al.*, 2018). Purification differences between our studies and these published reports are presumed to be due to the expression system and culture media used during fusion protein expression. The bacterial expressed RAMP-CLR fusion proteins in early studies were expressed intracellular, resulting in increased background binding during purification due to a higher amount of proteins in the sample (Hill and Pioszak, 2013; Booe *et al.*, 2015). The previous mammalian expressed RAMP-CLR fusion proteins contained FCS in the purification media, which leads to a vast amount of increased background binding, as demonstrated in this study (chapter 3, figure 3.10) (Roehrkasse *et al.*, 2018). These expression and cell culture differences would explain why a dual purification process was required to generate a pure sample in the earlier studies.

A limitation to this study is that we do not know the unit state of the expressed fusion protein (chapter 3, figure 3.16). Although it is clear the protein exists as a single state and is functional, it is not clear to whether this is a monomeric, dimeric or tetrameric subunit (figure 3.16). While comparison of the native page analysis (chapter 3, figure 3.16) to the previously expressed RAMP-CLR receptor ECD in mammalian cells would suggest the fusion protein is in a monomeric state, until SEC is carried out and the protein size calculated from the chromatogram, it remains unclear (Roehrkasse *et al.*, 2018). To my knowledge, there is no previously published data on the expression of the RAMP-CLR fusion protein in CHO-K1 mammalian cells and therefore the unit state of the fusion protein could change between mammalian cell types.

7.2 HTRF Assay Development

After successfully generating a pure protein sample, the RAMP-CLR fusion protein was used in drug binding assays by utilising an HTRF response (chapter 4). HTRF assays are being increasingly utilised in protein-protein or protein compound interactions as an initial screening process to determine hit compounds (Degorce *et al.*, 2009). The assays are advantageous over cell-based receptor interaction assays due to there likely being non-specific binding occurring to other proteins in the cell membranes (Degorce *et al.*, 2009). During HTRF assays, the protein receptor is purified before its addition into the reaction,

allowing direct interaction between the drug and receptor, reducing the non-specific background binding. While this may be advantageous and give the 'true' binding affinity of the compound, it may not be representative of the 'actual' binding constant to the receptor. The assay however offered high sensitivity with reliable and reproducible results during the displacement assays which corresponded closely with the initial cAMP assays carried out during high throughput screening (Avgoustou *et al.*, 2020).

The technology can be limited however with the size of the protein affecting signal transmission. If a very large protein or molecule is applied to the assay, excitation of the acceptor molecule may no longer be achievable. This was observed in this study, with the initial attempts at generating a HTRF response used an anti-MBP antibody and no acceptor signal was achieved (data not shown). It was fortunate that the histidine tag donor antibodies were available which were able to excite SHF-1257 and produce a response. Further limitations include the interactions between 2 molecules has to be a high affinity interaction and the ligand or compound being displaced needs to be tagged, which can affect compound affinity. The concentration range of the tagged drug is therefore limited as binding interference may be observed between the tag on the compound/peptide and the receptor, limiting the accuracy of the results gained. As observed in this study, tagged compounds and fragments which hold an affinity to the receptor (IC₅₀) above >200 nM could not be exploited in the assay due to large amounts of background binding blocking the 'actual' signal in the assay. This resulted in the functionality of the RAMP2-CLR fusion protein not being validated in this study (chapter 4). Further to this, during the fragment screening process, results generated may not have been as accurate as possible if a fragment labelled with the BODPY-FL tag could have been used.

The assay also provides no real insight into the kinetics (association and dissociation constants) of the compounds. Although this wasn't the focus of the experiment, it is important to know the binding kinetics of the compounds during candidate selection of pre-clinical and clinical trials. The protein of interest generated however could be applied to surface plasmon resonance (SPR), which provides association and dissociation constant data. The fusion protein purified in this study could easily be incorporated into an SPR experiment, with commercially available Ni-NTA chips which could capture the fusion protein based on its histidine tag. Previous studies which expressed the RAMP-CLR fusion proteins in bacteria to generate the RAMP2-CLR ECD structure also applied the protein to SPR for ligand association analysis (Kusano *et al.*, 2012). Nonetheless, for the purposes of

this study, the assay showed a functional fusion protein which retained its binding affinities for high and low affinity compounds.

7.3 Displacement of SHF-1257 Shows a Functional Fusion Protein Which Retains Precise Affinity for Novel Compounds

A BODIPY-FL labelled lead compound (SHF-1257, chapter 4, table 4.3) was used to assess the binding affinities of novel compounds to the receptor. Compounds were designed to primarily occupy CLR Trp72 and Thr122 on the CLR receptor with both regions being implemented to be important for peptide binding (Booe *et al.*, 2015). Compounds are then designed to extend further into the binding pocket to interact with the RAMP protein (chapter 5, figure 5.15) and enable selectivity between the receptor subtypes. Further contacts are made with other residues on the CLR protein to increase compound affinity but the overall affinity to the binding pocket is determined by the compound structure and how it fits in the binding pocket. A key example is noted with the SHF-638 enantiomers showing different binding affinities depending upon which racemate is applied to the receptor (see chapter 4 figure 4.22 for compound superimposition into the binding pocket). The compounds have been previously tested and show inhibition of the cAMP response on the full-length CGRP and AM₂ receptors (Avgoustou *et al.*, 2020). The responses (IC₅₀ values generated) were closely translated to the HTRF assay responses showing the fusion protein was functional and retained its binding capabilities for both higher and lower affinity compounds (Avgoustou *et al.*, 2020). This means the protein can be utilised for not only other functional displacement assays using fragments but can also be utilised for crystallography modelling.

While it might have been advantageous to use the receptors native ligand linked with a fluorescent tag as seen in previous studies to determine RAMP-CLR fusion protein functionality, key information would have been missed (Booe *et al.*, 2015; Roehrkasse *et al.*, 2018). The inner most residues which form contacts during ligand binding to the CGRP receptors N-terminus is situated to CLR Trp72, Thr122 and RAMP Trp84 (Booe *et al.*, 2015). CGRP antagonists, including novel and previously developed compounds (olcegepant and telcegepant) extend further into the binding pocket in comparison to the CGRP peptide to form further contacts and increase the compound affinity to the receptor (ter Haar *et al.*, 2010). This means when novel fragments targeting the RAMP region of the protein (see

chapter 5 section 5.3.1) were tested, they are unlikely to be directly competing with the CGRP peptide. It is most likely the RAMP binding region of the compound aids overall compound affinity to the receptor and therefore this region indirectly blocks peptide association (ter Haar *et al.*, 2010; Booe *et al.*, 2015). If a labelled peptide was used to investigate this fragment displacement, the binding affinities may not be observed. Nonetheless, a major limitation could be that although the fragments were binding at higher affinities, high compound/fragment inhibition does not necessarily mean receptor inhibition, which could have been determined if a labelled peptide was used instead of a labelled antagonist (Booe *et al.*, 2015) (Booe *et al.*, 2018).

Expressing the ECD of the RAMP1-3 protein proteins tethered to the CLR ECD, may not be limited to the CLR receptor. In theory, other class B GPCRs which couple to RAMP proteins may be expressed in this format, with the ECDs of the RAMP and GPCR receptor linked with a glycine-serine linker and tagged with an MBP protein and histidine tag. Other published studies show the generation of a functional CTR receptor ECD when coupled to RAMP proteins in a similar manner (Lee, Hay and Pioszak, 2016). A further project in our lab which links RAMP2 to the PTH receptor to form the PTH receptor 1 can produce pure fusion protein samples upon 1 purification attempt using the same conditions.

7.4 RAMP Fragments Show Selectivity Between RAMP-CLR Fusion Protein Receptors.

RAMP targeting fragments were designed to target either the RAMP1 or RAMP3 proteins in attempts to generate fragments that were selective for the RAMP3 protein. As fragments targeting this region do not directly compete with the receptors native peptide, it has previously been difficult using a cAMP assay to determine which fragments hold affinity (if any) to the protein at a applicable concentration. With the developed HTRF assay, this was now possible by displacing a tagged lead compound which holds a similar pocket occupancy and interactions in comparison to the fragments.

While there were significant differences of up to ~15% in selective binding for the RAMP3-CLR fusion protein (SHF-969) (chapter 5, figure 5.3 and 5.4), these differences were not as extreme as those observed between the CLR fragments (chapter 5, figures 5.8 and 5.9). This is most likely due to the fact the SHF- series of compounds hold more key contacts in the CLR region of the receptor, leading to less residue binding competition with the RAMP

fragments. Nonetheless, fragments generated did show preferential binding for either fusion protein.

This data is significant as RAMPs are not solely responsible for only trafficking CLR to the cell surface. RAMP proteins are responsible for trafficking other GPCRs to the cell surface (outlined in chapter 1 table 1.3) and these include the PTH receptor, amylin receptors and the calcium sensing receptor (Poyner *et al.*, 2002)(Christopoulos *et al.*, 2003)(Desai *et al.*, 2014). If these receptors, or others which associate to RAMP proteins become clinically relevant, these fragments could start as a building block to develop antagonists which can selectively target these receptors using FBDD or SBDD. Although the region of the fragment which forms interactions with the CLR protein may need to be edited to bind more appropriately to the associated GPCR protein, it could provide a good starting point to develop novel fragments and antagonists.

7.5 CLR Fragment Displacement Supports Theory of Allosteric Modulation of CLR by RAMPs

Recent studies are supporting the theory that RAMPs are allosterically modulating the CLR receptor (Booe *et al.*, 2018)(Liang *et al.*, 2020). Along with peptide changes which enable binding specifics between the CGRP, AM₁ and AM₂ receptors, a general conception is perceived that RAMPs may induce subtle conformational differences in the CLR receptor which further enables ligand selectivity (Liang *et al.*, 2020). The fragment displacement data generated supports this theory (chapter 5, figures 5.8 and 5.9), with some CLR fragments holding huge differences in affinity between the RAMP1-CLR and RAMP3-CLR fusion proteins. The CLR receptor is genetically identical across all 3 fusion proteins generated, which therefore implies any differences in fragment binding must be due to influences by the RAMP protein associating to the receptor, possibly through additional bond formation or changes upon the CLR conformation.

The differences between fragment binding affinity is most likely due to conformational changes in the region near CLR Thr122 and Trp72, the residues these fragments are designed to target (ter Haar *et al.*, 2010) (see chapter 5, figure 5.19 for SHF-638 antagonist binding pocket superimposition and CLR binding region identification). Although there are no clear differences observed between the CGRP and AM₁ ECD conformations in this region during their superimposition (chapter 1, section 1.12), it could be implying that the RAMP3 protein

could be heavily influencing CLR structure at this position to enable peptide binding. Whether this conformational change is located with CLR Thr122, the trp shelf (CLR Trp72) or another residue is still unclear.

Upon structural comparison of the CLR fragments with the previously developed CGRP receptor antagonists (olcegepant PDB: 3N7S and telcegepant PDB: 3N7R), it would imply a similar binding pocket occupancy between the fragments and compounds. As their structures are very similar, and presuming they occupy the binding pocket in an identical manner, there would be no additional bonding formed between the fragment and the RAMP receptor (ter Haar *et al.*, 2010). Differences between the CLR fragment affinity therefore would be a result of how the fragments bind to the CLR receptor and ultimately if the CLR conformation has been altered by the RAMP protein. Until the crystal images are developed however, this remains speculation. This data is significant as RAMPs are responsible for trafficking multiple GPCRs to the cell surface and altering their ligand pharmacology. RAMPs may therefore alter the conformation of other GPCR proteins which needs to be considered if SBDD or FBDD methods are being utilised for compound development.

A major limitation in this study is that it is unknown how antagonist binding would influence the entire GPCR conformation. As seen with olcegepant binding to the CGRP receptor ECD, small conformational differences are observed between the compound bound, and compound free structures, but it is unclear if olcegepant has any influence over the ECLs and transmembrane domains (ter Haar *et al.*, 2010). With only expressing the ECD in this study, structural differences are also limited to the receptors ECD and the membrane, loops and intracellular segments. With the ECL3 being conformationally changed in the full-length CGRP, AM₁ and AM₂ receptors upon peptide binding, it would be interesting to observe whether antagonist binding changes the conformation of other regions of the receptor (Liang *et al.*, 2020). As novel small molecules have been developed to target the N-terminus only of the CGRP and AM₂ receptors however, it was deemed more suitable to express the extracellular domains of the receptors, rather than the full-length receptor protein (ter Haar *et al.*, 2010)(Doods *et al.*, 2000)(Salvatore *et al.*, 2008)(Salvatore *et al.*, 2010)(Avgoustou *et al.*, 2020). While it would have been beneficial to express the entire RAMP-CLR complex for full structural analysis during crystallography screening, as this project was focused on compound occupation of the binding pocket, it was deemed not to be necessary.

7.6 Testing of Fragment Molecules Leads to the Production of Full-Length Antagonists.

Using the FBDD methodology, fragments were designed to target either the CLR protein in the receptor (CLR binding region) or RAMP/CLR region of the receptor (see chapter 5 figure 5.1 for reference). Multiple fragments were tested and regions of the fragments which appear to aid higher affinity for the RAMP1-CLR and/or RAMP3-CLR fusion protein were identified and combined for development of full-length antagonists.

The generated compounds from combining the fragment data offers structures which closely resemble the previously developed SHF- series of compounds (chapter 5, section 5.4.4). Apart from changing to the pyrroline centre and the extension of the benzene ring, the majority of the functional groups in the CLR binding region, the interface and the RAMP/CLR region are conserved within the structure (chapter 5, figure 5.22 and 5.23). Additional groups which have been added to the compound may provide extra affinity to the receptor binding pocket (see chapter 5, section 5.4). However, the additional groups and changes to the newly designed compounds results in a large molecule (molecular weight: >800 g/mol) and may propose issues when introducing the compound for clinical studies. A closely related example is CGRP antagonist olcegepant (870 Da) which resulted in poor bioavailability and meant intravenous administration was the only feasible application method (Doods *et al.*, 2000). Nonetheless, if compounds were able to antagonise the CGRP and AM₂ receptor at much higher potencies, this data could be utilised to target these regions of the binding pocket but with new compounds at smaller molecular weights. Until these compounds are synthesised and tested however, their potencies for the CGRP, AM₁ and AM₂ receptors remains unclear. Structural comparison to previous antagonists and data presented, suggests they should hold a high affinity for the CGRP and AM₂ receptors (Avgoustou *et al.*, 2020).

A possible limitation to this design method is correctly linking the two fragments together. As it is not specifically known where/how the compounds associate to the receptor binding pocket, if a linker is placed between the fragments, it may affect the association of the small molecules to the receptor by repositioning the functional groups (Kirsch *et al.*, 2019). These include regions specifically around CLR Trp72 and how effectively the compound could pack into the binding pocket (chapter 5, section 5.4.4 discusses in more detail linker addition and compound association). This may affect the compounds overall affinity to the receptor and therefore may not be as potent as expected from initial testing. Despite this, the SHF- series

of compounds have already produced potent full-length antagonists (SHF-638 and SHF-1036) and the linkers observed in these structures can be easily transferred to link the two fragments together (chapter 5, section 5.4.4).

7.7 cAMP Detection in Fragment/Compound-Receptor Inhibition

The TR-FRET cell-based cAMP assay kit was used to measure cAMP levels on cells which have been stimulated with peptides in competition with receptor antagonists. Our lab has used the kit to screen multiple compounds designed to target the CGRP and/or AM₂ receptors which produces high quality reproducible data of high sensitivity (Avgoustou *et al.*, 2020). cAMP assays are widely used in drug discovery with the majority of GPCRs predominantly activating adenylate cyclase which subsequently increases concentrations of cAMP as the secondary messenger (Gas pathway)(Hay *et al.*, 2003). This therefore can be used as a tool to determine the cellular response to peptides and the effectiveness of a compounds ability to antagonise a receptor and block the signalling response.

This study has also outlined the importance of the interface and linker region for overall increased compound affinity. When fragments were tested on receptor inhibition of cAMP, their potencies were much lower in comparison to the full-length antagonists despite holding the overall compound structure (RAMP/CLR binding region, interface and CLR binding region) (chapter 5, figure 5.14). It suggests the linking of the interface region may direct the correct positioning of the terminal functional groups which hold affinity to the receptor. This region therefore may be more important than just forming connections in the binding pocket but how the compound occupies its specific site (Archbold *et al.*, 2011).

However, a possible limitation is that signalling through the CGRP and AM receptors is not limited to only increasing cAMP concentrations. Numerous studies have shown an increase in intracellular Ca²⁺ ion concentration which is a non- cAMP depended pathway and the main characteristic of the G_{αq} pathway (Mellay and Marie, 1998)(Drissi *et al.*, 1999)(Schiess *et al.*, 2005). Measuring only cAMP levels after receptor inhibition therefore may not tell us exactly how effectively the compound is inhibiting the receptor.

The assay also does not provide details about the kinetic aspects (association and dissociation constants) of the compound and its mode of action. The assay informs details of how effectively the compound antagonises the receptor but does not inform the user of how long the compound remains bound to the receptor. This results in the use of a

combination of approaches during lead compound optimisation and development for clinical trials (Doods *et al.*, 2000)(Salvatore *et al.*, 2008)(Salvatore *et al.*, 2010). Nonetheless, for the purposes of this study, it was able to inform us how effectively the developed compounds and fragments were on antagonising CGRP, AM₁ and AM₂ receptor signalling.

7.8 Crystallography Modelling and Future Work

Previous efforts have developed the crystal structures of the full-length CGRP and AM₁ receptors with the structure of the receptors N-terminus being well reported (Kusano *et al.*, 2008)(ter Haar *et al.*, 2010)(Kusano *et al.*, 2012)(Booe *et al.*, 2015)(Liang *et al.*, 2019)(Liang *et al.*, 2020). While there has been the development of the full-length AM₂ receptor structure, those studies were focused on the transmembrane and ECL, with the N-terminus of the receptor remaining relatively unestablished (Liang *et al.*, 2020). While the data presented in this study does not show the crystal model of the AM₂ receptor ECD, it has shown it could be possible to crystallise the receptors ECD as a fusion protein expressed in mammalian cells, in its native state, deglycosylated and with a novel AM₂ antagonist (chapter 6, figures 6.7. 6.8 and 6.9). While it is presumed that Roehrkassee *et al.*, 2018 who expressed the AM₂ receptor ECD as a fusion protein in a very similar manner attempted to crystallise the ECD, there was no reported success or crystal generation. In the same study, they crystallised the RAMP1-CLR ECD bound to IMD from bacteria, suggesting there was no success in crystallography attempts from mammalian cell expressed fusion proteins (Roehrkassee *et al.*, 2018).

It is yet to be determine whether the crystals generated here contain macromolecular molecules, with the small possibility of salts in the crystalline solution, excess SHF-1036 compound or small impurities in the protein sample potentially forming the crystals. If the crystals do contain the protein of interest (RAMP3-CLR fusion protein co-crystallised with SHF-1036), it could suggest in order to successfully crystallise the AM₂ receptor ECD, an antagonist which targets this region is required. To date, we are the only lab group who possess such tools for this crystallisation process (Avgoustou *et al.*, 2020). These crystals are ready for collection and application to X-ray diffraction for structural determination.

As the crystals contain bound SHF-1036, structural determination would give more information than just the extracellular domain structure of the AM₂ receptor. While it would inform us of key residual positioning in the binding pocket, it also informs us of the receptor

site occupancy of the SHF-1036 compound. This is significant as it would identify the residues in the binding pocket that provide affinity for the compounds, which would aid future drug design to target the receptor. If the drug does not associate to the binding pocket as expected in the previous homology models (see chapter 4 for superimposition), the drug can be redesigned to be better tailored to fit the binding pocket. This could involve the addition or removal of functional groups. If the compound was not occupying the binding pocket as expected, RAMP3 allosterically modulating the CLR protein could be one of the most likely reasons for this observation, with data from fragment displacement suggesting RAMPs allosterically modulate the CLR ECD (chapter 5, figures 5.8 and 5.9).

The generated model would determine if RAMP3 allosterically modulates the CLR receptors N-terminus to aid ligand selectivity. As the crystallised protein in this study contains the ECD of the receptor only, any producible image would inform us of binding pocket structure and CLR conformation in relation to the RAMP protein. If the RAMP protein was allosterically modulating the CLR receptor, it could possibly propose a different site for antagonist targeting and binding. Although the CGRP and AM ligands occupy similar sites on the receptors, key residue interactions for receptor activation differ between each peptide (Booe *et al.*, 2015). These include interactions between CGRP and CLR Asp94 in the CGRP receptor and between AM and CLR Phe92 in the AM₁ receptor, which almost diminish ligand binding upon their mutagenesis (alanine substitution) (Booe *et al.*, 2015). The interactions have a vastly reduced effect on peptide stimulation on the opposing receptors (Booe *et al.*, 2015). Figure 7.1 outlines the two specified regions which could be targeted to aid compound selectivity on the CGRP and AM₂ receptors. Areas around these specific residues could be occupied to antagonise the receptor but compound specificity would be unclear until they are designed and tested. The conformation of the CLR receptor in the AM₂ receptor ECD and its key interaction points for peptide binding at this point are unclear, but if changes were seen it could aid antagonist design and occupy different regions of the binding pocket to selectively antagonise the receptor. As the structure of the AM₂ receptor ECD is unknown and there are currently no plans to antagonise these regions of the binding pocket mentioned and this theory remains purely speculative.

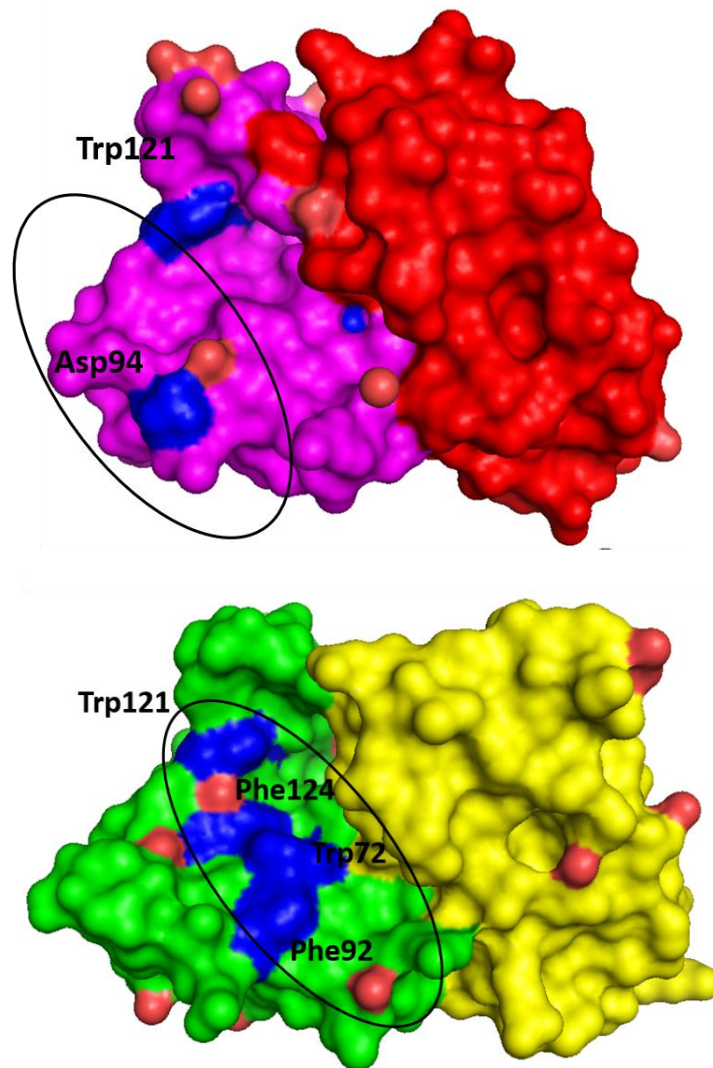


Figure 7.1: Potential sites of antagonist binding on the ligand free CGRP and AM₁ receptors ECD. Potential alternative sites for antagonist binding are outlines (black circle) on the ligand free CGRP ECD (PDB: 3N7R) and the AM₁ receptor ECD (PDB: 3AQF). Outlined are amino acids which are vital for ligand binding to the receptor and drastically reduce (almost diminish) peptide binding to the receptor after mutagenesis (Booe et al., 2015).

The discovery of the RAMP3 ECD structure is more substantial than just developing drugs to selectively target the AM₂ receptor. As stated in chapter 1, table 1.3, RAMPs are responsible for trafficking multiple GPCRs to the cell surface while altering their ligand pharmacology. The discovery of the RAMP3 structure would enable drug design to selectively bind to RAMP proteins in other receptors. Examples could include selectively agonising the AM₁ receptor over the CGRP and AM₂ receptor to treat sepsis or agonising the three AMY receptors (AMY₁, AMY₂ and AMY₃) to treat obesity and type 2 diabetes (Sandry and Drucker, 2013) (Marino et al., 2014) (Hay et al., 2015). If a receptor became clinically applicable due to its

signalling/expression characteristics and RAMP1, RAMP2, or RAMP3 proteins were involved in its surface trafficking and ligand binding, SBDD could easily be incorporated into the drug design procedure to target the RAMP protein within the receptor.

7.9 Conclusion

These data show the expression and purification of a RAMP-CLR ECD fusion protein from mammalian cells, with high sample purity (>95%) being achieved through one purification attempt with IMAC. The fusion protein can be utilised during fragment-based drug design to identify fragments which hold a high affinity for the receptor, which could be applied to any receptor which couples to the RAMP proteins to generate a selective compound. The displacement using CLR fragments is implying that RAMPs are not only responsible for making key contacts with the receptor peptides, but also could be allosterically modulating the CLR conformation which is aiding peptide selectivity between the receptors. This could lead to different sites of antagonist occupancy to block receptor signalling. The subsequent fragment analysis and their ability to inhibit a cAMP response also outlines the importance of the compound interface for increased drug potency to the receptor. In theory, this methodology of protein expression, purification and drug testing could be applied to any GPCR where the compound is targeting the receptor ECD. While the crystals generated have not yet been analysed with X-ray diffraction, data could be suggesting novel AM₂ antagonists were key to crystallise the ectodomain of the AM₂ receptor, which will provide key structural data in future SBDD studies.

References

- Aggarwal, G. *et al.* (2013) 'Adrenomedullin is Up-regulated in Patients With Pancreatic Cancer and Causes Insulin Resistance in β Cells and Mice', *Gastroenterology*. 143(6), pp. 1510–1517. doi: 10.1053/j.gastro.2012.08.044.Adrenomedullin.
- Aiyar, N. *et al.* (1996) 'A cDNA Encoding the Calcitonin Gene-related Peptide Type 1 Receptor', *Journal of Biological Chemistry*. 271(19), pp. 11325–11329.
- Aggarwal, G. *et al.* (2013) 'Adrenomedullin is Up-regulated in Patients With Pancreatic Cancer and Causes Insulin Resistance in β Cells and Mice', *Gastroenterology*. 143(6), pp. 1510–1517. doi: 10.1053/j.gastro.2012.08.044.Adrenomedullin.
- Archbold, J. K. *et al.* (2011) 'Structural insights into RAMP modification of secretin family G protein-coupled receptors: implications for drug development', *Trends in Pharmacological Sciences*. 32(10), pp. 591–600. doi: 10.1016/j.tips.2011.05.007.
- Asherie, N. (2004) 'Protein crystallization and phase diagrams', *Methods*. 34, pp. 266–272. doi: 10.1016/j.ymeth.2004.03.028.
- Avgoustou, P. *et al.* (2020) 'Discovery of a First-in-Class Potent Small Molecule Antagonist against the Adrenomedullin - 2 Receptor', *ACS Pharmacology and Translational Science*. doi: 10.1021/acspsci.0c00032.
- Bai, X., Yan, C., Yang, G. *et al.* (2015). 'An atomic structure of human γ -secretase'. *Nature*. 525, pp. 212–217 <https://doi.org/10.1038/nature14892>
- Baik, J. Y. *et al.* (2006) 'Initial transcriptome and proteome analyses of low culture temperature-induced expression in CHO cells producing erythropoietin', *Biotechnology and Bioengineering*. 93(2), pp. 361–371. doi: 10.1002/bit.20717.
- Baker, H. . *et al.* (1994) 'Enzymatic Deglycosylation as a Tool for Crystallization of Mammalian Binding Proteins', *Acta Crystallographica - Section D Biological Crystallography*. International Union of Crystallography, D50, pp. 380–384. doi: 10.1107/S0907444993013435.
- Barwell, J. *et al.* (2012) 'Calcitonin and calcitonin receptor-like receptors : common themes with family B GPCRs ?', *British Journal of Pharmacology*. 166(1), pp. 51–65. doi: 10.1111/j.1476-5381.2011.01525.x.
- Bell, I. M. *et al.* (2010) 'Discovery of MK-3207: A Highly Potent, Orally Bioavailable CGRP Receptor Antagonist', *ACS Medicinal Chemistry Letters*. 1(1), pp. 24–29. doi: 10.1021/ml900016y.
- Bell, M. R. *et al.* (2013) 'To fuse or not to fuse: What is your purpose ?', *Protein Science*. 22, pp. 1466–1477. doi: 10.1002/pro.2356.
- Berenguer-daiz, C. *et al.* (2013) 'Adrenomedullin Blockade Suppresses Growth of Human Hormone – Independent Prostate Tumor Xenograft in Mice', *Clinical Cancer Research*. 19(22), pp. 6138–6151. doi: 10.1158/1078-0432.CCR-13-0691.
- Besnard, J. *et al.* (2013) 'Automated design of ligands to polypharmacological profiles',

Nature. 492(7428). doi: 10.1038/nature11691.Automated.

Bijelic, A., Rompel, A. and Ppo, A. (2018) 'Polyoxometalates: more than a phasing tool in protein crystallography', *ChemTexts*. 4(3), pp. 1–27. doi: 10.1007/s40828-018-0064-1.

Blom, J. *et al.* (2012) 'Evidence for a functional adrenomedullin signaling pathway in the mouse retina', *Molecular Vision*. 18, pp. 1339–1353.

de Bold, M. K. *et al.* (2012) 'Regulatory Peptides Characterization of a long-acting recombinant human serum albumin-atrial natriuretic factor (ANF) expressed in *Pichia pastoris*', *Regulatory Peptides*. 175(1), pp. 7–10. doi: 10.1016/j.regpep.2012.01.005.

Bokhove, M. *et al.* (2016) 'Easy mammalian expression and crystallography of maltose-binding protein-fused human proteins', *Journal of Structural Biology*. 194(1), pp. 1–7. doi: 10.1016/j.jsb.2016.01.016.

Booe, J. M. *et al.* (2015) 'Structural Basis for Receptor Activity-Modifying Protein-Dependent Selective Peptide Recognition by a G Protein-Coupled Receptor', *Molecular Cell*. 58(6), pp. 1040–1052. doi: 10.1016/j.molcel.2015.04.018.

Booe, J. M. *et al.* (2018) 'Probing the Mechanism of Receptor Activity – Modifying Protein Modulation of GPCR Ligand Selectivity through Rational Design of Potent Adrenomedullin and Calcitonin Gene-Related Peptide Antagonists', *Molecular Pharmacology*. 93(4), pp. 355–367. doi: 10.1124/mol.117.110916.

Bornhorst, J. A. and Falke, J. J. (2000) 'Purification of Proteins Using Polyhistidine Affinity Tags', *Methods in Enzymology*. 326, pp. 245–254.

Carpenter, E. P. *et al.* (2008) 'Overcoming the challenges of membrane protein crystallography', *Current Opinion in Structural Biology*. 18(5), pp. 581–586. doi: 10.1016/j.sbi.2008.07.001.

Ceppa, E. P., Lennerz, J. K. and Ru, V. (2008) 'Calcitonin Receptor-Like Receptor (CLR), Receptor Activity-Modifying Protein 1 (RAMP1) and Calcitonin Gene-Related Peptide (CGRP) Immunoreactivity in the Rat Trigeminovascular System: Differences between Peripheral and Central CGRP Receptor Distri', *The Journal of Comparative Neurology*. 507(3), pp. 1277–1299. doi: 10.1002/cne.

Chandler, K. B. and Costello, C. E. (2017) 'Glycomics and Glycoproteomics of Membrane Proteins and Cell Surface Receptors: Present Trends and Future Opportunities', *Electrophoresis*. 37(11), pp. 1407–1419. doi: 10.1002/elps.201500552.Glycomics.

Chen, X., Zaro, J. and Shen, W.-C. (2014) 'Fusion Protein Linkers: Property, Design and Functionality', *Advanced Drug Delivery Reviews*. 65(10), pp. 1357–1369. doi: 10.1016/j.addr.2012.09.039.Fusion.

Cherezov, V. *et al.* (2008) 'High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor', *Science*. 318(5854), pp. 1258–1265. doi: 10.1126/science.1150577.High.

Chernov, A. A. (2003). 'Protein crystals and their growth'. *Journal of Structural Biology*. 142, 3–21.

Christopoulos, A. *et al.* (2003) 'Novel Receptor Partners and Function of Receptor Activity-modifying Proteins', *The American Society for Biochemistry and Molecular Biology*. 278(5), pp. 3293–3297. doi: 10.1074/jbc.C200629200.

Congreve, M., Dias, J. M. and Marshall, F. H. (2014) 'Progress in Structure-Based Drug Design for G Protein-Coupled Receptors', *Progress in Medicinal Chemistry*. 53, pp. 1-63 doi: 10.1016/B978-0-444-63380-4.00001-9.

Cook, R.M., Brown, A.J.H., Marshall, F.H. and Mason, J.S. (2015) 'Structures of G protein-coupled receptors reveal new opportunities for drug discovery'. *Drug Discovery Today*. 20(11) pp. 1355-1364.

Cowtan, K. (2001) 'Phase Problem in X-ray Crystallography and Its Solution', *Encyclopedia of Life Sciences*. pp. 1–5.

Deceglie, S. *et al.* (2012) 'A modified method for the purification of active large enzymes using the glutathione S-transferase expression system', *Analytical Biochemistry*. 421(2), pp. 805–807. doi: 10.1016/j.ab.2011.12.015.

Degorce, F. *et al.* (2009) 'HTRF : A Technology Tailored for Drug Discovery - A Review of Theoretical Aspects and Recent Applications', *Current Chemical Genomics*. 3, pp. 22–32.

Deng, B. *et al.* (2012) 'Adrenomedullin expression in epithelial ovarian cancers and promotes HO8910 cell migration associated with upregulating integrin $\alpha 5 \beta 1$ and phosphorylating FAK and paxillin', *Journal of Experimental & Clinical Cancer Research*. 31(1), p. 19. doi: 10.1186/1756-9966-31-19.

Desai, A. J. *et al.* (2014) 'Role of Receptor Activity Modifying Protein 1 in Function of the Calcium Sensing Receptor in the Human TT Thyroid Carcinoma Cell Line', *PLoS ONE*, 9(1). doi: 10.1371/journal.pone.0085237.

Dessau, M., Daniel, A. and Hirsch, J. A. (2006) 'crystallization communications Expression , purification and crystallization of a PCI domain from the COP9 signalosome subunit 7 crystallization communications', *Acta crystallographica Section F Structural Biology Communications*, 62(Pt 11), pp. 1138–1140. doi: 10.1107/S1744309106041959.

Dessau, M. A. and Modis, Y. (2011) 'Protein Crystallization for X-ray Crystallography', *Journal of Visualised Experiments*. 9(47), pp. 1–6. doi: 10.3791/2285.

Diener, H. C. (2003) 'RPR100893, a substance-P antagonist, is not effective in the treatment of migraine attacks', *Cephalalgia*. 23(3), pp. 183–185.

Dodick, D. *et al.* (2019) 'Ubrogepant for the Treatment of Migraine', *New England Journal of Medicine*. 381(23), pp. 2230–2241. doi: 10.1056/NEJMoa1813049.

Doods, H. *et al.* (2000) 'Pharmacological profile of BIBN4096BS , the first selective small molecule CGRP antagonist', *British Journal of Pharmacology*. 129, pp. 420–423.

Drissi, H. *et al.* (1999) 'CALCITONIN GENE-RELATED PEPTIDE (CGRP) INCREASES

INTRACELLULAR FREE Ca²⁺ CONCENTRATIONS BUT NOT CYCLIC AMP FORMATION IN CGRP RECEPTOR-POSITIVE OSTEOSARCOMA CELLS (OHS-4)', *Cytokine*. 11(3), pp. 200–207.

Dubowchik, G. M., Conway, C. M. and Xin, A. W. (2020) 'Blocking the CGRP Pathway for Acute and Preventive Treatment of Migraine: The Evolution of Success', *Journal of Medicinal Chemistry*. doi: 10.1021/acs.jmedchem.9b01810.

Ducruix, A. and Giege, R. (1992) Crystallization of nucleic acids and proteins. A practical approach, Oxford University Press, New York.

Dumont, J. *et al.* (2015) 'Human cell lines for biopharmaceutical manufacturing : history, status, and future perspectives', *Critical Reviews in Biotechnology*. 36(6), pp. 1110–1122. doi: 10.3109/07388551.2015.1084266.

Eftekhari, S. *et al.* (2010) 'DIFFERENTIAL DISTRIBUTION OF CALCITONIN GENE-RELATED PEPTIDE AND ITS RECEPTOR COMPONENTS IN THE HUMAN TRIGEMINAL GANGLION', *Neuroscience*. 169(2), pp. 683–696. doi: 10.1016/j.neuroscience.2010.05.016.

Ellies, L. G. *et al.* (1998) 'Core 2 Oligosaccharide Biosynthesis Distinguishes between Selectin Ligands Essential for Leukocyte Homing and Inflammation University of California San Diego', *Immunity*. 9(6), pp. 881–890.

Erlanson, D. A. (2012) 'Introduction to Fragment-Based Drug Discovery', *Topic in Current Chemistry*. 317, pp. 1–32. doi: 10.1007/128.

Errey, J. C. *et al.* (2015) 'Purification of Stabilized GPCRs for Structural and Biophysical Analyses', *G-protein Coupled Receptors in Drug Discovery*. 1335, pp. 1–15. doi: 10.1007/978-1-4939-2914-6.

Evans, P. and McCoy, A. (2008) 'An introduction to molecular replacement', *Acta Crystallographica Section D: Biological Crystallography*. International Union of Crystallography, 64(pt 1), pp. 1–10. doi: 10.1107/S0907444907051554.

Fan, X. *et al.* (2008) 'Randomized controlled trial of an oral CGRP receptor antagonist , MK-0974 , in acute treatment of migraine', *Neurology*. 70(16), pp. 1304–1312.

Flahaut, M. *et al.* (2003) 'N-Glycosylation and Conserved Cysteine Residues in RAMP3 Play a Critical Role for the Functional Expression of CRLR/RAMP3 Adrenomedullin Receptor', *American Chemical Society*. 42(34), pp. 10333–10341.

Flecha, F. L. G. (2017) 'Kinetic stability of membrane proteins'. *Biophysical Reviews*. 9(5) pp. 563–572. doi: 10.1007/s12551-017-0324-0.

Fruhmann, B. *et al.* (1995) 'A human orphan calcitonin like receptor structure', *Biochemical and Biophysical Research Communications*. 206(1), pp. 341–347.

Fox, S. R. *et al.* (2005) 'Active hypothermic growth: a novel means for increasing total interferon- γ production by Chinese-hamster ovary cells', *Biotechnology and Applied Biochemistry*. 41, pp. 265–272. doi: 10.1042/BA20040067.

Fujiyoshi Y. (2013) 'Low Dose Techniques and Cryo-Electron Microscopy: Electron Crystallography of Soluble and Membrane Proteins', *Methods in Molecular Biology (Methods*

and Protocols). 955. Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-62703-176-9_6

Gacasan, S. B., Baker, D. L. and Parrill, A. L. (2018) 'G protein-coupled receptors: the evolution of structural insight', *AIMS Biophysical*. 4(3), pp. 491–527. doi: 10.3934/biophy.2017.3.491.G.

Garayoa, M. *et al.* (2000) 'Expression in Human Tumor Cell Lines during Oxygen Deprivation : A Possible Promotion Mechanism of Carcinogenesis', *Molecular Endocrinology*. 14(6), pp. 848–862.

Gleiter, S. and Bardwell, J. C. A. (2008) 'Disulfide bond isomerization in prokaryotes', *Biochimica et biophysica acta*. 1783(4), pp. 530–534. doi: 10.1016/j.bbamcr.2008.02.009.Disulfide.

Goadsby, P. J. and Edvinsson, L. (1993) 'The Trigeminovascular System and Migraine: Studies Characterizing Cerebrovascular and Neuropeptide Changes Seen in Humans and Cats', *Annals of Neurology*. 33(1), pp. 48–56.

Grisshammer, R. (2009) 'Purification of recombinant G-protein-coupled receptors', *Methods in Enzymology*, 463(1), pp. 631–645. doi: 10.1016/S0076-6879(09)63036-6.Purification.

ter Haar, E. *et al.* (2010) 'Crystal structure of the ectodomain complex of the CGRP receptor, a class-B GPCR, reveals the site of drug antagonism', *Structure*. 18(9), pp. 1083–1093. doi: 10.1016/j.str.2010.05.014.

Hanson, M. A. *et al.* (2008) 'A specific cholesterol binding site is established by the 2.8 Å structure of the human beta2-adrenergic receptor', *Structure*. 16(6), pp. 897–905. doi: 10.1016/j.str.2008.05.001.A.

Hansen, J. M. *et al.* (2010) 'Calcitonin gene-related peptide triggers migraine-like attacks in patients with migraine with aura', *Cephalalgia*. 30(10), pp. 1179–1186. doi: 10.1177/0333102410368444.

Harikumar, K. G. *et al.* (2010) 'THE MOLECULAR BASIS OF ASSOCIATION OF RECEPTOR ACTIVITY-MODIFYING PROTEIN 3 WITH THE FAMILY B G PROTEIN-COUPLED SECRETIN RECEPTOR', *Biochemistry*. 48(49), pp. 11773–11785. doi: 10.1021/bi901326k.THE.

Hay, D. L. *et al.* (2003) 'CL/RAMP2 and CL/RAMP3 produce pharmacologically distinct adrenomedullin receptors: a comparison of effects of adrenomedullin 22-52, CGRP 8-37 and BIBN4096BS', *British Journal of Pharmacology*. 140(3), pp. 477–486. doi: 10.1038/sj.bjp.0705472.

Hay, D. L. *et al.* (2006) 'Determinants of 1-piperidinecarboxamide, N-[2-[[5-amino-1-[[4-(4-pyridinyl)-1-piperazinyl]carbonyl]pentyl]amino]-1-[(3,5-dibromo-4-hydroxyphenyl)methyl]-2-oxoethyl]-4-(1,4-dihydro-2-oxo-3(2H)-quinazoliny)] (BIBN4096BS) affinity for calcitonin gene-related peptide and amylin receptors--the role of receptor activity modifying protein 1', *Molecular Pharmacology*. 70(6), pp. 1984–1991. doi: 10.1124/mol.106.027953.Calcitonin.

Hay, D. L. *et al.* (2015) 'Amylin: Pharmacology , Physiology , and Clinical Potential', *ASPET:*

Pharmacological Reviews, 67(3), pp. 564–600.

Hay, D. L. and Pioszak, A. A. (2016) 'RAMPs (Receptor-Activity Modifying Proteins): New Insights and Roles', *Annual Review of Pharmacology and Toxicology*. 56, pp. 469–487. doi: 10.1146/annurev-pharmtox-010715-103120.RAMPs.

He, Z. X. (2004) 'Experimental Techniques of Biochemistry'. *Chemical Industry*. Beijing.

Hewitt, D. J. *et al.* (2011) 'Randomized controlled trial of the CGRP receptor antagonist MK-3207 in the acute treatment of migraine', *Cephalalgia*. 31(6), pp. 712–722. doi: 10.1177/0333102411398399.

Hill, H. E. and Pioszak, A. A. (2013) 'Bacterial expression and purification of a heterodimeric adrenomedullin receptor extracellular domain complex using DsbC-assisted disulfide shuffling', *Protein Expression and Purification*. 88(1), pp. 107–113. doi: 10.1016/j.pep.2012.11.019.

Ho, T. W. *et al.* (2008) 'Efficacy and tolerability of MK-0974 (telcagepant), a new oral antagonist of calcitonin gene-related peptide receptor, compared with zolmitriptan for acute migraine: a randomised, placebo-controlled, parallel-treatment trial', *The Lancet*. 372(9656), pp. 2115–2123. doi: 10.1016/S0140-6736(08)61626-8.

Ho, T. W., Connor, K. M. and Zhang, Y. (2014) 'Randomized controlled trial of the CGRP receptor antagonist telcagepant for migraine prevention', *American Academy of Neurology*. 83(11), pp. 958–966. doi: 10.1212/WNL.0000000000000771.

Hoare, S. R. (2005) 'Mechanisms of peptide and nonpeptide ligand binding to Class B G-protein-coupled receptors'. *Drug Discovery Today*. 10(6) pp.417-27

Holtmann, M. H., Hadac, E. M. and Miller, L. J. (1995) 'Critical contributions of amino-terminal extracellular domains in agonist binding and activation of secretin and vasoactive intestinal polypeptide receptors. Studies of chimeric receptors', *The Journal of Biological Chemistry*. 270(24), pp. 14394–8.

Howlett, G. J., Kemp, B. E. and Pountourios, P. (1998) 'Crystallization of a trimeric human T cell leukemia virus type 1 gp21 ectodomain fragment as a chimera with maltose-binding protein', *Protein Science*. 7(7) pp. 1612–1619.

Hunter, M. *et al.* (2018) 'Optimization of Protein Expression in Mammalian Cells', *Current Protocols in Protein Science*. 95(1), pp. 1–28. doi: 10.1002/cpps.77.

Ichikawa-Shindo, Y. *et al.* (2008) 'The GPCR modulator protein RAMP2 is essential for angiogenesis and vascular integrity', *Journal of Clinical Investigation*. 118(1), pp. 29–39. doi: 10.1172/JCI33022.

Jazayeri, A. *et al.* (2016) 'Extra-helical binding site of a glucagon receptor antagonist', *Nature*. 533(7602), pp. 274–277. doi: 10.1038/nature17414.

Jin, T. *et al.* (2017) 'Design of an expression system to enhance MBP-mediated crystallization', *Scientific Reports*. 7(40991), pp. 1–11. doi: 10.1038/srep40991.

Joshi, P. *et al.* (2014) 'Identification of potent CNS-penetrant thiazolidinones as novel CGRP

receptor antagonists', *Bioorganic & Medicinal Chemistry Letters*. 24(3), pp. 845–849. doi: 10.1016/j.bmcl.2013.12.089.

Kapust, R. B. and Waugh, D. S. (1999) 'Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused', *Protein Science*. 8(8), pp. 1668–1674.

Keleg, S. *et al.* (2007) 'Adrenomedullin is induced by hypoxia and enhances pancreatic cancer cell invasion', *International Journal of Cancer*. 121(1), pp. 21–32. doi: 10.1002/ijc.22596.

Khoshouei, M. *et al.* (2017) 'Cryo-EM structure of haemoglobin at 3.2 Å determined with the Volta phase plate', *Nature Communications*. Nature Publishing Group, 8(16099), pp. 1–6. doi: 10.1038/ncomms16099.

Kimple, M. E., Brill, A. L. and Pasker, R. L. (2015) 'Overview of Affinity Tags for Protein Purification', *Current Protocols in Protein Science*. 73(1), pp. 608–616. doi: 10.1002/0471140864.ps0909s73.Overview.

Kirsch, P. *et al.* (2019) 'Concepts and Core Principles of Fragment-Based Drug Design', *Molecules*. 24(23), p. 4309.

Kitamura, K. *et al.* (1993) 'Adrenomedullin a novel hypotensive peptide isolated from human pheochromocytoma', *Biochemical and Biophysical Research Communications*. 192(2), pp. 553–560.

Klein, K. R., Matson, B. C. and Caron, K. M. (2016) 'The expanding repertoire of receptor activity modifying protein (RAMP) function', *Critical Reviews in Biochemistry and Molecular Biology*. 51(1), pp. 66–71. doi: 10.3109/10409238.2015.1128875.

Koyama, T. *et al.* (2013) 'Vascular endothelial adrenomedullin-RAMP2 system is essential for vascular integrity and organ homeostasis', *Circulation*. 127(7), pp. 842–853. doi: 10.1161/CIRCULATIONAHA.112.000756.

Kusano, S. *et al.* (2008) 'Crystal structure of the human receptor activity-modifying protein 1 extracellular domain', *Protein Science*. 17, pp. 1907–1914. doi: 10.1110/ps.036012.108.defined.

Kusano, S. *et al.* (2012) 'Structural basis for extracellular interactions between calcitonin receptor-like receptor and receptor activity-modifying protein 2 for adrenomedullin-specific binding', *Protein Science*. 21(2), pp. 199–210. doi: 10.1002/pro.2003.

Kuwasako, K. *et al.* (2001) 'The Seven Amino Acids of Human RAMP2 (86-92) and RAMP3 (59-65) Are Critical for Agonist Binding to Human Adrenomedullin Receptors', *The Journal of Biological Chemistry*. 276(52), pp. 49459–49465. doi: 10.1074/jbc.M108369200.

Kuwasako, K., Kitamura, K., Nagoshi, Y., *et al.* (2003) 'Identification of the Human Receptor Activity-modifying Protein 1 Domains Responsible for Agonist Binding Specificity', *The Journal of Biological Chemistry*. 278(25), pp. 22623–22630. doi: 10.1074/jbc.M302571200.

Kuwasako, K., Kitamura, K., Uemura, T., *et al.* (2003) 'The function of extracellular cysteines in the human adrenomedullin receptor.', *Hypertension research: official journal of the*

Japanese Society of Hypertension. 26, pp. S25-31. doi: 10.1291/hypres.26.S25.

Kuwasako, K. *et al.* (2008) 'Biochemical and Biophysical Research Communications Functions of the extracellular histidine residues of receptor activity-modifying proteins vary within adrenomedullin receptors', *Biochemical and Biophysical Research Communications*. 377(1), pp. 109–113. doi: 10.1016/j.bbrc.2008.09.105.

Lagerstrom, M. C. and Schioth, H. B. (2008) 'Structural diversity of G protein-coupled receptors and significance for drug discovery'. *Nature Reviews Drug Discovery*. 7, pp. 339-357.

Larráyo, I. M. *et al.* (2014) 'Adrenomedullin and tumour microenvironment', *Journal of Translational Medicine*. 12(1), p. 339. doi: 10.1186/s12967-014-0339-2.

Lee, S.-M., Hay, D. L. and Pioszak, A. A. (2016) 'Calcitonin and Amylin Receptor Peptide Interaction Mechanisms: INSIGHTS INTO PEPTIDE-BINDING MODES AND ALLOSTERIC MODULATION OF THE CALCITONIN RECEPTOR BY RECEPTOR ACTIVITY MODIFYING PROTEINS', *Journal of Biological Chemistry*. 291(15), pp. 8686–8700. doi: 10.1074/jbc.M115.713628.

Lenhart, P. *et al.* (2014) 'G-protein-coupled receptor 30 interacts with receptor activity-modifying protein 3 and confers sex-dependent cardioprotection', *Journal of Molecular Endocrinology*. 51(1), pp. 191–202. doi: 10.1530/JME-13-0021.G-protein.

Li, X. X. *et al.* (2008) 'The Factors during Protein Crystallization : A Review', *Crystallography Reports*. 53(7), pp. 1261–1266. doi: 10.1134/S1063774508070286.

Liang, Y. *et al.* (2019) 'Cryo-EM structure of the active, Gs-protein complexed, human CGRP receptor', *Nature*. 561(7724), pp. 492–497. doi: 10.1038/s41586-018-0535-y.Cryo-EM.

Liang, Y. *et al.* (2020) 'Structure and Dynamics of Adrenomedullin Receptors AM1 and AM2 Reveal Key Mechanisms in the Control of Receptor Phenotype by Receptor Activity-Modifying Proteins', *ACS Pharmacology and Translational Science*. 3(2), pp. 263–284. doi: 10.1021/acspsci.9b00080.

Lilius, G. *et al.* (1991) 'Metal affinity precipitation of proteins carrying genetically attached polyhistidine affinity tails', *European Journal of Biochemistry*. 198(2), pp. 499–504.

Lin, C. Y. *et al.* (2015) 'Enhancing protein expression in HEK-293 cells by lowering culture temperature', *PLoS ONE*. 10(4), pp. 1–19. doi: 10.1371/journal.pone.0123562.

Lipton, R. B. *et al.* (2019) 'Effect of Ubrogepant vs Placebo on Pain and the Most Bothersome Associated Symptom in the Acute Treatment of Migraine The ACHIEVE II Randomized Clinical Trial', *Journal of the American Medical Association*. 322(19), pp. 1887–1898. doi: 10.1001/jama.2019.16711.

Lundin, L., Oth, H. B. S. and Breeding, A. (2003) 'The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints', *Molecular Pharmacology*. 63(6), pp. 1256–1272.

Mackie, D. I. *et al.* (2019) 'RAMP3 determines rapid recycling of atypical chemokine receptor-

3 for guided angiogenesis', *PNAS*. 116(48), pp. 24093–24099. doi: 10.1073/pnas.1905561116.

Mallee, J. J. *et al.* (2002) 'Receptor Activity-modifying Protein 1 Determines the Species Selectivity of Non-peptide CGRP Receptor Antagonists', *The Journal of Biological Chemistry*. 277(16), pp. 14294–14298. doi: 10.1074/jbc.M109661200.

Marco, A. De (2009) 'Strategies for Successful Recombinant Expression of Disulfide bond-dependent proteins in *Escherichia coli*', *Microbial Cell Factories*. 18(26). doi: 10.1186/1475-2859-8-26.

Marino, R. *et al.* (2014) 'Plasma adrenomedullin is associated with short-term mortality and vasopressor requirement in patients admitted with sepsis', *Critical Care*. 18(1), pp. 1–7. doi: 10.1186/cc13731.

Maru, Y. *et al.* (1996) 'The Dimerization Property of Glutathione S -Transferase Partially Reactivates Bcr-Abl Lacking the Oligomerization Domain', *Journal of Biological Chemistry*. 271(26), pp. 15353–15357.

McLatchie, L. M. *et al.* (1998) 'RAMPS regulate the transport and ligand specificity of the calcitonin- receptor-like receptor', *Nature*. 393(6683), pp. 333–339. doi: 10.1038/30666.

Mcpherson, A. (1985) 'Crysallisation of Macromolecules: Gernerall Principles', *Methods in Enzymology*. 114, pp. 112–120.

McPherson, A. (1999). 'Crystallization of Biological Macromolecules'. *Cold Spring Harbor*. New York.

McPherson, A. and Gavira, J. A. (2014) 'Introduction to protein crystallization', *Acta Crystallographica Section F:Structural Biology Communications*. International Union of Crystallography, 70(1), pp. 2–20. doi: 10.1107/S2053230X13033141.

Mellay, L. and Marie, P. J. (1998) 'Activation of Phospholipase C- β 1 via $G\alpha_{q/11}$ during Calcium Mobilization by Calcitonin Gene-related Peptide ', *Journal of Biological Chemistry*. 273(32), pp. 20168–20174.

Mesters, J. R. and Hilgenfeld, R. (2007) 'Protein Glycosylation, Sweet to Crystal Growth?', *American Chemical Society*. 7(11), pp. 18–20.

Milic, D. and Veprintsev, D. B. (2015) 'Large-scale production and protein engineering of G protein-coupled receptors for structural studies', *Frontiers in Pharmacology*. 6(66), pp. 1–24. doi: 10.3389/fphar.2015.00066.

Miller, P. S. *et al.* (2010) 'Biochemical and Biophysical Research Communications Non-peptidic antagonists of the CGRP receptor, BIBN4096BS and MK-0974, interact with the calcitonin receptor-like receptor via methionine-42 and RAMP1 via tryptophan-74', *Biochemical and Biophysical Research Communications*. 391(1), pp. 437–442. doi: 10.1016/j.bbrc.2009.11.076.

Moad, H. E. and Pioszak, A. A. (2013) 'Selective CGRP and adrenomedullin peptide binding by tethered RAMP-calcitonin receptor-like receptor extracellular domain fusion proteins',

Protein Science. 22(12), pp. 1775-1785 doi: 10.1002/pro.2377.

Mondal, M. *et al.* (2016) 'Drug Design Fragment Linking and Optimization of Inhibitors of the Aspartic Protease Endothiapepsin: Fragment-Based Drug Design Facilitated by Dynamic Combinatorial Chemistry', *Angewandte Chemie*. 55(32), pp. 9422–9426. doi: 10.1002/anie.201603074.

Moore, E. L. *et al.* (2009) 'Examining the binding properties of MK-0974: A CGRP receptor antagonist for the acute treatment of migraine', *European Journal of Pharmacology*. 602(2–3), pp. 250–254. doi: 10.1016/j.ejphar.2008.11.050.

Moore, E. *et al.* (2020) 'Characterization of Ubrogепant : A Potent and Selective Antagonist of the Human Calcitonin Gene – Related Peptide Receptors', *The Journal of Pharmacology and Experimental Therapeutics*, 373, pp. 160–166.

Moore, E. E. *et al.* (1995) 'Functionally Different Isoforms of the Human Calcitonin Receptor Result from Alternative Splicing of the Gene Transcript', *Molecular Endocrinology*. 9(8), pp. 959–968.

Moore, E. L. *et al.* (2010a) 'Mapping the CGRP receptor ligand binding domain: Tryptophan-84 of RAMP1 is critical for agonist and antagonist binding', *Biochemical and Biophysical Research Communications*. 394(1), pp. 141–145. doi: 10.1016/j.bbrc.2010.02.131.

Muhammed, T. and Aki-Yalcin, E. (2019) 'Homology modeling in drug discovery: Overview, current applications, and future perspectives', *Chemical Biology & Drug Design*. 93(1), pp. 12–20. doi: 10.1111/cbdd.13388.

Nanazashvili, M. *et al.* (2018) 'LRET Determination of Molecular Distances during pH Gating of the Mammalian Inward Rectifier Kir1.1b', *Biophysical Journal*, 114. pp. 88–97. doi: 10.1016/j.bpj.2017.10.044.

Oak, A., Jansen, G. and Chan, C. (2019) 'Expression and Purification of a Mammalian Protein : Cytosolic Domain of IRE1 α from Insect Sf21 Cells', *Cold Spring Harbour Laboratory*. pp. 1–20. doi: <https://doi.org/10.1101/750430>

Oehler, M. K. *et al.* (2001) 'Adrenomedullin inhibits hypoxic cell death by upregulation of Bcl-2 in endometrial cancer cells: A possible promotion mechanism for tumour growth', *Oncogene*. 20(23), pp. 2937–2945. doi: 10.1038/sj.onc.1204422.

Oehler, M. K. *et al.* (2003) 'Tissue and plasma expression of the angiogenic peptide adrenomedullin in breast cancer', *British Journal of Cancer*. 89(10), pp. 1927–1933. doi: 10.1038/sj.bjc.6601397.

Okada, T. *et al.* (2004) 'The Retinal Conformation and its Environment in Rhodopsin in Light of a New 2.2 Å Crystal Structure', *Journal of Molecular Biology*. 342(2), pp. 571–583. doi: 10.1016/j.jmb.2004.07.044.

Olsen, J. *et al.* (2004) 'Calcitonin Gene–Related Peptide Receptor Antagonist BIBN4096BS for the Acute Treatment of Migraine', *The New England Journal of Medicine*. 350(11), pp. 1104–1110.

- Overington J.P, Al-Lazikani, B. and Hopkins A.L. (2006) 'How Many Drug Target are There?', *Nature Reviews Drug Discovery*, 5(12), pp. 993-6
- Pan, D., Nakatsu, T. and Kato, H. (2013) 'Crystal structure of peroxisomal targeting signal-2 bound to its receptor complex Pex7p – Pex21p', *Nature Publishing Group*. 20(8), pp. 987–993. doi: 10.1038/nsmb.2618.
- Patrick, A. N. *et al.* (2013) 'Structure-function analyses of the human SIX1 – EYA2 complex reveal insights into metastasis and BOR syndrome', 20(4). pp. 447-453 doi: 10.1038/nsmb.2505.
- Paulsen, C. E. *et al.* (2015) 'Structure of the TRPA1 ion channel suggests regulatory mechanisms', *Nature*. 520(3), pp. 511–517. doi: 10.1038/nature14367.Structure.
- Pioszak, A. A. and Xu, H. E. (2008) 'Molecular recognition of parathyroid hormone by its G protein-coupled receptor', *PNAS*. 105(13), pp. 1–6.
- Poyner, D. R. *et al.* (2002) 'International Union of Pharmacology . XXXII . The Mammalian Calcitonin Gene-Related Peptides, Adrenomedullin, Amylin, and Calcitonin Receptors', *Pharmacological Reviews*. 54(2), pp. 233–246.
- Qi, T. *et al.* (2008) 'Identification of N-Terminal Receptor Activity-Modifying Protein Residues Important for Calcitonin Gene-Related Peptide, Adrenomedullin, and Amylin Receptor Function', *Molecular Pharmacology*. 74(4), pp. 1059–1071. doi: 10.1124/mol.108.047142.
- Qi, T. *et al.* (2011) 'Structure-function analysis of amino acid 74 of human RAMP1 and RAMP3 and its role in peptide interactions with adrenomedullin and calcitonin gene-related peptide receptors', *Peptides*. 32(5), pp. 1060–1067. doi: 10.1016/j.peptides.2011.03.004.
- Qi, T. and Hay, D. L. (2010) 'Structure-function relationships of the N-terminus of receptor activity-modifying proteins', *British Journal of Pharmacology*. 159(5), pp. 1059–1068. doi: 10.1111/j.1476-5381.2009.00541.x.
- Renaud, J.-P. *et al.* (2018) 'Cryo-EM in drug discovery: achievements, limitations and prospects', *Nature Reviews Drug Discovery*. Nature Publishing Group, 17, pp. 417–492. doi: 10.1038/nrd.2018.77.
- Reuten, R. *et al.* (2016) 'Maltose-binding protein (MBP), a secretion-enhancing tag for mammalian protein expression systems', *PLoS ONE*. 11(3), pp. 1–15. doi: 10.1371/journal.pone.0152386.
- Riggs, P. (2000) 'Expression and Purification of Recombinant Proteins by Fusion to Maltose-Binding Protein', *Molecular Biotechnology*. 15(1), pp. 51–63.
- Robertson, N. *et al.* (2011) 'The properties of thermostabilised G protein-coupled receptors (StaRs) and their use in drug discovery', *Neuropharmacology*. 60(1), pp. 36–44. doi: 10.1016/j.neuropharm.2010.07.001.
- Rocchi, P. *et al.* (2001) 'Expression of Adrenomedullin and Peptide Amidation Activity in Human Prostate Cancer and in Human Prostate Cancer Cell Lines 1'. *Cancer Research*, 61(3), pp. 1196–1206.

- Roehrkasse, A. M. *et al.* (2018) 'cro Structure – function analyses reveal a triple-turn receptor- bound conformation of adrenomedullin 2/intermedin and enable peptide antagonist design', *Journal of Biological Chemistry*. 293, pp. 15840–15854. doi: 10.1074/jbc.RA118.005062.
- Rosenbaum, D. M. *et al.* (2007) 'GPCR Engineering Yields High-Resolution Structural Insights into β_2 -Adrenergic Receptor Function', *Science*. 318(5854), pp. 1266–1273. doi: 10.1126/science.1150609.
- Rosenbaum, D. M., Rasmussen, S. G. F. and Kobilka, B. K. (2014) 'The structure and function of G-protein-coupled receptors', *Nature*. 459(7245), pp. 356–363. doi: 10.1038/nature08144.The.
- Rost, B. and Sander, C. (1996) 'Bridging the protein sequence-sequence gap by structure predictions', *Annual Review of Biophysical and Biomolecular Structures*. 25, pp. 113–136.
- Rucktooa, P. *et al.* (2018) 'Towards high throughput GPCR crystallography: In Meso soaking of Adenosine A_{2A} Receptor crystals', *Scientific Reports*. 8(41), pp. 2–8. doi: 10.1038/s41598-017-18570-w.
- Rudolf, K. *et al.* (2005) 'Development of human calcitonin gene-related peptide (CGRP) receptor antagonists. 1. Potent and selective small molecule CGRP antagonists. 1-[N₂-[3,5-dibromo-N-[[4-(3,4-dihydro-2(1H)-oxoquinazolin-3-yl)-1-piperidinyl]carbonyl]-D-tyrosyl]-l-lysyl]-4-(4-pyridinyl)piperazine: the first CGRP antagonist for clinical trials in acute migraine', *Journal of Medicinal Chemistry*. 48(19), pp. 5921–5931.
- Rupp, B. (2015) 'IYCr crystallization series Origin and use of crystallization phase diagrams IYCr crystallization series', *Acta crystallographica Section F Structural Biology Communications*. International Union of Crystallography.71(Pt 3), pp. 247–260. doi: 10.1107/S2053230X1500374X.
- Russell, F. A. *et al.* (2014) 'Calcitonin Gene-Related Peptide: Physiology and Pathophysiology', *Physiological Reviews*. 94(4), pp. 1099–1142. doi: 10.1152/physrev.00034.2013.
- Sabourin, M. *et al.* (2007) 'A flexible protein linker improves the function of epitope-tagged proteins in *Saccharomyces cerevisiae*', *Yeast*. 24(1), pp. 39–45. doi: 10.1002/yea.1431.A.
- Sadry, S. A. and Drucker, D. J. (2013) 'Emerging combinatorial hormone therapies for the treatment of obesity and T2DM', *Nature Reviews Endocrinology*. Nature Publishing Group, 9, pp. 425–433. doi: 10.1038/nrendo.2013.47.
- Salemme, F.R. (1972). 'A free interface diffusion technique for the crystallisation of proteins for X-ray crystallography'. *Arch Biochem Biophys*. (151) pp.533-539
- Schiöth, H. B. and Fredriksson, R. (2005) 'The GRAFS classification system of G-protein coupled receptors in comparative perspective', *General and Comparative Endocrinology*. 142(1–2), pp. 94–101. doi: 10.1016/j.ygcen.2004.12.018.
- Schmitt, J., Hess, H. and Stunnenberg, H. G. (1993) 'Affinity purification of histidine-tagged proteins', *Molecular Biology Reports*. 18, pp. 223–230.

- Salvatore, C. A. *et al.* (2008) 'Pharmacological Characterization of MK-0974 [N-[(3R,6S)-6-1-carboxamide], a Potent and Orally Active Calcitonin Gene-Related Peptide Receptor Antagonist for the Treatment of Migraine', *The Journal of Pharmacology and Experimental Therapeutics*. 324(2), pp. 416–421. doi: 10.1124/jpet.107.130344.
- Salvatore, C. A. *et al.* (2010) 'Pharmacological Properties of MK-3207, a Potent and Orally Active Calcitonin Gene-Related Peptide Receptor Antagonist', *Journal of Pharmacology and Experimental Therapeutics*. 333(1), pp. 152–160. doi: 10.1124/jpet.109.163816.
- Scapin, G. (2013) 'research papers Molecular replacement then and now research papers', *Acta Crystallographica Section D: Biological Crystallography*. 69(pt 11), pp. 2266–2275. doi: 10.1107/S0907444913011426.
- Schiess, M. C. *et al.* (2005) 'The effects of CGRP on calcium transients of dedifferentiating cultured adult rat cardiomyocytes compared to non-cultured adult cardiomyocytes: possible protective and deleterious results in cardiac function', *Peptides*. 26(3), pp. 525–530. doi: 10.1016/j.peptides.2004.10.020.
- Schindler, M. and Doods, H. N. (2002) 'Binding properties of the novel, non-peptide CGRP receptor antagonist radioligand, [3H] BIBN4096BS', *Journal of Pharmacology*. 442, pp. 187–193.
- Seras-franzoso, J. *et al.* (2012) 'Disulfide Bond Formation and Activation of Escherichia coli-Galactosidase under Oxidizing Conditions', *Applied and Environmental Microbiology*. 78(7), pp. 2376–2385. doi: 10.1128/AEM.06923-11.
- Sexton, P. M. *et al.* (2001) 'Receptor activity modifying proteins', *Cellular Signalling*. 13, pp. 73–83.
- Sezonov, G. and Ari, R. D. (2007) 'Escherichia coli Physiology in Luria-Bertani Broth', *Journal of Bacteriology*. 189(23), pp. 8746–8749. doi: 10.1128/JB.01368-07.
- Shoichet, B.K. and Kobilka, B.K. (2012). 'Structure-based drug screening for G-protein-coupled receptors'. *Trends in Pharmacological Sciences*. 33(5) pp. 268-272.
- Simms, J. *et al.* (2006) 'Characterization of the structure of RAMP1 by mutagenesis and molecular modeling.', *Biophysical journal*. 91(2), pp. 662–669. doi: 10.1529/biophysj.106.084582.
- Simms, J. *et al.* (2009) 'Structure-Function Analysis of RAMP1 by Alanine Mutagenesis', *Biochemistry*. 40(1), pp. 198–205.
- Smyth, M. S. and Martin, J. H. J. (2000) 'Review x Ray crystallography', *Molecular Pathology*. 53(1), pp. 8–14.
- Starkenstein, E. (1910). Ferment action and the influence upon it of neutral salts. *Biochemische Zeitschrift*. 24, pp. 210–218
- Stevens, R. C. *et al.* (2013) 'The GPCR Network: a large-scale collaboration to determine human GPCR structure and function', *Nature Reviews Drug Discovery*. 12, pp. 25–34. doi: <http://dx.doi.org/10.1038/nrd3859>.

- Stump, C. A. *et al.* (2009) 'Bioorganic & Medicinal Chemistry Letters The discovery of highly potent CGRP receptor antagonists', *Bioorganic & Medicinal Chemistry Letters*. 19(1), pp. 214–217. doi: 10.1016/j.bmcl.2008.10.106.
- Thomas, D.H., Rob A. and Rice, D. W. (1989). 'A novel dialysis procedure for the crystallisation of proteins'. *Protein Engineering*. (2) pp.489-491
- Thompson, R. F. *et al.* (2016) 'An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology', *Methods*. 100, pp. 3–15. doi: 10.1016/j.ymeth.2016.02.017.
- Tuteja, N. (2009) 'Signaling through G protein coupled receptors', *Plant Signaling & Behavior*. 4(10), pp. 942–947. doi: 10.4161/psb.4.10.9530.
- Uddman, R. *et al.* (1985) 'Innervation of the feline cerebral vasculature by nerve fibers containing calcitonin gene-related peptide: Trigeminal origin and co-existence with substance P', *Neuroscience Letters*. 62(1), pp. 131–136. doi: 10.1016/0304-3940(85)90296-4.
- Urh, M., Simpson, D. and Zhao, K. (2009) 'Affinity Chromatography : General Methods', *Methods in Enzymology*. 463, pp. 417-438. doi: 10.1016/S0076-6879(09)63026-3.
- Vekilov, P. G. & Chernov, A. A. (2002). 'The physics of protein crystallization'. *Solid State Physics*. 57, 2–147.
- Voss, T. *et al.* (2016) 'A phase IIb randomized, double-blind, placebo-controlled trial of ubrogepant for the acute treatment of migraine', *Cephalalgia*. 36(9), pp. 887–898. doi: 10.1177/0333102416653233.
- Wacker, M. *et al.* (2002) 'N-Linked Glycosylation in *Campylobacter jejuni* and Its Functional Transfer into *E. coli*', *Science*. 298(5599), pp. 1790–1794.
- Wasko, M. J. *et al.* (2015) 'A role for fragment-based drug design in developing novel lead compounds for central nervous system targets', *Frontiers in Neurology*. 6, pp. 1–11. doi: 10.3389/fneur.2015.00197.
- Watkins, H. A. *et al.* (2014) 'Receptor activity-modifying protein-dependent effects of mutations in the calcitonin receptor-like receptor: Implications for adrenomedullin and calcitonin gene-related peptide pharmacology', *British Journal of Pharmacology*. 171(3), pp. 772–788. doi: 10.1111/bph.12508.
- Weiû, H. M. and Grisshammer, R. (2002) 'Purification and characterization of the human adenosine A_{2a} receptor functionally expressed in *Escherichia coli*', *European Journal of Biochemistry*. 269(1), pp. 82–92.
- Weston, C. *et al.* (2016) 'Receptor activity-modifying protein-directed G protein signaling specificity for the calcitonin gene-related peptide family of receptors', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M116.751362.
- Weiû, H. M. and Grisshammer, R. (2002) 'Purification and characterization of the human adenosine A_{2a} receptor functionally expressed in *Escherichia coli*', *European Journal of*

Biochemistry. 269(1), pp. 82–92.

Wermuth, C. G. (2006) 'Selective optimization of side activities : the SOSA approach', *Drug Discovery Today*. 11(3), pp. 160–164.

Wootten, D. *et al.* (2013) 'Receptor activity modifying proteins (RAMPs) interact with the VPAC 2 receptor and CRF 1 receptors and', *British Journal of Pharmacology*. 168(4), pp. 822–34. doi: 10.1111/j.1476-5381.2012.02202.x.

Wulhfard, S. *et al.* (2008) 'Mild hypothermia improves transient gene expression yields several fold in Chinese hamster ovary cells', *Biotechnology Progress*. 24(2), pp. 458–465. doi: 10.1021/bp070286c.

Xiang, J. *et al.* (2016) 'Successful Strategies to Determine High-Resolution Structures of GPCRs', *Trends in Pharmacological Sciences*. 37(12), pp. 1055–1069. doi: 10.1016/j.tips.2016.09.009.

Yamauchi, A. *et al.* (2014) 'Journal of Molecular and Cellular Cardiology Functional differentiation of RAMP2 and RAMP3 in their regulation of the vascular system', *Journal of Molecular and Cellular Cardiology*. 77, pp. 73–85. doi: 10.1016/j.yjmcc.2014.09.017.

Zhang, D., Zhao, Q. and Wu, B. (2015) 'Structural Studies of G Protein-Coupled Receptors', *Molecules and Cells*. 38(10), pp. 836–842.

Zhang, Y., Devries, M. E. and Skolnick, J. (2006) 'Structure Modeling of All Identified G Protein – Coupled Receptors in the Human Genome', *PLoS One Computational Biology*. 2(2), p. e23. doi: 10.1371/journal.pcbi.0020013.

Zudaire, E. *et al.* (2006) 'Adrenomedullin Is a Cross-Talk Molecule that Regulates Tumor and Mast Cell Function during Human Carcinogenesis', *American Journal of Pathology*. 168(1), pp. 280–291. doi: 10.2353/ajpath.2006.050291.

S. Supplementary Section

S.1 Supplementary Section 1: Signal peptide, MBP and RAMP-CLR ECD construct Sequences

S.1.1 RAMP1-CLR Receptor ECD 10x Histidine Sequence

G AAT TCG ACT GCC TGC CAG GAG GCT AAC TAC GGT GCC CTC CTC CGG
GAG CTC TGC CTC ACC CAG TTC CAG GTA GAC ATG GAG GCC GTC GGG
GAG ACG CTG TGG TGT GAC TGG GGC AGG ACC ATC AGG AGC TAC AGG
GAG CTG GCC GAC TGC ACC TGG CAC ATG GCG GAG AAG CTG GGC TGC
TTC TGG CCC AAT GCA GAG GTG GAC AGG TTC TTC CTG GCA GTG CAT
GGC CGC TAC TTC AGG AGC TGC CCC ATC TCA GGC AGG GCC GTG GGA
AGC GAA GGA AGC GAA GGA AGC GAA GAG GAC TCA ATT CAG TTG GGA
GTT ACT AGA AAT AAA ATC ATG ACA GCT CAA TAT GAA TGT TAC CAA
AAG ATT ATG CAA GAC CCC ATT CAA CAA GCA GAA GGC GTT TAC TGC
AAC AGA ACC TGG GAT GGA TGG CTC TGC TGG AAC GAT GTT GCA GCA
GGA ACT GAA TCA ATG CAG CTC TGC CCT GAT TAC TTT CAG GAC TTT
GAT CCA TCA GAA AAA GTT ACA AAG ATC TGT GAC CAA GAT GGA AAC
TGG TTT AGA CAT CCA GCA AGC AAC AGA ACA TGG ACA AAT TAT ACC
CAG TGT AAT GTT AAC ACC CAC GAG AAA GTG AAG ACT GCA CTA AAT
TTG TTT TAC GGA GGC GGA GGC GGC GGC CAT CAT CAC CAT CAC CAC
CAT CAT CAC CAC TAG GGT AC

S.1.2 RAMP2-CLR Receptor ECD 10x Histidine Sequence

G AAT TCG GGG GGG ACG GTG AAG AAC TAT GAG ACA GCT GTC CAA TTT
TGC TGG AAT CAT TAT AAG GAT CAA ATG GAT CCT ATC GAA AAG GAT
TGG TGC GAC TGG GCC ATG ATT AGC AGG CCT TAT AGC ACC CTG CGA
GAT TGC CTG GAG CAC TTT GCA GAG TTG TTT GAC CTG GGC TTC CCC
AAT CCC TTG GCA GAG AGG ATC ATC TTT GAG ACT CAC CAG ATC CAC
TTT GCC AAC TGC TCC CTG GTG CAG CCC ACC TTC TCT GGA AGC GAA
GGA AGC GAA GGA AGC GAA GAG GAC TCA ATT CAG TTG GGA GTT ACT
AGA AAT AAA ATC ATG ACA GCT CAA TAT GAA TGT TAC CAA AAG ATT
ATG CAA GAC CCC ATT CAA CAA GCA GAA GGC GTT TAC TGC AAC AGA
ACC TGG GAT GGA TGG CTC TGC TGG AAC GAT GTT GCA GCA GGA ACT
GAA TCA ATG CAG CTC TGC CCT GAT TAC TTT CAG GAC TTT GAT CCA

TCA GAA AAA GTT ACA AAG ATC TGT GAC CAA GAT GGA AAC TGG TTT
 AGA CAT CCA GCA AGC AAC AGA ACA TGG ACA AAT TAT ACC CAG TGT
 AAT GTT AAC ACC CAC GAG AAA GTG AAG ACT GCA CTA AAT TTG TTT
 TAC GGA GGC GGA GGC GGC GGC CAT CAT CAC CAT CAC CAC CAT CAT
 CAC CAC TAG GGT ACC

S.1.3 RAMP3-CLR Receptor ECD 10x Histidine Sequence

G AAT TCG GGC TGC AAC GAG ACA GGC ATG TTG GAG AGG CTG CCC CTG
 TGT GGG AAG GCT TTC GCA GAC ATG ATG GGC AAG GTG GAC GTC TGG
 AAG TGG TGC AAC CTG TCC GAG TTC ATC GTG TAC TAT GAG AGT TTC
 ACC AAC TGC ACC GAG ATG GAG GCC AAT GTC GTG GGC TGC TAC TGG
 CCC AAC CCC CTG GCC CAG GGC TTC ATC ACC GGC ATC CAC AGG CAG
 TTC TTC TCC AAC TGC ACC GTG GAC AGG GTC CAC TTG GAG GAC CCC
 CCA GAC GAG GTT GGA AGC GAA GGA AGC GAA GGA AGC GAA GAG GAC
 TCA ATT CAG TTG GGA GTT ACT AGA AAT AAA ATC ATG ACA GCT CAA
 TAT GAA TGT TAC CAA AAG ATT ATG CAA GAC CCC ATT CAA CAA GCA
 GAA GGC GTT TAC TGC AAC AGA ACC TGG GAT GGA TGG CTC TGC TGG
 AAC GAT GTT GCA GCA GGA ACT GAA TCA ATG CAG CTC TGC CCT GAT
 TAC TTT CAG GAC TTT GAT CCA TCA GAA AAA GTT ACA AAG ATC TGT
 GAC CAA GAT GGA AAC TGG TTT AGA CAT CCA GCA AGC AAC AGA ACA
 TGG ACA AAT TAT ACC CAG TGT AAT GTT AAC ACC CAC GAG AAA GTG
 AAG ACT GCA CTA AAT TTG TTT TAC GGA GGC GGA GGC GGC GGC CAT
 CAT CAC CAT CAC CAC CAT CAT CAC CAC TAG GGT ACC

S.1.4 RAMP3-CLR Receptor ECD 6x Histidine Sequence

G AAT TCG GGC TGC AAC GAG ACA GGC ATG TTG GAG AGG CTG CCC
 CTG TGT GGG AAG GCT TTC GCA GAC ATG ATG GGC AAG GTG GAC GTC
 TGG AAG TGG TGC AAC CTG TCC GAG TTC ATC GTG TAC TAT GAG AGT
 TTC ACC AAC TGC ACC GAG ATG GAG GCC AAT GTC GTG GGC TGC TAC
 TGG CCC AAC CCC CTG GCC CAG GGC TTC ATC ACC GGC ATC CAC AGG
 CAG TTC TTC TCC AAC TGC ACC GTG GAC AGG GTC CAC TTG GAG GAC
 CCC CCA GAC GAG GTT GGA AGC GAA GGA AGC GAA GGA AGC GAA GAG
 GAC TCA ATT CAG TTG GGA GTT ACT AGA AAT AAA ATC ATG ACA GCT
 CAA TAT GAA TGT TAC CAA AAG ATT ATG CAA GAC CCC ATT CAA CAA

GCA GAA GGC GTT TAC TGC AAC AGA ACC TGG GAT GGA TGG CTC TGC
TGG AAC GAT GTT GCA GCA GGA ACT GAA TCA ATG CAG CTC TGC CCT
GAT TAC TTT CAG GAC TTT GAT CCA TCA GAA AAA GTT ACA AAG ATC
TGT GAC CAA GAT GGA AAC TGG TTT AGA CAT CCA GCA AGC AAC AGA
ACA TGG ACA AAT TAT ACC CAG TGT AAT GTT AAC ACC CAC GAG AAA
GTG AAG ACT GCA CTA AAT TTG TTT TAC CAT CAT CAC CAT CAC CAC
TAG GGT ACC

Key

Black: CLR ECD gene sequence

Green: RAMP ECD gene sequence

Red: Linker sequences

Gold: Histidine tag sequence

Yellow: Stop codon

Blue: Restriction Sequence

S.1.5 Signal Peptide Gene String Sequence

GACACACCCGACACACCCGCCAGCGCCGCAGGAGGTACTCACGATGGCGATGAGGGCCTG
GATCTTCTTTCTCCTTTGCTGGCCGGGAGGGCTCTGGCAGCCCCGCTAGCTGAAGAAGGT
AAACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGA
CACACCC

Key :

Blue: Restriction site

Green: BM40 signal peptide sequence

S.1.5 MBP Amino-Acid sequence

EEGKLVIIWINGDKGYNGLAEVGGKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWA
HDFRGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSIIYNKDLLPN

PPKTWEEI PALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGKYDIKDVGVND
 AGAKAGLTFVLVDLIKKNHMNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGTV
 LPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSY
 EEELAKDPRIAATMENAQKGEIMPNI PQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNS
 SS

S.2. Supplementary Section 2: pSF-CMV-Puro-NH₂-MBP Vector Map

