UNIVERSITY of BRADFORD Library

University of Bradford eThesis

This thesis is hosted in Bradford Scholars – The University of Bradford Open Access repository. Visit the repository for full metadata or to contact the repository team

© University of Bradford. This work is licenced for reuse under a Creative Commons Licence.

SYNTHESIS AND CHARACTERISATION OF DUAL CCR7 / CXCR4 ANTAGONIST

M.C. IZIDRO

PHD

2019

Synthesis and Characterisation of Dual CCR7/CXCR4 Antagonists

Mario Carlos IZIDRO

Submitted for the Degree of
Doctor of Philosophy

Faculty of Life Sciences

Institute of Cancer Therapeutics

University of Bradford

2019

Abstract

Mario Carlos Izidro

Synthesis and Characterisation of Dual CCR7/CXCR4 Antagonists

Key words: cancer, metastasis, chemokines, receptor, CCR7, CXCR4, antagonist, GPCR, medicinal chemistry, calcium flux assay.

Metastasis is a major cause of death in cancer patients but currently there are no drugs available for its treatment. Hence there is an urgent clinical need for identifying and developing anti-metastatic drugs. The activation of CC chemokine receptor 7 (CCR7) and C-X-C chemokine receptor type 4 (CXCR4) plays an important role in lymph node metastasis in a variety of cancers. Indeed, in patients with tumours which are positive for CCR7 and/or CXCR4 expression, prognosis and survival are poorer than those whose tumours are negative for these receptors. CCR7 and/or CXCR4 activation, in addition to being involved in inducing invasive phenotypes in cancer cells, promotes tumour cell growth and survival. Our group has previously identified a series of sulfonamides as CCR7 antagonists. This project aims to extend on those studies and to develop a dual CCR7 and CXCR4 antagonist to reduce metastasis in cancer.

Novel potent biaryl sulfonamide CCR7 antagonists were synthesised and assessed by calcium flux assay. Several potential dual CCR7 and CXCR4 biaryl sulfonamide antagonists have been synthesised, these are hybrid compounds incorporating features from CCR7 antagonists of this project, and from known sulfonamide CXCR4 antagonists. The most potent of such compound was able to inhibit CCR7 activation in calcium flux assay (95% inhibition at 1 μM), however, the relative potency of these compounds as CXCR4 antagonists was low. Molecular docking was used to investigate the binding mode of the synthesised compounds in CCR7 and CXCR4. The generated docking poses were able to rationalise some of the calcium flux assay results.

Table of contents

	Abstract	i
	Table of contents	ii
	List of abbreviations	iv
	List of figures	vii
	List of tables	xii
	Acknowledgement	xiv
1.0	**Introduction**	
1.1	What is metastasis?	1
1.2	Chemokines and chemokine receptors	2
1.3	CCR7 and CXCR4 in cancer	7
1.4	Known CXCR4 and CCR7 antagonists	11
1.5	Rationale for a CCR7 and CXCR4 small molecule dual antagonist	18
2.0	**Developing a small molecule dual antagonist**	
2.1	Shifting to multi-target drugs from selective single target drugs	20
2.2	Lead generation for multi-target drugs	21
2.3.1	Case study: Dual aromatase steroid sulfatase inhibitors	23
2.3.2	Case study: the discovery of CCR2 and 5 dual antagonists	24
2.3.3	Case study: the discovery of endothelin and angiotensin II receptor dual antagonist	25
2.4.0	Lead generation for multi-target drugs with screening methods	27
2.4.1	High-throughput screening	27
2.4.2	Fragment-based drug design	29
2.4.3	Virtual ligand screening	35
3.0	**Project background and objectives**	
3.1	Previous work on CCR7 in our group	37
3.2	Dual CXCR4/CCR7 antagonists	38
3.3	Aims and objectives	40
4.0	**Chemistry: results and discussion**	

4.1	Modifications to ICT13069 at fragments C and A	41
4.2	Metabolically stable analogues of ICT13069	43
4.3	Synthesis of reverse sulfonamides	45
4.4	Synthesis of CXCR4 and dual CXCR4/CCR7 antagonists	56
4.5	Synthesis of reverse sulfonamides with modifications at positions A and C	65
5.0	**Molecular docking**	
5.1	Introduction	71
5.2	Construction of homology models vs crystal structures	72
5.3	Preparation for molecular docking	78
5.4	Handling and analysing docking results on MOE	81
5.5	Replicating binding mode of Cmp2105 and Vercirnon	83
5.6	Molecular docking: discussion and results	85
5.6.1	CCR7 binding: extracellular and transmembrane domain	86
5.6.2	CCR7 binding: intracellular domain	91
5.6.3	CXCR4 binding: extracellular-transmembrane domain	100
5.6.4	CXCR4 binding: Intracellular domain	108
5.6.5	CCR7 binding: extracellular and transmembrane domain – homology model	110
5.7	Correlating docking scores with IC_{50} from calcium flux assay	123
6.0	**Biological investigation of CXCR4 and dual CXCR4/CCR7 antagonists**	
6.1	Cytotoxicity assay	127
6.2	Calcium flux assay: Results and discussion	128
7.0	**Summary and conclusion**	138
8.0	**Experimental**	142
9.0	**References**	183
Appendix 1: Structures and IC_{50} (μM) of 34 CCR7 antagonists used in molecular docking		200
Appendix 2: MTT curves		203

List of abbreviation

AI	Non-steroidal aromatase inhibitor
AT_1	Angiotensin II subtype 1
Bcl-2	B-cell lymphoma 2
BCLL	B-cell chronic lymphocytic leukaemia
B-Raf	Serine/threonine-protein kinase B-Raf
CCL	Chemokine C-C motif ligand
CCR	C-C chemokine receptor
CDK1	Cyclin-dependent kinase 1
CXCL	Chemokine C-X-C motif ligand
CXCR	C-X-C chemokine receptor
DASI	Dual aromatase steroid sulfatase inhibitors
DCM	Dichloromethane
DMF	Dimethylformamide
EMT	Epithelial mesenchymal transition
ErbB1/2	Receptor tyrosine kinase ErbB1/2
ERK1/2	Extracellular signal-regulated kinases 1/2
ETa	Endothelin subtype A
FBDD	Fragment based drug design
FDA	Food drug administration
FLT3	Receptor tyrosine kinase FLT3
G2 phase	Growth 2 phase
GPCR	G-protein coupled receptor
GSK3B	Glycogen synthase kinase-3 beta
HER	Human epidermal growth factor receptor
HMBC	Heteronuclear multiple bond correlation

HIV	Human immunodeficiency virus
IR	Infrared (spectroscopy)
IC_{50}	Half maximal inhibitory concentration
kcal/mol	Kilocalorie per mole
KIT	Receptor tyrosine kinase KIT
Ki	Inhibitory constant
MOE	Molecular Operating Environment
m.p.	Melting point
M phase	Mitosis phase
mg	Milligram
mL	Millilitre
mM	Millimolar
µM	Micromolar
MIF	Macrophage migration inhibitory factor
MMP	Matrix metallopeptidase
MS	Mass spectroscopy
m.w.	Molecular weight
NK_1	Neurokinin 1 receptor
nM	Nanomolar
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser effect spectroscopy
NSCLC	Non-small-cell lung carcinoma
PDGFR	Platelet-derived growth factor
Pin1	Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1
PDB	Protein Data Bank
RET	Receptor tyrosine kinase RET

r.t.	Room temperature
SAR	Structure activity relation
SERT	Serotonin transporter
STS	Steroid sulfatase
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TM	Transmembrane domain
VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor

List of figures

Figure		Page
1	The metastatic process	1
2	The four sub groups of chemokines	2
3	Classification of chemokines and their receptors in a "chemokine wheel"	3
4	Disease free survival rate and overall survival rate of cervical cancer patients	4
5	Structure of Cosalane, Cmp2105 and Navarixin	11
6	Chemical structures of a selection of CXCR4 antagonists	13
7	Structure of AMD11070 and the general structure of the tetrahydroquinoline based CXCR4 antagonists	16
8	General structure of the sulfonamide CXCR4 antagonists	17
9	The three ways of joining two drug molecules together to form a multi-targeting drug	21
10	The development of the most promising dual aromatase steroid sulfatase inhibitors	22
11	Two CCR2 antagonists developed by Carter et al.	23
12	Development of the endothelin and angiotensin II receptor dual antagonist	25
13	NK1 and SERT dual antagonist found from high-throughput screening. Further optimization gave a more potent dual ligand	28
14	The lead drug molecule from the high throughput screening might not have the desirable balance of activity for the targets	28
15	The difference in binding between drug like molecules and fragments	30
16	Fragment growing	30
17	Fragment linking	31
18	Computational modelling of fragment molecule binding to RET	32
19	Creation of the dual pan-RET/VEGFR2 kinase inhibitor Pz-1	33

20	The development history of ICT13069	37
21	Compound **30** (ICT13069) representing the general structure of the small molecule antagonist of this project	38
22	Sequence alignment of CCR7 and CXCR4	39
23	Combining structural features between the CXCR4 antagonists	39
24	Synthesis of compounds **30** and **36**-**40**	41
25	Proposed mechanisms for the formation of sulfonamides	42
26	Synthesis of compound **41**	42
27	Synthesis of compound **42**	43
28	Reaction mechanism leading to the formation of compound **42**	44
29	Proposed preparation of reverse sulfonamides analogues of ICT13069.	45
30	Synthesis of compound **46**	45
31	Reaction scheme of 3-hydroxybenzoic acid to reverse sulfonamide	46
32	One of the ways to differentiate the para and ortho isomer through NOESY correlation	47
33	Reaction scheme of 3-methoxybenzonitrile to reverse sulfonamide	48
34	Reaction scheme of methyl and t-*butyl* ester of 3-methoxybenzoate to reverse sulfonamide of ICT13069	49
35	Sulfonation of 3-methoxy benzylpyrrolidine	50
36	Proposed mechanism of reductive amination of **58**	50
37	Using Sandmeyer reaction in attempt to synthesize the reverse sulfonamide of ICT13069	52
38	Reaction mechanism of the diazonium ion formation used for the Sandmeyer reactions	53
39	Reaction scheme for the synthesis of compound **64** (ICT13303)	53
40	Synthesis of **65**-**67**	55
41	Reaction scheme of **6**, **72** and **73**	56
42	Synthesis of **74**	57

43	Synthesis of **78**	58
44	Synthesis of **79**	58
45	The general reaction scheme of rearrangement reaction which utilize directed ortho metalation	59
46	Proposed reaction mechanism of the rearrangement reaction	60
47	Synthesis of **81**	61
48	Attempted synthesis of **84**	62
49	Synthesis of **86** and **87**	64
50	Reverse sulfonamides with modifications at fragment A and C	65
51	Synthesis of **92** and **93**	66
52	Structure of ICT13312	66
53	Synthesis of **96**-**98**	67
54	Synthesis of **107**	68
55	Mechanism of Heck coupling reaction	69
56	Synthesis of **108**-**110**	70
57	Structure of **111,** potent adenosine A2A receptor antagonist	71
58	Alignment of CCR7 and CCR9 amino sequence	73
59	Structure of CCR7/6QZH with Cmp2105 docked at the intracellular site	75
60	Superposition all orthosteric ligands binding mode observed in the crystal structure for all class A GPCRs	76
61	CCR9 5LWE Vercirnon bound to the intracellular allosteric site	77
62	3D representation of the two binding sites in CCR7 calculated by the site finder function in MOE	79
63	Example work flow of repeating the docking of ICT13069 to acquire a docking pose with higher score. Initial docking gives ligand	81
64	Superposition of two Cmp2105 docked at the CCR7 intracellular site	82

65	Superposition of two Vercirnon docked at the CCR9 intracellular site	83
66	Superposition of best docking poses for compounds similar to ICT13069	85
67	An unlikely docking pose of an analogue of ICT13069 with a nitro substituent	86
68	Docking pose of compound **41** to the extracellular-transmembrane domain of CCR7, and its ligand interaction map generated by MOE	87
69	Superposition of CCR7 (6QZH, **Red**) and CXCR4 (3ODU, **Green**), comparing the two conformations	88
70	Superposition of CCR7 (6QZH) with Cmp2105 (pink), and CCR9 (5LWE) with Vercirnon (yellow)	90
71	Ligand interaction map of Vercirnon within the intracellular allosteric site of CCR7, after superimposing CCR9 and CCR7 then delete the CCR9 structure.	91
72	Docking pose of compound **41** at the intracellular site of CCR7, generated from altering Vercirnon to **41**	92
73	Docking pose of **41** at the CCR7 intracellular site and its ligand interaction map	93
74	Best docking pose of **67** at the CCR7 intracellular allosteric site, and the ligand interaction map generated by MOE	94
75	Top docking pose for compound **97** at the CCR7 intracellular site, and the ligand interaction map generated by MOE	95
76	Top docking pose of **6** at the CCR7 intracellular site, and its ligand interaction map generated by MOE	96
77	Superposition of the docking pose for pyrrolidine compounds with 4-*tert*butylbenzene	97
78	Top docking pose of **79** at the CCR7 intracellular site, and its ligand interaction map generated by MOE	98
79	Docking pose of **6** published by Mooring et al.	100
80	Best docking pose of **6** at CXCR4 transmembrane site	101

81	Top docking pose of compound **72** at CXCR4 transmembrane site and its ligand interaction map generated by MOE	101
82	General structure of pyrrolidine compounds with 4-*tert*butylbenzene	102
83	Docking pose of compound **72** and its ligand interaction map generated by MOE	103
84	Superposition of the top scoring docking poses of the pyrrolidine compounds with 4-*tert*butylbenzene	103
85	Best docking pose of compound **78** at the CXCR4 transmembrane site, and its ligand interaction map generated by MOE	104
86	Best docking pose of **87** at the CXCR4 transmembrane site	105
87	Best docking pose of **79** at the CXCR4 transmembrane site, and its ligand interaction map generated by MOE	105
88	Best docking pose of **67** at the CXCR4 transmembrane site, and its ligand interaction map generated by MOE	106
89	Superposition of CXCR4/3ODU (Green) to the crystal structures of CCR2,5,7,9	108
90	3D representation of the binding site calculated by MOE	110
91	The Ramachandran plot and other properties used for evaluating the CCR7 homology model	111
92	Possible docking pose of compound **67** at the transmembrane site of the CCR7 homology model, and the ligand interaction map generated by MOE	112
93	Docking pose of compound **42** at the transmembrane site of the CCR7 homology model	113
94	Interaction maps of extended chain compounds generated by MOE	114
95	Possible docking pose of compound **6** at the transmembrane site of the CCR7 homology model	115
96	Superposition of CCR7 antagonist docking pose with the pyrrolidine compound docking pose	116

97	Top docking pose of compound **79** at the transmembrane site of the CCR7 homology model, and the ligand interaction map generated by MOE	117
98	Second top docking pose of compound **79** at the transmembrane site of the CCR7 homology model	118
99	Top pose of compound **87** at the transmembrane site of the CCR7 homology model	119
100	The superposition of CCR7 homology model with a CCR7 antagonists, and CXCR4 with compound **6**	121
101	Correlation plot between IC_{50} of 27 CCR7 antagonists and their top docking scores at the transmembrane site for each scoring function available to MOE	124
102	Correlation plot between IC_{50} of 27 CCR7 antagonists and their top docking scores at the intracellular site for each scoring function available to MOE	125
103	Correlation plot between IC_{50} of 27 CCR7 antagonists and their post energy minimisation docking scores at the transmembrane and intracellular site	126
104	The Conversion of MTT to formazan	127
105	The principles of the calcium flux assay	128
106	Fluorescent response versus time in addition of different antagonists	129
107	Ligand interaction map generated by MOE	133

List of tables

Table		Page
1	A list of chemokine receptors, their functions and tumour expression	4
2	A selection of multi-targeting Tyrosine kinase inhibitors	20
3	R^2 of correlation plots between IC_{50} and various docking scores at the CCR7 transmembrane site.	123
4	R^2 of correlation plots between IC_{50} and various docking scores at the CCR7 intracellular site.	123
5	Relative potencies of molecules using calcium flux assay	130

Acknowledgement

Dr Kamyar Afarinkia and Dr Victoria Vinader performed the MTT assay and the calcium flux assay in this project.

Dr Mohaned Saad Abdalla Ahmed synthesised the 34 CCR7 antagonists used in the molecular docking study of this project, and performed the calcium flux assay on these compounds.

1.0 Introduction

1.1 What is metastasis?

Metastasis is one of the hallmarks of cancer. It is the process of cancer cells migrating from the original site, and growing into new tumours in other parts of the body [1][2]. Metastasis is a complex process involving a number of distinct steps. The first of these steps involves tumour cells acquiring invasive properties through the down regulation of cell adhesion molecules, gaining cell motility and breaching the basement membrane. At this stage tumour cells become malignant and are able to intravasate into either lymphatic or blood vessels. The cancer cells are then transported through the circulation in the blood or lymphatic vessels around the body. Eventually, circulating cancer cells that are able to survive in the blood vessels adhere to the micro vessels of distant organs and extravasate, forming a micro metastasis and subsequently form a new cancer cell colony [2] (**Figure 1**).

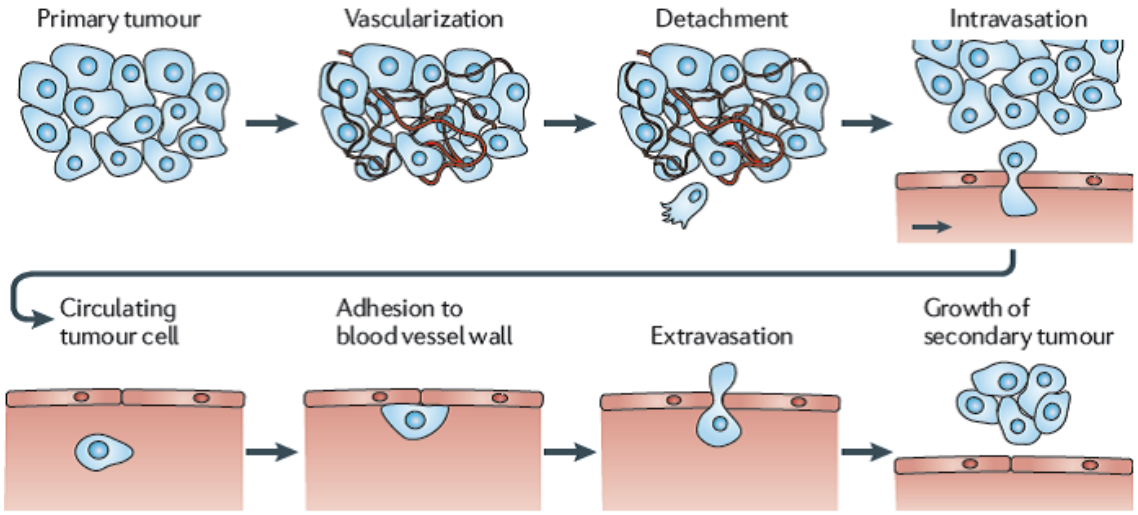


Figure 1: *The metastatic process. The primary tumour grows new blood vessels. Cancer cells detached from the primary tumour and enter the blood vessels and circulate in the vascular system. Some of these circulating cancer cells adhere to blood vessel walls and are able to extravasate and migrate into the local tissue, thus forming a secondary tumour [2].*

Metastasis has serious consequences as the newly established tumour will disrupt the function of the colonized organs. In addition, having multiple tumours throughout the body makes the disease more difficult to treat. In fact, the most frequent cause of death for cancer patients is the complication arising from metastasis, rather than the primary tumour itself [3]. Therefore, preventing and suppressing metastasis should reduce the lethality of the disease and increase

the survival rate of patients. However, whilst many drugs treat primary tumours, there are no drugs that target metastasis. Hence, there is a therapeutic need for anti-metastasis agents [2].

1.2 Chemokines and chemokine receptors

Chemokines are a group of 49 chemotactic cytokines whose main functions are in mediating immune responses and inflammatory processes, mostly through the regulation of leukocyte trafficking [34]. Chemokines are categorized into 4 subgroups; XCL, CCL, CXCL, and CX3CL. This is based on the spacing of the cysteine residue found near the NH_2 terminus of the chemokine protein (**Figure 2**).

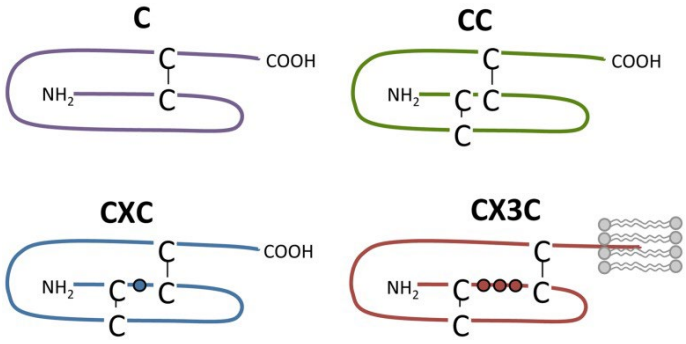


Figure 2: *The four sub-groups of chemokines, C, CC, CXC and CX3C [5].*

Chemokines can also be categorized based on their biological role as homeostatic or inflammatory, although some fall into both categories. Homeostatic chemokines are constitutively expressed. They are generally involved in lymphocyte trafficking, immune surveillance and co-localisation of lymphocytes with antigen-presenting cells (such as dendritic cells) in the lymphatic compartments. Inflammatory chemokines, however, are only expressed in the response to inflammation, tissue damage or other physiological abnormalities.

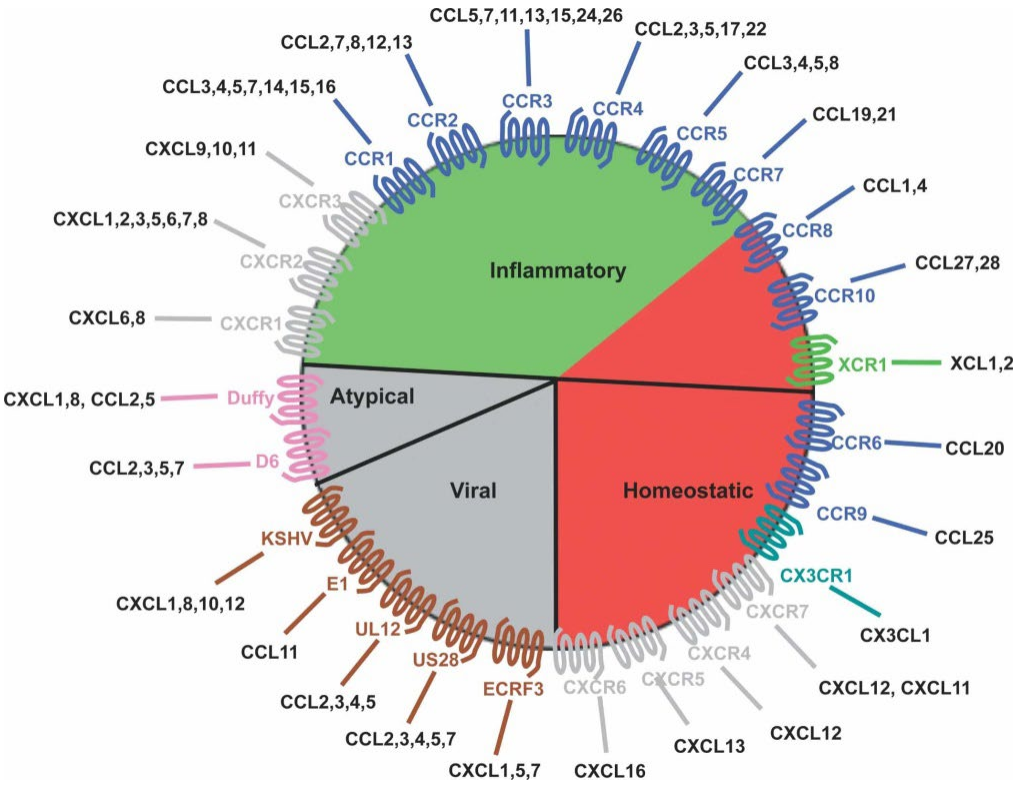


Figure 3: *Classification of chemokines and their receptors in a "chemokine wheel" (Balkwill, 2012). This cartoon shows major chemokine receptors and the chemokine ligands. The 'inflammatory' chemokines are inducible and involved in all aspects of the immune response. The 'homeostatic' chemokines are involved in development and normal physiology. Atypical chemokine receptors are generally silent and act as negative regulators of the systems, by 'mopping up' chemokines. Viral chemokines and receptors allow the pathogens to modulate immune responses to viral infection. Chemokine receptor CXCR7 is now shown to be an atypical chemokine receptor [6]*

Chemokines exert their biological activity by binding to their corresponding chemokine receptors. Chemokine receptors are part of the seven transmembrane G protein-coupled receptor (GPCR) superfamily and are grouped into XCR, CCR, CXCR and CX3CR subgroups based on the chemokines they primarily bind with [7]. Generally speaking, there is a significant level of promiscuity between chemokines and their receptors meaning that some chemokines bind to multiple receptors and that the same receptor is activated by different chemokines (**Figure 3**). For example, chemokine CCL2 binds receptors CCR2 and CCR4, whilst CCR2 is activated not only by CCL2 but also by CCL8 and CCL16. However, there is no cross-reactivity between receptors of one subgroup with ligands on another, meaning for example that CXCR's only bind CXCL's. **Table 1** shows a list of selected chemokine receptors and their general function and tumour expression, CCR7 and CXCR4 will be discussed separately.

Chemokine receptor	Function	Tumour expression	Ref.
CXCR1	Neutrophil migration; innate immunity; acute inflammation	Melanoma, prostate, breast, liver, lung, ovarian, colorectal, nasopharyngeal carcinoma, thyroid	8, 9, 10, 11
CXCR2	Neutrophil migration; innate immunity; acute inflammation; angiogenesis	Melanoma, liver, pancreatic, ovarian, prostate, lung, colorectal, renal cell carcinoma (RCC), glioblastoma, thyroid	8, 9, 10, 11
CXCR3	T cell migration; adaptive immunity; Th1 inflammation	Breast, thyroid, colorectal, melanoma, gastric acute lymphocytic leukaemia, colon, B-cell chronic lymphocytic leukaemia, glioblastoma	8, 11 12
CXCR5	B-cell trafficking; lymphoid development	Head and neck, lung, breast, gastric, prostate	8, 13
CXCR6	T cell migration	Prostate, lung, breast	8, 14, 15
CXCR7	Cardiac development	Breast, lung, prostate, thyroid	8, 11
CCR1	T cell and monocyte migration; innate and adaptive immunity; inflammation	Colorectal, multiple myeloma, breast	8, 16
CCR2	T cell and monocyte migration; innate and adaptive immunity; Th1 inflammation	Multiple myeloma, prostate, breast, colorectal,	8, 17
CCR3	Eosinophil, basophil, and T cell migration; allergic inflammation	RCC, thyroid, glioblastoma, cutaneous T-cell lymphoma (CTCL)	8, 11
CCR4	T cell and monocyte migration; allergic inflammation	Breast, adult T-cell leukaemia/lymphoma, CTCL	8
CCR5	T cell and monocyte migration; innate and adaptive immunity	Breast, gastric, head and neck, melanoma, prostate, glioblastoma, oesophageal, multiple myeloma, pancreatic, lymphocytic leukaemia, Hodgkin lymphoma	8, 18

CCR6	Dendritic cell migration	Colorectal, pancreatic, multiple myeloma, thyroid, prostate, lung, gastric, ovarian, RCC	8, 11, 19
CCR8	T cell trafficking	Kaposi sarcoma	8
CCR9	T cell homing to gut	Melanoma, breast, ovarian, prostate, lung, liver, colon, nasopharyngeal carcinoma,	8, 20
CCR10	T cell homing to skin	Melanoma, breast	8, 21
CX3CR1	T cell and NK cell trafficking and adhesion; innate and adaptive immunity; Th1 inflammation	Neuroblastoma, prostate	8

Table 1: *A list of chemokine receptors, their functions and tumour expression.*

As stated above, chemokine systems (comprising chemokines and their receptors) regulate the trafficking of leukocytes to the sites of inflammation [7]. This works because chemokine(s) are released at the sites of inflammation and diffuse to the surrounding tissue and to general circulation. The leukocytes, which express the cognate chemokine receptor(s), then migrate against the chemokine gradient and travel towards the site of inflammation. Two of the most important chemokine axes are CXCR4 (comprising of receptor CXCR4 and its cognate ligand CXCL12) and CCR7 (comprising of receptor CCR7 and its cognate ligands CCL19 and CCL21).

C-C chemokine receptor type 7 (CCR7) is involved in the homing and trafficking of T cells and mature dendritic cells to the lymph nodes during immune and inflammatory response. CCR7 is expressed on naïve, central memory and regulatory T cells, B cells, natural killer (NK) cells, and mature dendritic cells. CCR7 controls their migration to secondary lymphoid organs such as the lymph node [7]. The chemokines CCL21 and CCL19 are the only known chemokine ligands for CCR7 [7]. CCL21 and CCL19 are abundantly expressed by lymphatic endothelial cells and the paracortical regions in the lymph node, which supports its important role in the homing of cells into lymph nodes [22].

The reason there are two different chemokine ligands for CCR7 is not yet fully known. The two chemokines are structurally dissimilar in two respects. One is

that CCL21 has a longer chain containing a heparansulfate binding domain. This allows the chemokines to non-covalently bind on certain endothelial cells lining the entrance to secondary lymphoid organs rather than be soluble, and thus acts as a niche to attract T cells. The second is that the N-terminus, which is responsible for binding inside the receptor and causing conformational changes that lead to G-protein signalling, is quite different between CCL19 and CCL21. This can mean that each chemokine ligand can induce a distinct receptor conformation change in CCR7. This can explain why both CCL19 and CCL21 can activate the same signalling pathways, they can also each stimulate pathways that are unique to just one of the chemokines. For example, CCL19 is able to efficiently desensitize CCR7 by inducing β-arrestin-dependent ERK1/2 activation but CCL21 cannot [23].

C-X-C chemokine receptor type 4 (CXCR4) is one of the more researched chemokine receptors due to its extensive and varied roles in the biology of the cell. The CXCR4/CXCL12 axis is necessary for cell migration during embryonic haematopoiesis, organogenesis, vascularization and organ homeostasis. Like CCR7, it is involved in lymphocyte trafficking during inflammation and for tissue repair or regeneration.

CXCR4 is highly expressed in a variety of colourless blood cell types, including lymphocytes, endothelial, epithelial and hematopoietic stem cells. CXCL12, which is expressed in lymph nodes, lung, liver and bone marrow, is the only known chemokine ligand for CXCR4. The cytokine Macrophage Migration Inhibitory Factor (MIF) is another ligand that binds to CXCR4. The MIF-CXCR4 axis is able to promote T cell chemotaxis and regulate endothelial progenitor cell migration [24]. CXCL12 and MIF are both promiscuous ligands, CXCL12 is able to bind to CXCR7, and MIF is able to bind to CCR2 and CD74 [25]. CXCR4 is the first chemokine receptor for which a crystal structure was reported [26].

The similarity between leukocyte trafficking and the metastatic process, prompted the speculation that chemokines may also play a role in cancer metastasis. Muller et al. [27] hypothesized that the chemokine axes are misused during metastasis and this can explain site-specificity of metastasis. It is already known that the metastatic process is highly organ specific. For example, breast cancer most commonly metastasises to brain, bone, liver and lungs [28]. Muller et al. found

that the chemokine receptor CXCR4 is highly expressed in human breast cancer cells, whereas CXCL12, which is the chemokine ligand for this receptor is particularly highly expressed in brain, bone, liver and lungs which are the common sites of metastasis for breast cancer [27]. Muller et al. also suggested that the CCR7 uniquely controls metastasis to the lymph nodes since its two chemokine ligands, CCL19 and CCL21 are abundant in the secondary lymphoid organs.

1.3 CCR7 and CXCR4 in cancer

As described above, it has been hypothesised that cancer cells gain the expression of chemokine receptors in order to enable them to migrate away from the tumour body. Indeed, both CCR7 and CXCR4 are individually reported to be expressed in patients in many different cancers, including breast [27], colorectal carcinoma [29][30], oesophageal squamous cell carcinoma [31][32], gastric carcinoma [33][34], non-small cell lung cancer [35][36], melanoma [37][30] and mantle cell lymphoma [38]. Kodama et al. [39] have shown that the disease-free survival and overall survival rates of cervical cancer patients (n=174) exhibiting both CXCR4 and CCR7 expression were significantly lower compared to those with CCR7 or CXCR4 negative tumours (**Figure 4**), suggesting that the two axes may work in tandem in promoting metastasis. Individual studies of gene knockdown of CCR7 and CXCR4 in cancer cells have both demonstrated inhibition of migration [40-42].

More recently, it has been shown that both of these chemokine axes have wider roles in cancer by promoting tumour growth and angiogenesis as well as metastasis and that they often share common downstream pathways that regulate these functions. For example, both the ERK1/2 and Akt pathways are known for promoting cancer cell survival, proliferation, angiogenesis, invasion and metastasis [43] and are shown to be on the downstream pathway of CCR7 and CXCR4 in different cancer cell types [24,44]. Studies have shown that CCR7 activation promotes human NSCLC proliferation by increasing G2/M phase progression which is due to the upregulation of cyclin A, cyclin B1 and CDk1 via the ERK pathway [35]. CXCR4 activation can promote proliferation in pancreatic cancer cells by increasing GSK3B expression via the Akt pathway [45].

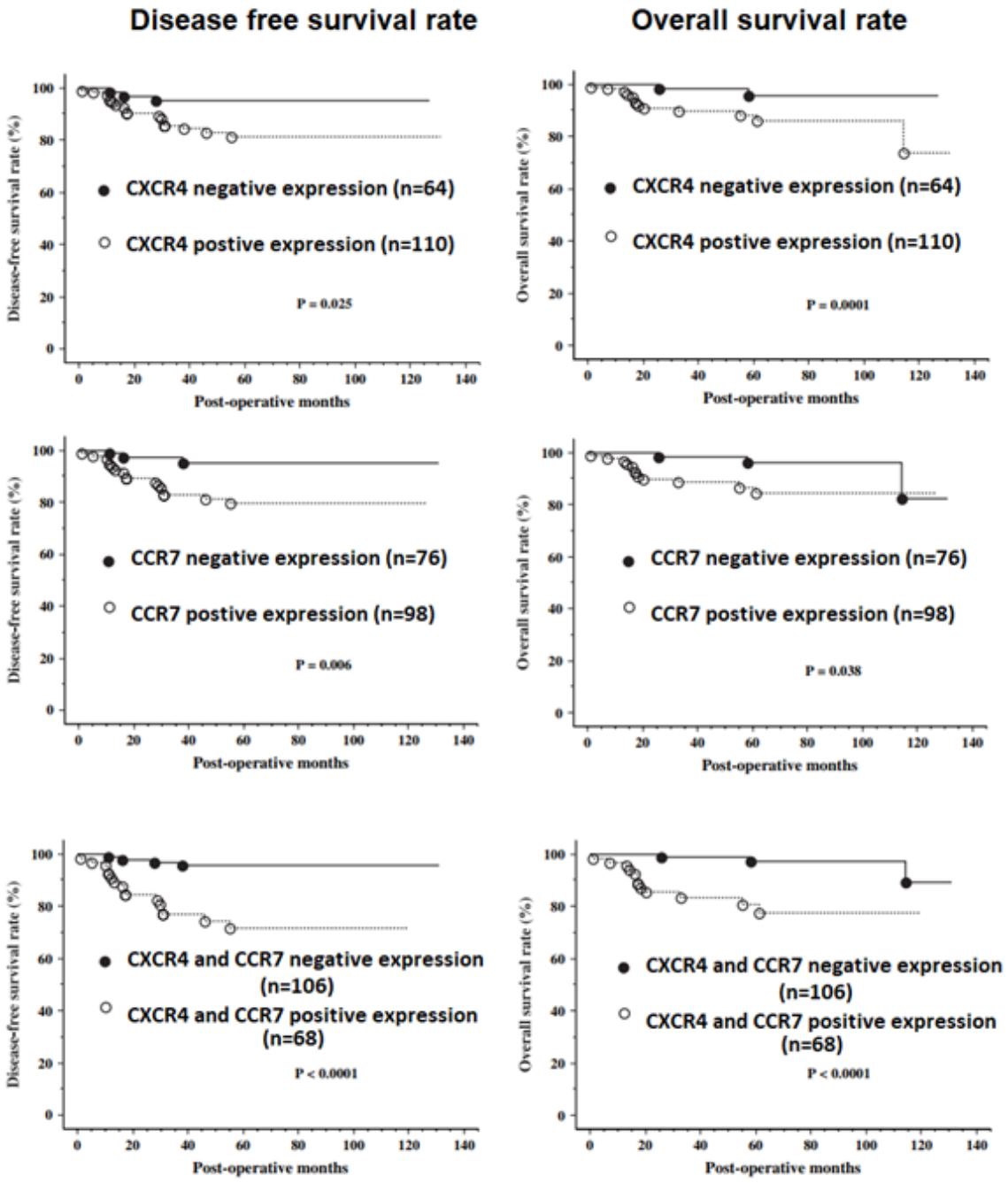


Figure 4: *Disease free survival rate and overall survival rate of cervical cancer patients. [40]*

CCR7 activation can prevent the apoptosis of human NSCLC by the upregulation of anti-apoptotic protein BCL-2 and downregulation of the pro-apoptotic proteins bax and caspase-3, but not p53 [35]. In breast cancer, CCR7 activation is able to increase the expression of VEGF-C to promote lymphangiogenesis via the PI3K/Akt pathway [46]. Lymphangiogenesis is the sprouting of new lymphatic vessels and they provide channels for tumour cells to migrate to lymph nodes, VEGF-C is known to promote lymphangiogenesis [47]. The CXCL12-CXCR4 pathway was also able to promote lymphangiogenesis but independent of the VEGF-C/VEGFR-3 pathway. Wei et al. [48] have shown that the CXCR4 antagonist AMD3100 and a CXCR4 antibody were able to significantly reduce CXCL12 induced lymphangiogenesis in a Matrigel plug assay. A VEGFR-3 antibody was unable to significantly reduce CXCL12 induced lymphangiogenesis, but in the absence of CXCL12, a VEGFR-3 antibody was able to significantly reduce VEGF-C induced lymphangiogenesis. The assay involved injecting Matrigel containing CXCL12, VEGF-C, antibodies or AMD3100 subcutaneously into the abdomen of immunodeficient mouse. After 8 days, the plugs containing lymphatic vessels were dissected, then subjected to immunofluorescent analysis to assess the degree of lymphatic vessel formation [48].

Epithelial mesenchymal transition (EMT) is the process of carcinoma cells acquiring the characteristic of mesenchymal cells. It is one of the key steps in metastasis and it allows carcinoma to gain invasive characteristics. There are studies which indicate both CCR7/CCL21 and CXCR4/CXCL12 interaction are involved in EMT and promote motility in cancer cells [27, 49, 50]. Both chemotaxis axes are able to down regulate the cell-cell adhesion molecule E-cadherin [49]. The loss of adhesion molecule like E-cadherin is associated with EMT [51]. Both interactions also up-regulate vimentin, [49] N-cadherin and Slug. [52] Vimentin and N-cadherin are both markers of EMT [49, 52]. The upregulation of Slug was reported to repress E-cadherin expression and induce EMT [52]. The CXCL12/CXCR4 axis was reported to activate the canonical Wnt/β-catenin pathway in colorectal cancer. The activation of this pathway increases free cytoplasmic β-catenin, which leads to EMT and subsequent cancer cell invasion and metastasis [53].

The CCR7/CCL21 axis was reported to upregulate matrix metalloproteinase-9 (MMP-9) in B-cell chronic lymphocytic leukaemia (BCLL) cells via the ERK1/2 pathway [54]. The upregulation of MMP-9 has a huge implication in cancer due

to its role in invasion and metastasis in various cancers. MMP-9 is able to break down the extracellular matrix, thus enabling tumour cells to invade and spread [55]. Lin et al. [55] found that treating cells from a human epithelial carcinoma cell line with MMP-9 led to the loss of intact E-cadherin while inducing EMT-like morphological changes. The knockout of CCR7 did not reduce the expression of MMP-9 but it was able to significantly inhibit the CCL21 induced upregulation of MMP-9 [54]. The CXCR4/CXCL12 axis increasing MMP-9 and MMP-2 secretion was reported in endometrial cancer [56]. MMP-2 is capable breaking down extracellular matrix just like MMP-9 [57].

Both CCL21/CCR7 and CXCL12/CXCR4 interactions are able to trigger intracellular actin polymerization and the formation of pseudopodia in cancer cells [27]. Actin polymerization is a prerequisite process for motility and migration in cells [58]. High levels of actin polymerization are required for the formation of pseudopodia, which is the process of temporary cytoplasmic extension of cell membranes. Formation of pseudopodia is necessary for cancer cells to gain motility for the invasion and metastasis [59].

The process of how CCR7 expressing tumour cells access the lymphatic vessels and travel to the lymph node has been studied. Shield et al. have found a CCR7 dependent mechanism best described as “autologous” chemotaxis [60]. Four human cancer cell lines (three breast cancers, one melanoma) were found to be able to secret CCL21 and CCL19 to create an autocrine gradient towards the lymphatic vessels under the influence of interstitial flow. The chemokine gradient then is able to guide the cancer cells which express CCR7 to enter the lymphatic vessels [60]. This autologous chemotaxis was observed in human glioma and hepatocellular carcinoma cell lines which express CXCR4. For the glioma cell line, the mechanism is CXCR4 dependent, as CXCR4 inhibition was able to abrogate the interstitial fluid flow promoted invasion. For the hepatocellular carcinoma cell line, it was not entirely CXCR4 dependent, as CXCR4 inhibition significantly reduced the cancer cell invasion but did not eliminate it [61].

1.4 Known CCR7 and CXCR4 antagonists

CCR7 antagonists

Currently there are no clinically approved CCR7 antagonists. In 2018, two CCR7 antagonists have been reported in the literature, cosalane (**1**) [62] and cmp2105 (**2**) [63].

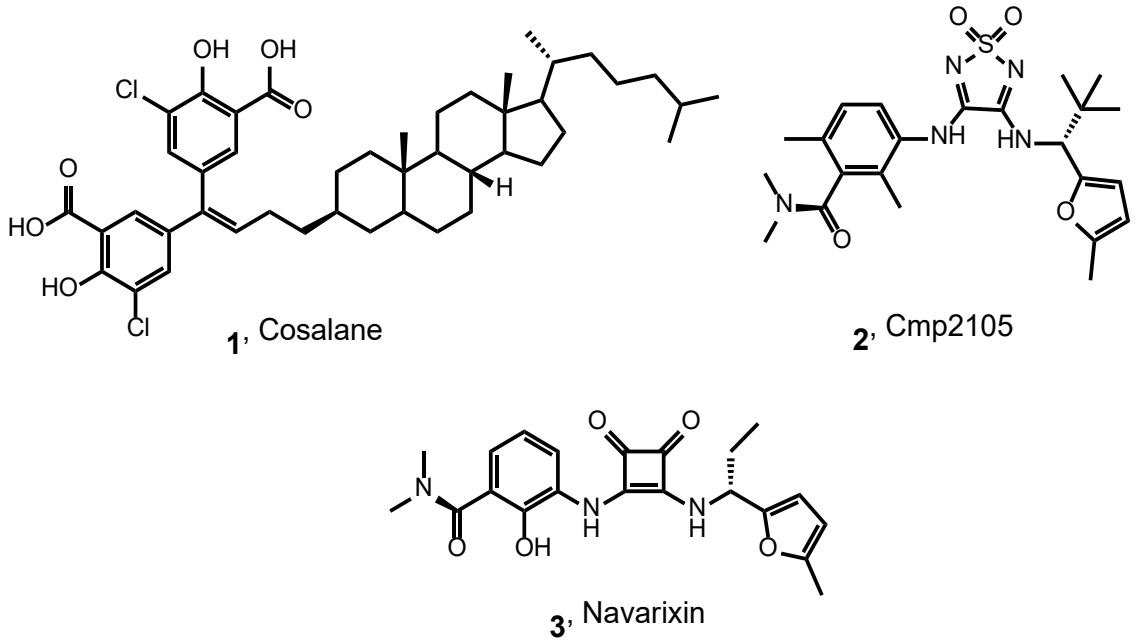


Figure 5*: Structure of Cosalane, Cmp2105 and Navarixin*

Cosalane is an aurintricarboxylic acid derivative containing cholestane. Cosalane was developed as human immunodeficiency virus therapeutic [64], however, it was never used clinically, mainly due to very poor oral bioavailability, and it accumulates in high concentration in the liver after intravenous administration [65]. High throughput screening by Hull-Ryde et al. [62] revealed the compound to be a CCR7 antagonist. A study suggests cosalane does not bind to CCR7, the molecule instead imbeds itself into the cell membrane, subsequently affecting CCR7 activation. Cosalane was able to impair CCL19 or CCL21 induced CCR7 dependent chemotaxis in a migration assay using human H9 cells, a cutaneous T-cell lymphoma cell line [62].

Jaeger et al. [63] published the X-ray crystal structure of CCR7, 6QZH. Cmp2105 was bound at the intracellular allosteric site of CCR7, thus showing Cmp2105 is an intracellular allosteric ligand of CCR7. Cmp2105 consists of a thiadiazole-dioxide core motif with two amine-linked substituents. Jaeger et al. searched for

compounds structurally similar to Cmp2105 and found Navarixin (**3**), a potent CXCR1 and CXCR2 antagonist [66]. *In vitro* assays by Jaeger et al. suggest that Cmp2105 and Navarixin are CCR7 antagonists. Both Cmp2105 and Navarixin are able suppress arrestin binding in response to CCL19 induced CCR7 activation. Cmp2105 was shown to be more potent than Navarixin, with an IC_{50} of 7.3 μM compared to IC_{50} of 33.9 μM [63]. The structure of Navarixin is similar to Cmp2105, it has cyclobutene-dione in the center of the molecule instead of thiadiazole-dioxide, it has almost identical two amine-linked substituents.

Data on CCR7 inhibition in cancer are based on knockdown studies *in vivo* and *in vitro* [40, 41], and the use of inhibitory monoclonal antibodies *in vitro* [41]. *In vivo* studies in which the animals were inoculated by cancer cells transfected with CCR7, have shown that increased CCR7 expression correlates with tumourgenesis and lymph node metastasis [67, 68]. So far, few studies have examined the effects of CCR7 inhibition *in vivo*. In one example, mice with xenografted human mantle cell lymphoma were treated with a CCR7 inhibiting monoclonal antibody. [38] The antibody was able to cause a significant delay to the tumour growth rate, and the infiltration of tumour cells to the lymphoid and distant organs was reduced. The survival of the mice was significantly increased. No side effects were observed in the tumour bearing mice. However, as the mice used in this and other studies are immunodeficient, signs of immunodeficiency from the CCR7 inhibition would not have been observed [38]. These studies show that CCR7 is a target worth pursuing, with the caveat that any potential immunological effects should be monitored. A small molecule CCR7 antagonist will be a valuable tool for further elucidating the effects of CCR7 antagonism in cancer. In addition, it could potentially be useful in treating the metastasis of CCR7 positive tumours.

CXCR4 antagonists

Currently the most prominent CXCR4 small molecule antagonist is AMD3100 (**4**) (**Figure 6**), an FDA approved drug used in combination with granulocyte-colony stimulating factor to mobilize hematopoietic stem cells from bone marrow for collection prior to radiotherapy, and subsequent autologous transplantation in non-Hodgkin's lymphoma and multiple myeloma [69]. AMD3100 was discovered by accident as part of a mass screening campaign to discover anti-HIV agents [70]. AMD3100 has very poor oral bioavailability and in clinical practice is directly injected into patient's bone to avoid that problem [69]. Nonetheless, AMD3100 was shown to inhibit lung metastasis and improve survival in a mouse oral squamous cell carcinoma xenograft model [69-71].

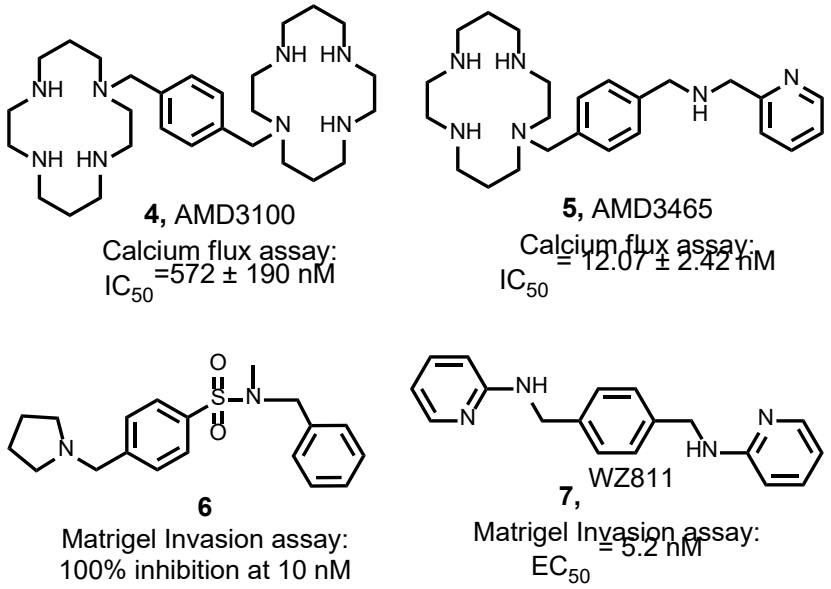


Figure 6: *Chemical structures of a selection of CXCR4 antagonists*

Fricker et al. [72] used the calcium flux assay to demonstrate that AMD3100 did not inhibit calcium flux against cells expressing CCR7 and other chemokine receptors when stimulated with their respective chemokines, indicating that AMD3100 is a selective CXCR4 antagonist.

AMD3100 features two cyclam rings connected by *para*-xylylenediamine linker. AMD3100 derivatives that possess only a single cyclam ring have been investigated. Among these compounds, AMD3465 is one of the more prominent examples of this series. AMD3465 demonstrated that a single cyclam ring is also able to induce potent selective CXCR4 inhibition (IC_{50} = 12.07 ± 2.42 nM in calcium flux assay) [73]. AMD3465 (**5**) was reported to inhibit breast cancer

growth and metastasis *in vivo* [74]. As was the case with AMD3100, AMD3465 does not inhibit chemokine stimulated calcium flux in cells expressing CCR7 and a few other chemokine receptors [74].

Whilst the cyclam ring appears to be an important functional group for selective CXCR4 inhibition, it has been shown that smaller nitrogen containing aromatic heterocycles that mimic features of the cyclams can substitute these macrocycles. Non-cyclam derivatives of AMD3100 had been developed as CXCR4 antagonists (**Figure 6**).

The diamine in these compounds initially aimed to mimic the three-carbon chain spacing in the cyclam ring. Many of these non-macrocycle diamine compounds with various substituted aromatic heterocycles were synthesized [75]. WZ811 (**7**) is one of the notable examples of these compounds. It was reported to be 8-10 times more potent than AMD3100 in CXCR4 induced Matrigel invasion [76]. It features two pyrimidin-2-amine moieties instead of two cyclam rings. The central *para*-xylylenediamine that was carried from AMD3100 was found to be important for high CXCR4 binding affinity. Attempts to modify the central benzene ring have mostly resulted in reduced binding affinity [77]. WZ811 being a CXCR4 antagonist might be questionable. In a different study, WZ811 was tested in various *in vitro* studies and shown to be not a CXCR4 antagonist, or a very weak one [77]. In another study, WZ811 is able to inhibit cell proliferation, cell motility and colony formation of two chronic lymphocytic leukaemia (CLL) cell lines, both of which have high CXCR4 expression. Additionally, in a xenograft model with human CLL, WZ811 exhibited tumour growth inhibition [78].

AMD11070 (**8, figure 7**) (AMD070 in some literature) is part of the tetrahydroquinoline based CXCR4 antagonists developed by Skerlj et el. [79]. Some of the more potent compounds of this type have nanomolar activity in calcium flux assay. AMD11070 has an IC_{50} of 12 nM, it is a selective CXCR4 antagonist and not active in CCR7 and a number of other chemokine receptors. The general key features of this series of compounds include a tetrahydroquinoline, benzoimidazole and a nitrogen containing side chain, all of which connect to a methyl amine (**Figure 7**). The *S*-enantiomer of AMD11070 is far more active than the *R*-enantiomer. Many derivatives have been made for this class of compound and they generally don't deviate from this structure. The

structure of this series of compounds appear to be unrelated to the cyclam ring CXCR4 antagonists, but the inclusion of the tetrahydroquinoline came from SAR studies that aimed to alter the cyclam rings of the early AMD series (**Figure 7**), the research finally leads to compounds without macrocycles. The CXCR4 antagonists developed by Skerlj et al. were mostly evaluated for anti-HIV activity and not directly for their CXCR4 inhibition. However, for some of their compounds that were tested in both CXCR4 inhibition and anti-HIV activity, potent anti-HIV activity appears to correlate with potent CXCR4 inhibition (calcium flux assay). Compound **9** is analogous to AMD3465 with a tetrahydroquinoline instead of the cyclam ring. Its anti-HIV activity indicates that compound **9** is likely to be a less potent CXCR4 antagonist compare to AMD3465. This indicates that tetrahydroquinoline is not an optimal substitute for the cyclam ring even though it gave good results in this series of CXCR4 antagonist. The addition of benzoimidazole to compound **9** resulted in **10**, which increased its anti-HIV activity and presumably its activity on CXCR4 inhibition too. Lastly, an aliphatic amine side chain replaced the pyridin-2-ylmethanamine from AMD3465, along with the central benzene ring that have been inherited since AMD3100 [79].

To summarize, nitrogen containing heterocycles appear to be important moieties for CXCR4 inhibition. The cyclam ring is an important functional group, but simpler heterocycles that can mimic the binding interaction of cyclam have demonstrated their effectiveness in CXCR4 inhibition as seen in WZ811. Lastly, the heterocyclic moiety that gave good results in one series of CXCR4 antagonists might not translate into good results in another series, as seen in compound **9.**

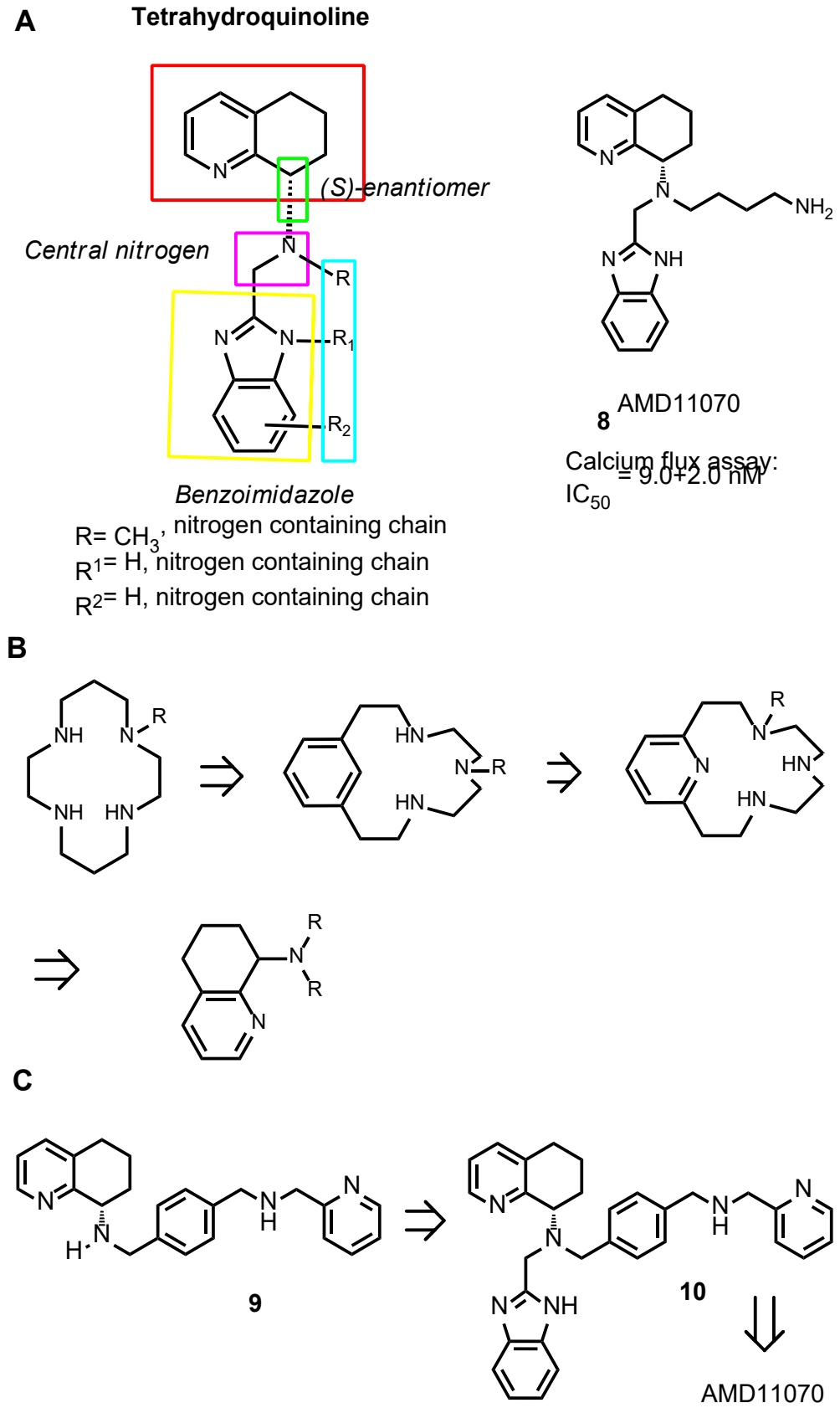


Figure 7: *A) structure of AMD11070 and the general structure of the tetrahydroquinoline based CXCR4 antagonists. B) Redesign evolution leading to compounds without macrocycle functional groups. C) Design progression of AMD11070, from an AMD3100 derivative to a new series of compounds.*

More recently, Mooring et al, reported biaryl sulfonamides of general structure **6** as potential CXCR4 antagonists [80] Their most potent compound was compound **6** (**Figure 6**). Compound **6** has 100% inhibition at 10 nM in a Matrigel invasion assay involving human breast cancer cells. The general structure of this series of CXCR4 antagonists is mostly similar to compound **6**. **Figure 8** highlights the regions (A and B) modified for gathering initial SAR data. The notable trend is that having an aromatic ring at region A reduces invasion inhibition, with the same observed with having substituents on the benzyl ring at region B.

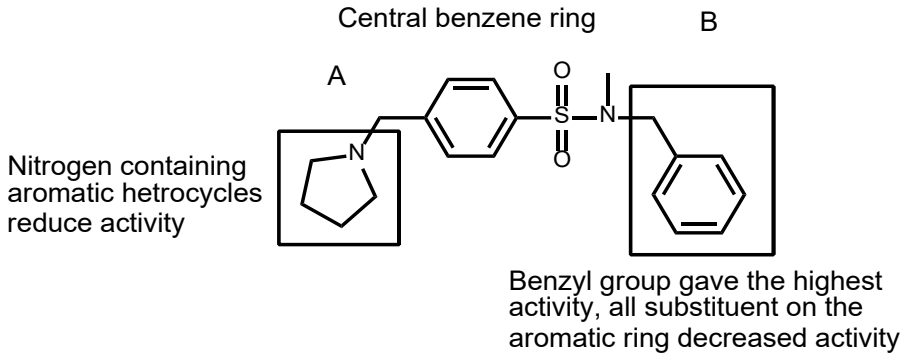


Figure 8: *General structure of the sulfonamide CXCR4 antagonists developed by Mooring et al. [80]*

It should be noted that this series of sulfonamides were originally analogues of WZ811 (**Figure 6**). The central benzene ring from AMD3100 derivatives was kept due its importance in binding affinity. The pyridin-2-ylmethanamine group from AMD3465 was included in one of the analogues of **6**, which only has 48% inhibition at 10 nM in CXCR4 induced Matrigel invasion assay [80]. It is likely that this sulfonamide series of CXCR4 antagonist have a very different binding mode compared to the non-macrocycle AMD3100 derivatives mentioned previously.

1.5 Rationale for a CCR7 and CXCR4 small molecule dual antagonist

There are significant overlaps between the mechanisms of invasion and metastasis utilizing CCR7 and CXCR4. For cancer cells that express both CCR7 and CXCR4, the inhibition of only one of the two receptors might not be sufficient in stopping the metastasis. If only one is inhibited, the remaining uninhibited receptor is likely to carry on inducing the cancer cells to metastasize; therefore, inhibiting both receptors simultaneously might be more beneficial than only inhibiting only one. In addition, as mentioned earlier, the disease-free survival and overall survival rate of cervical cancer patients with both CXCR4 and CCR7 expression were significantly lower compared to those with CCR7 or CXCR4 negative tumours.

The concept of dual antagonism of chemokine receptors has already been recognised in a variety of diseases and there are now a number of examples where antagonism of more than one chemokine receptor is considered advantageous as a treatment strategy [81-84]. For example, for the treatment of rheumatoid arthritis, CCR2 and CCR5 are expressed on different cells in a complementary manner, hence dual targeting of CCR2 and CCR5 is thought to be more advantageous than targeting only one of the two targets [81].

The ways of inhibiting both receptors simultaneously would be either with a drug combination of CCR7 and CXCR4 antagonists or a dual CCR7 and CXCR4 antagonist. There are advantages of a dual antagonist over a drug combination; Optimizing the dosage and sequences of a drug combination is a difficult task. Also, developing a single potent dual antagonist might be cheaper than developing two potent single target antagonists. The down side of attempting the development of a single potent and selective dual antagonist is that it likely to be significantly more difficult or might not be plausible.

The chemokine receptors are structurally similar. The selectivity of the CCR7 antagonists against other chemokine receptors has not been tested. It is possible that the CCR7 antagonists from this group or any potential CCR7 and CXCR4 antagonists also bind to other chemokine receptors. A number of other chemokine receptors were reported to have roles in cancer. For example, in breast cancer, other than CCR7 and CXCR4; CCR10 [24] and CXCR7 [85] are also associated with metastasis (see **Table 1** for a list of chemokine receptors

and their expression in cancer). Inhibiting off target chemokine receptors might be beneficial but at the same time increasing the risk of off target toxicity.

2.0 Developing a small molecule dual antagonist

2.1 Shifting to multi-target drugs from selective single target drugs

The development of selective, single target agents has been the prominent approach in cancer drug development for many years [84]. The goal of this approach has been to provide effective cancer treatment without severe toxicity through selective modulation of a single target. However, nearly all cancer types are polygenic and involve dysregulation in multiple pathways which lead to cancer cell proliferation and survival [84]. Single target drugs may not be able to deal with these different pathways. Furthermore, drug resistance remains a challenge in cancer. To combat these resistance mechanisms, it may be necessary to modulate multiple targets in cancer treatment. In fact, combinations of drugs with different mechanisms of action have long been in use to combat drug resistance. Indeed, many clinically useful cancer drugs, like the tyrosine kinase inhibitors, can be considered multi-targeting drugs. The potency of these drugs on “relevant off targets” probably contributes to their effectiveness, showing the importance of blocking multiple targets and pathways [84]. To illustrate this, **Table 2** shows a number of FDA approved tyrosine kinase inhibitors and their “multiple” targets [85].

Tyrosine kinase inhibitor	Tyrosine kinase targets	Cancer type
Sunitinib	VEGFR type 1 and 2, PDGFR-α and PDGFR-β, KIT, FLT3, RET, CSF1R	Gastrointestinal stromal tumours
Sorafenib	B-Raf, VEGFR2/3, PDGFR, Flt3, KIT, RAF, RET	Advance renal cell cancer
Pazopanib	VEGFR1/2/3, PDGFR, KIT	Renal cell cancer
Lapatinib	ErbB1 (Her-1), ErbB2 (Her-2)	Breast
Imatinib	Bcr-Abl, KIT, PDGFRβ	Chronic myeloid leukaemia Gastrointestinal stromal tumours

Table 2: *A selection of multi-targeting Tyrosine kinase inhibitors [85]*

In fact, many supposedly “single” target drugs are likely to affect multiple targets. According to Boran et al. [86], data analysis from drugbank (a database which generally lists the intended therapeutic target of a drug) and WOMBAT (a database which compiles the affinities of drugs for any protein target published in literature), revealed that the average number of biological targets per drug is 2.7. If the data set is extended by adding hypothetical drug targets predicted *in silico*, the number jumps to 6.3 targets per drug [86]. In recent years many researchers have begun to look into a polypharmacological approach, as an alternative to single targeting in drug development. Polypharmacology does not have a fixed definition but it is commonly referred to as the effect or pharmacological consequences of a single drug binding to multiple targets. The polypharmacology approach takes into consideration that drugs are generally likely to be promiscuous and aims to develop a single drug to modulate multiple relevant targets for a therapeutic effect. For this polypharmacology approach, researchers have the challenge of rationally developing a mutli-target drug from early stages of drug development. This approach utilises screening techniques to find a lead compound that shows activity at two (or more) different targets simultaneously, alternatively, a complimentary approach involves the modification of an existing single target compound in an attempt to promote its potent activity on another target. This is of course not an easy task; the idea is not to completely eradicate

selectivity, which would yield a useless drug molecule, but to “selectively” affect more than one biological target.

The development of a dual antagonist is not a novel idea. However, the challenges of the approach mean that is not necessarily widely pursued. In fact, many existing multi-targeted drugs, like the tyrosine kinase inhibitors mentioned earlier, are discovered serendipitously rather than being rationally developed as multi targeting agents [87].

2.2 Lead generation for multi-target drugs

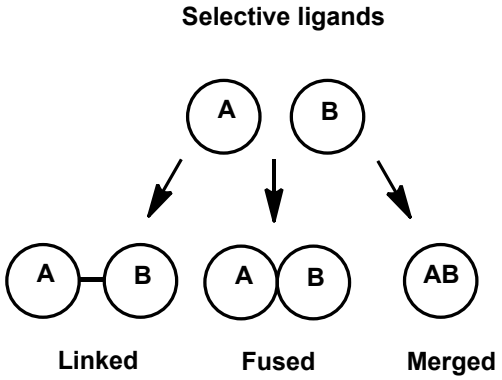


Figure 9: *The three ways of joining two drug molecules together to form a multi-targeting drug [87]*

One lead generation strategy for obtaining a dual antagonist is to create a molecule by linking two existing molecules with selectivity towards each of their own targets. [87] (**Figure 9**). The second strategy involves two molecules attached to each other *via* one of their functional groups, rather than a defined linker, in both these approaches, the two initial molecules remain clearly defined. Multi-target drugs developed by the two methods above tend to have large molecular weight and have other physiochemical properties non-compliant with the Lipinski’s rule of five. For example, they are often very lipophilic and, in many cases, they lack favourable pharmacokinetic properties. For molecules that are structurally similar, it is possible to extract structural features from both molecules and design a “hybrid” molecule with activities towards both targets. These merged multi-target molecules are more likely to comply with the rule of five and thus this approach is more desirable, yet it is clearly more challenging [88].

The discovery of many multi-target drugs is sometimes a continuation of research into single target drugs. The SAR data generated from the single target drug development largely aided the optimization of these multi-target ligands [88]. The

conversion of a potent selective single target antagonist to a potent dual antagonist is difficult. Nevertheless, there are existing cases of potent small molecule dual antagonists. How this was achieved will be examined in the following three case studies.

2.3.1 Case study: Dual aromatase steroid sulfatase inhibitors (DASI)

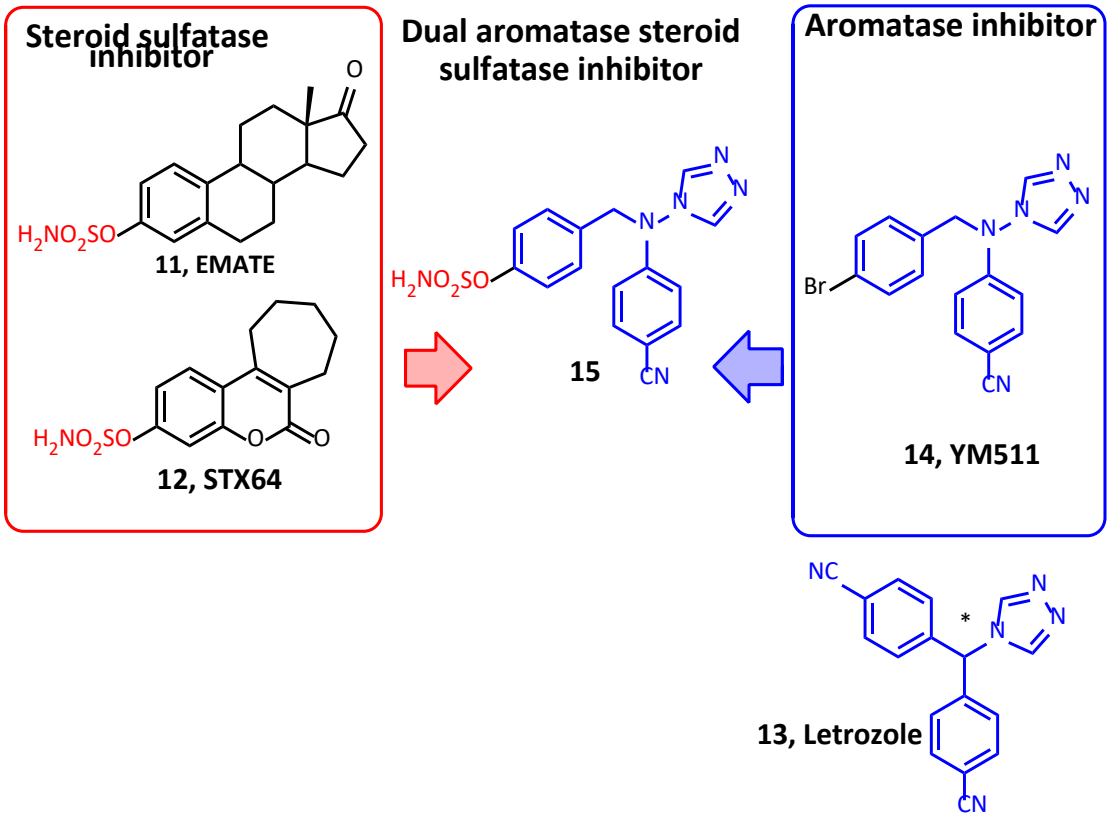


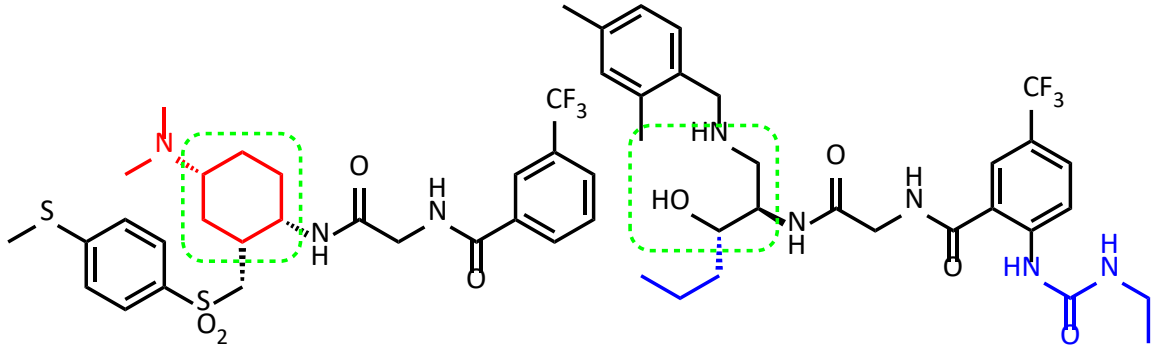
Figure 10: *The development of the most promising dual aromatase steroid sulfatase inhibitors from Woo et al. [89].*

Aromatase inhibitors inhibit the conversion of androgens to oestrogens, which is essential in hormone-dependent breast cancer. It has been established that inhibition of steroid sulfatase (STS) coupled with aromatase inhibition can further enhance the deprivation of oestrogens and other steroids with oestrogenic properties in patients. All known STS inhibitors such as EMATE (**11**), and STX64 (**12**), share a common pharmacophore, a phenyl sulfamate [89].

Woo et al. intended to develop dual aromatase steroid sulfatase inhibitors (DASI) by incorporating the phenyl sulfamate moiety onto an aromatase inhibitor. FDA approved non-steroidal aromatase inhibitors (AI) such as Letrozole (**13**) contain a prochiral centre (marked with an asterisk). Therefore, the synthesis of a dual aromatase steroid sulfatase inhibitor based on Letrozole would lead to the

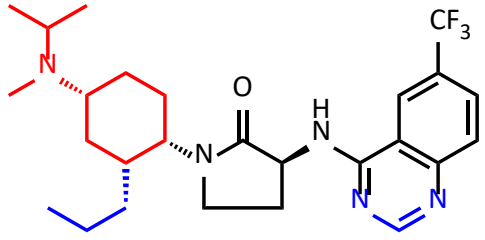
complication of having to separate enantiomers, or to develop an enantioselective synthesis. Therefore, YM511 (**14**), a compound that bears no chiral centres, was chosen as the preferred AI skeleton. In addition, the bromine substituent was thought to be a convenient position to introduce the sulfamate group present in the STS inhibitors [89]. A series of YM511 analogues which possess the sulfamate ester moiety at the *para* or *meta* positions were synthesised. Each compound was tested for the inhibition of both aromatase and STS. The most promising analogue was compound **15** with IC_{50} (aromatase) = 0.82 nM and IC_{50} (STS) = 39 nM [89].

2.3.2 Case study: the discovery of CCR2 and CCR5 dual antagonists



16 (cyclic series)
CCR2 binding: 0.3 nM;
Chemotaxis blockade - 1.9 nM
CCR5 binding: ~1000 nM

17 (acyclic series)
CCR2 binding: 3 nM;
Chemotaxis blockade - 16 nM
CCR5 binding: 86 nM



18
CCR2 binding: 0.7 nM
CCR5 binding: 2.4 nM

Figure 11: *Two CCR2 antagonists developed by Carter et al. representing two series of CCR2 antagonist. The two series were differentiated based on the ring highlighted in red; one cyclic, one acyclic. IC_{50} for receptor binding and chemotaxis blockade are displayed. [81]*

There are an increasing number of literature reports supporting the case for dual chemokine CCR2 and CCR5 receptor antagonists for the treatment of

cardiovascular and autoimmune diseases, and intimal hyperplasia [81]. For that reason, Carter et al. decided to develop a dual CCR2 and CCR5 antagonist starting from their existing work on small molecule CCR2 antagonists. Carter's group began the work of their dual antagonists from two series of cyclic (referring to the cyclohexane ring, e.g. **16**) and acyclic CCR2 antagonists developed previously by the group, e.g. **17**. Compound **17** is less potent in the reducing CCR2 driven chemotaxis of peripheral blood mononuclear cells, when compared to compound **16**, however its binding to CCR5 (IC_{50} of 86 nM), is superior to that of compound **16** (~1 μM).

Carter's group performed site directed mutagenesis studies and carried out homology modelling that informed the hypothesis that the cyclic and acyclic compounds bind to CCR2 in a very similar way. This suggested that the moieties responsible for CCR2 binding for both the cyclic and acyclic compounds, are interchangeable and should give rise to compounds with similar biological activities. To test this hypothesis, they prepared a new series of acyclic compounds that contain similar functional groups to the ones present in the cyclic compounds. The observed potency of these compounds supported their hypothesis.

From the point of view of CCR5 antagonism SAR, Carter's group knew that the propyl group from the acyclic series was responsible for the CCR5 selectivity and decided to incorporate it into the more potent cyclic series. The next series of cyclic antagonists containing the *N*-propyl substituent was indeed active for CCR2, albeit its potency was lower than the previous cyclic compound **16**. In an attempt to increase the oral bioavailability and CCR2 potency of this compound, Carter et al. referred to their previous SAR studies and decided to replace the glycinamide motif with a γ-lactam. The benzamide from the acyclic series was replaced with an anthranilic amide moiety and these changes were incorporated into the structure of the cyclic series. Lastly, quinazoline was introduced as an alternative anthranilic acid mimetic. The combination of these changes resulted in compound **18**, a potent dual CCR2 and CCR5 antagonist with IC_{50} of 0.7 nM for CCR2, and IC_{50} of 2.4 nM for CCR5 binding respectively. [81]

2.3.3 Case study: the discovery of endothelin and angiotensin II receptor dual antagonist.

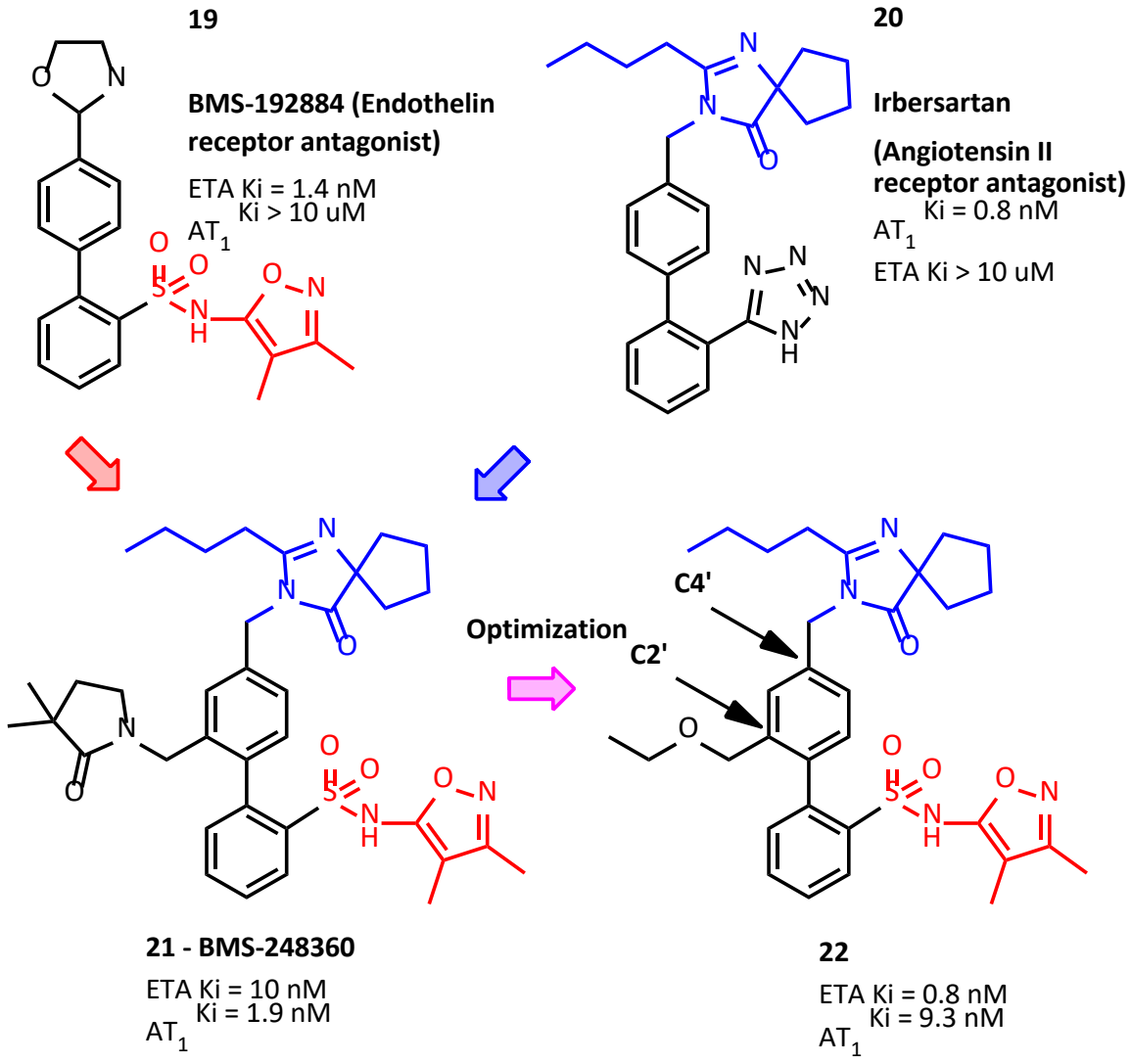


Figure 12: *Development of the endothelin and angiotensin II receptor dual antagonist* ***22*** *by Murugesan et al. BMS-248360 incorporate structural features from angiotensin II antagonist Irbesartan and endothelin subtype A antagonist BMS-193884. BMS-193884 was optimized into compound* ***22*** *[90].*

Murugesan et al. [90] developed a potent dual antagonist of endothelin subtype A (ET_A) and angiotensin II subtype 1 (AT_1) receptors (human ETA Ki=0.8 nM; rat AT1 Ki=9.3 nM) for the treatment of hypertension [90]. Murugesan et al. aimed to create the dual antagonists by merging certain structural features from the potent angiotensin II antagonist irbesartan, and their own potent ET_A antagonist BMS-193884. The resulting compound BMS-248360 (**20**) was designed based on the existing knowledge of the SAR for ET_A and AT_1 antagonists.

BMS-193884 (**19**) came from a series of ET_A antagonists by Murugesan et al. The biphenyl sulfonamide was present in the entire series of their antagonists, therefore the biphenyl sulfonamide was kept unchanged and it was chosen as the key scaffold of their dual antagonist programme. Based on the SAR data, three requirements were required to create the biphenyl sulfonamide dual antagonist. 1) The hetero aryl sulfonamide must be conserved, as is required for ET_A antagonism. 2) A hydrogen bond donating heterocycle at the C4' position of biphenyl is required to ensure good AT_1 receptor binding affinity. 3) Substitution at C2' of biphenyl group often led to an increase in ET_A receptor affinity. One of the more potent C2'groups was the (3,3-dimethyl-2-oxopyrrolidine)methyl moiety, which was incorporated into BMS-248360 (**21**). Efforts to further optimise BMS-248360 **20** by altering the substituent at C2' afforded compound **22** [90].

In these three cases, the incorporation of a certain moiety into an antagonist from another antagonist facilitated the conversion of a selective antagonist into a potent dual antagonist. In the case of the DASI, adding the sulfamate moiety to YM511 was sufficient to give the compound nanomolar activity at both targets (aromatase and steroid sulfatase). Previous knowledge of SAR for each target was crucial in the development of dual antagonist in this approach. The SAR data generated from the development of each of the single target antagonists pointed out towards which moiety was important in the binding of the drug and aided in the optimization of the dual antagonist. Sometimes this SAR information is not fully available for both targets and therefore screening methods are required.

2.4.0 Lead generation for multi-target drugs with screening methods

Screening in drug discovery fundamentally involves assessing the biological activity of a library of drug like compounds against a particular target. Compounds with activity within acceptable range will be selected as a lead compound. These methods are typically for selective single target drugs, but researchers have modified their screening methods to generate leads for multi-target drugs [87].

2.4.1 High-throughput screening

High-throughput screening utilizes automation to biologically assess a large library of compounds. The number of compounds that can be tested can go up to millions depending on the scale of the screening facility. Due to its significant cost, it is mostly used in the pharmaceutical industry and is unavailable to most academic groups, although medium-small scale throughput screening is a viable option [87].

For the purpose of finding compounds with multiple targets, one approach would be to conduct a normal high-throughput screening campaign to identify some hit compounds for one of the targets of interest. The hit compounds can then be screened against a second target of interest and so on. In this process, data analysis and validation become a major factor that dictates the success of the campaign, as false positives can lead to erroneous lead molecules [88]. However, a key advantage of this approach is that previously unknown chemotypes can be identified. Alternatively, a biased library of compounds known to have activity at one of the targets of interest can be compiled and screened against a second target of interest. Although this strategy is unlikely to afford novel chemotypes, it is designed to increase the chance of success. An example of the compounds developed through this strategy is a neurokinin 1 receptor (NK_1) and serotonin transporter (SERT) dual targeting compound for treatment of depression [87]. The high-throughput screening of NK_1 and SERT led to the identification of a molecule, **23**, with modest binding affinity towards both targets. The resulting screening hit compound had a balanced activity for both targets and further optimisation afforded greater potency for both targets, affording compound **24** [87].

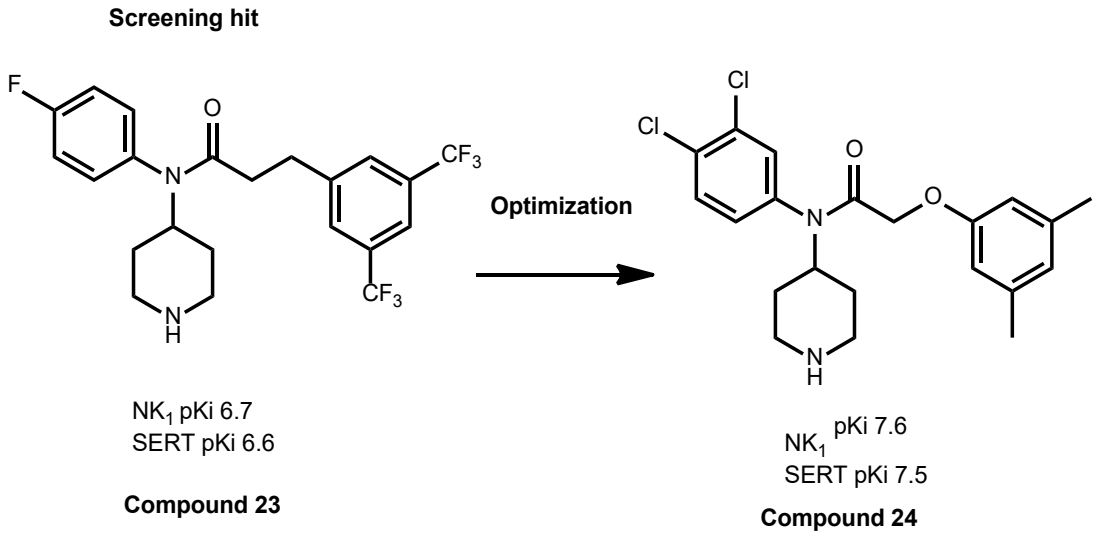


Figure 13: *NK1 and SERT dual antagonist found from high-throughput screening. Further optimization gave a more potent dual ligand [87].*

Of course, the resulting dual-targeting hit molecules might not have similar SAR profiles, meaning that potency of the optimised lead molecule against the two chosen targets may not be the same. As the example in **Figure 12** illustrates, the drug molecule from the screening hit shows potency for the two desired targets, endothelin and angiotensin II receptor. However, in comparison it is more potent towards angiotensin II and less potent towards endothelin. This can be the desired profile, or may also be a necessary compromise, as the level of potency for the two targets is very difficult to balance and optimize when the activity is combined in one molecule. An unbalanced activity profile can lead to complications in drug dosage if it is not desirable. A drug dosage that can sufficiently inhibit one target might not be sufficient to inhibit the second target, and increasing the dosage of the drug might not be an option as both levels of activities is set within one molecule [87]

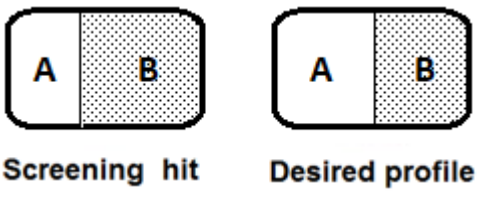


Figure 14: *The lead drug molecule from the high throughput screening might not have a desirable balance of activity for the targets*

2.4.2 Fragment-based drug design

Fragment-based drug design (FBDD) is a method that uses virtual computational screening to identify low molecular weight molecules (fragments) that bind within a pocket of a target protein. The hit fragments are then combined to give larger and more potent molecule. Fragment screening usually has a higher hit rate compared to high throughput screening [91]. This is mainly because molecules used in high throughput screening may have higher molecular weight and complexity, and therefore the chance of matching their shape to any given binding sites is decreased. However, fragments are capable of binding to small parts of a binding site, thus increasing the hit rate. Therefore, FBDD serves as a less costly alternative to high throughput screening in the pharmaceutical industry for the development of dual or selective single target drugs [92,93]. In one very illustrative example, the pharmaceutical company Pfizer screened a million compounds against Pin1 (Peptidyl-prolyl cis-trans isomerase NIMA-interacting 1) without obtaining any hits that could be verified [93,94], whereas Potter et.al. [95] managed to obtain two hits from a fragment library of nine hundred against Pin-1. Although this approach is mainly used for designing selective single target small molecule drugs and computational methods greatly facilitate the design process. With the increasing interest in polypharmacology, the prospect of utilizing FBDD in multi-target drugs development has been discussed in the literature [92, 93].

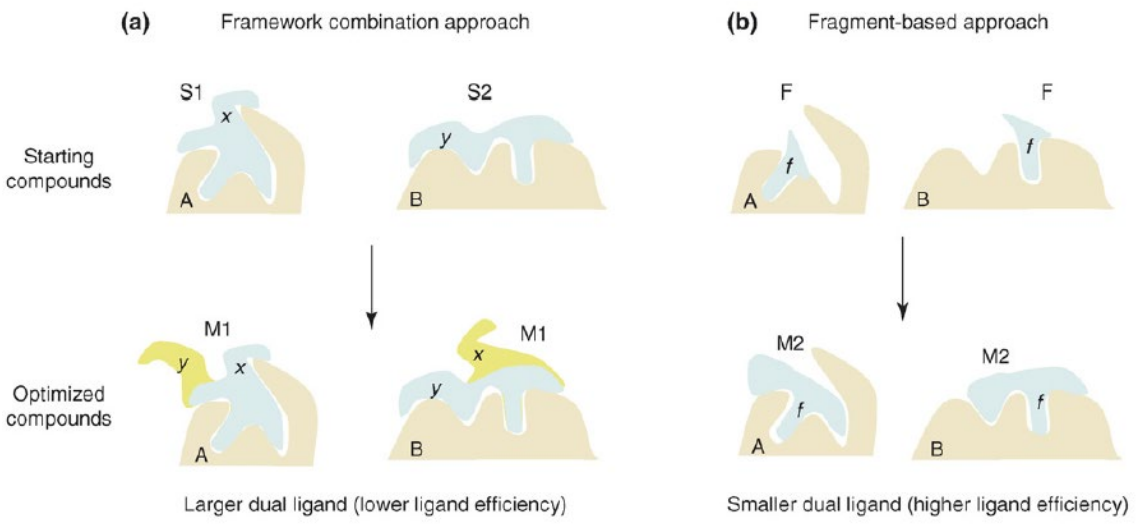


Figure 15: *The difference in binding between drug like molecules and fragments. The small size of fragments has a higher chance of binding to target molecule compared to the drug like small molecules typically used in high through-put screening [93].*

In the context of dual pharmacology, the screening strategy for the search of lead fragments for a dual antagonist will be largely similar to high throughput screening. Once the fragments are identified from the screening, they can be modified to improve their potency [93].

One of the more popular approaches to do this is “fragment growing” (**figure 16**). Fragment growing begins with examining the X-ray crystal structure or molecular model of the binding fragment. Additional functional groups and structures are added to the fragment to access the binding sites near to the fragment, to increase the potency of the compound [75].

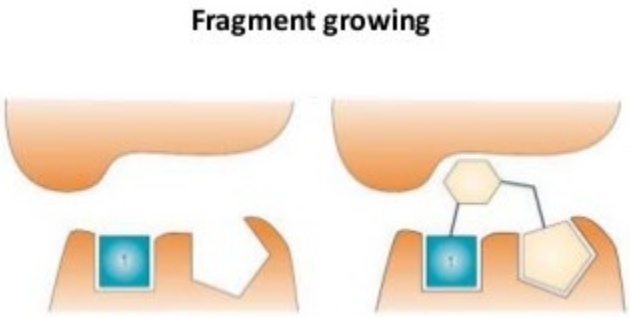


Figure 16: *Fragment growing: The lead fragment binds to the receptor at one site. Additional functional groups are added to the fragment to make contact with other pockets of the molecule [93].*

Another strategy involves the merging of a fragment with an existing molecule that shares structural similarities to increase its potency. For several selective single target molecules, the addition of another moiety can facilitate the acquisition of activity towards another target. Interestingly, merging a fragment and a drug molecule from different targets with similar structures could potentially create a lead molecule for selective dual targeting drug molecules [93].

Figure 17: *Fragment linking: Fragment 1 and 2 happens to bind to receptor pockets that are close to each other. The two fragments are joined together by a linking group. [93]*

In the scenario of multiple fragments binding to the adjacent sub-pockets of each other, the fragments can be merged to form a single molecule; this approach is referred to as fragment linking. Such binding would reveal the preferred binding orientation of the fragments at the binding site. The next step would be adding a flexible chemical linker to join the two fragments together [93]. This strategy requires the condition of multiple fragments binding to the adjacent sub-pockets of each other on the same target, it will not be viable to adopt this strategy for multi-target drugs. For FFBD, fragment growing appears to be the most reliable strategy for the development of multi-target drugs.

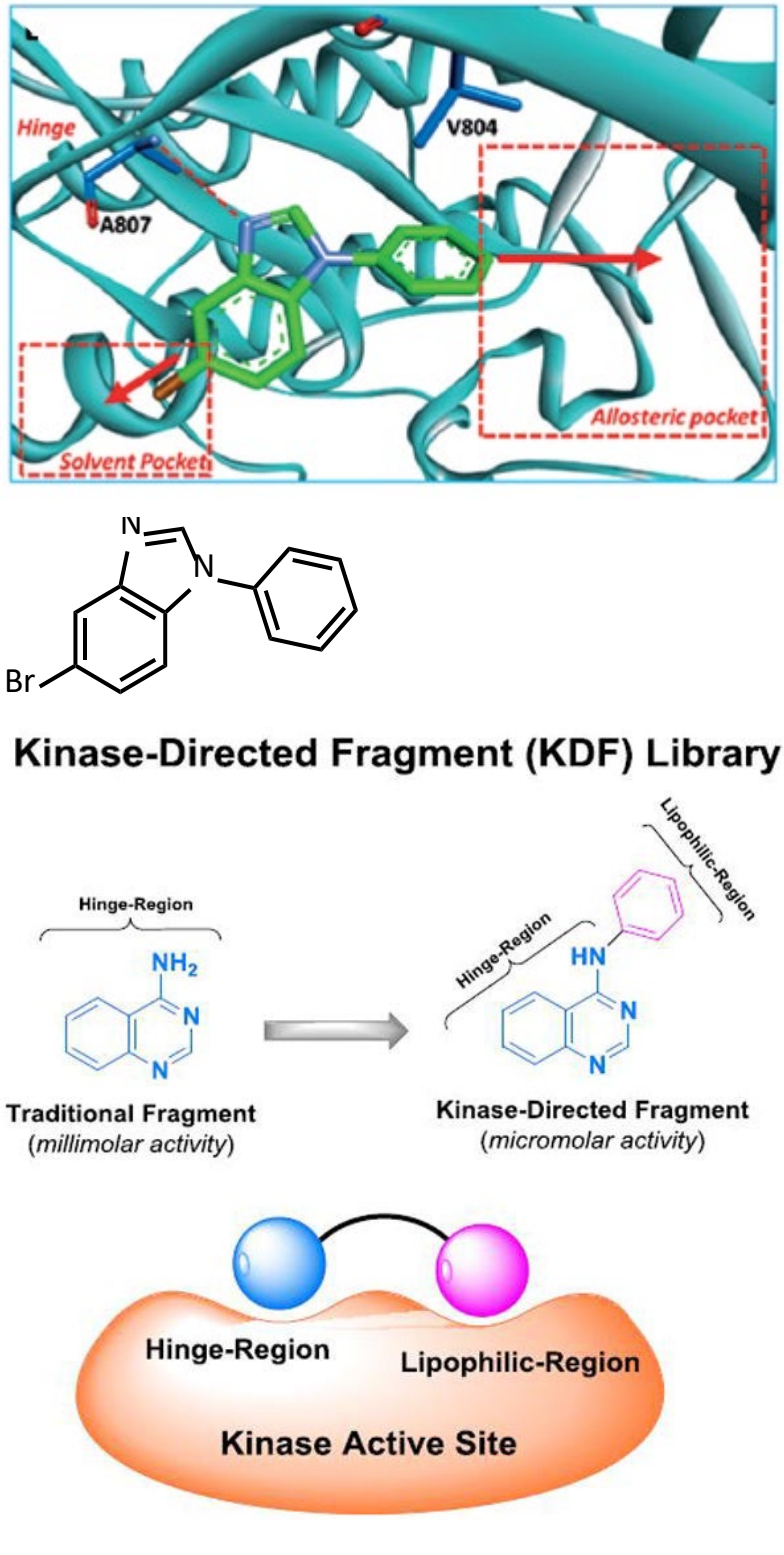


Figure 18: *Top: computational modelling of a fragment molecule binding to RET. Bottom: Simplified graphic representation of computational modelling of fragment molecule binding to RET [97].*

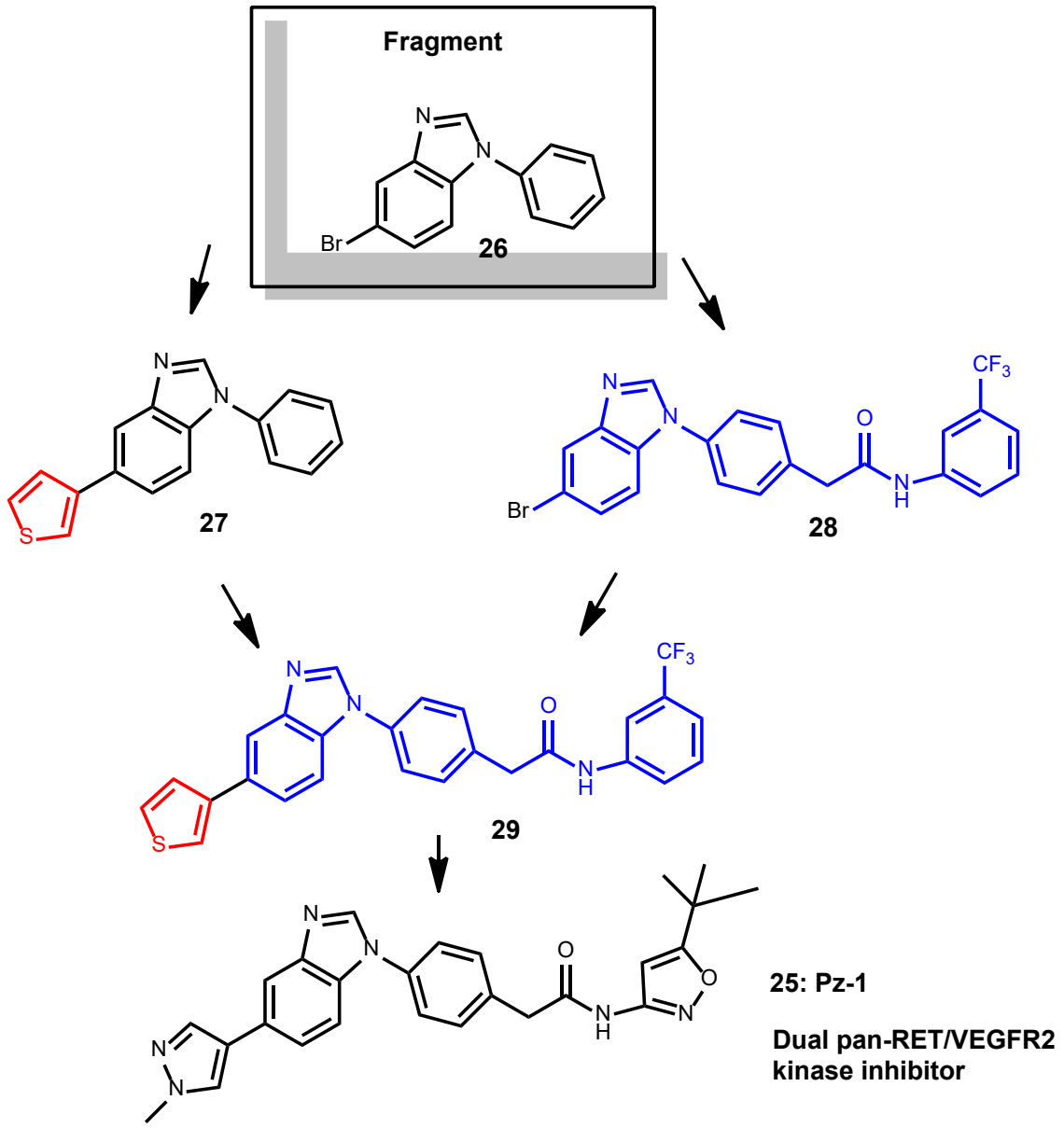


Figure 19: *Creation of the dual pan-RET/VEGFR2 kinase inhibitor Pz-1. The molecule was created starting from the lead fragment from the screening hit, compound **26**. Compound **26** was modified to **27** and **28**. Compound **29** was designed to have structural features from compound **27** and **28**. Further optimization gave compound **25**, Pz-1 [97].*

Some literature precedents of dual antagonists developed through FBDD are available. For example, Frett et. al [97] reported the development of a dual pan-RET/VEGFR2 kinase inhibitor Pz-1 (**25**) using a fragment-based drug design. RET (rearranged during transfection tyrosine kinase) and VEGFR2 (vascular endothelial growth factor receptor 2) are both molecular targets in cancer. Fragment screening on RET and VEGFR2 identified compound **26**, a

benzimidazole analogue as a promising lead molecule. Computational modelling of the fragment on the RET crystal structure identified the solvent pocket and the allosteric pocket as two possible points for optimization (**Figure 18**).

Modification to the fragment was made in attempt to access these two pockets. The bromine of the fragment was changed to thiophene to access the solvent pocket, and the *meta*-trifluoromethylphenyl moiety was added to access the allosteric pocket. The resulting analogues compounds **27** and **28**, showed increased potency towards RET (**Figure 18**). Compound **29**, which possesses the two moieties, was synthesized, and Kd values of 6 nM for RET and 18 nM for VEGFR2 were achieved. Further optimization resulted in Pz-1 **26**, which completely blocked RET-driven tumour formation with no detectable toxicity. The design of Pz-1 was largely guided by computational modelling using AutoDock tool and SWISS-MODEL, which are freely available [97].

It should be noted that in this example, the binding pockets on RET and VEGFR2 where Pz-1 binds are very similar. The binding conformation of Pz-1 is also very similar in both receptors. The optimization of the compound would have been more complex if the binding pockets from the two receptors are more dissimilar. Pz-1 nevertheless demonstrates that it is possible to use fragment-based drug design to create potent and selective dual targeting drugs for structurally similar targets [97].

2.4.3 Virtual ligand screening

Another avenue that can be used to identify possible lead molecules for dual antagonists is virtual ligand screening. Virtual ligand screening is used as an alternative and complementary screening method to high-throughput screening and offers a low-cost method in searching for lead molecules. It is widely used in predicting the binding of both drug-like molecules and fragment molecules to a target protein using computational tools [98]. In a typical virtual screening campaign, a library of molecules will be screened against the target using molecular modelling. Each compound can be docked into the target by a large number of different possible binding modes. The compounds are scored, typically by binding affinity. Compounds with highest scores can then be selected and tested by *in vitro* studies [98].

In recent years the prospect of utilizing virtual screening for multi-target drugs development has been discussed [98], although literature detailing virtual screening for the discovery of a multi-target lead molecule remains uncommon. Two notable examples have come from Wei et al. [99] and Bottegoni et al [100].

Wei et al. [99] discovered three lead molecules with low micromolar activity against human leukotriene A_4 hydrolase (LTA4H-h) and human non-pancreatic secretory phospholipase A2 (hnps-PLA2). Firstly, Wei et al. identified the common pharmacophore of the two enzymes. The binding conformation of the common pharmacophore then was predicted by molecular docking. Molecules from the screening library that did not have the similar binding conformation as the common pharmacophore were excluded. The binding conformations of the remaining results were inspected to compare with the common pharmacophore. Nine compounds were purchased and tested to produce three dual targeting lead molecules. Although this method was successful, identifying the common pharmacophore of the two enzymes is of course a complex task; Wei et al used software developed by themselves to aid the process [99] so this method of conducting virtual screening would be difficult to be reproduced by other researchers.

Finally, Bottegoni et al. [100] identified a fragment with low micromolar activity in two structurally unrelated enzymes; β-secretase (BACE-1) and glycogen synthase kinase 3β (GSK-3β). The two targets were screened by their own

individually assigned fragment library. The results were filtered and scored according to their ability to form important hydrogen bonds to the targets. The remaining results underwent visual inspection. 24 fragments were selected to find compounds with similar structures on the *Available Chemicals Directory*. This led to a new fragment library of 400 molecules. Screening was performed again, after further filtering based on the formation of important hydrogen bonds to both targets, followed by visual inspection on the filtered results, 27 compounds were purchased and tested. Lastly, the results led to the identification of a dual targeting lead fragment [100]. This virtual screening campaign had the advantage of knowing the important hydrogen bonding for the two targets from available homology models of an earlier fragment-based drug design project, showing virtual screening can complement and benefit from existing drug design projects.

Virtual screening is a complex and difficult task with many pitfalls. The success of virtual screening is highly dependent on the protocols used for the screening. One of the challenges of virtual screening is many receptors have no experimentally determined structure by X-ray crystallography, especially for GPCRs [88]. Homology modelling is therefore needed to predict the structure of receptors. Homology modelling predicts the unknown structure of a protein based on the assumption that a related protein with similar amino acid sequence will have similar 3D structure. The reliability and accuracy of the predicted structure are highly dependent on the protein sequence similarities of the templates used for the modelling. According to Vyas et al. [101], homology models built over 50% sequence similarities are accurate enough for drug discovery applications.

3.0 Project background and objectives

3.1 Previous work on CCR7 in our group

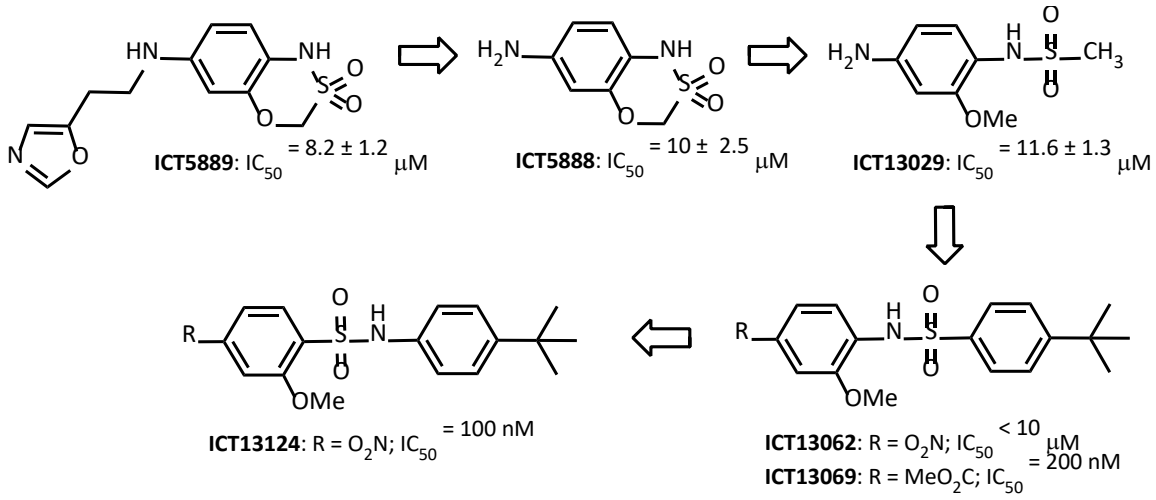


Figure 20: *The development history of ICT13069, the structure of the previously synthesised CCR7 antagonists and their IC_{50} values for calcium flux assay.*

This project is a continuation of an ongoing programme in our group for the development of a CCR7 small molecule antagonist suitable for *in vivo* investigation. ICT5889 was originally identified in our group as a moderately potent CCR7 antagonist (**Figure 20**). We later discovered that an analogue, ICT5888, without the oxazole sidechain, was similarly potent. Synthesis of a series of open chain sulfonamides, which no longer have the cyclic sulfonamide ring (for example ICT13029) also gave active compounds. This eventually led to ICT13062, and then to the first sub-micromolar potent compound ICT13069 (**30, figure 21**). ICT13062 lacks solubility in aqueous media and it was difficult to determine its IC_{50} accurately, but it was below 10 μM. ICT13069 on the other hand was relatively potent with an IC_{50} below 0.2 μM. Although ICT13069 is relatively potent, it is unsuitable for *in vivo* work as it is likely to rapidly hydrolyse under physiological conditions.

Following the identification of ICT13069 further work in our group focused on changing the fragments highlighted in **Figure 21**. This included changing the fragment A (the functional group *para* to the nitrogen of the sulfonamide) and fragment B (the functional group *para* to the sulfonyl of the sulfonamide). Prior to this investigation, fragment C was mostly kept as a methoxy group, and very few analogues of this methyl ether had been made. This meant the SAR for this

fragment was unknown. Interestingly, ICT13124, in which the sulfonamide group was reversed compared to ICT13062, whilst still poorly soluble, was more potent. This led us to the hypothesis that the corresponding “reverse sulfonamide” of ICT13069 should also be significantly more potent.

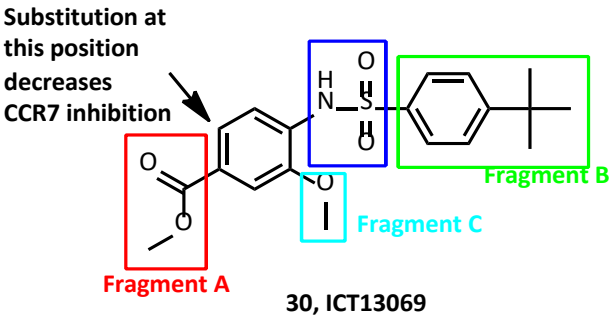


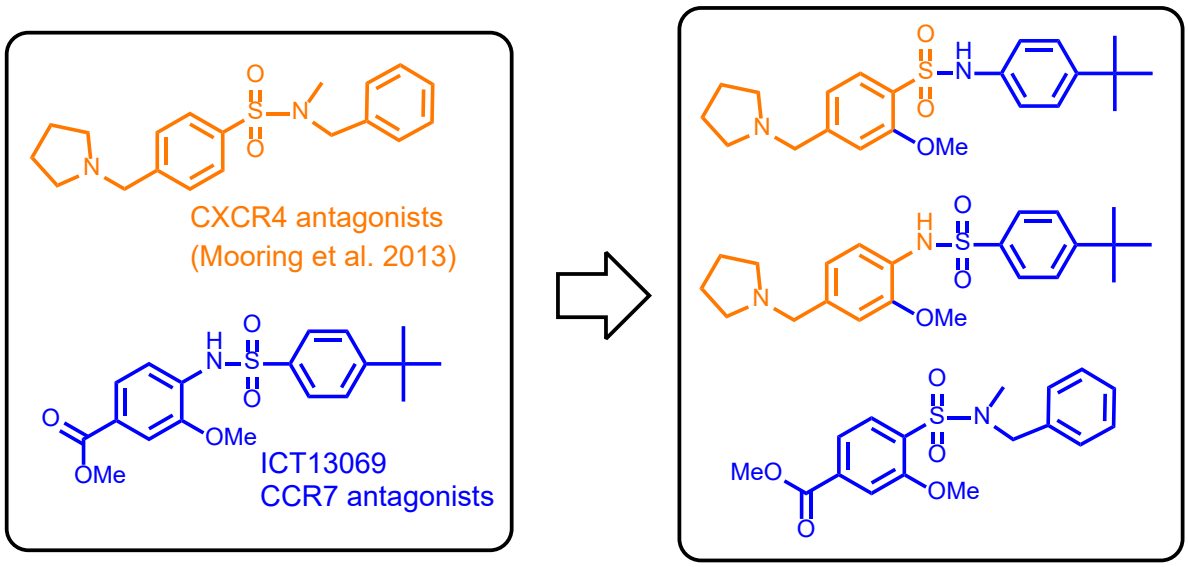
Figure 21: *Compound **30** (ICT13069) representing the general structure of the small molecule antagonist of this project*

3.2 Dual CXCR4/CCR7 antagonists

As discussed above, both the CXCR4 and CCR7 axes appear to contribute to expansion of tumours and co-expression of both CXCR4 and CCR7 receptor leads to poorer prognosis [24]. Therefore, a dual CXCR4/CCR7 antagonist would be advantageous over a CXCR4 antagonist. In line with the earlier discussion, our group constructed homology models of both CXCR4 and carried out vHTS of both targets. A key component of the design of dual CCR7/CXCR4 antagonists is that it is possible to compare the binding of dual antagonists inside both CCR7 and CXCR4 receptors and predict binding affinities. The alignment of the peptide sequence between CCR7 and CXCR4 shows that the two receptors are quite similar (**figure 22**).

```
CCR7_HUMAN     1   MDLGKPMKSVLVVALLVIFQVCLCQDEVTDDYIGDNTTVDYTLFESLCSKKDVRNFKAWF   60
CXCR4_HUMAN    1   ----------------------MEGISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIF   40
                                         :   .  .:*:*  :  : **  ::. * :::  **:  *

CCR7_HUMAN    61   LPIMYSIICFVGLLGNGLVVLTYIYFKRLKTMTDTYLLNLAVADILFLLTLPFWAYSAAK  120
CXCR4_HUMAN   41   LPTIYSIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVA  100
                   ** :**** :.*::*****:*.  * *:*::***.* *:*:***:**::****** .*.

CCR7_HUMAN   121   SWVFGVHFCKLIFAIYKMSFFSGMLLLLCISIDRYVAIVQAVSAHRHRARVLLISKLSCV  180
CXCR4_HUMAN  101   NWYFGNFLCKAVHVIYTVNLYSSVLILAFISLDRYLAIVHATNSQ--RPRKLLAEKVVYV  158
                   .* ** .:** :..**.:.::*.:*:*  **:***:***:*..::  * * ** .*:  *

CCR7_HUMAN   181   GIWILATVLSIPELLYSDLQRSSSEQAMRCSLITEHVEAFITIQVAQMVIGFLVPLLAMS  240
CXCR4_HUMAN  159   GVWIPALLLTIPDFIFANVSEADD--RYICDRFYPNDLWVVVFQFQHIMVGLILPGIVIL  216
                   *:** * :*:**:::::::..:..     *. :  :   .:.:*. ::::*:::* :.:

CCR7_HUMAN   241   FCYLVIIRTLLQARNFERNKAIKVIIAVVVVFIVFQLPYNGVVLAQTVANFNITSSTCEL  300
CXCR4_HUMAN  217   SCYCIIISKLSHSKGHQKRKALKTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEF  276
                    ** :** .* :::..::.**:*. : :::.*:.  ***   :  ::.  ::* .. **:

CCR7_HUMAN   301   SKQLNIAYDVTYSLACVRCCVNPFLYAFIGVKFRNDLFKLFKDLGCLSQEQLRQWSSCRH  360
CXCR4_HUMAN  277   ENTVHKWISITEALAFFHCCLNPILYAFLGAKFKTSAQHALTSVSRGSS--LKILSKGKR  334
                   .: ::   .:* :** .:**:**:****:*.**:..  : :..:.  *.  *:  *. ::

CCR7_HUMAN   361   IRRSSMSVEAETTTTFSP                                          378
CXCR4_HUMAN  335   GGHSSVSTESESSSFHSS                                          352
                     :**:*.*:*::: .*
```

Figure 22: *Sequence alignment of CCR7 and CXCR4.*



Known CCR7 and CXCR4 antagonists

Potential dual CCR7/CXCR4 antagonists

Figure 23*: Combining structural features between the CXCR4 antagonists by Mooring et al. and the CCR7 antagonist from this project can give rise to ideas of a dual CXCR4 and CCR7 antagonist.*

For this research project, the most viable option to create small molecule dual CCR7 and CXCR4 antagonists is by incorporating structure features from existing CXCR4 antagonists into the CCR7 antagonists of this project, as was the case in the three examples shown in section 2. The sulfonamide CXCR4 antagonists from Mooring et al. [80] have similar structures as the CCR7 antagonists of this project (**Figure 23**). The most notable functional group to add from these CXCR4

antagonists would be the methyl pyrrolidine group, incorporating this functional group to the CCR7 antagonists of this project might yield potential dual CCR7 and CXCR4 antagonists.

3.3 Aims and objectives

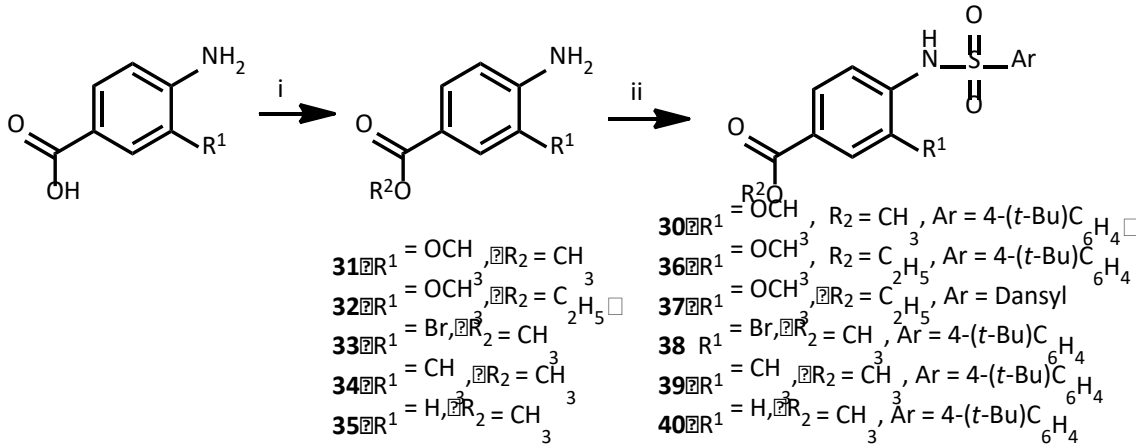
Based on what is discussed above in section 3.1, we set ourselves the following objectives:

1. To explore different functional groups at fragment C (**Figure 21**), to gather more SAR data for the CCR7 antagonists.
2. To synthesise metabolically stable analogues of ICT13069.
3. To synthesise analogues of ICT13069 with the sulfonamide group in reverse.
4. To prepare a series of dual CXCR4 and CCR7 antagonists which are hybrid structures from ICT13069 and compound **6** reported by Mooring et al. [80] (**Figure 23**).
5. Assess the potency of all compounds in biological assays against CCR7 and CXCR4.
6. Use molecular docking to investigate the binding mode of the antagonists.

4.0 Results and discussion

4.1 Modifications to ICT13069 at positions C and A (Figure 21)

Compound **30** and **36-40** were synthesized using the reaction scheme outlined in **Figure 24**. The general synthesis starts with esterification of the carboxylic acids, follow by reacting the resulting anilines **31-35** with the corresponding aryl sulfonyl chloride to afford compound **30** and **36-40**.



i) Hydrogen chloride, R^2OH, reflux; ii) $ArSO_2Cl$, Pyridine, DCM, r.t.

Figure 24: *Synthesis of compounds **30** and **36-40***

Compound **36** and **37** had been previously prepared in our group, whilst compound **38-39** are novel and are used to test the importance of the methoxy group to the activity of ICT13069 and to generate SAR data for functional groups on that position. Previous SAR studies had shown the absence of the methoxy group (**37**) made the compound inactive, indicating it is required to have an ether functional group at this position for CCR7 inhibition. We also attempted to synthesise an analogue with the trifluoromethyl group as R^1 using methyl 4-amino-3-(trifluoromethyl)benzoate, however, no reaction occurred using the same reaction condition outlined (**Figure 24**). Different methods involving trimethylamine or sodium hydride in THF or DMF also failed to afford the sulfonamide. Presumably, the presence of the trifluoromethyl at the *ortho* position to nitrogen renders this group too weak a nucleophile to attack the sulfonyl group and hence, we were unable to obtain a sulfonamide from this aniline. Further work in our group is ongoing to explore the range of ether groups that is tolerated

at this position. A mechanism for the formation of sulfonamides is outlined below (**Figure 25**).

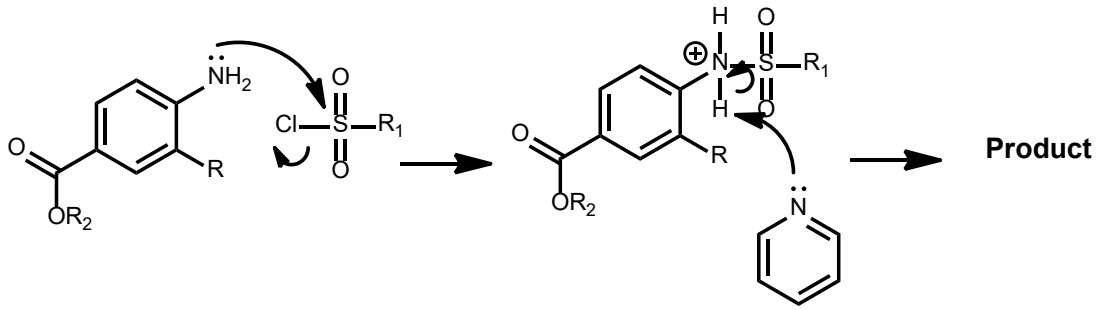
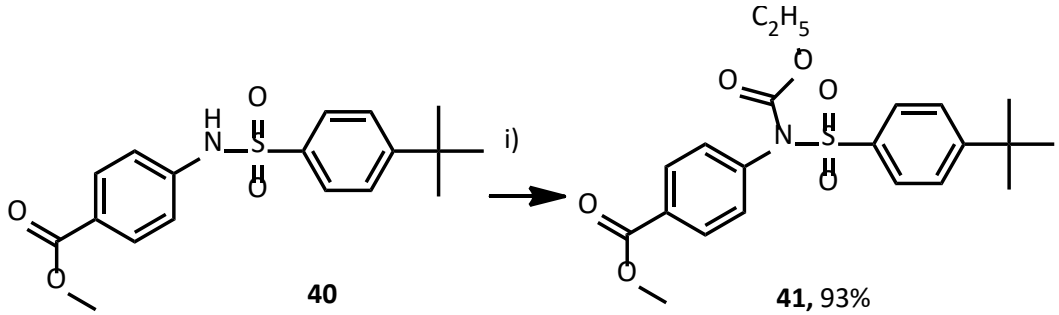


Figure 25: *Proposed mechanisms for the formation of sulfonamides*



i) THF, 60% NaH, ethyl chloroformate, r.t.

Figure 26: *Synthesis of compound **41***

Compound **41** was synthesized to test the effect of having an ester on the sulfonamide. Compound **41** was synthesized by treating **40** with ethyl chloroformate in THF and 60% sodium hydride.

4.2 Metabolically stable analogues of ICT13069

We then turned our attention to the synthesis of metabolically stable analogues of ICT13069 by modifications to fragment A. An obvious modification is to make a corresponding methyl amide analogue of the ester, compound **42**. Amides are less prone to aqueous (non-enzymatic) hydrolysis so this compound would be expected to have a longer half-life *in vivo*.

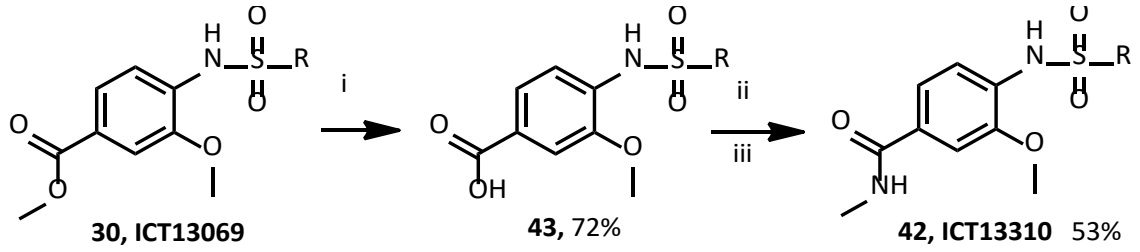


Figure 27: *Synthesis of compound **42***

Compound **42** and was synthesized using ICT13069 (**30**) as the starting material (**Figure 27**). The synthesis starts with hydrolysis of ICT13069 with potassium hydroxide to convert the ester to carboxylic acid (compound **43**), follow by reflux in thionyl chloride to create the acid chloride which was used directly in the next step without purification. The acid chloride was treated with methylamine in ethanol at -10 °C (acetone and ice bath) to prevent the acid chloride from reacting with the ethanol instead of the methylamine. The reaction was left in cold storage under 0 °C for one day to give the methyl amide **42**. A mechanism for the sequence leading to compound **42** is shown in **Figure 28**.

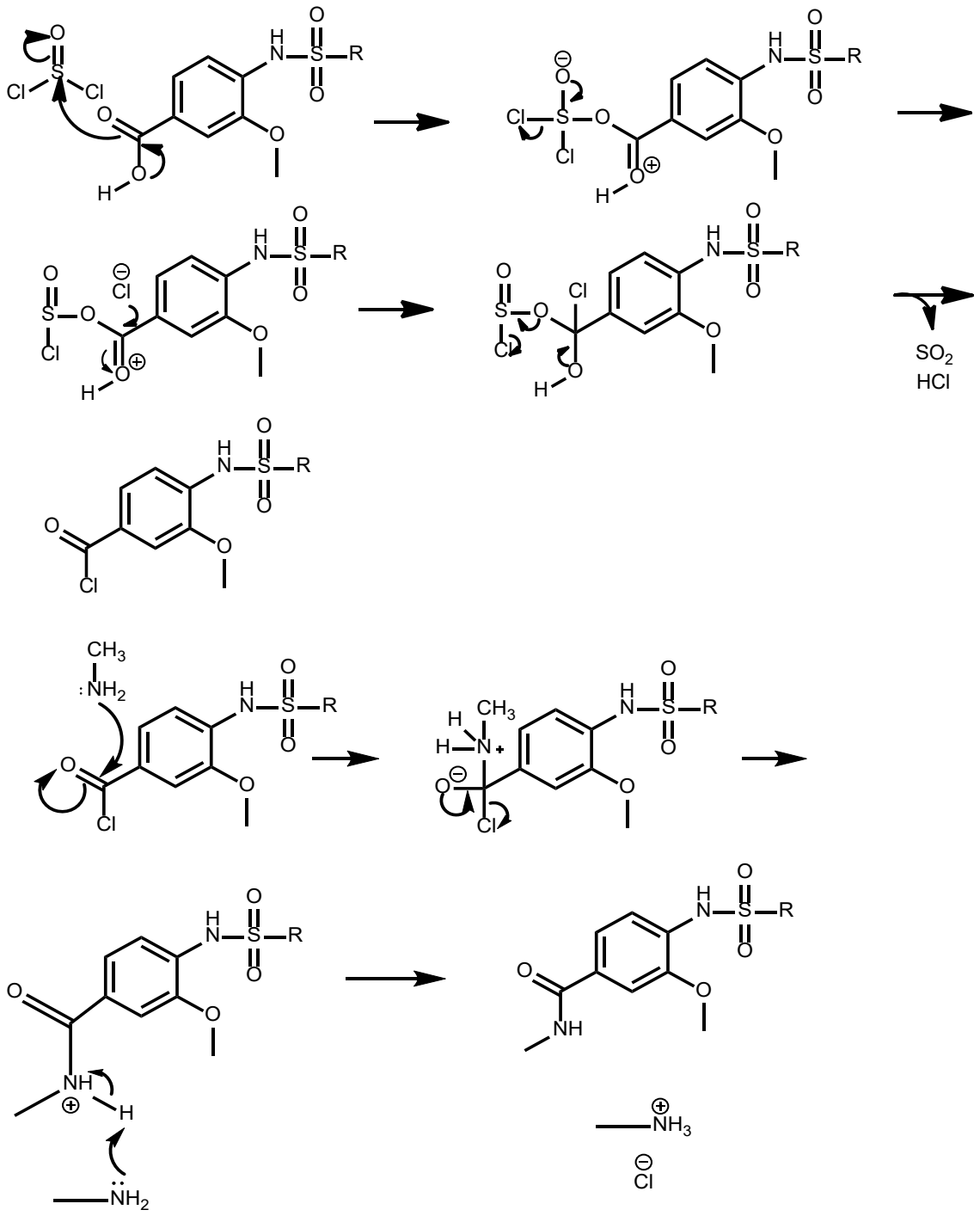


Figure 28: *Reaction mechanism leading to the formation of compound **42***

Primary amide **42** was tested and shown to have similar potency to ICT13069 (see later). We concluded that compound **42**, may be more useful for assessing the role of CCR7 antagonism *in vivo*.

4.3 Synthesis of reverse sulfonamides

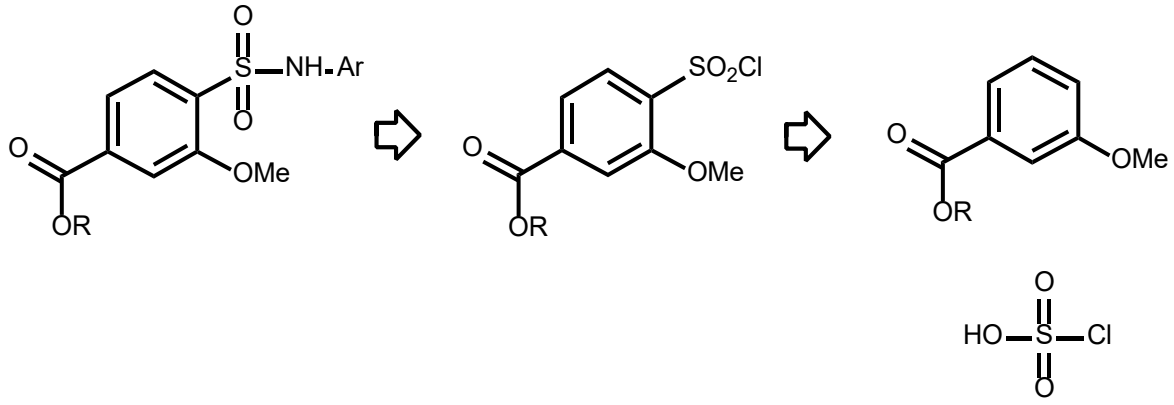


Figure 29: *Proposed preparation of reverse sulfonamides analogues of ICT13069*

Our initial strategy (**Figure 29**) to prepare the reverse sulfonamide of ICT13069 was using methyl 4-(chlorosulfonyl)-3-methoxybenzoate as the starting material, but it is not commercially available and therefore needed to be synthesized. Our approach was to use chlorosulfonic acid for the sulfonation of methyl 3-methoxybenzoate. Sulfonation using chlorosulfonic acid under anhydrous condition can afford the aryl sulfonyl chloride in one step and it is unlikely to hydrolyse the methyl ester. The position where the sulfonations take place depends on the substituents in the aromatic ring. The methoxy group is a strong electron donating group and thus strongly *ortho* and *para* directing. Therefore, both *ortho* and *para* isomers should be expected but we hoped they can be separated as part of the purification process.

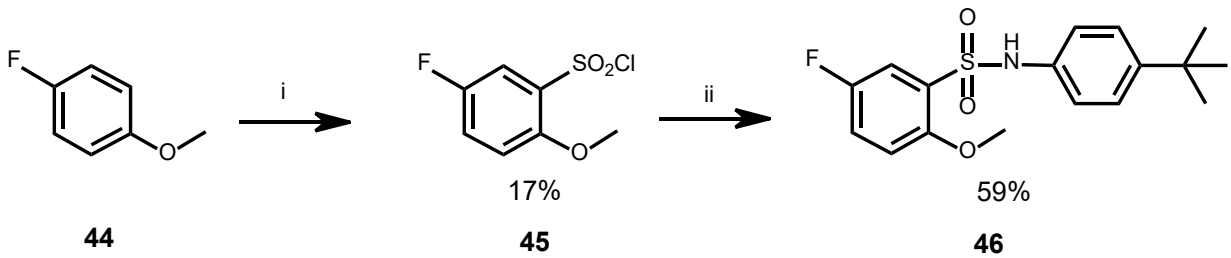
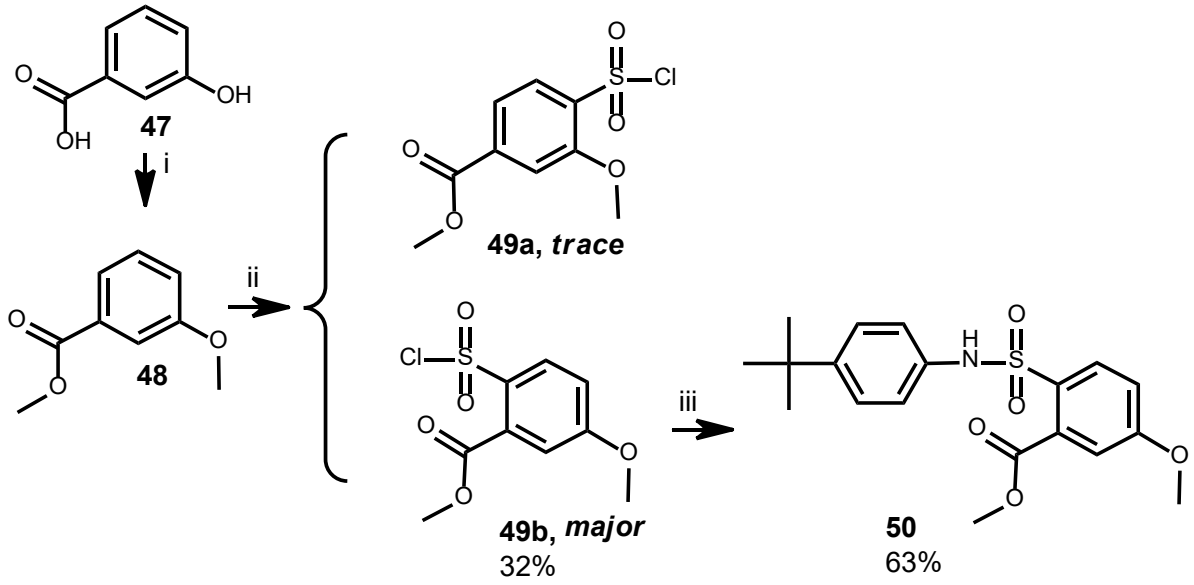


Figure 30: *Synthesis of compound **46**.*

The synthetic route was trialled on 4-flouroanisole, **44** (**Figure 30**). By blocking the *para* position, the sulfonation is forced to take place at the *ortho* position. The sulfonation with anisole is reported to occur in room temperature [102] but for 4-flouroanisole heating was required. The yield of this reaction (17%) was notably lower compare to the ~60% yield that reported for anisole [103], possibly due to the heating leading to some decomposition. The sulfonyl chloride product, **45** was purified by flash column chromatography and was taken further to react with 4-*tert*-butylaniline to afford **46**. Sulfonation with chlorosulfonic acid required the presence of chloroform or dichloromethane, the same condition using tetrahydrofuran/diethyl ether or with no solvent gave less than 10% yield. In spite of the poor yield of the reaction, we were sufficiently encouraged to attempt the synthesis of the target analogues.



i) MeI, K_2CO_3, acetone, reflux; ii) $ClSO_3H$, $CHCl_3$, reflux; iii) *tert*-butylaniline, DCM, r.t

Figure 31: *Reaction scheme of 3-hydroxybenzoic acid to reverse sulfonamide*

3-Hydroxybenzoic acid **47** was converted to methyl 3-methoxybenzoate **48** by treatment with iodomethane in the presence of potassium carbonate and acetone (**Figure 31**). We expected the sulfonation of **48** to occur at positions *ortho* and *para* to the activating methoxy group, affording **49a** and **49b**. However, we hoped that the presence of carbomethoxy group would hinder sulfonation at the *para* and one of the *ortho* positions to the methoxy group thus favouring the desired sulfonation product, **49a**. Even if the desired product was not the major isomer,

we hoped to synthesize it in sufficient quantities to allow us a straight forward synthesis of the required sulfonyl chloride.

The sulfonation of methyl 3-methoxybenzoate occurred under room temperature and afforded essentially a single compound **49b** with only traces of the desired isomer, **49a**, detectable in the crude ^{1}H NMR spectrum. Altering the temperature and stoichiometry of chlorosulfonic acid showed no significant change in the ratio (**Figure 31**). The mixture was treated with 4-*tert*-butylaniline to afford compound **50**.

As the product in this reaction was not crystalline and we could not obtain an X-ray, we had to rely on NMR spectroscopy to confirm the position of the sulfonyl chloride. Both expected products are 1,2,4-substituted benzenes and hence splitting patterns in ^{1}H NMR are expected to be similar. The chemical shifts for the methoxy group and the methyl ester are close, but they are easily identified by using HMBC. The carbonyl carbon has the highest chemical shift in ^{13}C NMR, and in HMBC it correlated to the methyl ester hydrogens. Having identified the methyl ether and methyl ester chemical shifts in ^{1}H NMR, we then set out to identify the spatial relationship of ring hydrogens with respect to the methyl protons. In the NOESY spectra, the hydrogen from methoxy ether group had a correlation to two hydrogens, whereas the hydrogen from the methyl ester had a correlation to only one. This is only possible for the unwanted isomer (**Figure 32**).

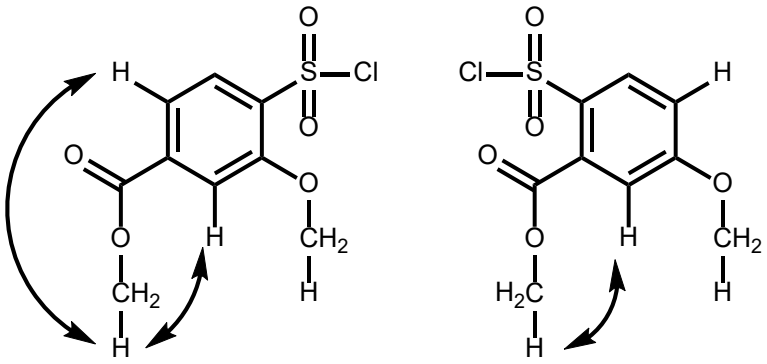


Figure 32: *One of the ways to differentiate the para and ortho isomer through NOESY correlation.*

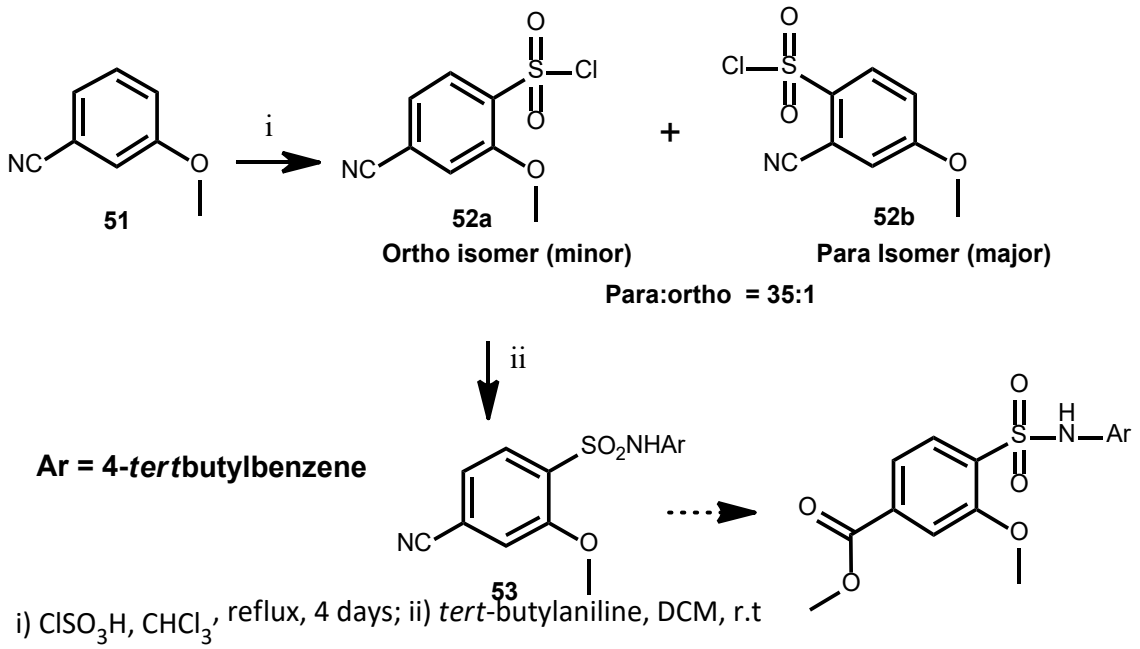


Figure 33: *Reaction scheme of 3-methoxybenzonitrile to reverse sulfonamide*

We next attempted sulfonation with 3-methoxybenzonitrile (**51**). The nitrile group is much more electron withdrawing than the methyl ester, which might decrease the *para* directing effect of the methoxy group and improve the ratio. If the reaction is a success, the nitrile group can be converted to carboxylic acid by hydrolysis, subsequent esterification will lead to the reverse sulfonamide of ICT13069. In the reactions of 3-methoxy benzoates, the sulfonation reaction using chlorosulfonic acid was complete within a day, but for 3-methoxybenzonitrile, the reaction was, as expected, sluggish and took at least 2 days of stirring and heating at 75 °C for the sulfonyl chloride product to even appear on thin layer chromatography (TLC). Interestingly, on TLC, the starting material disappeared after 3 hours whilst the sulfonyl chloride product was produced over several days, suggesting that a hitherto unknown intermediate was produced in the reaction mixture. Stopping the reaction at day 4 gave both products **52a** and **52b** at the combined yield of 1.7%. Stopping the reaction at day 2 and 3 gave even lower yield. Allowing the reaction to continue for more than 4 days did not improve the yield. Although the reaction was very slow, our prediction of a more favourable selectivity did materialise, albeit very slightly. The ratio increased to 1:35 in favour of the undesired product. It appears that the electron withdrawing nitrile group increased the activation energy of the reaction

and made the sulfonation much slower and slightly less selective. The long reaction time and the temperature might possibly provide thermodynamic reaction control to the reaction, thus the formation of the *ortho* isomers was increased. When the reaction was carried on a large scale, it allowed the two isomers to be isolated in workable quantities following flash column chromatography. The few milligrams of the required isomer were taken further to react with 4-*tert*butylaniline to give the sulfonamide products **53**. However, as the sulfonation of 3-methoxybenzonitrile is very inefficient, this route was not pursued any further (**Figure 33**).

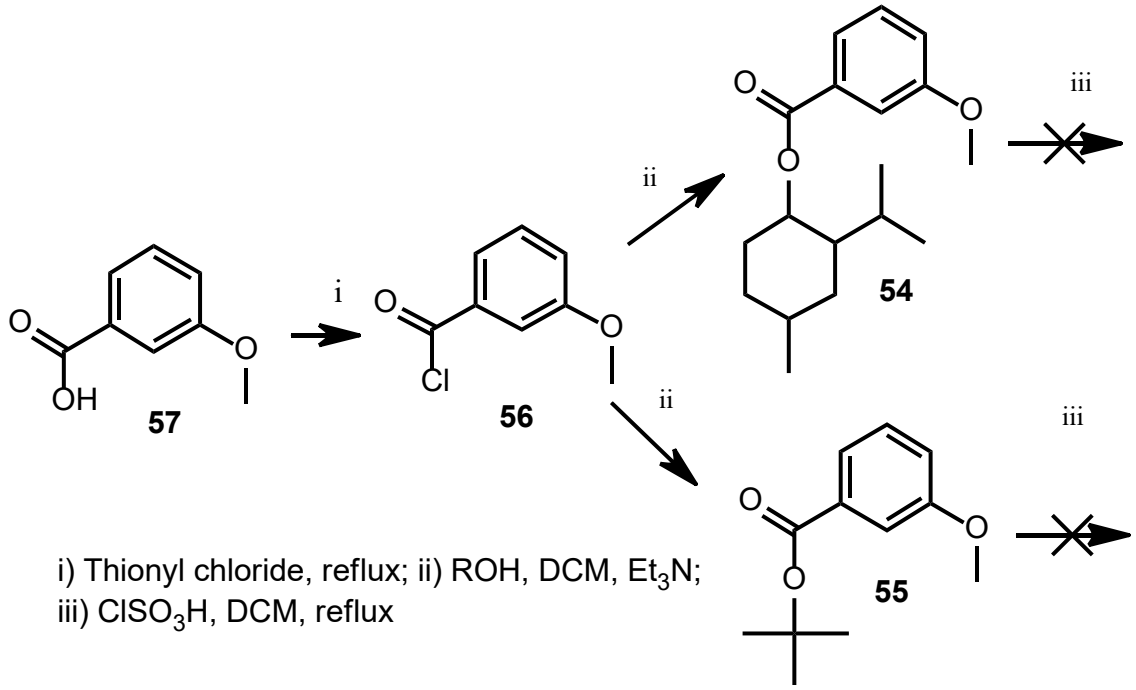
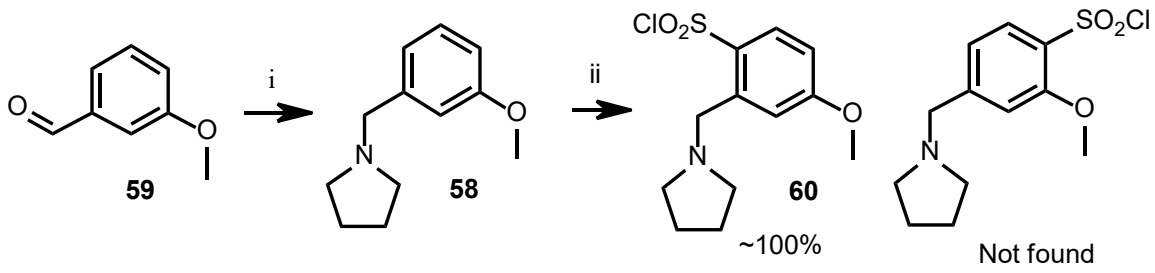


Figure 34: *Reaction scheme of methyl and t-butyl ester of 3-methoxybenzoate to reverse sulfonamide of ICT13069*

The next idea was the sulfonation of *tert*-butyl and menthyl esters of 3-methoxybenzoate, **54** and **55** respectively (**Figure 34**). These were prepared from acid chloride **56** which in turn was obtained from treating **57** with thionyl chloride. *tert*-Butyl and menthyl esters are bulkier and more sterically demanding than a methyl ester, their bulk can cause steric hindrance, possibly preventing sulfonation at their carboxylate *ortho* position, thus forcing the sulfonation to take place at the position ortho to the methoxy group instead. The reaction for these two esters also required heating to occur. However, heating the reaction for several days at 75 °C, as was the case with 3-methoxybenzonitrile, did not afford any sulfonation product.



i) Pyrrolidine, r.t.; then Methanol, $NaBH_4$, r.t. ii) $ClSO_3H$, $CHCl_3$, reflux, 4 days

Figure 35: *Sulfonation of 3-methoxy benzylpyrrolidine*

The next attempt for sulfonation was carried out on a benzyl pyrrolidine compound (**Figure 35**). The purpose was to create a CCR7 antagonist with a moiety from the CXCR4 antagonist developed by Mooring et al. [80]. The benzyl pyrrolidine compound **58** was made from the corresponding aldehyde **59** through reductive amination. The proposed mechanism is shown in **Figure 36**. We argued that the pyrrolidine could possibly cause steric hindrance to sulfonation at its *ortho* position, but as it is not an electron withdrawing group, the reactions can proceed efficiently and theoretically it might increase the ratio of the desired isomer. However, results from the NOESY spectrum showed that the CH_2 of the pyrrolidine correlates with only one proton from the aromatic ring. This indicates the product **60** was the undesired compound and sulfonation had occurred at the *para* position to the methoxy group. The *meta* isomer was not observed in the ^1H NMR spectrum.

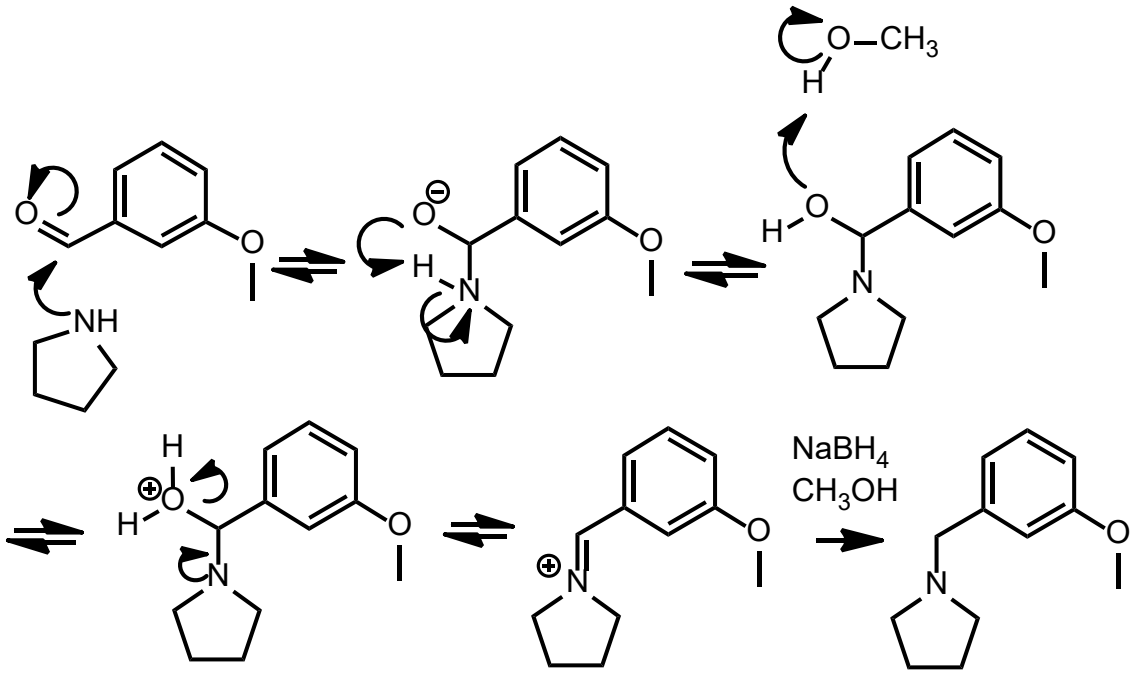


Figure 36: *Proposed mechanism of reductive amination of **58***

Sandmeyer reaction as alternative strategy

The Sandmeyer reaction can be used as an alternative strategy to synthesize the reversed sulfonamide of ICT13069 (**Figure 37**). In the Sandmeyer reaction an amine is converted to the diazonium cation, which can then be substituted by nucleophiles by using the appropriate copper compound and concentrated acid. The mechanism of the reaction is outlined in **Figure 38**. We envisaged two different approaches. Approach A (**Figure 37A**) involved conversion of an amino group to sulfonyl chloride (via the diazonium salt), following a literature procedure [103]. In this approach, sulfur dioxide generated by the highly exothermic reaction between thionyl chloride and water acts as a nucleophile. The literature reported that the reaction does not work well for electronically neutral and electron-rich rings. Due to the presence of both the methoxy and carboxyl groups in the starting material, the efficacy of the reaction was doubtful. In the event, the reaction failed to work and after the work up, only the starting material could be detected (by TLC and NMR).

Approach B (**Figure 37B**) involved conversion of an amino group to a nitrile (via diazonium salt). The first attempt at the reaction was successful but the reaction gave a mixture of the corresponding chlorobenzene and the nitrile products. The chlorobenzene analogue was presumably obtained due to the chloride anions from hydrochloric acid used in the formation of diazonium salt competing with cyanide anions as nucleophile. Therefore, we switched to using sulfuric acid instead which eliminated the formation of the chlorobenzene. There were two other problems with this reaction. Firstly, the yield of this reaction was low (~30%) although it can afford sufficient amount of the product to be used in the next step. The other problem with this reaction is that it requires potassium cyanide in excess. Potassium cyanide is vital to this reaction as copper cyanide is insoluble in organic solvent and in sulfuric acid. The poor yield of this reaction means that a larger scale reaction will be necessary. Due to the hazardous nature of working with large amount of potassium cyanide, and the relative inefficiency of this reaction, we decided to seek an alternative approach.

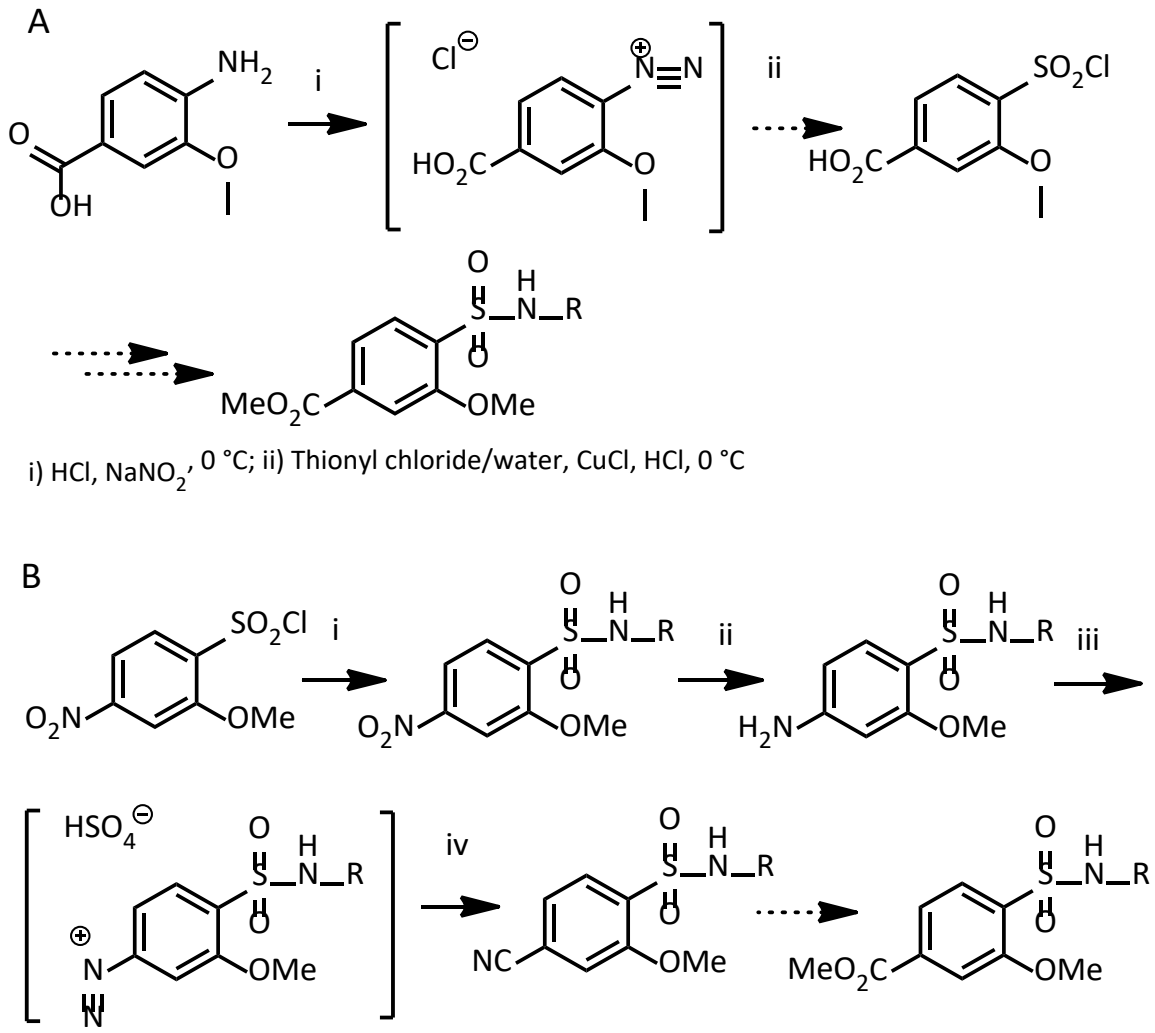


Figure 37: *Using Sandmeyer reaction in attempt to synthesize the reverse sulfonamide of ICT13069. A: conversion of amine to sulfonyl chloride. B: conversion of amine to nitrile*

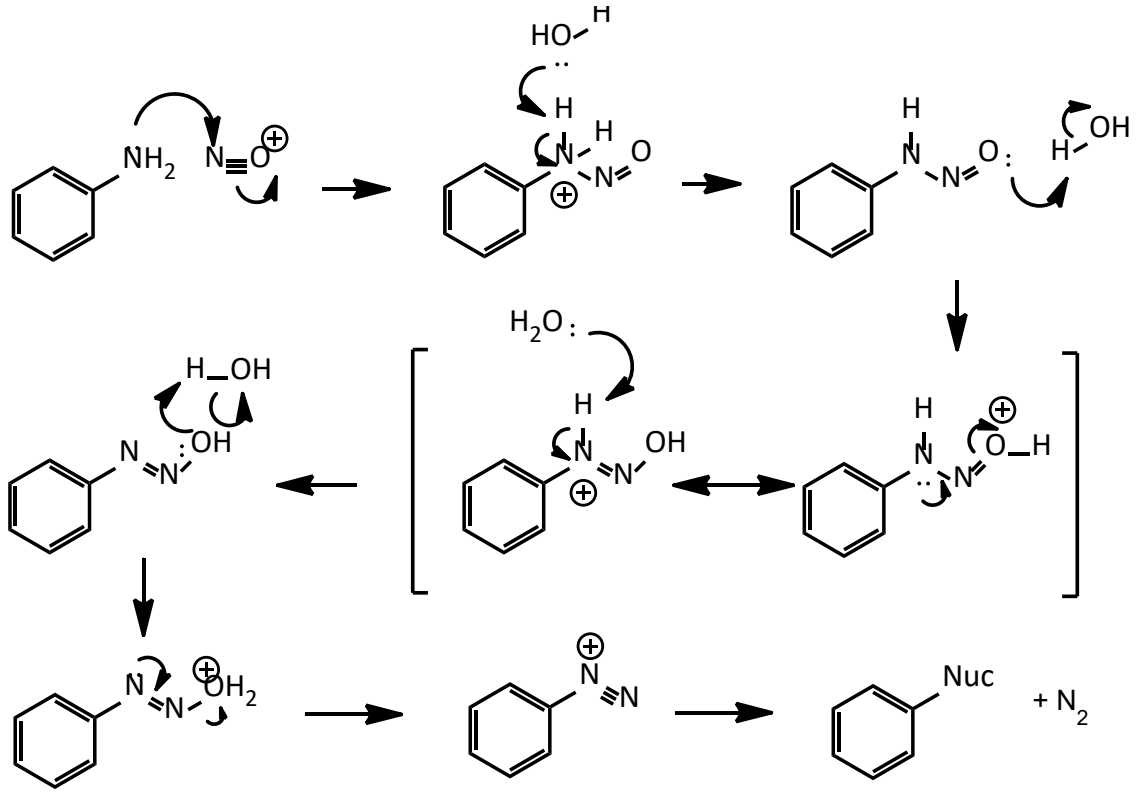
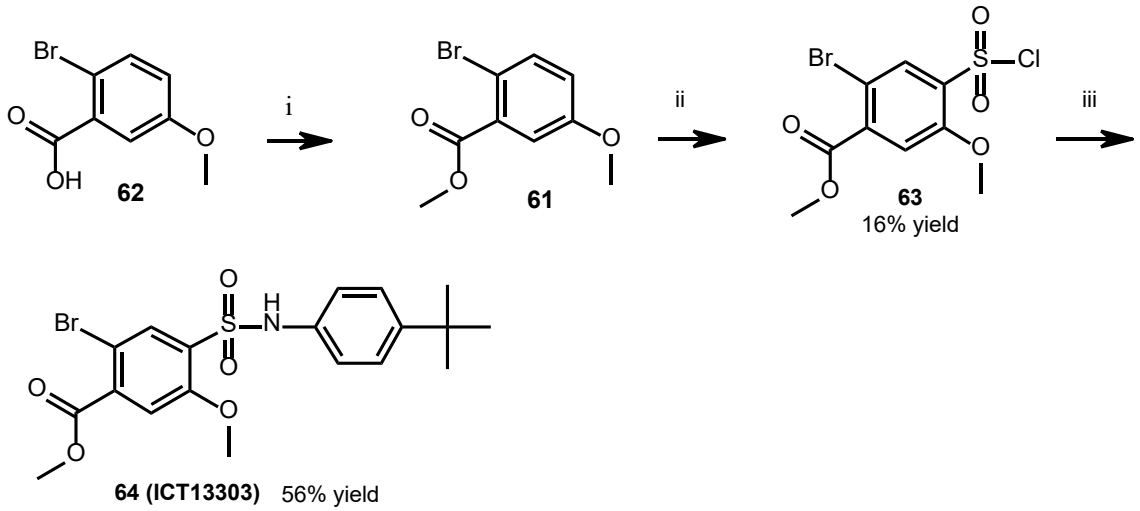


Figure 38: *Reaction mechanism of the diazonium ion formation used for the Sandmeyer reactions*



i) AcCl, MeOH, reflux; ii) $ClSO_3H$, $CHCl_3$, reflux; iii) *tert*-butylaniline, DCM, r.t

Figure 39: *Reaction scheme for the synthesis of compound **64** (ICT13303)*

Since by now it was obvious that sulfonation on a benzene ring with methoxy group is very selective in favour of the undesired product, it was decided to block the position *para* to the methoxy group. This would arguably allow the sulfonation to occur at the desired position; however, this approach requires a step to remove the blocking group. Sulfonation was carried out on methyl 2-bromo-4-methoxybenzoate, **61** which was prepared from the corresponding acid, **62**. Similar to 4-flouroanisole, the position *para* to the methoxy group is blocked, thus forcing the sulfonation to occur at the position *ortho* to the methoxy group. Heating was required for the starting material to react with chlorosulfonic acid. The reaction was successful affording **63** which was taken further to react with 4-butylaniline and afford the reverse sulfonamide **64** (ICT13303) (**Figure 39**).

Palladium catalysed dehalogenation

A palladium catalysed dehalogenation method was adopted from Chen et al. [104] to remove the bromine in **64** to synthesise the reverse sulfonamide of ICT13069. This procedure involves the early stage of the Heck reaction mechanism. Adopting the procedure from Chen et al [104], **64** was dissolved in anhydrous *N*-butanol and treated with palladium acetate, triphenylphosphine, and potassium carbonate. The reaction mixture was heated under nitrogen at 100 °C with stirring for 16 hours. The reverse sulfonamide of ICT13069 was expected but this reaction only gave the carboxylic acid **65** as the product in a moderate yield of 64%, no signs of the methyl ester product was observed. Simple esterification of **65** gave **66**, the reverse sulfonamide of ICT13069. For the synthesis of **67**, the methyl amide analogue, **65** was treated with thionyl chloride, the acid chloride crude was treated with methylamine in ethanol by using an acetone ice bath, and the reaction was left for 24 hours in a freezer, the same way **42** was synthesized.

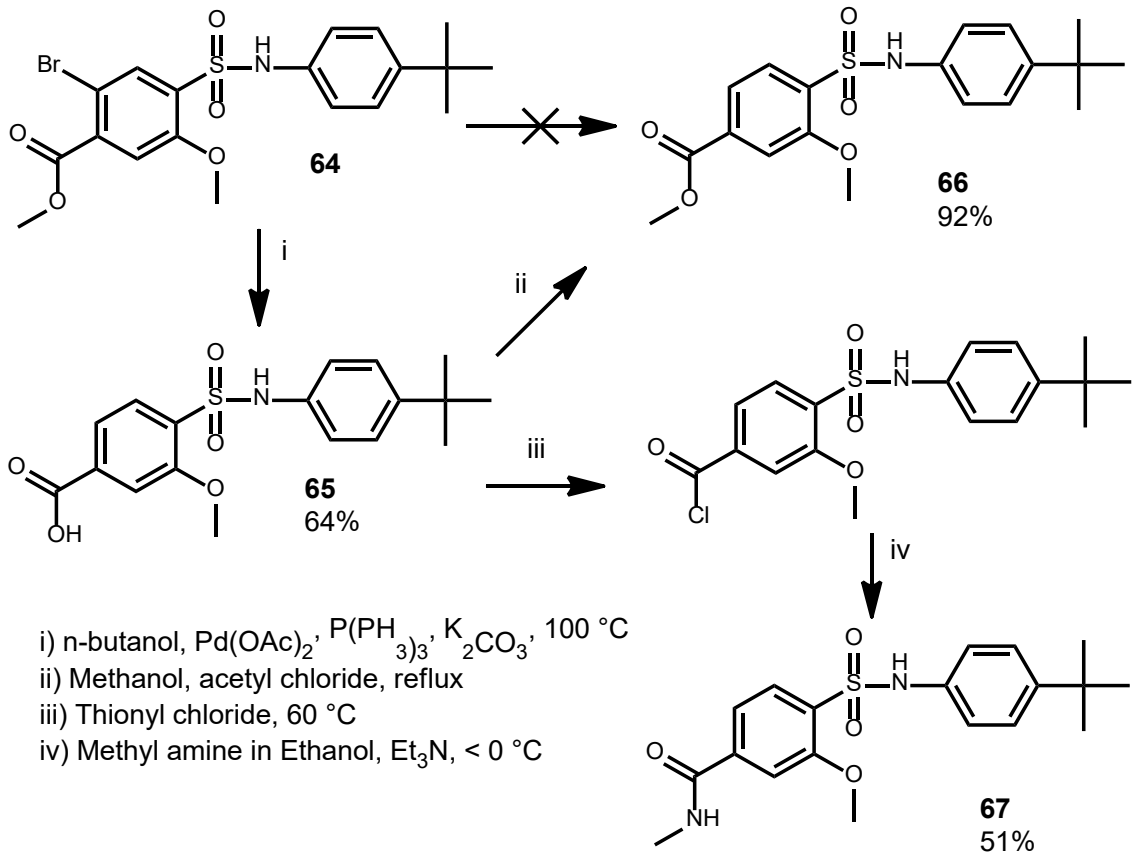
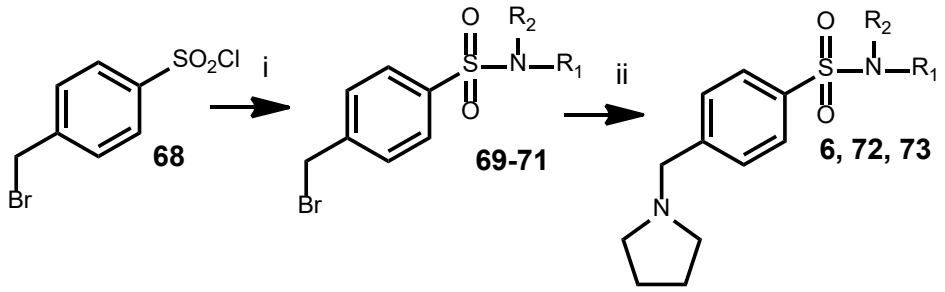


Figure 40: *Synthesis of 65-67*

This dehalogenation method wasn't found until the very late stage of this research, therefore the number of analogues synthesised through this method was limited to only two compound, **66** and **67**. Nonetheless this should have been sufficient to determine whether these "reverse" sulfonamide series CCR7 antagonists are more potent, before embarking on the synthesis of more analogues.

4.4 Synthesis of CXCR4 and dual CXCR4/CCR7 antagonists

As indicated previously (**Figure 23**), we planned to prepare a series of molecules with hybrid structures of CCR7 and CXCR4 antagonists. The sulfonamide CXCR4 antagonists from Mooring et al. [80] only has data on binding affinity and Matrigel assays. It was then decided to remake compound **6**, the most active sulfonamide CXCR4 antagonist from Mooring et al. [80] plus a series of analogues so that their potency can be determined in our own assays.



i) CH_2Cl_2, R_1NR_2H r.t., ii) pyrrolidine, potassium carbonate, acetonitrile

Figure 41: *Reaction scheme of **6, 72** and **73***

Compound	R_1	R_2	Yield
69 **6**	Benzyl	CH_3	72% 39%
70 **72**	Benzyl	H	68% 35%
71 **73**	4-*tert*-Butyl	H	71% 42%

The synthesis involves 4-(bromomethyl)benzenesulfonyl chloride, **68**, dissolved in CH_2Cl_2 and treated with *N*-methyl-*N*-benzylamine, benzylamine and 4-*tert*butylamine, to afford compounds **69-71**. The reaction mixture was left to stir at room temperature overnight (**Figure 41**). The sulfonamide product was purified, then dissolved in acetonitrile, pyrrolidine (1 eqv.) and potassium carbonate (2 eqv.) were added. The reaction mixture was left to stir at room temperature overnight to give the putative CXCR4 antagonists **6, 72 and 73**. Two novel analogues of this compound were made with 4-*tert*-butylaniline, and benzyl amine instead of *N*-methylbenzylamine. The reaction with benzylamine forms

large amount of white precipitate during the reaction. A large quantity of solvent is needed to prevent the reaction mixture from turning into white precipitate.

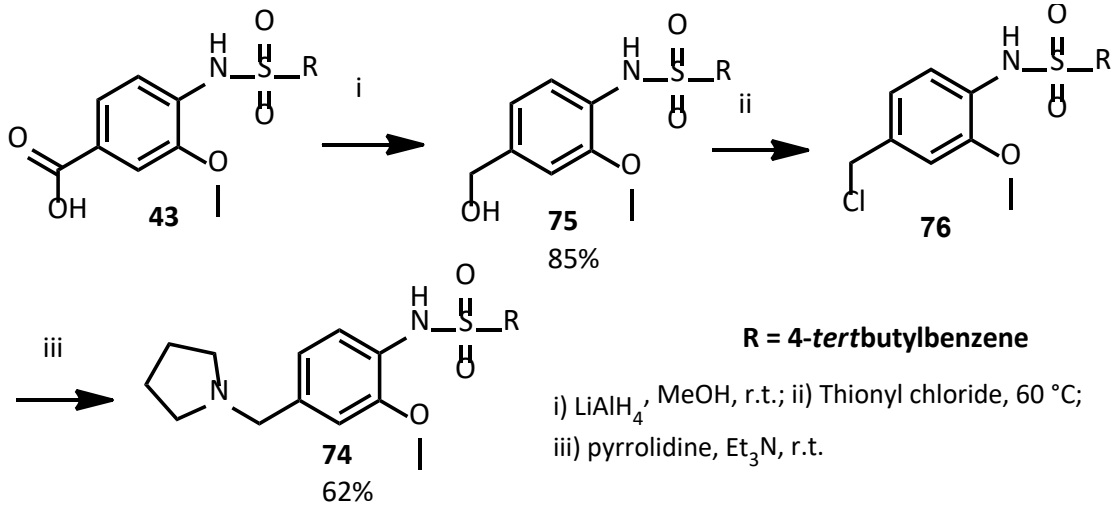
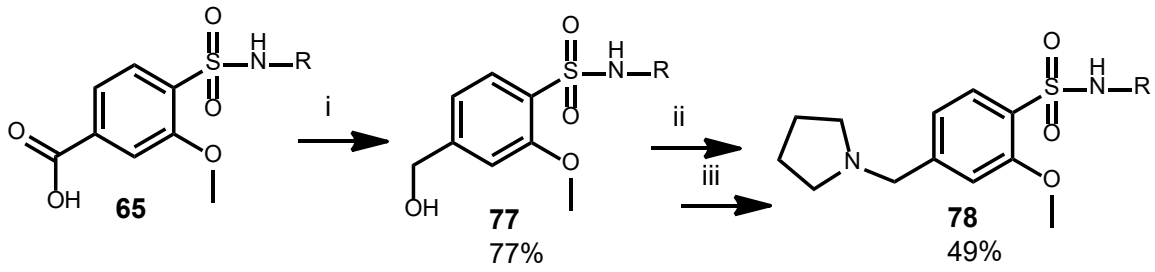


Figure 42: *Synthesis of **74***

Compound **74** was one of the analogues that incorporate methylene pyrrolidine substituent from the sulfonamide CXCR4 antagonists to the CCR7 antagonists of this research. The carboxylic acid compound **43** was reduced to the hydroxylmethyl **75** by using lithium aluminium hydride. **75** was treated with thionyl chloride to synthesise the benzyl chloride **76**. After removal of the excess thionyl chloride, excess pyrrolidine was added to the crude material. Excess pyrrolidine ensures any remaining thionyl chloride was neutralized, the reaction mixture was stirred at room temperature for 16 hours to give **74** after work up and purification (**Figure 42**).

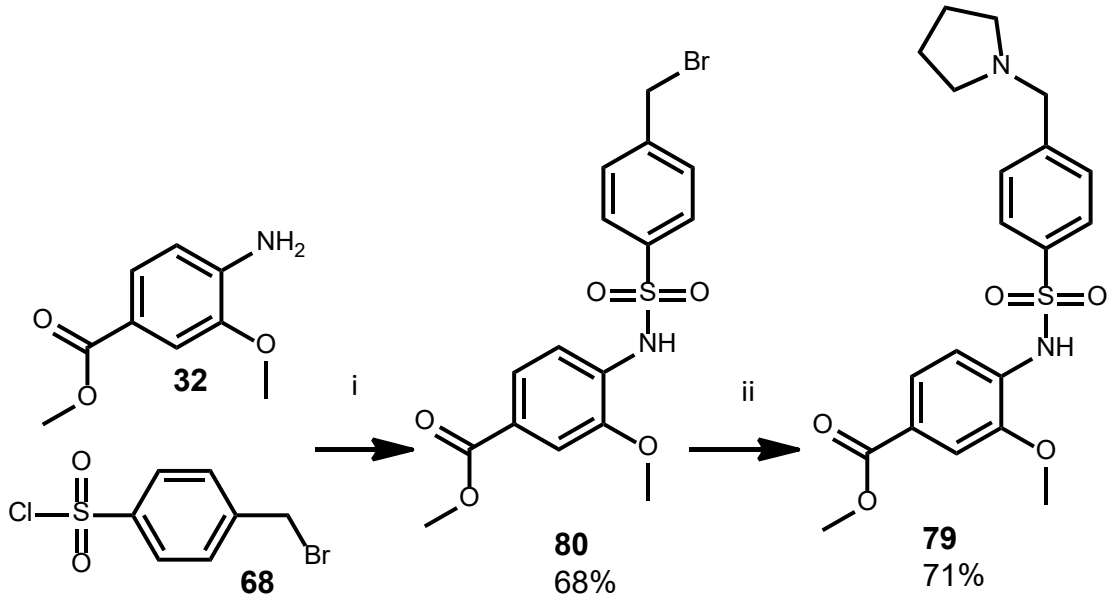
The reverse sulfonamide analogue of **74** was synthesized by using compound **65** as the starting material (**Figure 43**), compound **77** was treated with the same conditions to give **78** in moderate yield after purification.



i) $LiAlH_4$, MeOH, r.t.; ii) Thionyl chloride, 60°C;
iii) pyrrolidine, Et_3N, r.t.

Figure 43: *synthesis of **78***

Compound **79** (**Figure 44**) is another compound with the methylene pyrrolidine group incorporated to CCR7 antagonists. The synthesis starts with treating **32** and **68** with pyridine (3 eqv.) in CH_2Cl_2**,** which forms the sulfonamide product **80**. Without purification, the sulfonamide crude product was treated with potassium carbonate and pyrrolidine in acetonitrile to give **79** (**Figure 47**).



i) pyridine, CH_2Cl_2, r.t.; ii) acetonitrile, potassium carbonate, pyrrolidine, r.t.

Figure 44: *Synthesis of **79***

Rearrangement reaction with directed ortho metalation reaction

During this project, our group set up a collaboration with Dr William Martin (University of Bradford) to explore a synthesis to biarylsulfonamides where the *ortho* position was substituted with a carbomethoxy group. The Martin group had earlier shown that directed *ortho* metalation of biaryl sulfonamides where the nitrogen bears a formate group results in a rearrangement to give 2-carbomethoxy sulfonamides and that this protocol is more efficient than trapping the *ortho*-lithiated sulfonamides with a chloroformate [105] (**Figure 45**). Thus, the synthetic route starts with adding a chloroformate compound to the nitrogen of the sulfonamide to create the tertiary sulfonamide compound, then uses the directed *ortho* metalation reaction followed by rearrangement reaction to synthesize the target compounds.

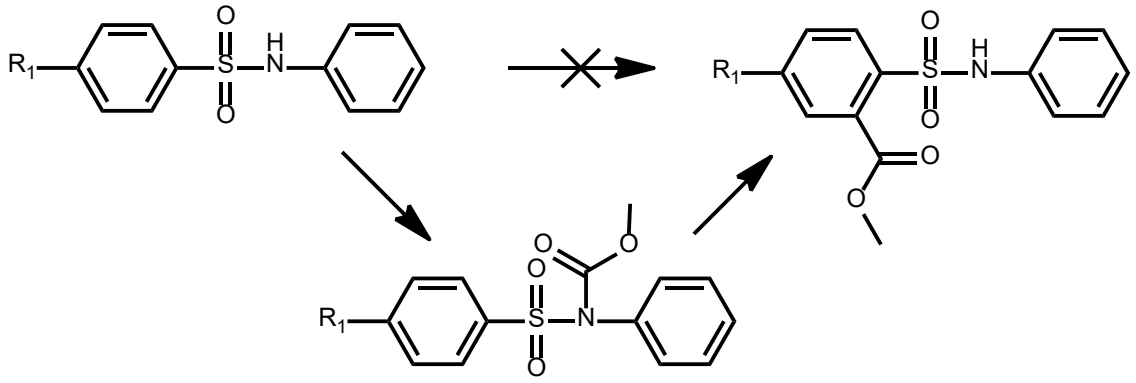


Figure 45: *The general reaction scheme of rearrangement reaction which utilize directed ortho metalation*

This rearrangement reaction was shown to be an intramolecular reaction (**Figure 46A**) by Dr William Martin's group. Two different sulfonamide compounds with different ester groups were mixed into a single solution. The rearrangement reaction was carried out on the mixture. After isolating and analysing the products, only two rearrangement products were found, which indicated an intramolecular reaction. If the reaction was an intermolecular reaction, there would be rearrangement products that possess an ester different from the starting material. The proposed mechanism of the intramolecular reaction is shown below (**Figure 46B**)

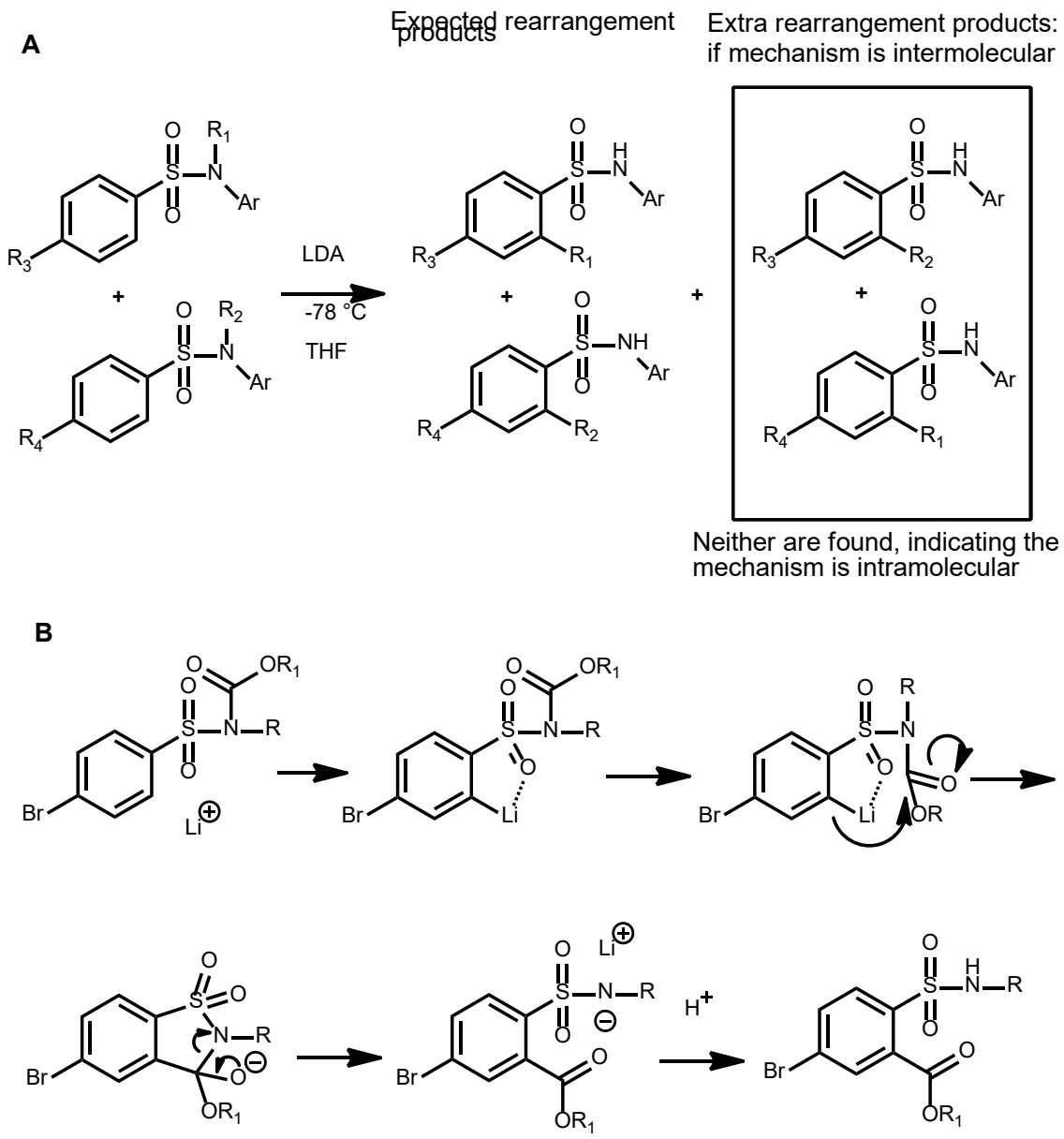
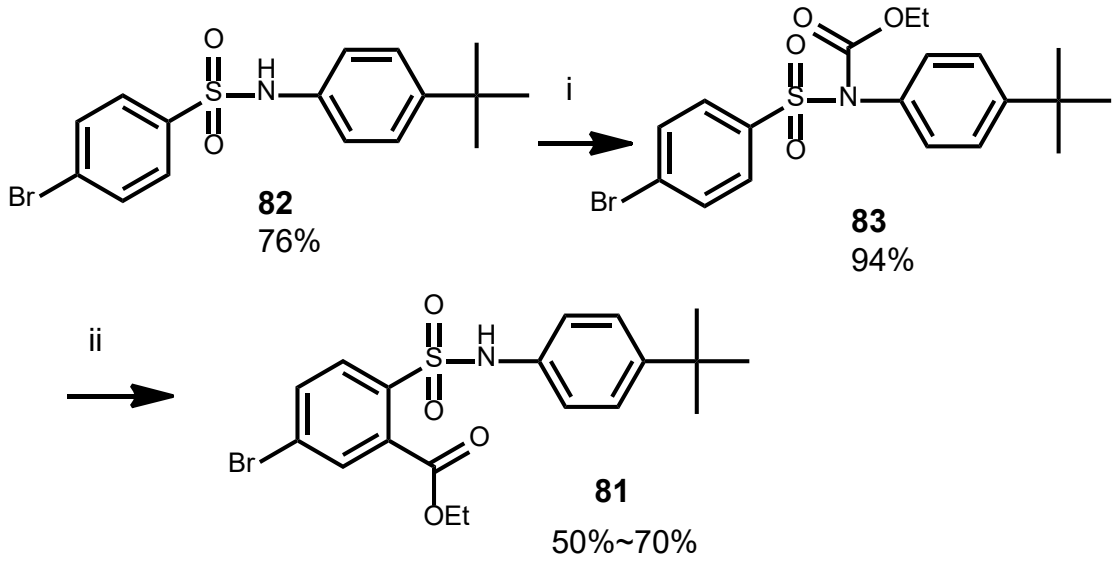


Figure 46: *A) experiment to test whether the rearrangement reaction mechanism was intra- or inter- molecular. Intra-molecular reaction would only give two products. Inter-molecular reaction would give more than two. B) Proposed reaction mechanism of the rearrangement reaction.*



i) THF, 60% NaH, ethyl chloroformate, r.t. ii) LDA, THF, -78 °C

Figure 47: *synthesis of **81***

This reaction was first trialled by attempting the synthesis of **81** (**Figure 47**). For the formation of the tertiary sulfonamide, compound **82** was treated with ethyl chloroformate, 60% sodium hydride and THF at room temperature for an hour. **83** was obtained in high yield (94%) and taken to the next step directly. The condition for the directed *ortho* metalation rearrangement reaction is similar to an orthodox directed *ortho* metalation reaction. Compound **83** was dissolved in anhydrous THF, the stirred solution was maintained under anhydrous nitrogen, and kept at -78 °C by using an acetone and dry ice bath. 5 equivalents of lithium diisopropylamide in THF (LDA) was freshly prepared and then transferred to the flask containing compound **83** with a syringe or a needle cannula depending on the scale of the reaction. This reaction goes to completion in approximately 3 minutes. Afterwards the reaction was worked up by addition of dilute aqueous acid (dilute aqueous citric acid in this procedure) immediately or the rearrangement product will be lost. The yield of this reaction can be difficult to reproduce, and was generally between 50% to ~70%. This reaction was found to be very time sensitive, and as the timing of transferring the LDA can be difficult to control and this is likely the cause for the reproducibility.

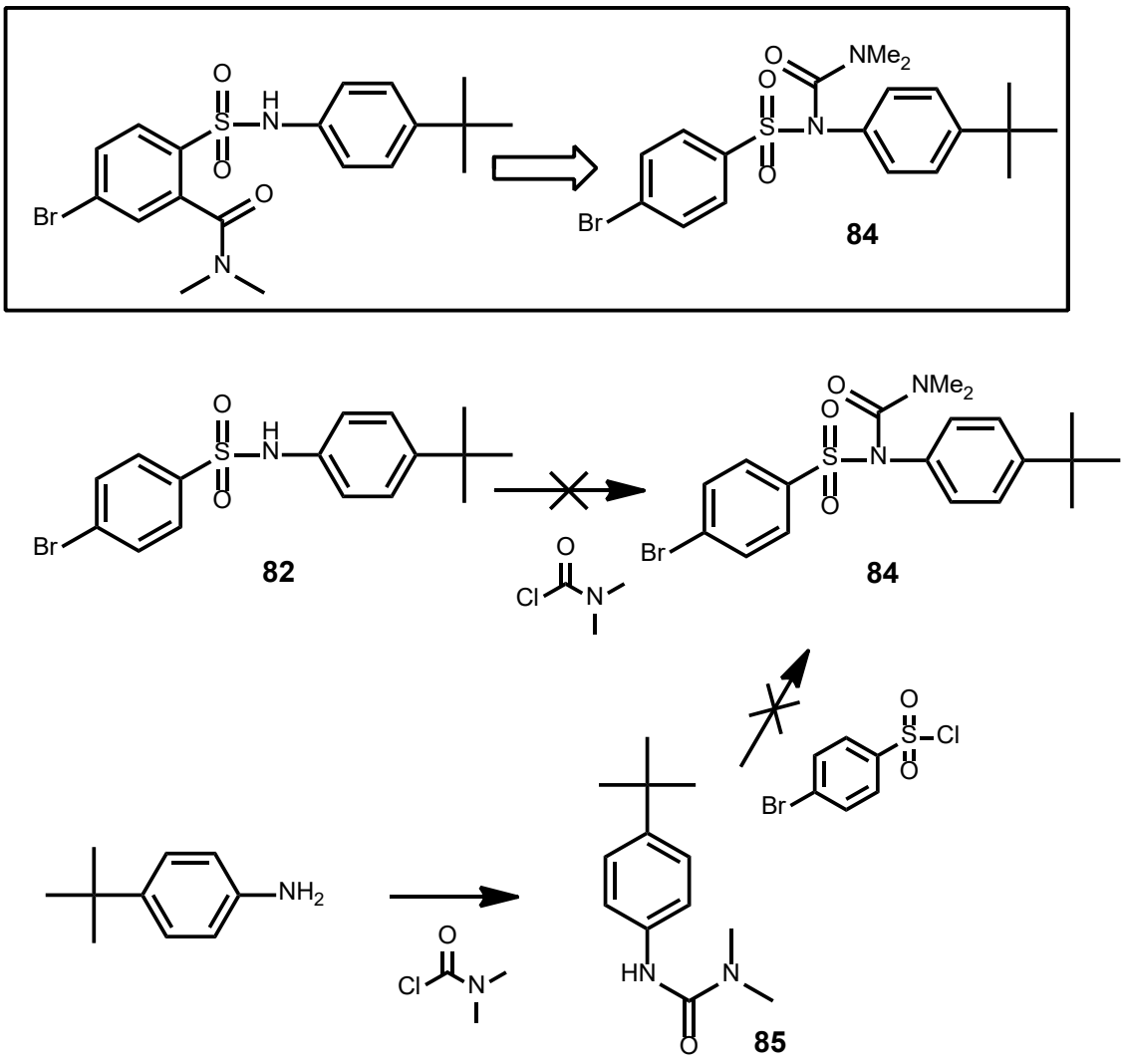


Figure 48: *Attempted synthesis of **84***

Attempts were made to synthesize **84** (**Figure 48**), as the rearrangement reaction of biphenyl sulfonylurea compound had not been attempted. Unfortunately, all attempts to synthesize compound **84** failed. These included treating **82** with dimethylcarbamoyl chloride under basic conditions. The following basic conditions have been attempted: 1) K_2CO_3, acetonitrile (r.t./reflux), 2) THF, NaH (r.t./reflux), 3) dimethylformamide, NaH (r.t./100 °C/reflux). The second approach was synthesising **85**, then treating it with 4-bromobenzenesulfonyl chloride under basic conditions. Compound **85**, 1-(4-(*tert*-butyl)phenyl)-3-methylurea was synthesized by treating 4-*tert*buylaniline with dimethylcarbamoyl chloride, and stirring at room temperature without solvent. The basic conditions attempted were 1) pyridine, r.t./100 °C, and 2) dimethylformamide, NaH (r.t./100 °C/reflux), but these failed to give the desired compound.

Finally, we set out to use the rearrangement reaction on the previously made CXCR4 antagonists (compound **72** and **73**) from Mooring et al. [80] to synthesize possible dual CXCR4 and CCR7 antagonists (compounds **86** and **87, figure 49**).

For the synthesis of **88** and **89**, the tertiary sulfonamides required for the rearrangement reaction, we used the reaction conditions previously employed for **82** (ethyl chloroformate, and sodium hydride in THF) on **70** and **71**. This reaction condition did not work as intended and it resulted in removal and replacement of the pyrrolidine group to give the chloromethyl product compound **90.** This was confirmed by mass spectrometry. The reaction condition then was changed to the combination of acetonitrile and use of the weaker base potassium carbonate instead. The ethyl chloroformate (1 eqv) was added slowly over a period of an hour. Under this reaction condition, we were able to synthesize the tertiary sulfonamide compounds, but still obtained the methylene chloride compound as a minor by-product. The poor yields in these reactions were partly because after 24 hours of reaction, substantial quantities of the starting material had remained. Additional ethyl chloroformate may have increased the yield, but due to the concurrent formation of the methylene chloride compound in this reaction, it was decided that it was better to recover the starting material and the product instead of allowing more of the methylene chloride product to be synthesized. The rearrangement reaction utilized the same reaction conditions used for **82** and worked as expected to give a moderate yield of the desired compounds (**Figure 49**).

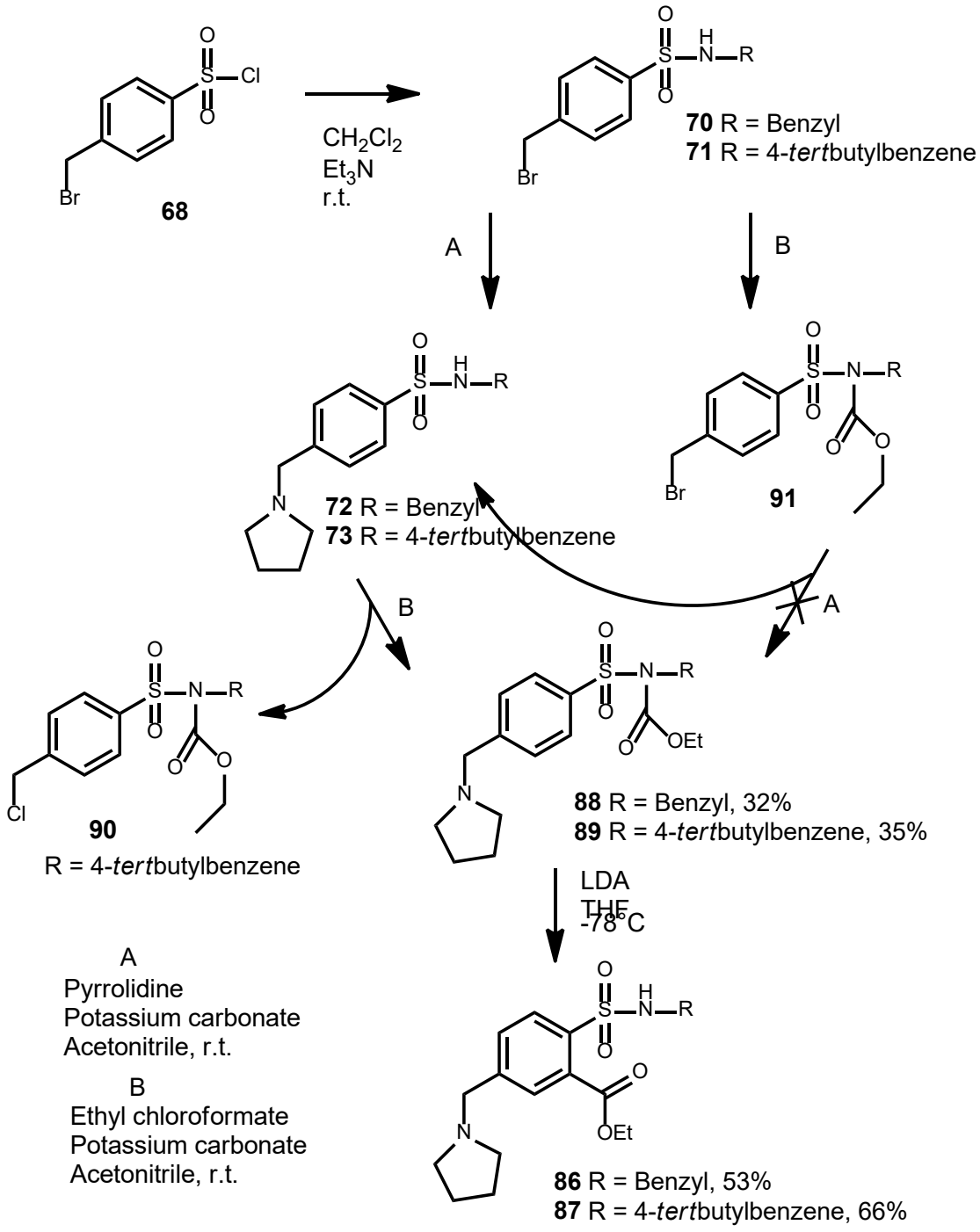


Figure 49: *Synthesis of **86** and **87***

To obtain a more efficient synthesis of target compounds, an alternative synthesis route was attempted which avoid the synthesis of **91** (**Figure 49**). This route involves synthesizing the tertiary sulfonamide **91** first and then adding the pyrrolidine afterward. However, this route was unsuccessful because treating **91** with pyrrolidine removes the ethyl ester from the sulfonamide and give **72**/**73** as the only products.

4.5 Synthesis of reverse sulfonamides with modifications at positions A and C

Having successfully developed a synthesis of the "reverse sulfonamide" series, we returned to investigating modifications in the A and C fragments as outlined below (**Figure 50**).

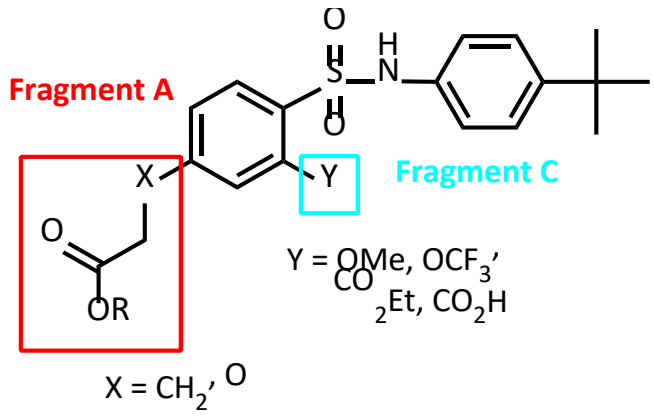
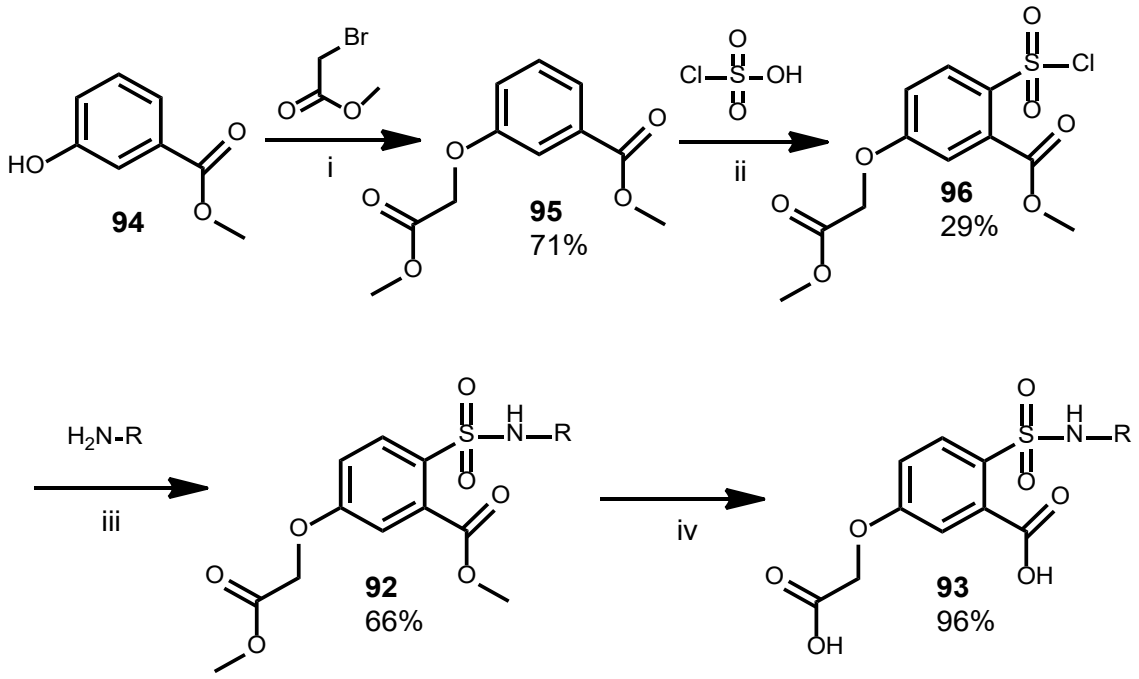


Figure 50: *Reverse sulfonamides with modifications at positions A and C.*

The above reaction scheme was designed to take advantage of the selectivity of the sulfonation reaction to synthesize reverse sulfonamide compounds that are suitable for testing. For the synthesis of **92** and **93 (Figure 51)**, methyl 3-hydroxybenzoate, **94**, was dissolved in tetrahydrofuran, followed by 60% sodium hydride and methyl 2-bromoacetate. The reaction mixture was heated at 60 °C at room temperature with stirring for 16 hours to give **95**. After purification, **95** was treated with CH_2Cl_2 and 3 equivalents of chlorosulfonic acid and stirred at room temperature for two days to give **96** in low yield (29%) as usual. As expected, the NOSEY NMR spectrum of **96** confirms that the *para* product was the major isomer. The *ortho* product was not observed. The product was taken further to react with 4-*tert*butyaniline to give the sulfonamide **92**. Lastly, **92** was hydrolysed to give **93 (Figure 51).**



i) THF, NaH, 60 °C, ii) CH_2Cl_2, r.t., iii), Pyridine, DCM, r.t., iv) NaOH, H_2O, 100 °C

R = 4-*tert*butylbenzene

Figure 51: Synthesis of **92** and **93**

Compounds made with Heck coupling reactions

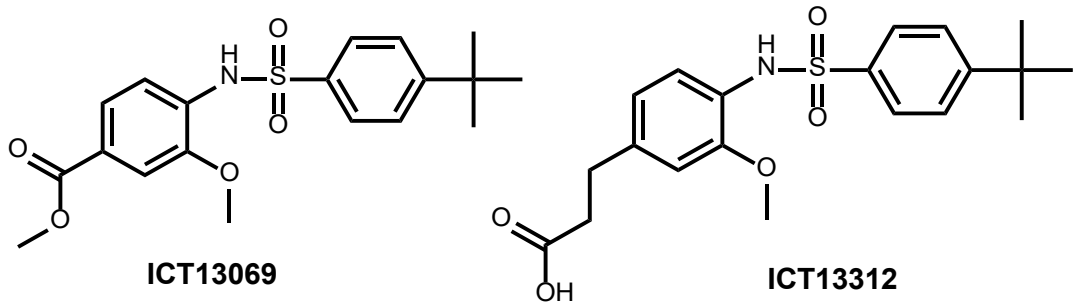


Figure 52: *Structure of ICT13312*

During this research project, another member of our group had synthesized ICT13312 (**Figure 52**), an analogue of ICT13069 with a chain extension at the methyl ester position. ICT13312 was shown to be more potent than ICT13069 in the calcium flux assay, hence we decided to prepare three reverse sulfonamide analogues of ICT13312, compounds **96**, **97** and **98** (**Figure 53**). The synthesis for the 3 compounds started with 4-bromo-2-(trifluoromethoxy)benzenesulfonyl chloride **99**. It was the only commercially available sulfonyl chloride compound that can be used in the synthesis reverse sulfonamide analogues of ICT13312 in

minimal number of steps. **99** was treated with 4-*tert*butylaniline, CH_2Cl_2 and pyridine (3 eqv.) to synthesise **100**. For the Heck coupling reaction, **100** was dissolved in the minimum amount of DMF. Palladium (II) acetate and triphenylphosphine were used for the palladium catalyst, triethylamine, methyl acrylate or *N, N*-dimethylacrylamide were added. The reaction mixture was heated at 100 °C under nitrogen with stirring for 16 hours. After purification, both the *trans* and *cis* isomers **101a** and **101b** could be isolated as a mixture. Both isomers underwent hydrogenation together by using 10% palladium on carbon and methanol to give **96** and **97** in moderate yield (48%). The crude of **96** was hydrolysed in 3M sodium hydroxide and heated at 100 °C for 2 hours, after neutralisation and purification. **98** was obtained from this synthetic route in moderate yield (51%).

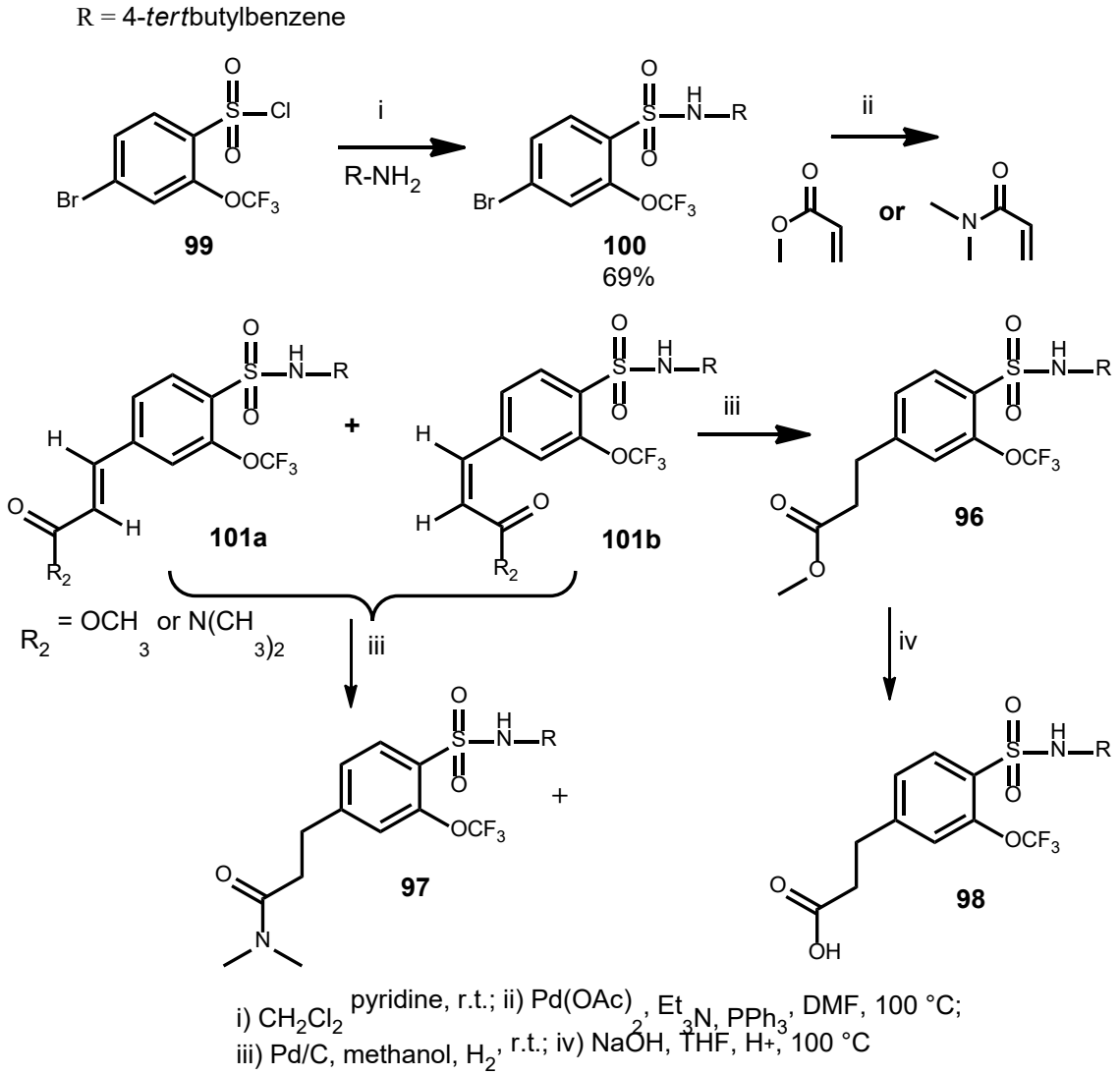
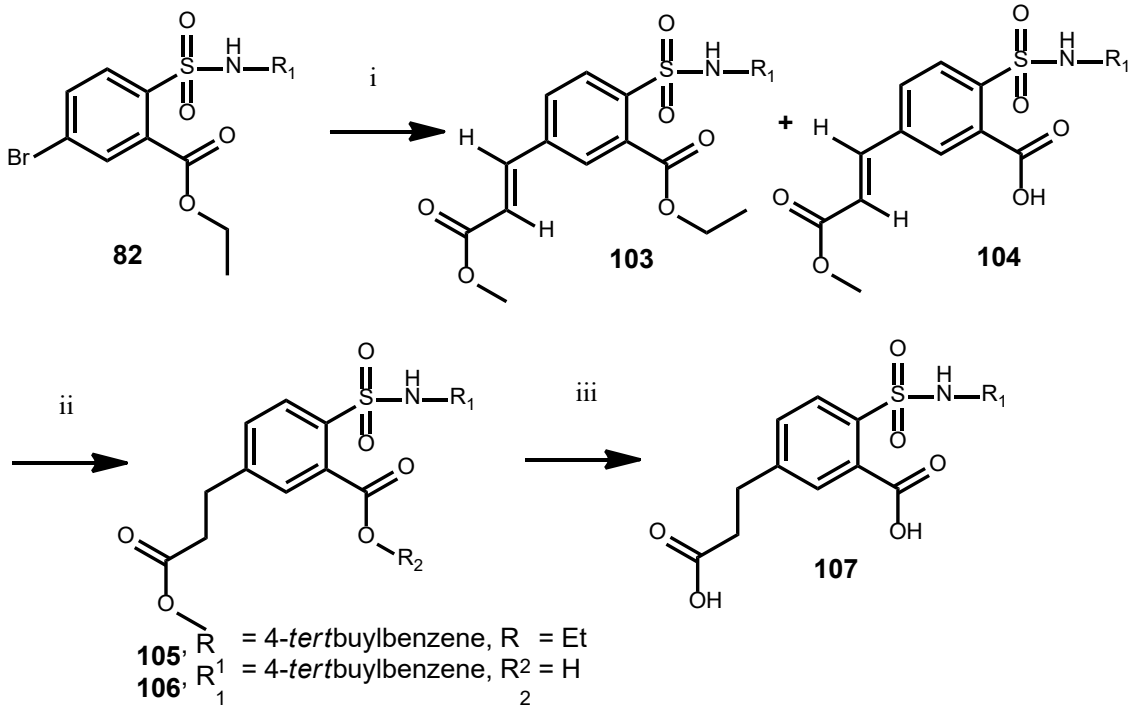


Figure 53: *Synthesis of **96**-**98***

Compound **82** from the rearrangement reaction was also used to synthesize reverse sulfonamide analogue of ICT13312 (**Figure 54**). The same Heck coupling reaction condition used previously was employed for **82**. After purification, the two largest fractions were analysed. One was the expected *trans* isomer, **103**, the other was expected to be the *cis* isomer but it was the Heck coupling product with its ethyl ester hydrolysed to a carboxylic acid, **104**. Surprisingly, the Heck coupling reaction had hydrolysed the ethyl ester but not the methyl ester, although the reason for this is unclear. A similar hydrolysis of a methyl ester to carboxylic acid was previously observed by us during palladium catalysed dehalogenation of compound **64** (**Figure 40**). This observation suggests that the conversion occurs during the Heck reaction and the conversion ceased after the completion of reaction. A small investigation was carried out by allowing the reaction to continue for a week. if the conversion can occur after the formation of the product due to general condition of the reaction, the major Heck coupling products wouldn't have the methyl ester. The methyl ester is still present after a week, suggesting the general condition of the reaction cannot change the methyl ester into the carboxylic acid after the completion of the reaction. This is not directly relevant to the overall aims of project, so this wasn't studied further.



i) $Pd(OAc)_2$, Et_3N, PPh_3, methyl acrylate, DMF, 100 °C, ii) Pd/C, methanol, H_2, r.t.; iii) NaOH, 100 °C

Figure 54: *Synthesis of **107***

The *trans* isomers **103** and **104** are likely both present, and instead of trying to isolate them, it was decided that it is preferable not to purify the Heck coupling and hydrogenate the crude product instead. After the hydrogenation to a mixture of **105** and **106** and the subsequent hydrolysis and neutralisation, the crude was finally purified by flash column chromatography to give compound **107** in moderate yield (52%) from the synthetic route. A mechanism for the Heck coupling reaction is proposed below (**Figure 55**).

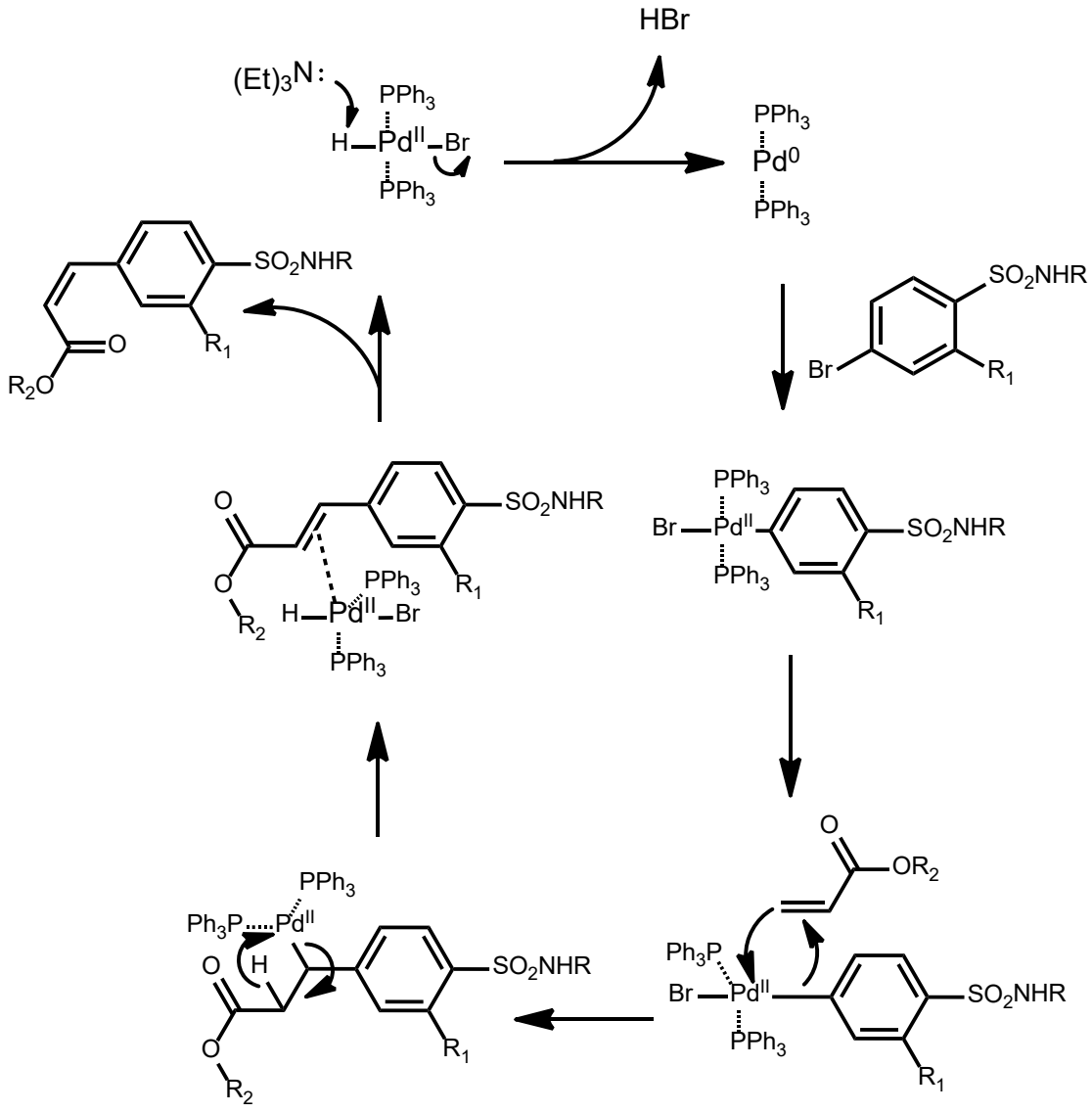


Figure 55: *Mechanism of Heck couplng reaction*

Miscellaneous reverse sulfonamide

Compound **108** and **109** (**figure 56**) aimed to investigate the effect of extra chain length between the sulfonamide and the benzene ring. 2-Methoxy-4-nitrobenzenesulfonyl chloride was treated with benzylamine or phenyl hydrazine in the presence of pyridine. The reaction with benzyl amine (2 eqv.) to give compound **108** was straightforward. The reaction with phenylhydrazine was troublesome. The reaction generated many side products which made the yield of **109** very low and the product difficult to purify. In addition, the compound has very poor solubility even in methanol. Hydrogenation was carried out on this compound to convert the nitro group into an amine (compound **110**), the solubility slightly increased as expected from a nitro group to an amine, but the solubility is still very poor. Both compound **109** and **110** are not likely to be soluble enough to be used in biological testing, therefore both were abandoned.

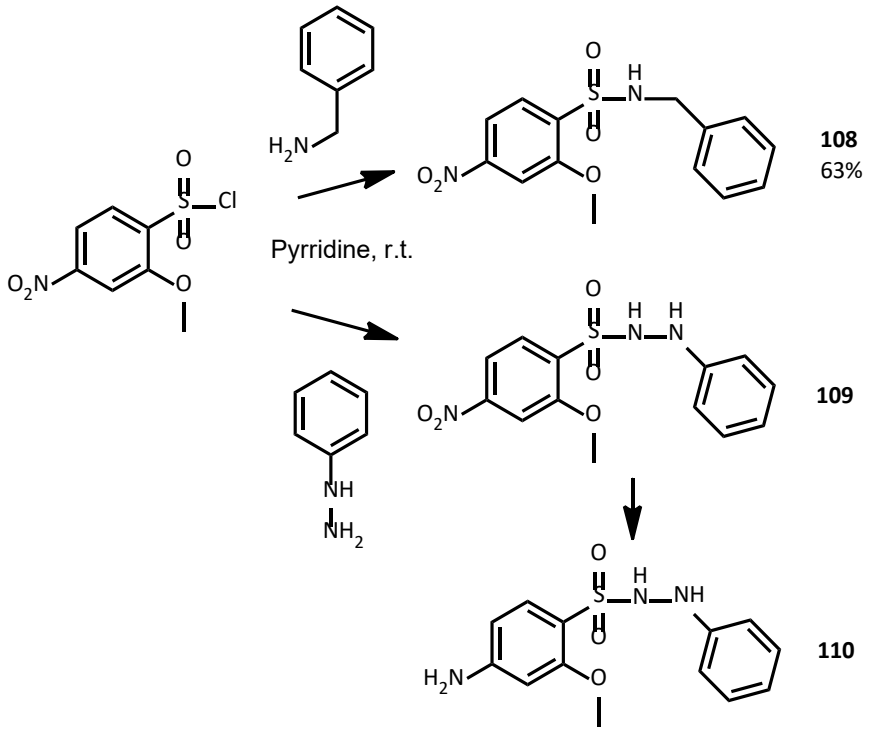


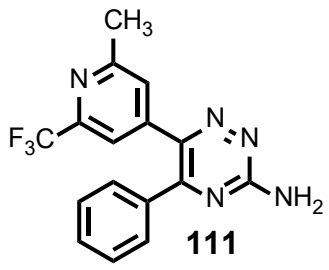
Figure 56: *Synthesis of **108-110***

5.0 Molecular docking

5.1 Introduction

In drug discovery, molecular docking is an *in-silico* method to the predict the binding mode of a drug ligand when bound to the target protein, in order to provide an understanding of the molecular interactions between the drug molecule with its target protein at an atomic level. Molecular docking aims to achieve this by generating a series of docking poses, and calculate docking scores to the docking poses. The term binding mode refers to the conformation and the orientation of a ligand when bound to the receptor protein, docking poses refer to the *in-silico* generated predictions of the binding mode. Docking score refers to a quantitative method of determining the binding free energy of a docking pose. As mentioned in chapter 2.4.3, the docking scores are used in a virtual ligand screening campaign to rank a library of compounds to identify the lead molecules.

Molecular docking allows the characterisation of the docking poses and its interactions with the residues within the receptor, this information coupled with the docking score is used for determining which of the docking poses are most likely to be the accurate binding mode. The inspection of the binding site can provide information which can aid rational drug design, and the optimisation of the drug molecule. One such example is the development of a potent small molecule antagonist for the GPCR adenosine A_{2A} receptor ($A_{2a}R$). Congreve et al. [106] used homology modelling in their virtual screening campaign to identify 1,2,4-triazine derivatives as their lead compounds. Their binding mode revealed a region of the orthosteric site which is normally occupied by adenosine, the natural ligand of $A_{2a}R$. Expanding the lead compound 1,2,4-triazine derivatives into this region gave a more potent series of antagonists. The predicted binding mode was validated when the authors produced the X-ray crystal structures of $A_{2a}R$ in complex with their compounds. Further SAR study gave rise to compound **111** (**Figure 57**).



111

Figure 57: *Structure of **111,** potent adenosine A_{2A} receptor antagonist [106].*

For this project, we attempted to use molecular docking to elucidate the binding mode of novel sulfonamide antagonists of CCR7 and CXCR4. We hoped this may give insights as to how the sulfonamide and other moieties interact with the receptor. This could possibly lead to improving the efficacy of the antagonists. This project also attempted to find correlation between data generated from molecular docking and the experimental data (IC_{50} from calcium flux assay), both to validate the model and as a tool to predict the potency of the antagonists. The software Molecular Operating Environment (MOE) 2015.10 and the online server tool SWISS MODEL were used for the docking studies.

The correlation is typically drawn between the predicted and the experimental binding affinity, but the antagonists of this project are not assessed with binding affinity so no experimental binding affinity data are available. However, Kataria et al. [107] created a correlation plot with the IC_{50} of 13 small molecule antagonists and their docking scores. The target protein was a urease. The correlation indicated that compound with greater docking scores had better IC_{50}, the R^2 was 0.7, confirming a strong correlation and validates their docking pose. All 13 molecules had the same pharmacophore; which led to the assumption that IC_{50} and docking score correlation may predict the potency of structurally similar compounds based on their docking score.

5.2 Construction of homology models vs crystal structures

Molecular docking requires the 3D structure of the target protein. CXCR4 has several X-ray crystals structures, 3ODU has the best resolution at 2.5 Å, thus it was chosen for the docking study. The X-ray crystal structure of CCR7 wasn't solved for most of the duration of this project, therefore a homology model of CCR7 was constructed. Homology modelling predicts the unknown structure of a protein based on the assumption that a related protein with similar amino acid sequence will have similar 3D structure. Homology modelling requires a template, which is ideally a solved structure with high sequence identity with the target protein. The reliability and accuracy of the predicted structure are highly dependent on the sequence similarities/identity. In the case of CCR7, the online server BLAST (Basic Local Alignment Search Tool) [108] identified human CCR9 (protein database code: 5LWE, resolution: 2.80 Å [109]) to have the highest sequence identity of 43%, indicating it is the most suitable template.

A

```
CCR7_HUMAN   1   ---MDLGKPMKSVLVVALLVIFQVCLCQDEVTDDYIGD-----NTTVDYTLFESLCSKKD   52
CCR9_HUMAN   1   MTPTDFTSP-----------------IPNMADDYGSESTSSMEDYVNFNFTDFYCEKNN   42
                    *: .*                    ::**** .:     :  *::.: :  *.*::

CCR7_HUMAN  53   VRNFKAWFLPIMYSIICFVGLLGNGLVVLTYIYFKRLKTMTDTYLLNLAVADILFLLTLP  112
CCR9_HUMAN  43   VRQFASHFLPPLYWLVFIVGALGNSLVILVYWYCIRVKTMTDMFLLNLAIADLLFLVTLP  102
                 **:* : *** :* :: :** ***.**:*.* * .*:***** :*****:**:***:***

CCR7_HUMAN 113   FWAYSAAKSWVFGVHFCKLIFAIYKMSFFSGMLLLLCISIDRYVAIVQAVSAHRHRARVL  172
CCR9_HUMAN 103   FWAIAAADQWKFQTFMCKVVNSMYKMNFYSCVLLIMCISVDRYIAIAQAMRAHTWREKRL  162
                 *** :**..* * ..:**:: ::***.*:* :**::***:***:**.**: **  * : *

CCR7_HUMAN 173   LISKLSCVGIWILATVLSIPELLYSDLQRSSSEQAMRCSLITEH---VEAFITIQVAQMV  229
CCR9_HUMAN 163   LYSKMVCFTIWVLAAALCIPEILYSQIKEESGIAI--CTMVYPSDESTKLKSAVLTLKVI  220
                 * **: *. **:**:.*.***:***::::.*.     *:::      .:   :: . :::

CCR7_HUMAN 230   IGFLVPLLAMSFCYLVIIRTLLQARNFERNKAIKVIIAVVVVFIVFQLPYNGVVLAQTVA  289
CCR9_HUMAN 221   LGFFLPFVVMACCYTIIHTLIQAKKSSKHKALKVTITVLTVFVLSQFPYNCILLVQTID  280
                 :**::*::.*: ** :**:**:**:: .::***:** *:*:.**:: *:*** ::*.**:

CCR7_HUMAN 290   NFNITSSTCELSKQLNIAYDVTYSLACVRCCVNPFLYAFIGVKFRNDLFKLFKDLGCLSQ  349
CCR9_HUMAN 281   AVAMFISNCAVSTNIDICFQVTQTIAFFHSCLNPVLYVFVGERFRRDLVKTLKNLGCISQ  340
                  : :  *.* :*.:::*.::** ::* .:.*:**.**.*:* :**.**.* :*:***:**

CCR7_HUMAN 350   EQLRQWSSCRHI-RRSSMSVEAETTTTFSP                               378
CCR9_HUMAN 341   AQWVSFTRREGSLKLSSMLLETTS-GALSL                               369
                  *  .::  .   : *** :*: :  ::*
```

B

Poor rotamers	5	1.79%	Goal: <0.3%
Favored rotamers	266	95.00%	Goal: >98%
Ramachandran outliers	4	1.28%	Goal: <0.05%
Ramachandran favored	299	95.83%	Goal: >98%
Rama distribution Z-score	1.56 ± 0.46		Goal: abs(Z score) < 2
Cβ deviations >0.25Å	3	0.99%	Goal: 0
Bad bonds:	3 / 2563	0.12%	Goal: 0%
Bad angles:	22 / 3481	0.63%	Goal: <0.1%
Cis Prolines:	0 / 6	0.00%	Expected: ≤1 per chain, or ≤5%

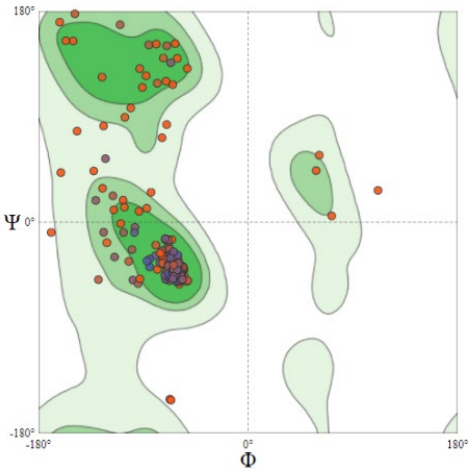


Figure 58: *Top) Alignment of CCR7 and CCR9 amino sequence. Bottom) The Ramachandran plot and other properties used for evaluating the CCR7 homology model based on CCR9/5LWE generated by SWISS MODEL.*

Initially, the online server tool SWISS-MODEL was used to generate the homology model. SWISS-MODEL also identified CCR9 (5LWE) as the most suitable template with 44.69% sequence identity. The model generated by SWISS-MODEL [110] was assessed by Ramachandran plot and the online server tool MolProbity, most of the assessed properties are below the ideal values. Ramachandran plot provided by MolProbity [111] shows 4 outliers and 95% residues at the favoured region (**Figure 58**). Similarly, MOE 2015 was used to generate several homology models. Assessment with Ramachandran plot and MolProbity [111] shows either slightly or significantly worse results, thus they were discarded.

In September 2019, 6QZH, the X-ray crystal structure of CCR7 with the resolution 2.10 Å was published by Jaeger et al. [63]. The CCR7 homology model was considered obsolete and 6QZH was used for the molecular docking studies. In the crystal structure of CCR7 (pdb code 6QZH), Cmp2105 was found at the intracellular allosteric site. The sulfonyl group in Cmpd2105 interacts with two residues at the C-terminus: Gly330 and Phe333, the two amine bridges interact with Asp94, and carbonyl interacts with Thr93 (**Figure 59A/B**). Jaeger et al. searched for compounds similar to Cmp2105 and found Navarixin, their docking study shows Navarixin is able to interact with Asp94 with its two amine bridges, but only one of the ketones interacts with Lys332 (**Figure 59C**). These multiple ligand interactions of Cmp2105 might explain why Cmp2105 is more potent than Navarixin in inhibiting arrestin binding in an assay as mentioned in chapter 1.3 [63].

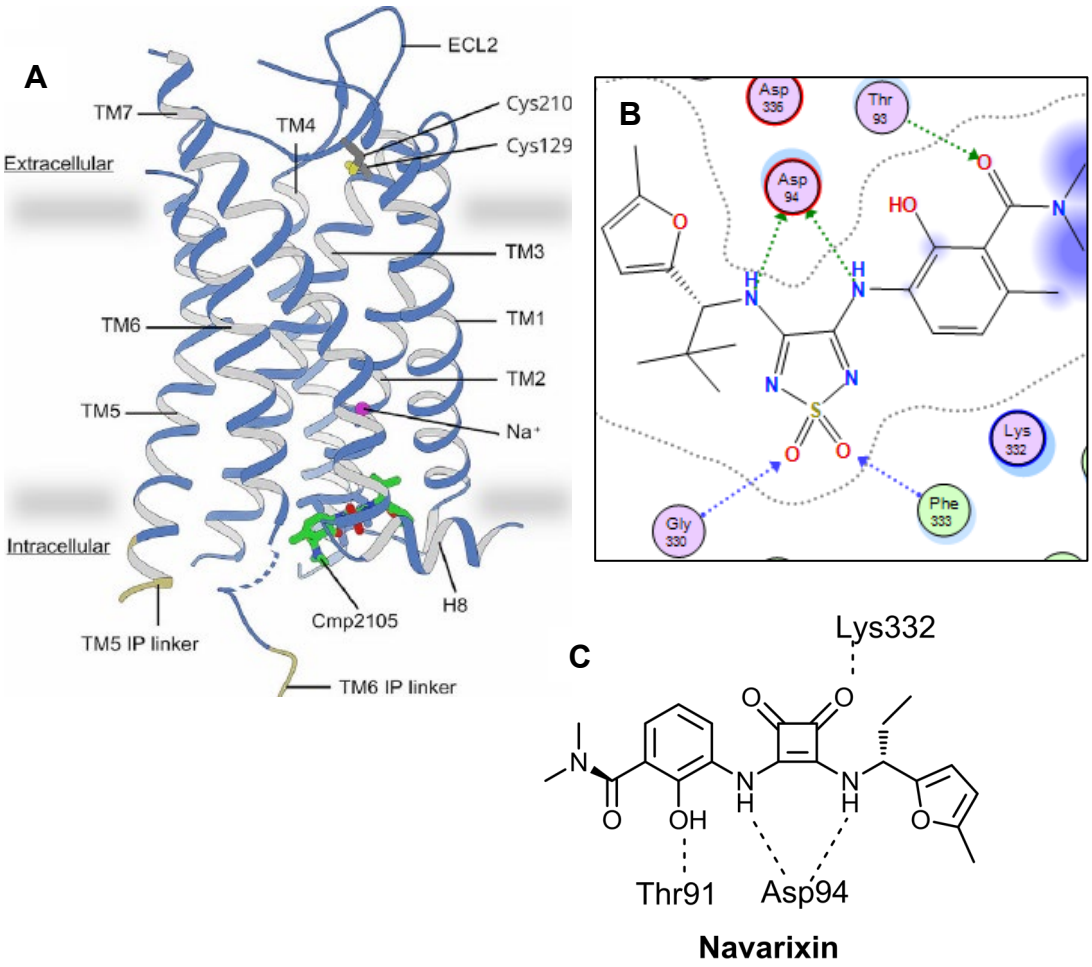


Figure 59: *A) Structure of CCR7/6QZH with Cmp2105 docked at the intracellular allosteric site [63], and its ligand interaction map generated by MOE. B) ligand interactions of Navarixin [63]*

This project had access to crystal structures of both CXCR4 (pdb code 3ODU) and CCR7 (pdb code 6QZH) to carry out the *in-silico* investigation. However, there is a fundamental difference between the two crystal structures. 3ODU contains a ligand bound in the orthosteric (transmembrane) domain, whilst in 6QZH, the ligand bound at the intracellular allosteric site rather than the orthosteric site. It is not clear if binding at the orthosteric site can squeeze the space in the allosteric site or *vice versa*.

Most of the currently approved drugs which target class A GPCRs (to which both CCR7 and CXCR4 belong), bind at the orthosteric site [112]. The traditional orthosteric site is close to the extracellular region of the receptors, between a highly conserved residue in transmembrane helix 6 (residue 6.48) and extracellular loop 2. Chan et al. superposed all orthosteric ligands binding mode observed in the crystal structure for all class A GPCRs [112] (**Figure 60**). Very few class A GPCR small molecule ligands bind at the extracellular or the

intracellular allosteric site. Two of such extracellular binding ligands are MRS2500 [113] and LY2119620 [114]. Neither share similarities to the antagonists from this project. Another intracellular binding ligand besides Cmp2105 is vercirnon, a CCR9 antagonist bound at the X-ray crystal structure of CCR9 (5LWE), which has significant homology to CCR7. Vercirnon is a compound with 4-*tert*butylbenzene sulfonamide similar to the compounds of our group (**Figure 61**). This raises the question of whether the sulfonamide CCR7 antagonists in this project bind to the intracellular allosteric site. Therefore, it is worth investigating the molecular docking of the CCR7 antagonists of this project to both the orthosteric and intracellular allosteric site.

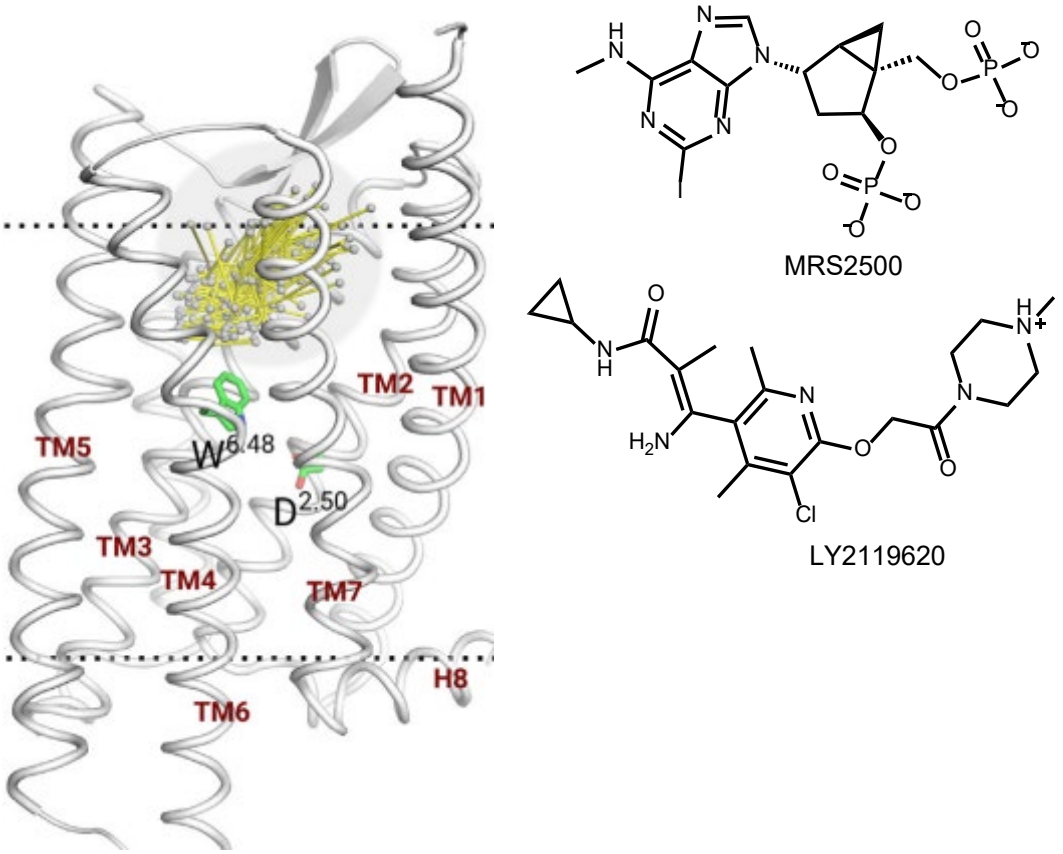


Figure 60*: A) superposition all orthosteric ligands binding mode observed in the crystal structure for all class A GPCRs [112]. B) structure of MRS2500 and LYS2119600.*

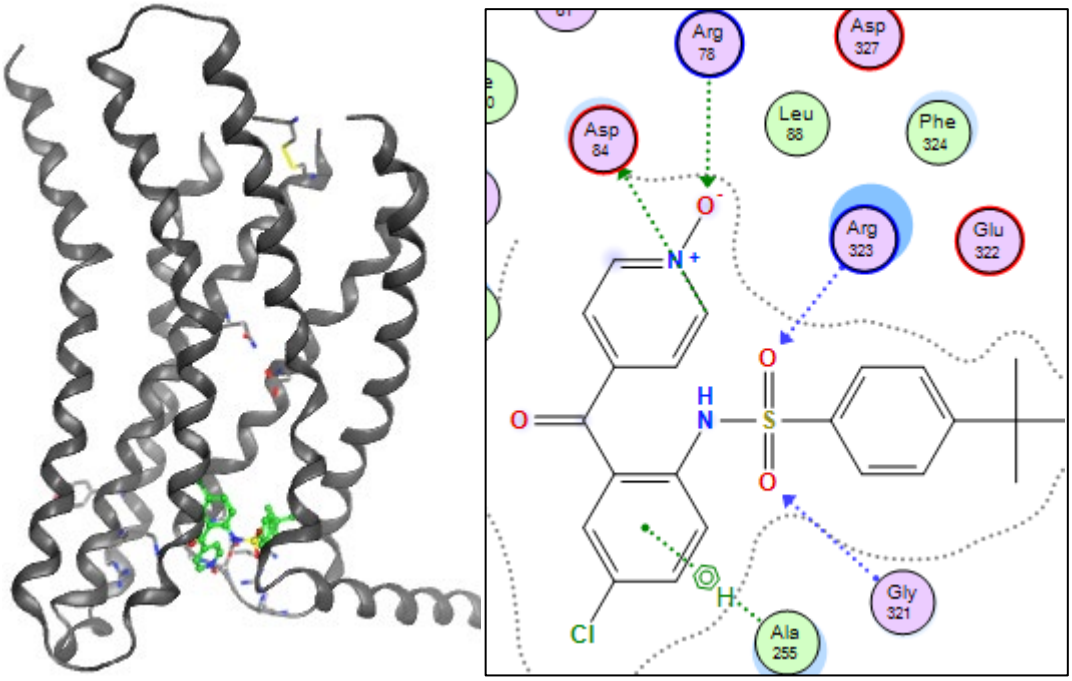


Figure 61: *CCR9 5LWE Vercirnon bound to the intracellular allosteric site, and ligand interaction map of Vercirnon at the allosteric binding site, generated by MOE.*

5.3 Preparation for molecular docking

The following are needed for a basic molecular docking on MOE: 1) a database file containing the 3D structure of the compounds and their different conformations, 2) the receptor protein prepared for the docking, 3) a method to identify and nominate a binding site for the ligands to dock to the receptor protein.

1) Preparation of the compound database file

The structures of the compounds used for the docking study were inputted and compiled into a single database file. The database then went through MOE's conformation search application and created a second database file, which contained multiple low energy 3D atomic configurations for each compound. This was used for the docking study. The conformation was searched by the systematic search method setting in MOE, which was intended for small molecules such as the compounds in this project. According to the MOE manual description, the systematic search method generates molecular conformations by rotating non-ring bonds in a molecule by discrete dihedral increments. All possible combinations of the dihedral increments are produced. Each such dihedral angle combination is subjected to dihedral angle energy minimization to relieve gross strain due to atom overlap, followed by an all-atom energy minimization. The resulting conformation is saved to the output database provided it meets the energetic and geometric criteria [115].

2) Preparation of the receptor protein

The 3D structure obtained from protein databank required preparation prior to docking. In the case of 6QZH, the inserted Sialidase NanA used for facilitating the crystallisation of CCR7 was removed. 3ODU is a CXCR4 homodimer, one of the CXCR4 units was removed. Ligands and water molecules on the crystal structure were removed. The 3D structures went through the structure preparation application in MOE, which attempts to correct any errors detected in the structure. For 6QZH and 3ODU, it was mostly completing residues with missing atoms.

3) Generating a binding site

Molecular docking requires a binding site. For our docking study, there are two methods to generate the binding site. The first method is to utilise the MOE's site

finder application, this feature calculates and identifies a number of residues as the possible binding sites of a receptor protein. These residues will be selected by MOE, and the docking simulation will occur at the space around these residues.

For example, in the case of CCR7/6QZH. Two binding sites were identified, one encompasses the residues at the parts of the transmembrane domain, one at the intracellular allosteric site. The region of the two binding sites are highlighted in **Figure 62**.

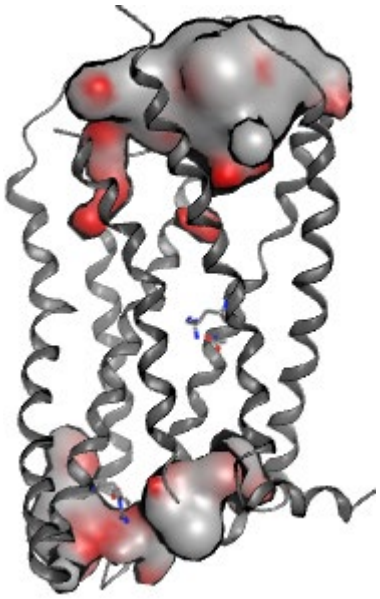


Figure 62*: 3D representation of the two binding sites in CCR7 calculated by the site finder function in MOE.*

The second method involves using an existing ligand-receptor complex. The binding mode of a ligand at a receptor protein can be selected as the binding site, this feature allows the residues surrounding the ligand to be used as the binding site. For example, the ligand-receptor complex of Cmp2105 and CCR7/6QZH can be used to dock our CCR7 antagonists to the intracellular allosteric site of CCR7. This method allows the resulting binding poses to occupy similar space to Cmp2105.

4) Setting up the docking protocol

Most of the docking protocol options were left to the default setting. The docking simulation began with the placement of the ligand to the binding site. Several placement methods were available. The default placement method was “Triangle matcher”, which generates docking poses by aligning ligand triplets of atoms on triplets of alpha spheres [115]. It is recommended by MOE for docking small ligands. The placements of the ligands were assessed by the default scoring function, London ΔG. The docked ligand went through a refinement stage and it was re-scored by a different scoring function, GWVI/WSA ΔG was the default option. The score calculated at this stage became the final score of the docking pose. The receptor was set to rigid throughout the docking studies, flexible receptor docking was not available.

MOE possesses 5 different scoring functions. ASE, Affinity ΔG, GWVI/WSA ΔG, London ΔG and Alpha HB. ASE and Alpha HB are shape-based scoring methods. Other scoring functions attempt to estimate the free energy of binding. A lower number indicates higher score and a favourable pose. It should be noted it is a predicted value and it should only be used as a guidance.

5.4 Handling and analysing docking results on MOE

The completion of the docking experiments generated a database file of the results. The file contains multiple top scoring docking poses for each compound, and the docking scores calculated for each individual docking pose.

For the visual inspection, the individual docking pose from the database, and protein structure used for the docking study were loaded to MOE's 3D rendering viewer, thus allowing the ligand-receptor complex to be viewed. This ligand-receptor complex then was subjected to energy minimization function (induced fit model). This changes the conformation of both the ligand and the receptor protein, meaning it will create a new docking pose different from the database file, and the receptor will also have a different conformation from the one used for the initial docking. Applying energy minimisation allows ligand interactions to optimise, and it lowers the free energy of binding which is calculated by the GWVI/WSA ΔG scoring function. The option to calculate the value with other scoring functions isn't available. The new score will be calculated after the completion of the energy minimisation. The receptor protein is not a static structure, the docking of a ligand should change its conformation, the energy minimisation compensates for the lack of rigid receptor during the docking studies. Theoretically, this should produce a more favourable docking pose.

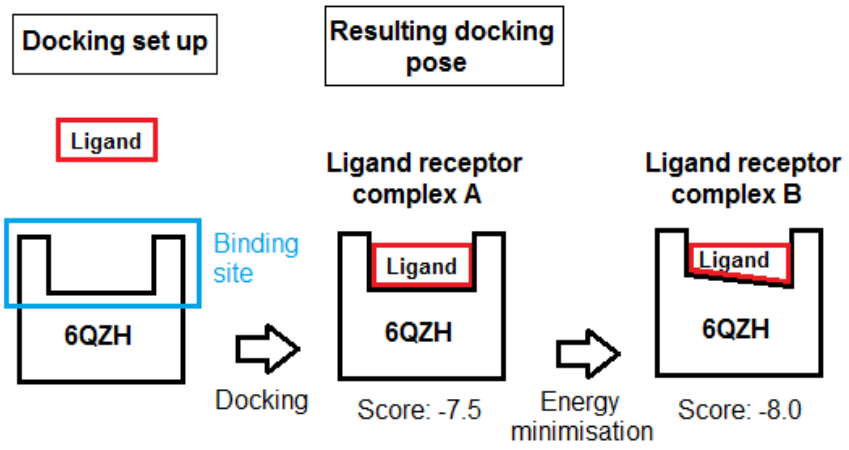
The second method of selecting the binding site involves using an existing ligand-receptor complex. The energy minimised receptor-ligand complex obtained during visual inspection can serve as the binding site and the protein structure for further docking studies. We have found that repeated or further docking studies by using the energy minimised ligand-receptor complex as the binding site and the receptor protein, the docking scores are higher, indicating more favourable docking poses. An example of this seemingly optimised docking is illustrated in **Figure 63**.

Initial docking:

Receptor: 6QZH after initial preparation

Ligand: ICT13069

Binding site: Searched with site finder function



Repeated docking:

Receptor: 6QZH with new conformation from ligand receptor complex B

Ligand: ICT13069

Binding site: docking pose from ligand receptor complex B

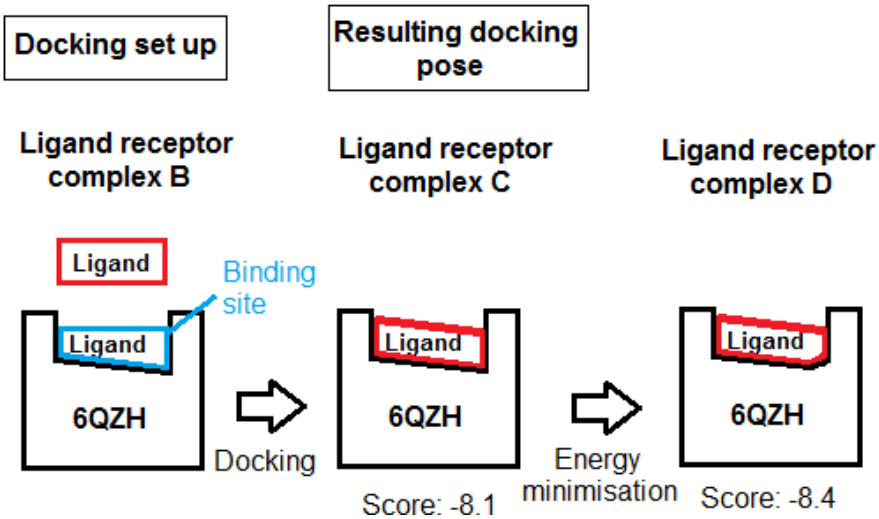


Figure 63: *Example work flow of repeating the docking of ICT13069 to acquire a docking pose with higher score. Initial docking gives ligand-receptor complex A. Applying energy minimisation gives ligand-receptor complex B, conformation of ligand and receptor changed, docking score increased. Ligand receptor complex B was used in the repeated docking of ICT13069, which gives ligand complex C. Applying energy minimisation gives ligand complex D with docking score of -8.4.*

5.5 Replicating binding mode of Cmp2105 and Vercirnon

In order to test MOE's capability in generating an accurate binding pose. Attempts were made to replicate the binding mode of Cmp2105 found in CCR7/6QZH. The first attempt utilised the site finder function. All of the docking poses generated are significantly different from the crystal structure binding mode. Afterward, the binding mode of Cmp2105 was used as the binding site of the docking, in theory, this should allow MOE to replicate the binding mode. This attempt failed to replicate the binding mode exactly, however the predicted docking pose with the highest score is similar to the binding mode, showing a number of interactions to the same residues. The thiodiazole-dioxide scaffold, the *tert*-butyl and the furan moiety are docked almost identical to the binding mode. The superposition of the binding mode and the top docking pose is shown in **Figure 64** below. GWVI/WSA ΔG docking score is -10.44 kcal/mol. The other calculated docking poses are very different from the actual crystal structure binding mode, and they have failed to show interactions with Asp94 and Phe333, which shown interactions with the thiadiazole and the bridging amine groups of Cmp2105.

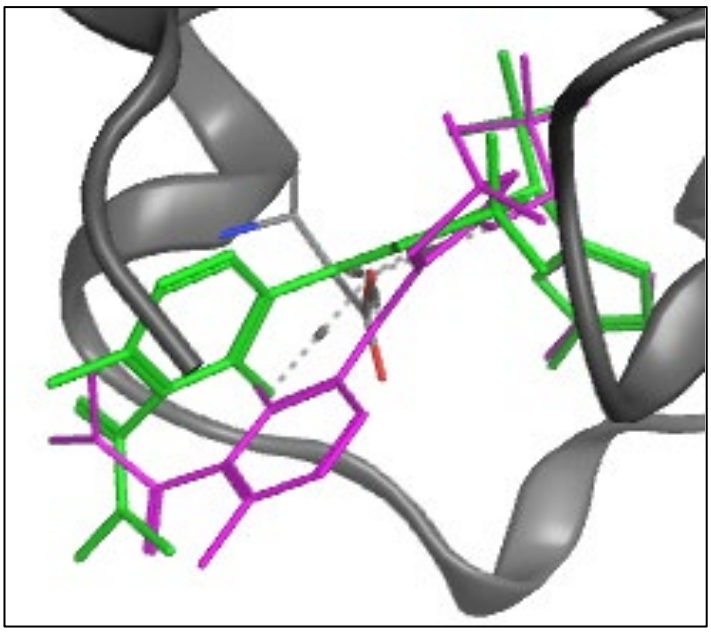


Figure 64*: Superposition of two Cmp2105 docked at the CCR7 intracellular site. Green: binding mode of Cmp2105 found in the crystal structure, magenta: the docking pose Cmp2105 generated in MOE.*

In contrast, the attempt to replicate Vercirnon's binding mode in 5LWE was successful. The superposition of the binding mode and the generated docking pose is shown below (**Figure 65**). Docking using the site finder function generated a predicted docking pose that is almost identical to the crystal structure binding mode. The GWVI/WSA ΔG docking score is -9.98 kcal/mol. The success appears to be attributed to the Vercirnon having fewer single bonds for rotation compared to Cmp2105. This indicates molecules that are more flexible will have less reliable docking poses. This also establishes the docking score of -10.44 and -9.98 kcal/mol are high scores for the GWVI/WSA ΔG scoring function.

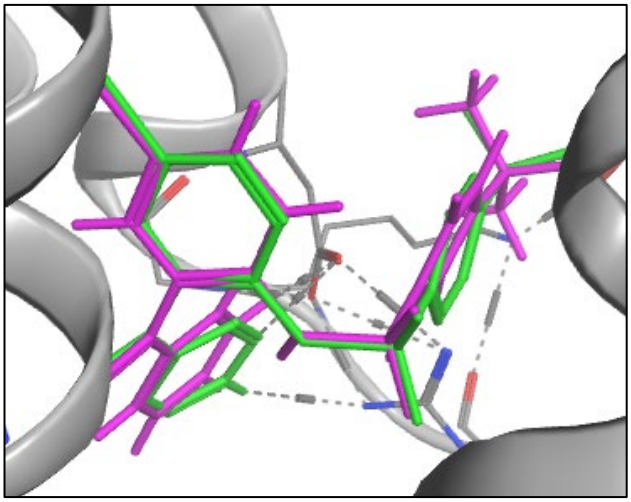
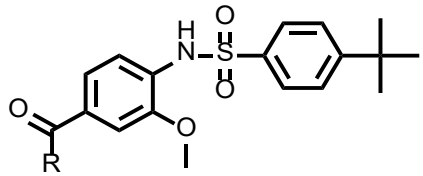


Figure 65*: Superposition of two Vercirnon docked at the CCR9 intracellular site. Green: binding mode of Vercirnon in the CCR9 crystal structure. Magenta: docking pose of Vercirnon generated by MOE.*

5.6 Molecular docking: discussion and results

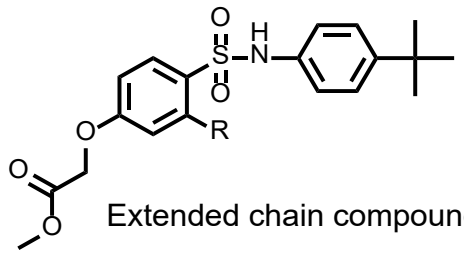
For the discussion of the docking poses. The compounds will be split into 3 categories.

Group A: The first category will be compounds similar to ICT13069, which are the biphenyl sulfonamide compounds with small functional groups such as a methyl ester. 34 sulfonamide compounds synthesised prior to this project by another member of our group will also be part of the investigation, their IC_{50} from calcium flux assay are available.



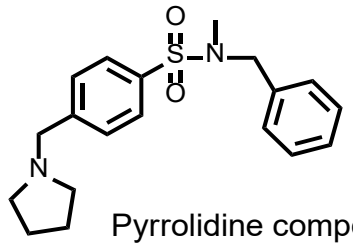
Compounds similar to ICT13069

Group B: The second category will be the extended chain compounds, these compounds are similar to ICT13069, the group para to the sulfonamide is a short side chain instead of a small ester or other small moieties.



Extended chain compounds

Group C: The third category will be the compounds containing the pyrrolidine group, which includes compound **6** from Mooring et al. and other pyrrolidine compounds containing features of our CCR7 antagonists, such as 4-*tert*butylbenzene.



Pyrrolidine compounds

5.6.1 CCR7 docking: extracellular and transmembrane domain

The initial docking study largely generated docking poses at the extracellular region near the transmembrane domain. **Figure 66A** highlighted the extracellular residues of CCR7. **Figure 66B** shows the superposition of the top scoring docking poses for the compounds similar to ICT13069 (group A), which are above the transmembrane site close to the extracellular loop.

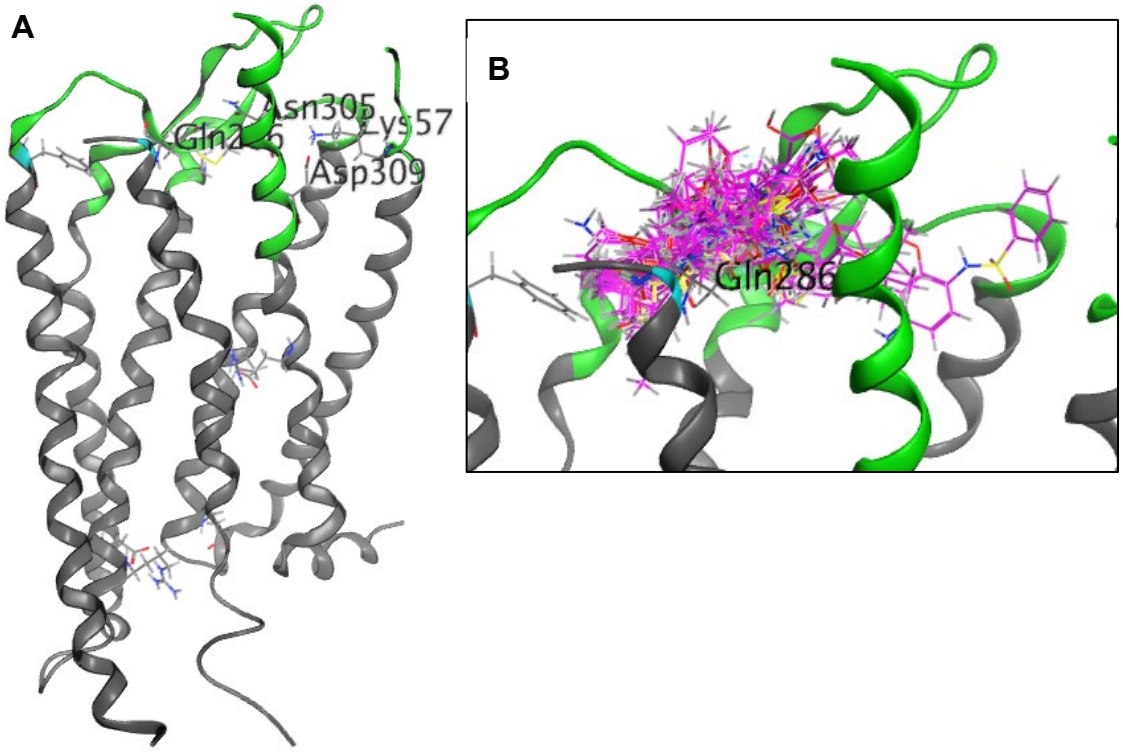


Figure 66*: A) Crystal structure of CCR7/6QZH, extracellular residues highlighted in green. B) Superposition of best docking poses for compounds similar to ICT13069*

Many different top docking poses show hydrogen bonding between the sulfonyl group and amine of Gln286, a residue at the transmembrane region. This includes docking poses which appear to be incompatible with the experimentally observed potency, **Figure 67** shows one of the top docking poses for a CCR7 antagonist with IC_{50} below 1 μM. The molecule is docked at the extracellular part of the molecule, the nitrobenzene ring is exposed.

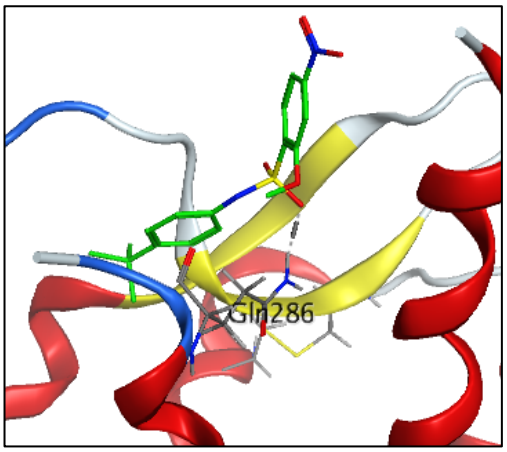


Figure 67*: An unlikely docking pose of an analogue of ICT13069 with a nitro substituent.*

This highlights the unreliability of MOE generating docking poses at the extracellular-transmembrane domain of 6QZH for our docking study. For many poses, the amide of the sulfonamide interacts with Leu212, if carbonyl group is present at the polar end of the molecule, it shows interaction with Lys57. **Figure 68** shows one of the top reasonable docking poses for compound **41**, with the aforementioned ligand interactions. The GWVI/WSA ΔG score is -7.54 kcal/mol after energy minimisation.

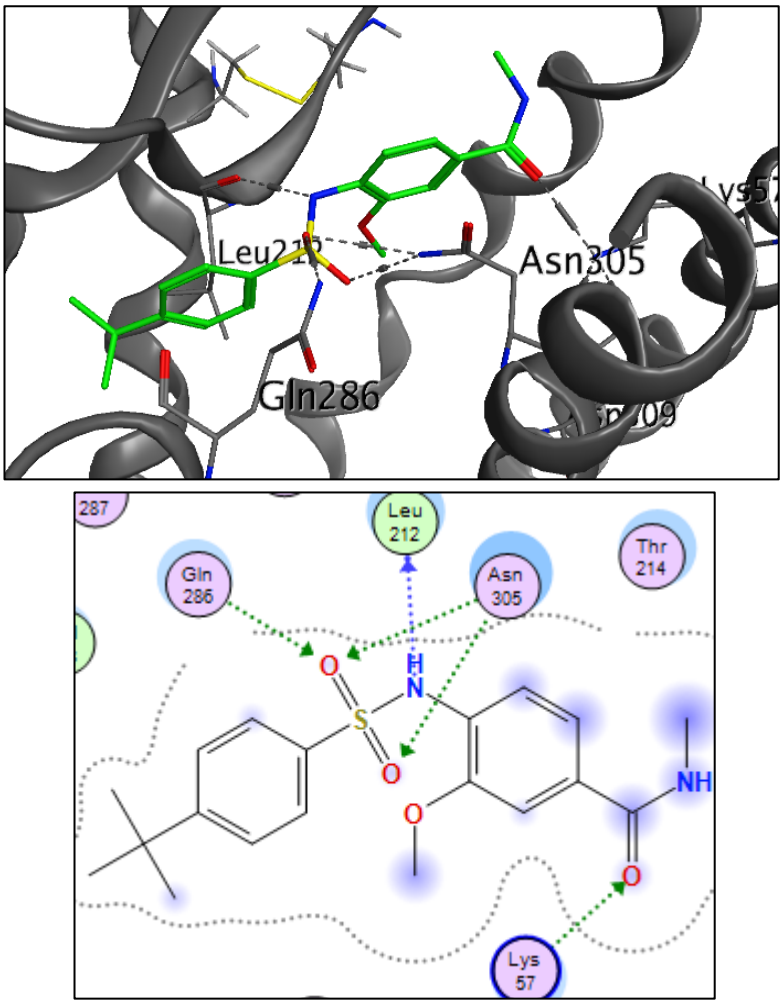


Figure 68: *Docking pose of compound **41** to the extracellular-transmembrane domain of CCR7, and its ligand interaction map generated by MOE.*

We believe that the reason is that binding of Cmpd2105 in the allosteric pocket may have constrained the transmembrane binding pocket in the crystal structure model. Jaeger et al. shown that Cmp2105 was able to inhibit the binding of CCL19, and Cmp2105 was able to stabilise the inactive CCR7 conformation [63, 115]. The binding of Cmp2105 possibly have changed the conformation of the orthosteric site and made it less exposed, preventing the ligands to dock at the orthosteric site.

To test this hypothesis, we compared the conformation between CCR5 (4MBS), CCR2 (5T1A), CXCR4 (3ODU), CCR7 (6QZH) and CCR9 (5LWE). The crystal structure of CCR5, CCR2 and CXCR4 have a ligand bind at the orthostatic site, whereas the ligand binds at the intracellular site for CCR7 and CCR9. Between

these 5 structures, two different receptor conformations at the extracellular transmembrane domain are clearly identifiable. The first conformation is observed in CCR5, CCR2 and CXCR4 (**Figure 69A**), in which the extracellular end of TM6 point outward, and TM5 points inward. This appears to be the conformation of the structure when a small molecule ligand binds at the orthosteric site. The other conformation observed in CCR7 and CCR9 (**Figure 69B**) is the opposite, the TM6 points inward, and TM5 points outward, which appears to be the conformation when a small molecule ligand binds at the intracellular site. **Figure 69C** shows the difference between the two conformation at TM5 and TM6. The conformation of the TM6 appears to determine the accessibility of the orthosteric site. Since 6QZH doesn't have this conformation, our antagonists are unable to dock at the orthosteric site.

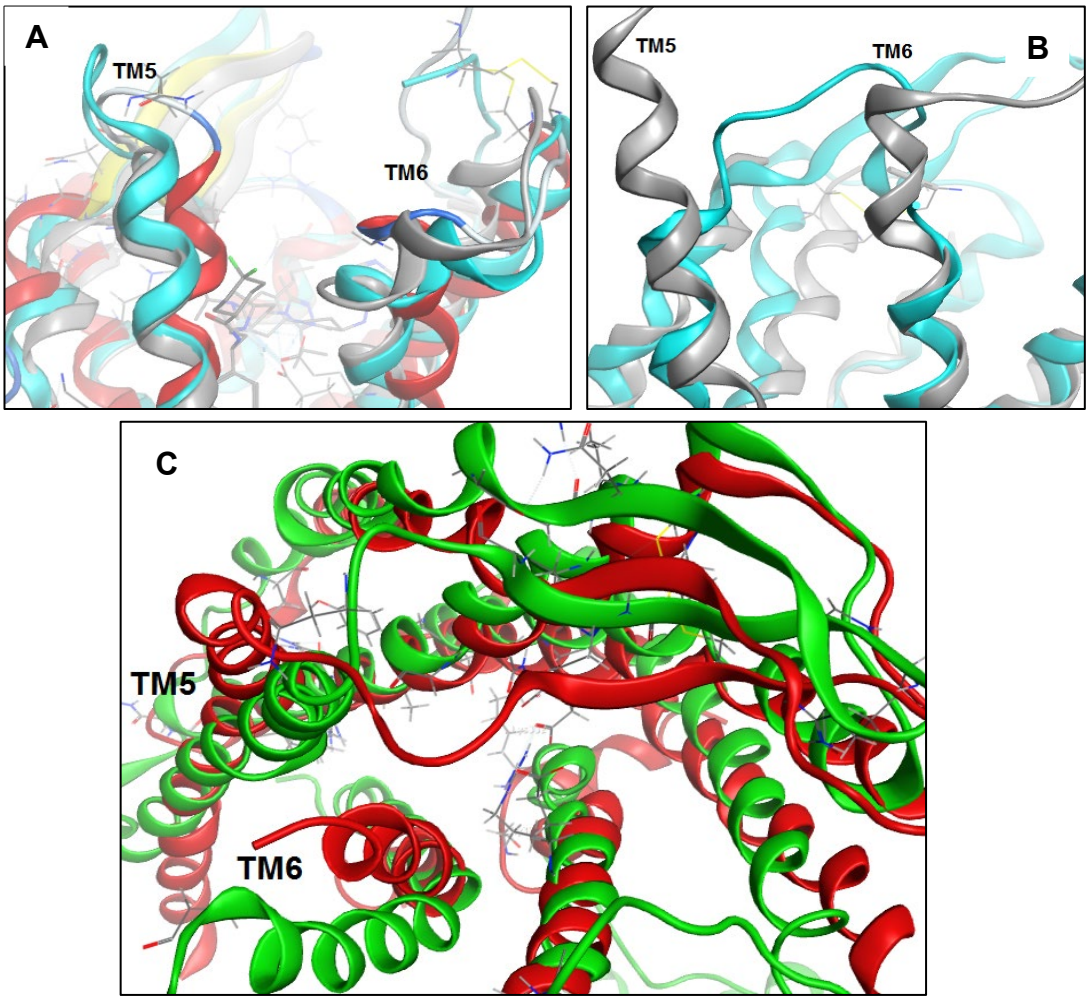


Figure 69: *A) Superposition of CCR5 (4MBS), CCR2 (5T1A) and CXCR4 (3ODU), TM5 and TM6 have similar conformation. B) Superposition of CCR7 (6QZH) and CCR9 (5LWE), TM5 and TM6 have similar conformation. C) Superposition of CCR7 (6QZH,* ***Red****) and CXCR4 (3ODU,* ***Green****), comparing the two conformations.*

We attempted to widen the pocket. CCR7/6QZH was superposed to CXCR4/3ODU and the pyrrolidine CXCR4 antagonist compound **6** docked at the transmembrane site of 3ODU (docking of **6** and 3ODU will be discussed at a later section). 3ODU was removed, leaving **6** in the transmembrane region of CCR7/6QZH. Applying the energy minimisation function to 6QZH and **6** forces the structure to widen. This only opened a small pocket, docking poses generated by this method show little to no interaction between the functional groups and the receptor, they are very likely to be unreliable.

In summary, attempts at docking molecules into the transmembrane domain of models based on crystal structure of CCR7 were not successful. This is mainly because in this model the binding pocket in the transmembrane domain of the receptor is constrained due to conformation changes imposed on it due to presence of a ligand in the intracellular binding domain of the crystal structure.

5.6.2 CCR7 binding: intracellular domain

There is significant sequence similarity between CCR7 and CCR9. The superposition of CCR7 and CCR9 with Vercirnon (the allosterically bound ligand in CCR9 crystal structure) and Cmp2105 shows that their binding mode are located at similar region within the intracellular site (**figure 70**).

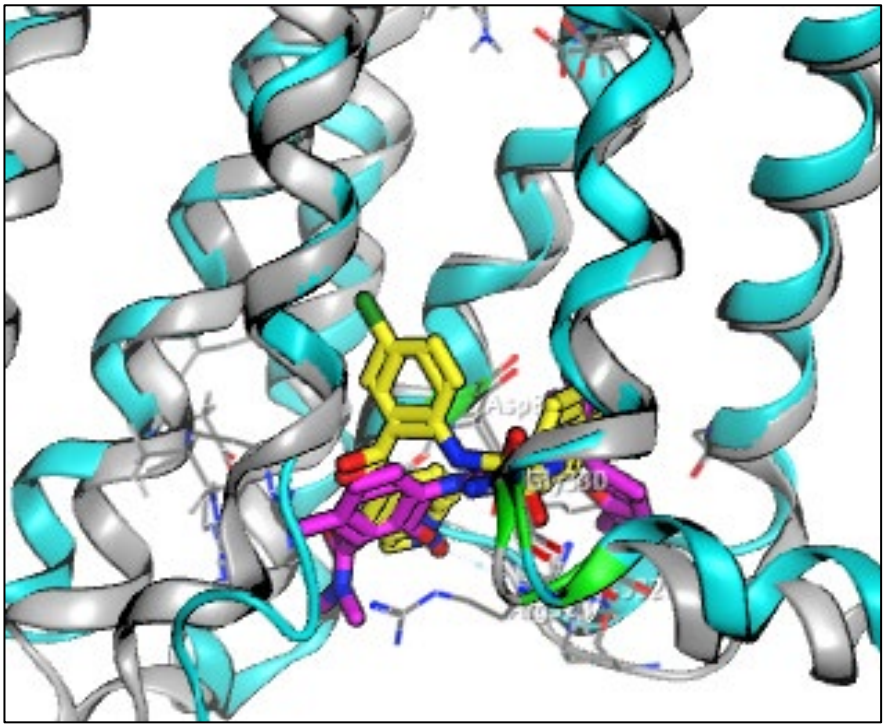


Figure 70*: Superposition of CCR7 (6QZH) with Cmp2105 (pink), and CCR9 (5LWE) with Vercirnon (yellow).*

Vercirnon possesses a 4-*tert*butylbenzene sulfonamide similar to our antagonists, we hypothesized that if our antagonists bind at the intracellular site, this moiety could bind to CCR9 and CCR7 in a similar hydrophobic pocket at the allosteric binding domain. We superimposed CCR7/6QZH to a Vercirnon docked CCR9/5LWE. 5LWE was deleted, leaving Vercirnon in the intracellular site of 6QZH. The whole system was then energy minimised to obtain a model of Vercirnon bound inside a putative CCR7 allosteric binding domain. The ligand interaction map (**Figure 71**) shows that the sulfonyl group interacts with Lys332 and Phe333 of CCR7. Lys332 is the corresponding residue to Arg323 of CCR9, which interacts with Vercirnon in 5LWE (**Figure 71**).

On a side note, Vercirnon is not active in CCR7, it did not show significant inhibition in a binding assay at 10 µM [117]. We decided to conduct a docking study of Vercirnon in CCR7 in attempt to explain this. In CCR9, the pyridine-*N*-

oxide of Vercirnon is able to form interactions with Arg78 of intracellular vestibule 1 and Asp84 in CCR9. The binding mode of Vercirnon bound inside CCR7 was used as the binding site and the receptor for the docking. Additionally, we also used site finder and a normal prepared 6QZH for the docking. While we were able to replicate the binding pose of Vercirnon in CCR9, we were unable to generate a docking pose in which the pyridine-*N*-oxide forms interactions with surrounding residues. Arg88 is the corresponding residue to Arg78 in CCR9. Arg78 in CCR9 faces inward to the binding site, Arg88 faces outward, thus the interaction cannot be formed. The pyridine-*N*-oxide unable to form ligand interaction in the CCR7 intracellular site might be why Vercirnon is not active in CCR7.

Receptor	CCR9	CCR7	CCR7
Ligand	Vercirnon	Vercirnon	Cmp2105
Corresponding residues and interactions	Gly321 $^{8.47x47}$ (sulfonyl)	Gly330 $^{8.47x47}$	Gly330 $^{8.47x47}$ (sulfonyl)
	Arg323 $^{8.49x49}$ (sulfonyl)	Lys332 $^{8.49x49}$ (sulfonyl)	Lys332 $^{8.49x49}$
	Phe324 $^{8.50x50}$	Phe333 $^{8.50x50}$ (sulfonyl)	Phe333 $^{8.50x50}$ (sulfonyl)
	Arg78 $^{12.49x49}$ (pyridine-*N*-oxide)	Arg88 $^{12.49x49}$	Arg88 $^{12.49x49}$
	Asp84 $^{2.40x40}$ (pyridine-*N*-oxide)	Asp94 $^{2.40x40}$	Asp94 $^{2.40x40}$ (amine bridge)

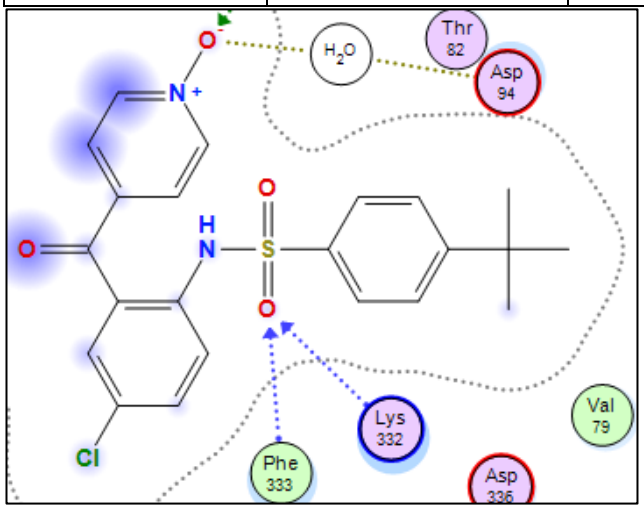


Figure 71: *Ligand interaction map of Vercirnon within the intracellular allosteric site of CCR7, after superimposing CCR9 and CCR7 then delete the CCR9 structure. The table shows the corresponding residues in CCR9 and CCR7.*

We attempted to generate the docking pose of compound **41** by using the previous generated Vercirnon-CCR7 complex. Vercirnon in CCR7 was gradually altered, by keeping the 4-*tert*butylbenzene sulfonamide moiety intact, but by changing the rest of the molecule in a stepwise fashion so it becomes compound **41**, one of our CCR7 antagonists. The whole system was energy minimised after each alteration. The docking pose, shown in **Figure 72A**, was generated from this approach. The GWVI/WSA ΔG score is -8.45 kcal/mol. The docking pose **Figure 72A** below has the carbonyl of the amide forms hydrogen bond with Arg154, and sulfonyl group forms hydrogen bond with Lys332. The 4-*tert*butylbenzene appears to be in a hydrophobic pocket consist of Val79, Tyr83, Asp94, Leu98, Tyr326.

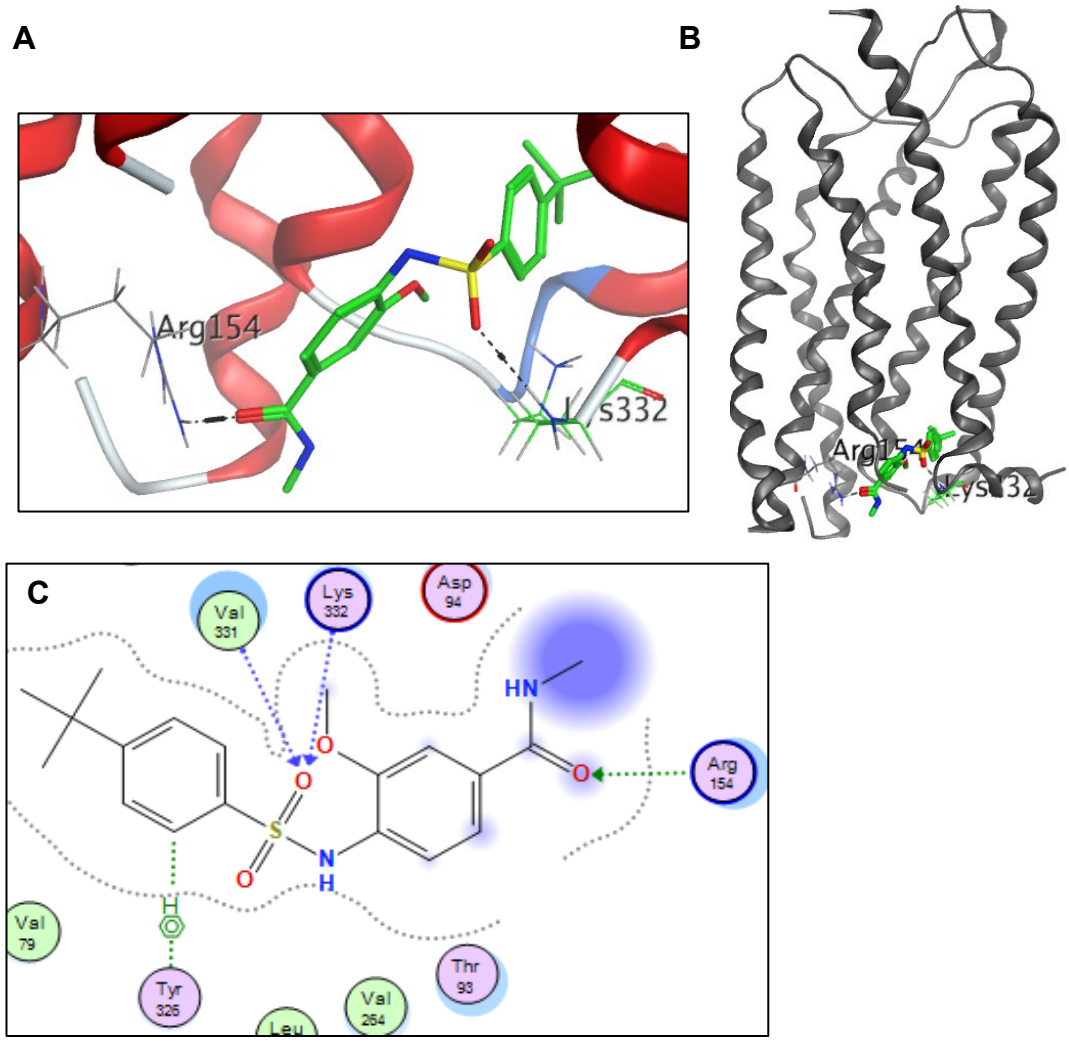


Figure 72: *A) Docking pose of compound **41** at the intracellular site of CCR7, generated from altering Vercirnon to **41**. B) Side view of **41** docked at the CCR7 intracellular site. C) Ligand interaction map of the docking pose generated by MOE*

A second approach to docking Compound **41** to the intracellular site was by using the site finder function as described previously for the orthosteric binding site. The top scoring docking pose is similar to docking pose created from altering Vercirnon. It shows the same the ligand interaction with Lys332 and Arg154 (**Figure 73**). GWVI/WSA ΔG score after energy minimisation is -7.32 kcal/mol. The intracellular binding site calculated by MOE is smaller, thus the docking poses show little variations. The similarity between the docking poses from different approaches, gives us confidence as to the validity of the docking pose for further *in silico* investigation.

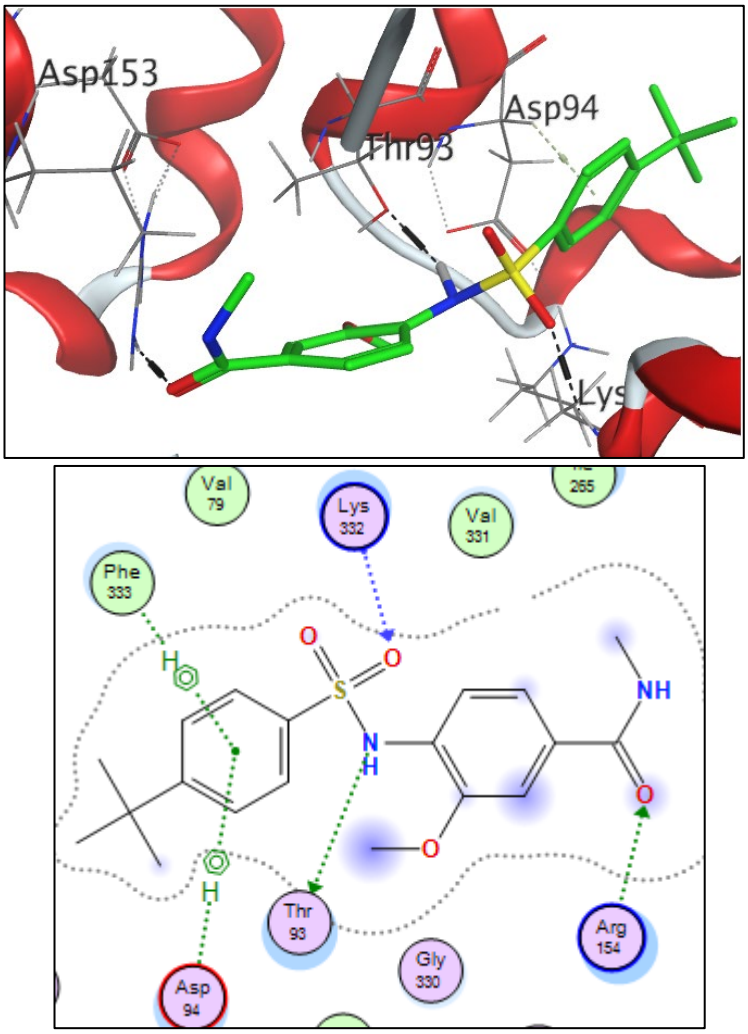


Figure 73: *Docking pose of **41** at the CCR7 intracellular site and its ligand interaction map.*

The energy minimised ligand-receptor complex of **41** shown in **Figure 73** was used as the binding site and the receptor protein for the repeated docking of **41** and for further docking study of compounds similar to ICT13069. The repeated docking of **41** yielded a new docking pose with -9.54 kcal/mol after energy minimisation. This docking pose is mostly similar and retained interaction with Lys332, Arg154 and Thr93. As mentioned in chapter 5.4, the docking poses generated by energy minimised ligand-receptor complex have higher scores compared to the poses generated via site finder and unchanged 6QZH. The range of the docking scores prior to energy minimisation is -6.5 to -9.3 kcal/mol, elevated from -5.5 to -8.3 kcal/mol. Compound **67** is the "reverse" sulfonamide of **41**, in which the order of the sulfonamide is different**.** The docking pose is similar and retained interaction with Lys332, Thr93, Arg154 (**Figure 74**). The main difference is the methyl ether faces toward the binding site. GWVI/WSA ΔG score is -9.34 kcal/mol after energy minimisation. In terms of energy difference, docking pose and ligand binding, the difference between **67** and **41** are very small.

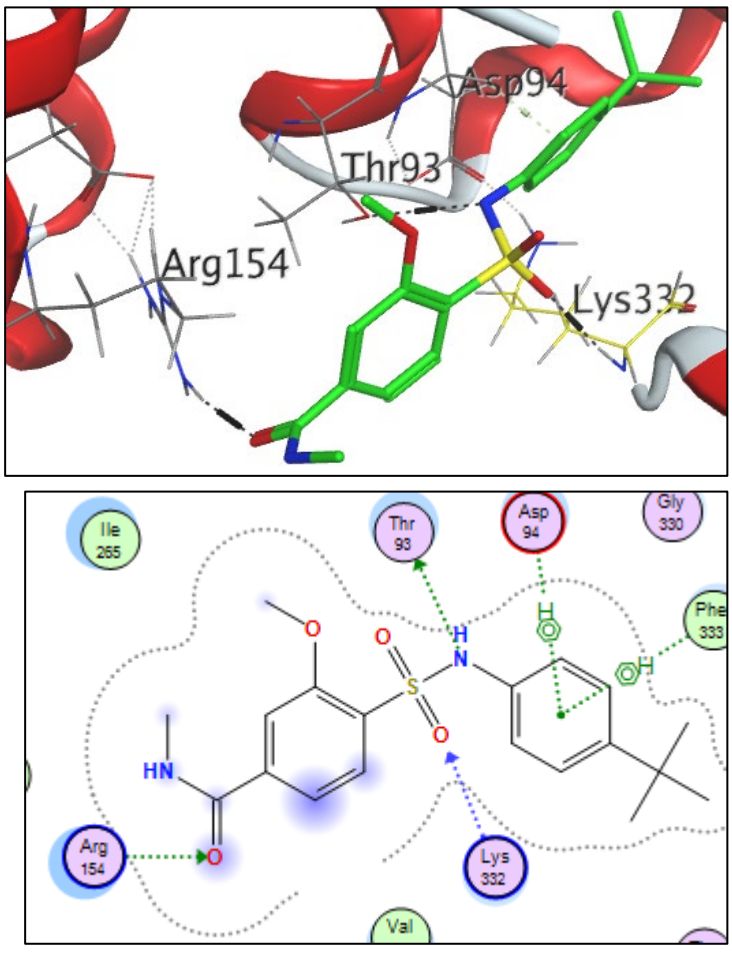


Figure 74*: Best docking pose of* ***67*** *at the CCR7 intracellular allosteric site, and the ligand interaction map generated by MOE.*

Docking of extended chain compounds

The antagonists with extended chain shown very similar docking poses and ligand interactions to that of compound **67** and others in group A. The 4-*tert*butylbenzene ring docks at the hydrophobic pocket, the sulfonyl forms a hydrogen bond with Lys332. The extended chain was thought to possibly have greater reach at the binding site. The extended chain compounds all have a carbonyl group at the end of the chain, the carboxyl forms hydrogen bond with Arg154 similar to compound **67**. The docking pose of compound **97** (**Figure 75**) shows the general docking pose of the extended chain compounds.

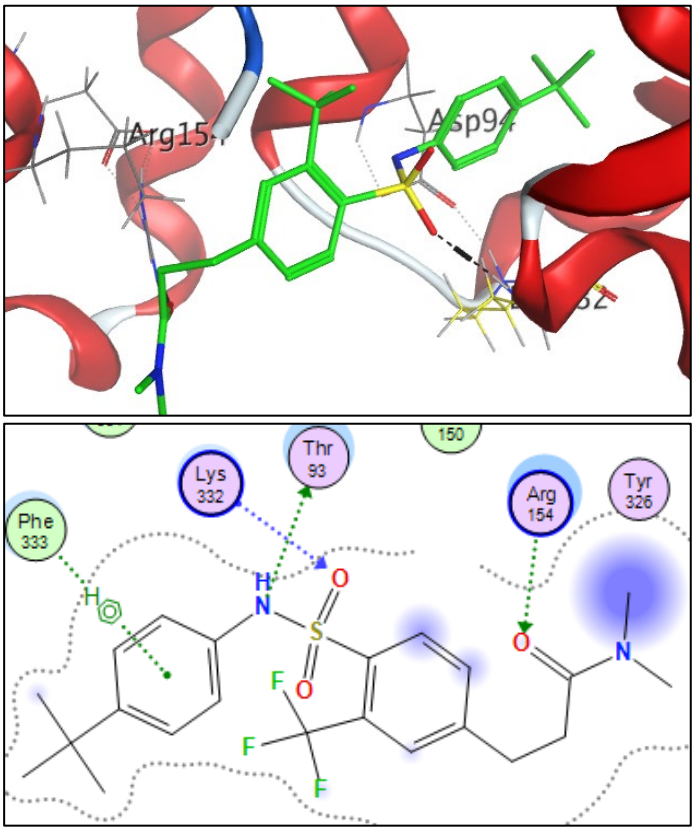


Figure 75: *Top docking pose for compound **97** at the CCR7 intracellular site, and the ligand interaction map generated by MOE.*

Docking of pyrrolidine compounds

We attempted to dock the pyrrolidine compounds to the CCR7 intracellular site. The pyrrolidine compounds dock similarly to the sulfonamide CCR7 antagonists. For Mooring et al.'s compound **6**, the protonated pyrrolidine docks at the pocket near Asp94 similar to 4-*tert*butylbenzene of the CCR7 antagonists, and forms a salt bridge with Asp94, the sulfonyl forms hydrogen bond with Lys332. The best docking pose of compound **6** at the CCR7 intracellular site is shown below. The GWVI/WSA ΔG score is -8.41 kcal/mol after energy minimisation (**Figure 76**).

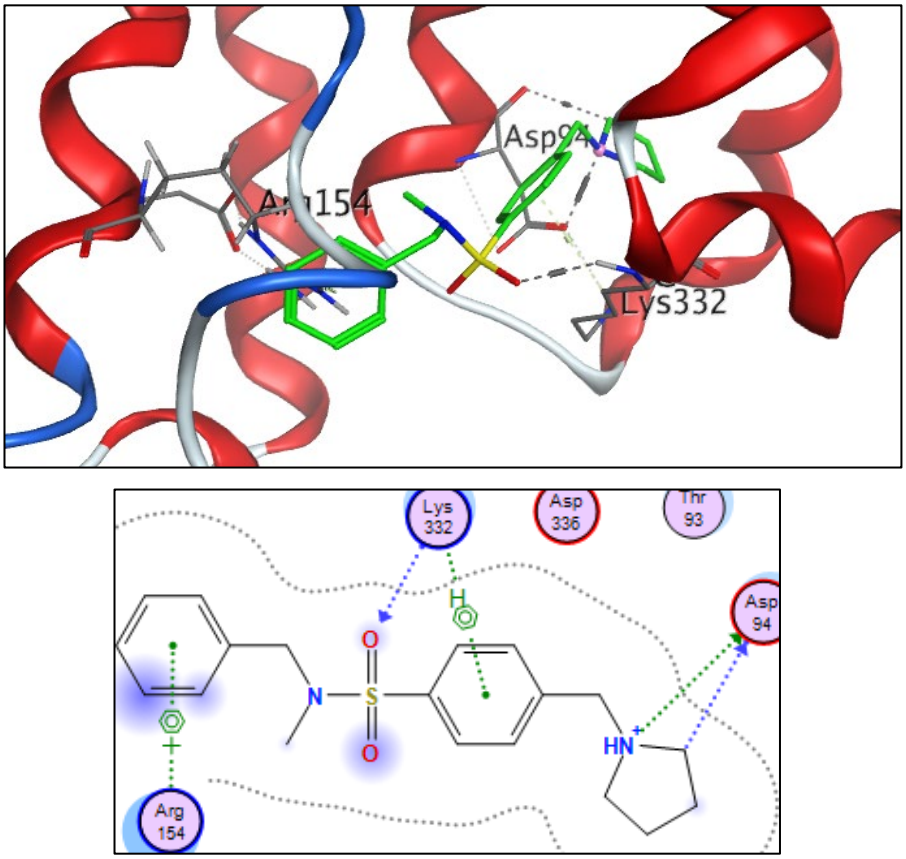


Figure 76: *Top docking pose of* ***6*** *at the CCR7 intracellular site, and its ligand interaction map generated by MOE.*

For pyrrolidine compounds with 4-*tert*butylbenzene (**73**, **74**, **78**, **89**), the pyrrolidine and the 4-*tert*butylbenzene ring appears to compete for the pocket near Asp94 in the docking simulation, thus leaving either the pyrrolidine or the 4-*tert*butylbenzene exposed to the extracellular space. The superposition of these two types of docking poses is shown below (**Figure 77**). Although the pyrrolidine is able to form salt-bridge with Asp94, making the docking pose possibly more favourable, the docking scores of the two docking poses are similar. This possibly indicates the 4-*tert*butylbenzene ring and pyrrolidine in the same molecule are not compatible.

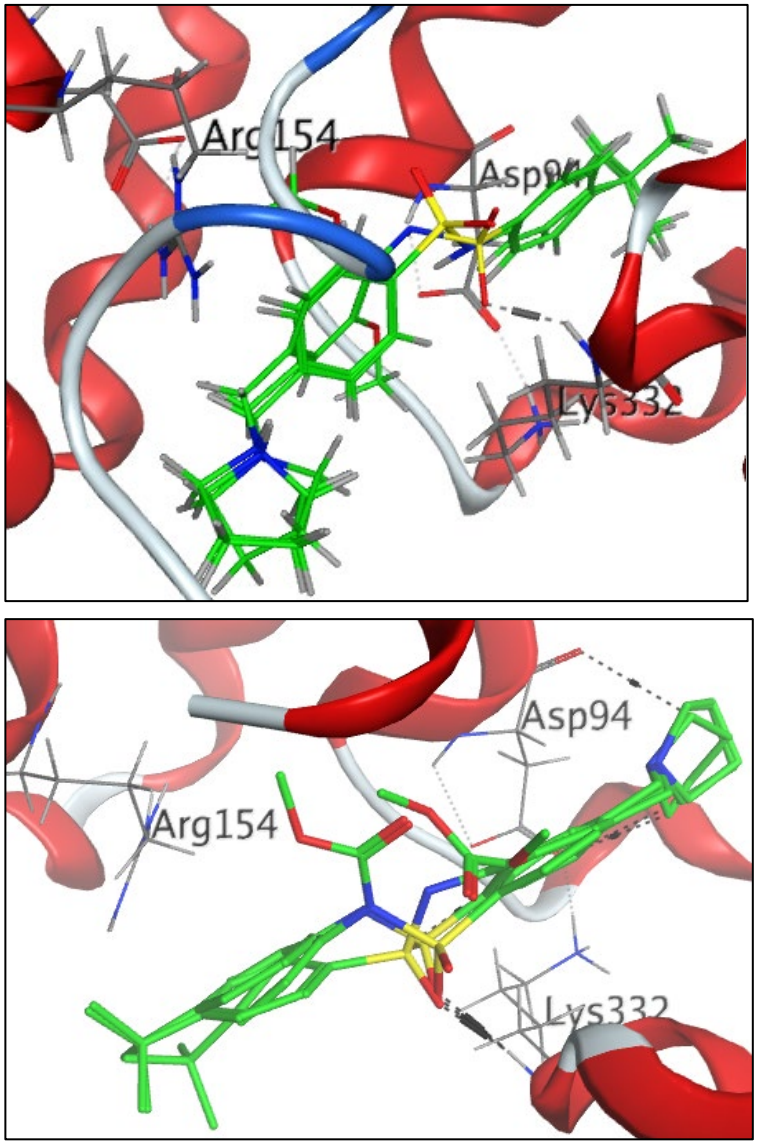


Figure 77*: Superposition of the docking pose for pyrrolidine compounds with 4-tertbutylbenzene. A) docking poses in which the 4-tertbutylbenzene docks at the pocket near Asp93. B) Docking poses in which the pyrrolidine docks at the pocket near Asp93.*

Compound **79** is the pyrrolidine compound with methyl methoxybenzoate of ICT13069. Replacing the benzyl of compound **6** with this polar ring appeared to have yielded favourable docking results in the intracellular site. Similar to other pyrrolidine compounds, the pyrrolidine and the sulfonyl show the same interaction with Asp94 and Lys332. The carbonyl of the methyl ester forms hydrogen bonding with Arg154 (**Figure 78**). GWVI/WSA ΔG score after energy minimisation is -9.99 kcal/mol.

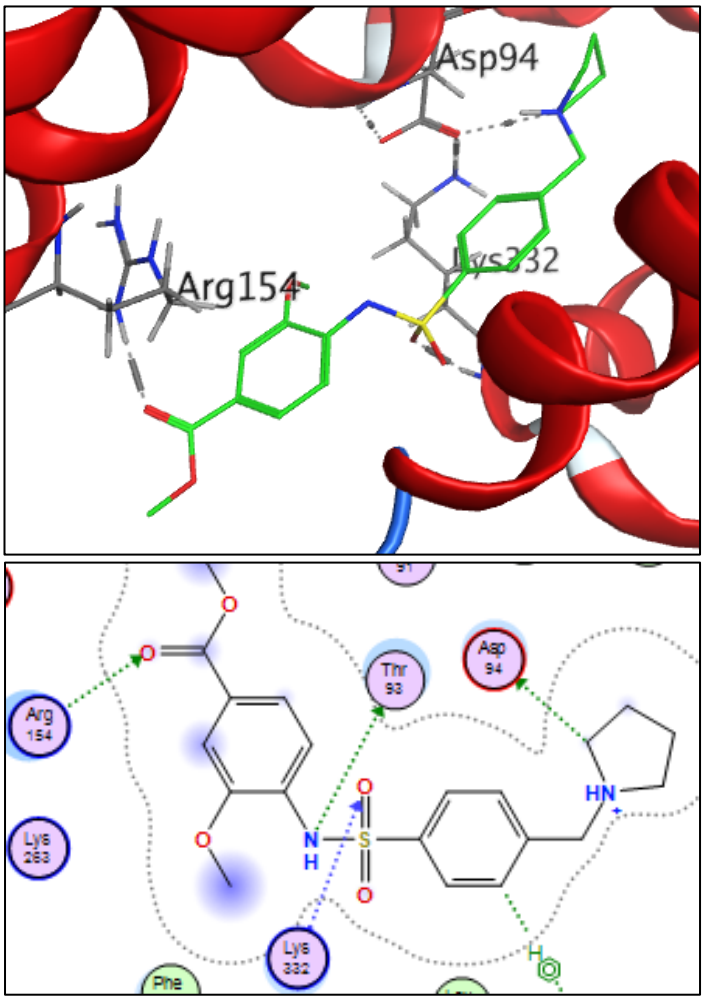


Figure 78*: Top docking pose of* ***79*** *at the CCR7 intracellular site, and its ligand interaction map generated by MOE.*

Summary and key findings of docking at the CCR7 intracellular site

- The 4-*tert*butylbenzene of the CCR7 antagonists dock at the hydrophobic pocket near Asp94. The sulfonyl groups form hydrogen bond with Lys332. If a polar group is presents in the position para to the sulfonamide, it forms hydrogen bond with Arg154.
- The order of the sulfonamide in the CCR7 compounds don't affect the ligand interaction and orientation of the docking pose, it doesn't indicate whether the "normal" or "reverse" sulfonamide form more favourable interactions.
- Pyrrolidine compounds have a similar docking pose as the other CCR7 antagonists. The protonated pyrrolidine forms salt bridge with Asp94.
- The docking poses of pyrrolidine compounds with 4-*tert*butylbenzene appear to be unfavourable, indicating they are not likely to bind at the intracellular site if they are active in CCR7.
- Compound **79** is able to form salt bridge with Asp94, and able to form ligand interaction with residues observed in other CCR7 antagonists, showing favourable docking pose in the intracellular site.

5.6.3 CXCR4 binding: extracellular-transmembrane domain

Mooring et al. [80] has shown the top docking pose for compound **6**, their most potent CXCR4 antagonist, at the CXCR4 transmembrane site (**Figure 79A**). Attempting to replicate this docking pose of compound **6** was decided to be a good starting point for the docking study at the CXCR4 orthosteric site. The pyrrolidine moiety of the compound was predicted to be protonated and forms a salt bridge to Asp97 [80]. The sulfonyl forms hydrogen bond with the guanidine of Arg188 at extracellular loop 2. One of the top docking poses was similar to the docking pose shown by Mooring et al. This docking pose (**Figure 79B)** occupies the same space at the transmembrane domain, and it has the same interaction with Arg188 and Asp97. The GWVI/WSA ΔG score is -7.64 kcal/mol. **Figure 79C** shows a wider view of the docking pose.

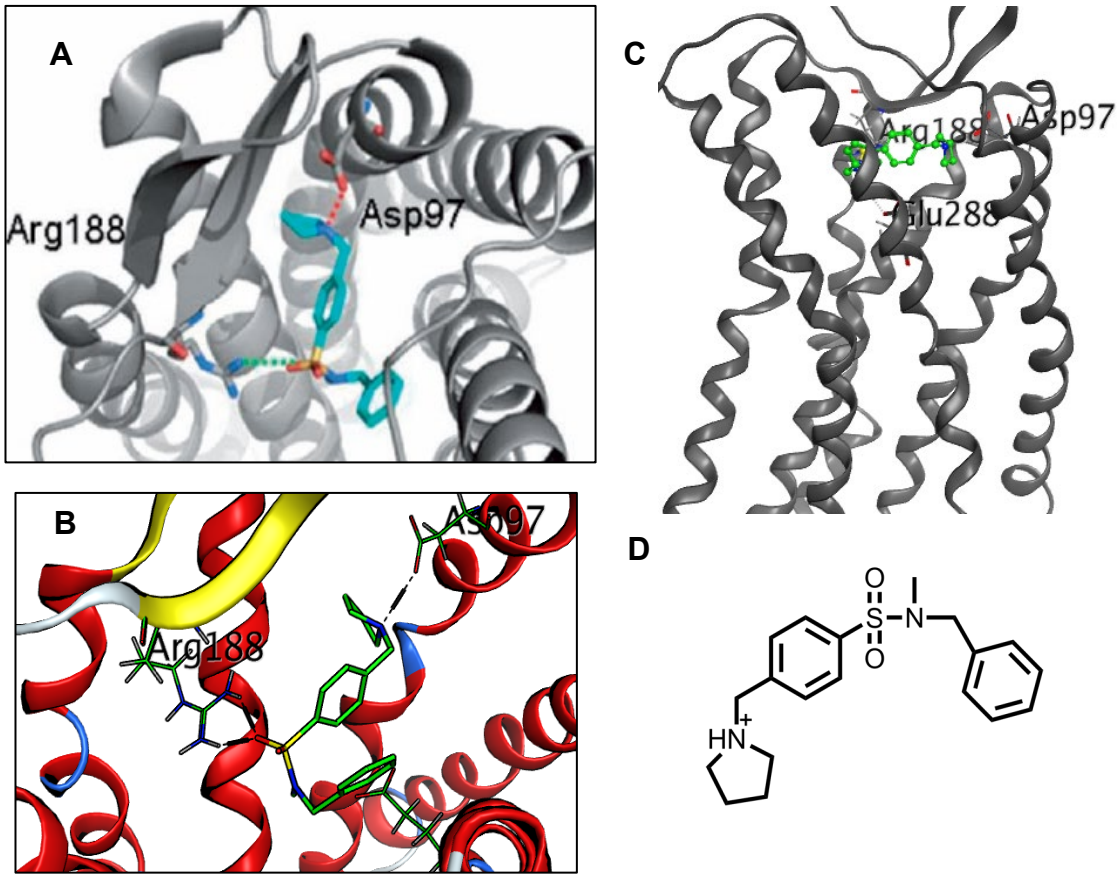


Figure 79*: A) docking pose of **6** published by Mooring et al. [80]. B/C) best docking pose of **3** generated by MOE. D) Structure of compound **6**.*

The ligand-receptor complex in **Figure 79B** was used for repeating the docking of compound **6** and other pyrrolidine compounds. The new top docking pose for **6** has the GWVI/WSA ΔG score of -8.37 kcal/mol after energy minimisation kcal/mol, indicating a more favorable docking pose. It shares the same ligand interaction; the benzyl group binds at a different pocket (**figure 80**).

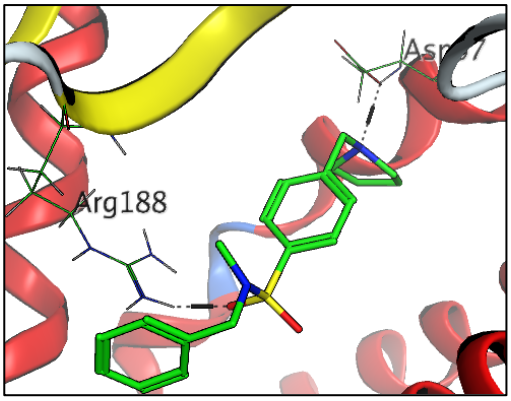


Figure 80*: Best docking pose of **6** at CXCR4 transmembrane site.*

Compound **72** is an analogue of **6** without the *N*-methyl group on sulfonamide. The compound was synthesised to investigate the role of the methyl group, the *N*-methyl sulfonamide is present in all of Mooring et al.'s CXCR4 antagonists with the benzyl moiety, but its role was not discussed. The docking pose of **72** is similar and the ligand interaction with Asp97 and Asp187 are the same (**Figure 81**). In the absence of the methyl group, the proton of the amide forms a hydrogen bond with Asp187, a residue required for CXCL12 binding with CXCR4. This possibly indicates **72** binds more strongly to CXCR4 and possibly a more potent antagonist than **6**. GWVI/WSA ΔG score after energy minimisation is -8.11 kcal/mol. The score is slightly lower than the score observed in **6**, it is unlikely that having the additional ligand interaction with Asp187 makes this docking pose less favourable than the docking pose observed for **6**.

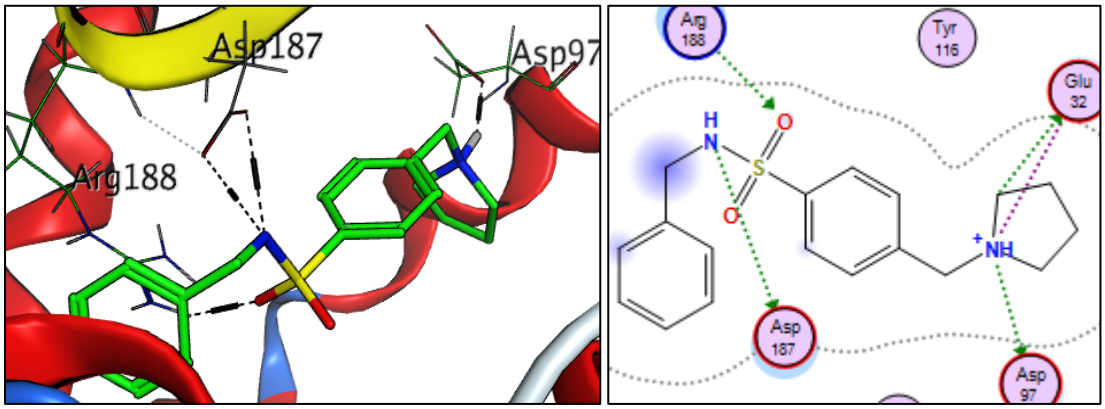


Figure 81: *Top docking pose of compound **72** at CXCR4 transmembrane site and its ligand interaction map generated by MOE.*

The pyrrolidine compounds with 4-*tert*butylbenzene (general structure at **Figure 82A**) are the hybrid compounds incorporating features from the CCR7 and CXCR4 antagonists, and they are intended to be dual CCR7 and CXCR4 antagonists. The top scoring docking poses for these compounds generally have the same ligand interaction with Asp97 and Asp187. **Figure 82B** shows the superposition of the docking poses of these compounds from a side view. The 4-*tert*butylbenzene docks at the extracellular region close to the extracellular loop 2. For some of the top docking poses with lower scores, the 4-*tert*butylbenzene occupies the pocket near Asp97 instead, and the pyrrolidine docks at the extracellular region.

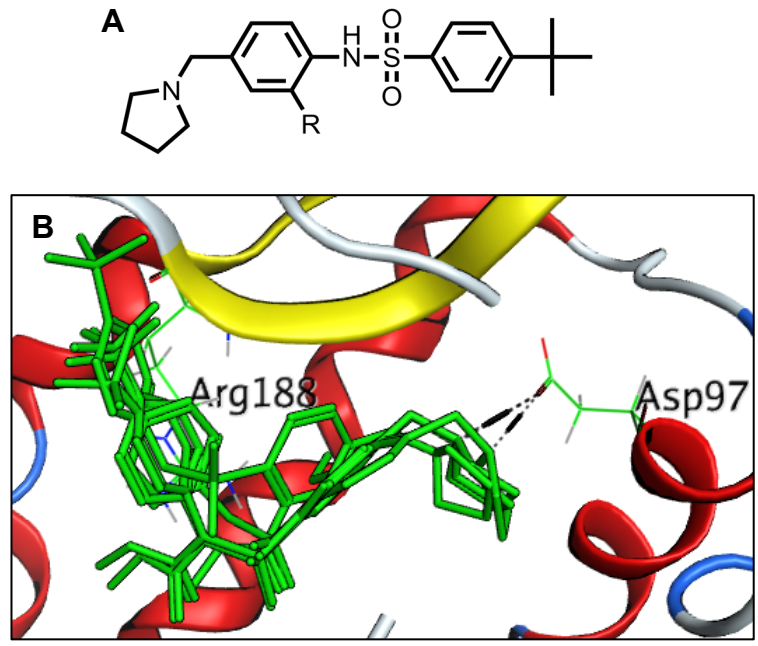


Figure 82: *A) general structure of pyrrolidine compounds with 4-tertbutylbenzene. B) Superposition of the top scoring poses of pyrrolidine compounds with 4-tertbutylbenzene.*

Compound **72** is one of the pyrrolidine compounds with 4-*tert*butylbenzene. One of the best scoring docking poses for compound **72** (**Figure 83**) showed similarity to the docking pose of **6**. The pyrrolidine interacts with Asp97, the 4-*tert*butylbenezene docked at the sub-pocket where the benzyl ring of compound **6** and **72** docked, which is different from the docking poses shown in **Figure 82**. The GWVI/WSA ΔG score after energy minimisation is -9.64 kcal/mol. This docking pose appears to be more favourable compared to the docking poses at **Figure 82**; hence the ligand-receptor complex of this docking pose was used for repeating the docking of the pyrrolidine compounds, in attempt to make the rest of the pyrrolidine compounds with 4-*tert*butylbenzene adopt this docking pose.

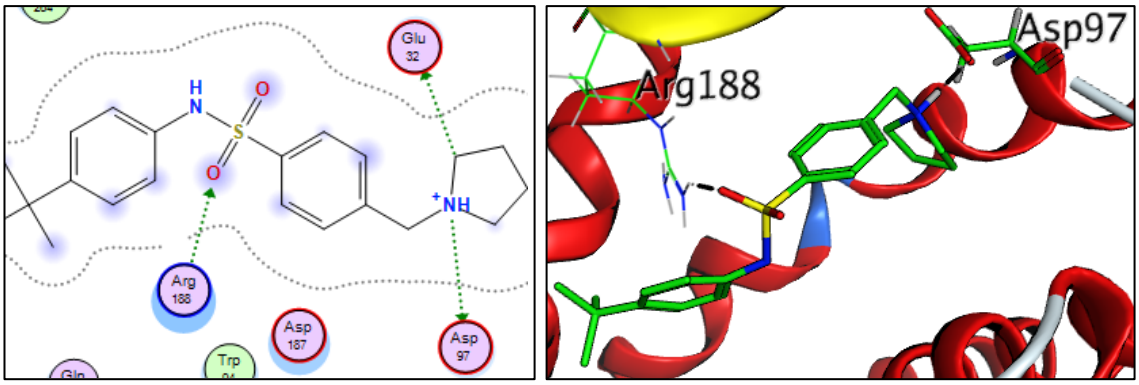


Figure 83*: Docking pose of compound* **72** *and its ligand interaction map generated by MOE.*

The new sets of docking poses were able to adopt the docking pose, and saw improvement in docking score. The overall top docking scores prior to energy minimisation were elevated for all compounds. The range of the scores prior to energy minimisation for GWVI/WSA ΔG is -7.88 to -10.12 kcal/mol, elevated from -7.14 to -7.61 kcal/mol. The superimposition of these top scoring docking poses is shown below (**figure 84**).

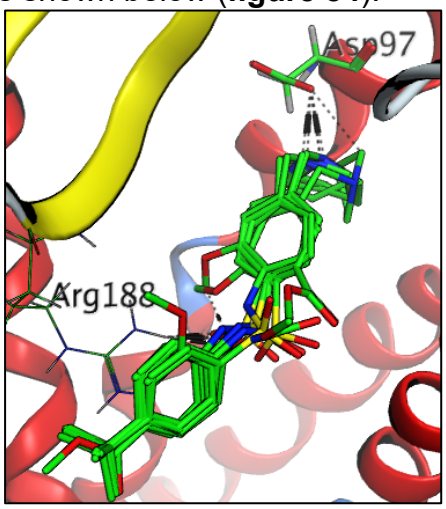


Figure 84*: Superposition of the top scoring docking poses of the pyrrolidine compounds with 4-tertbutylbenzene.*

Compound **78** is similar to **73** but with a methoxy group *ortho* to the sulfonamide. In addition to the ligand interaction with Arg188 and Asp97. The proton of the sulfonamide forms hydrogen bond with Asp187 (**Figure 85**). The methoxy group forms hydrogen bond with Ser285. The additional ligand interaction possibly indicate it is more potent than **72**, if they are active in CXCR4. The GWVI/WSA ΔG score after energy minimisation is -10.40 kcal/mol.

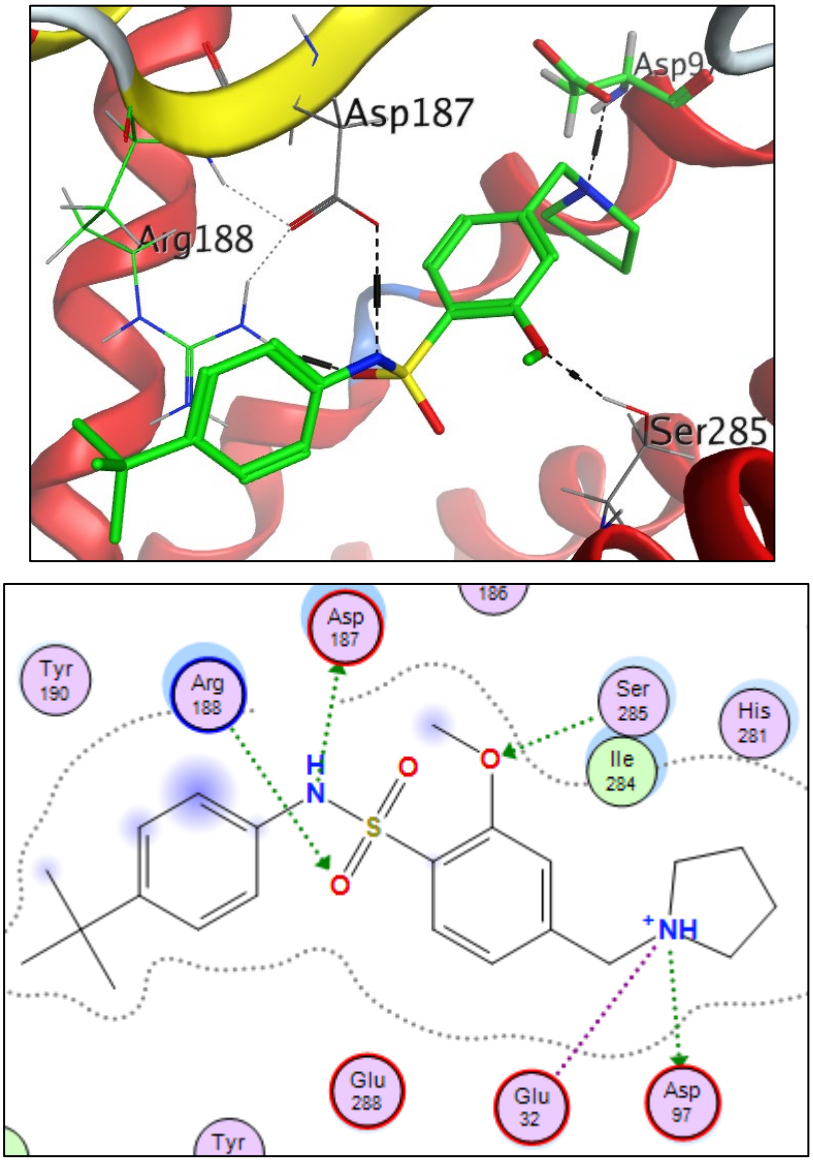


Figure 85*: Best docking pose of compound **78** at the CXCR4 transmembrane site, and its ligand interaction map generated by MOE.*

Compound **87** is similar to **78** but has an ethyl ester instead of a methoxy group. The proton of the sulfonamide also forms hydrogen bonding with Asp187 (**Figure 86**). It appears in this class of compounds, the sulfonamide is able to form ligand interactions with both Arg188 and Asp187. The ethyl ester is also able to form interaction with Ser285. It might be beneficial for the pyrrolidine CXCR4 antagonists of Mooring et al. to have a substituent *meta* to the pyrrolidine side chain. None of their compounds have substituents at that position.

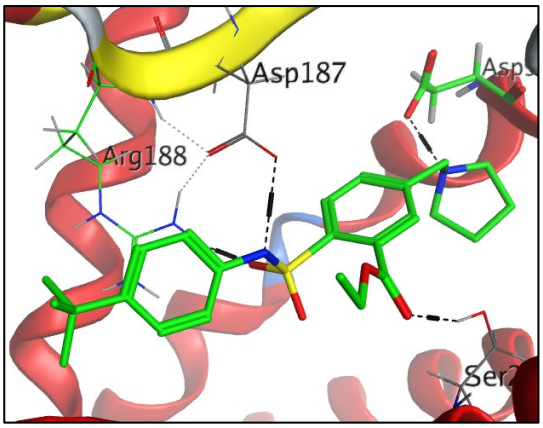
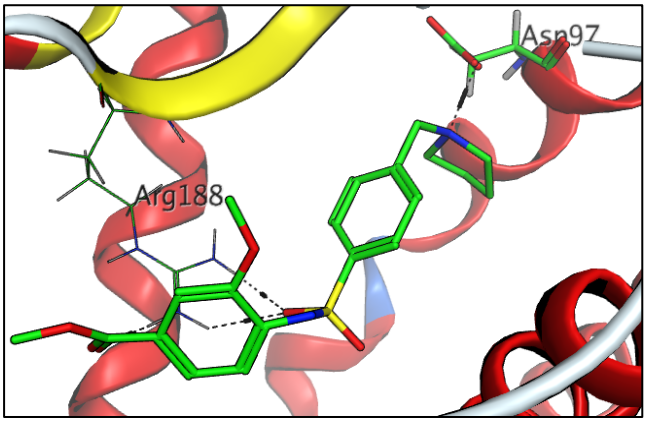
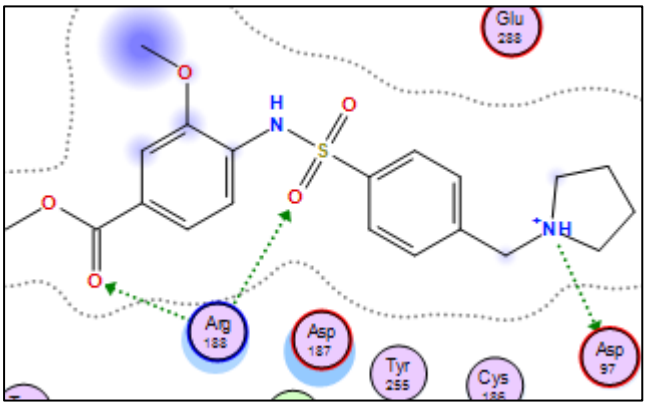


Figure 86*: Best docking pose of **87** at the CXCR4 transmembrane site*





For compound **79 (Figure 87)**, the pyrrolidine and sulfonyl show the same interaction with Asp97 and Arg188. The carbonyl of the ester also forms a hydrogen bond with Arg188. The GWVI/WSA ΔG score after energy minimisation is -10.19 kcal/mol. While the ester is able to form a ligand interaction with Arg188, the benzene ring containing hydrophilic substituents occupies a pocket that is occupied by the lipophilic benzyl group, indicating it is not favourable.

Figure 87*: Best docking pose of **79** at the CXCR4 transmembrane site, and its ligand interaction map generated by MOE.*

Docking of compound 67 and extended chain compounds

The best docking pose of **78** and its receptor pair were used for the docking of compounds similar to ICT13069, and the compounds with extended chains. Most of our CCR7 compounds and the extended chain compounds have carboxylic acid or methyl ester at the other end of compound, thus they are unable to form interaction with Asp97. For most of the docking poses of the extend chain compounds, the extended chain was not able to form ligand interaction with nearby residues.

Compounds **67** and **42** have the methyl amide moiety. These are the very few CCR7 antagonists that can form ligand interaction with the Asp97. The top scoring docking pose (**Figure 88**) of **67** is very similar to **78** (**Figure 85**). The sulfonamide is able to form hydrogen bond with both Arg187 and Arg188. The proton of the methyl amide interacts with Asp97. GWVI/WSA ΔG score after energy minimisation is -9.00 kcal/mol. Replicating similar ligand interaction in **40** was difficult. The order of the sulfonamide in **67** appears to affect the conformation of the compound, making it easier to form ligand interaction with both Arg188 and Asp197.

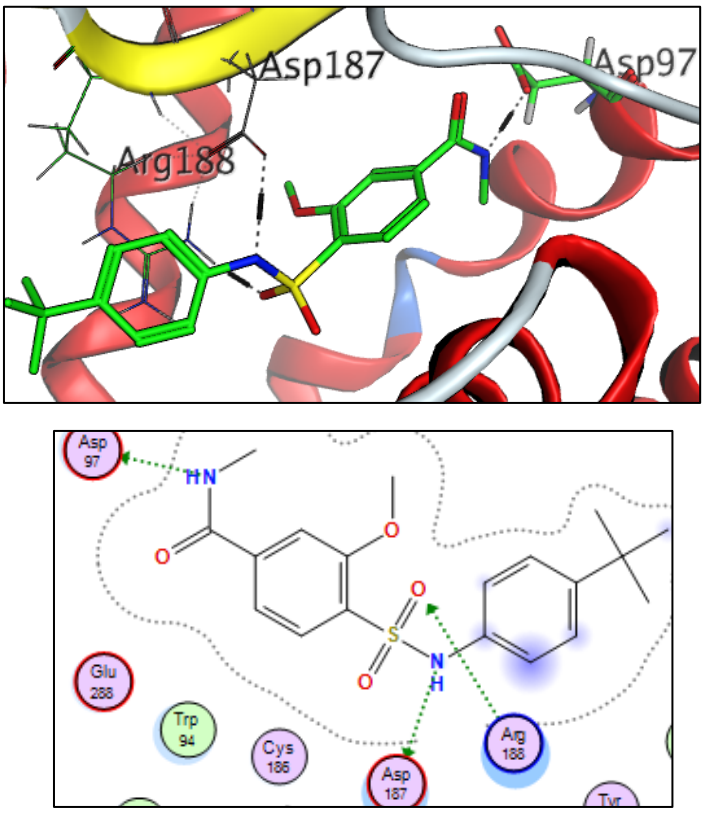


Figure 88*: Best docking pose of* ***67*** *at the CXCR4 transmembrane site, and its ligand interaction map generated by MOE.*

Summary and key findings of docking at the CXCR4 transmembrane site

- Ligand interactions of compound **3** were replicated in our own docking pose. The sulfonyl group forms a hydrogen bond with Arg188, the protonated pyrrolidine forms salt bridge with Asp97, a residue important for CXCL12 binding.
- The benzyl of compound **3** sits at a pocket consist of Gln200, Arg188, Tyr255, His203. Ile284
- Our 4-*tert*butylbenzene sulfonamide compounds such as **67** are able to form a hydrogen bond with Asp187, one of the residues required for CXCL12 binding. The removal of the methyl group on the sulfonamide of Mooring et al.'s CXCR4 compounds might improve their potency.
- Our 4-*tert*butylbenzene sulfonamide pyrrolidine compounds with hydrogen bond acceptor substituents *para* to the pyrrolidine side chain are able to form ligand interaction with Ser285, indicating having a small hydrogen bond acceptor substituent at that position might improve the potency of this class of CXCR4 antagonists.
- Our CCR7 antagonist **67** is able to form ligand interaction with both Asp97, Arg188 and Asp187, indicating it might be active in CXCR4.
- The "normal" sulfonamide such as ICT13069 and **42** are possibly unable to form interaction with both Asp97 and Arg188, indicating the order of the sulfonamide is important in this class of sulfonamide antagonists for the binding of CXCR4, and possibly less potent than the "reverse" sulfonamide analogues if they are active and binds to the transmembrane site in CXCR4.

5.6.4 CXCR4 binding: Intracellular domain

Although no intracellularly allosteric antagonist has been discovered for CXCR4, we also explored the possibility that the docking of sulfonamides might occur at that site. This would be in contrast to what was suggested by Mooring et al. [80]. CXCR4/3ODU was superposed onto CCR7/6QZH with compound **6** docked at its intracellular site, 6QZH then was deleted. The molecule was outside of the intracellular site of 3ODU, thus can't be used for docking. Docking studies of **6** using the site finder application yielded docking poses with very low GWVI/WSA ΔG scores: the highest was -5.99 kcal/mol after energy minimisation. The 3ODU intracellular site appears to be unsuitable for docking studies at the intracellular site due to the conformation of the 3ODU C-terminus. In CCR7 and CCR9, residues at the intracellular loops (which are situated below TM7) are important to the docking of Cmp2105, Vercirnon and the antagonists of this project. Superposition of CXCR4/3ODU to CCR2/5/7/9 crystal structures in **Figure 89** below shows the 3ODU C-terminus has a very different conformation compare to the others. Gly306 and Lys308 are the corresponding residues to the residues involved in the binding of Vercirnon and others, and they are far away from TM7, thus unable to form ligand interaction with our compounds. Other CXCR4 crystal structures cannot be used for the docking as they have no C-terminus.

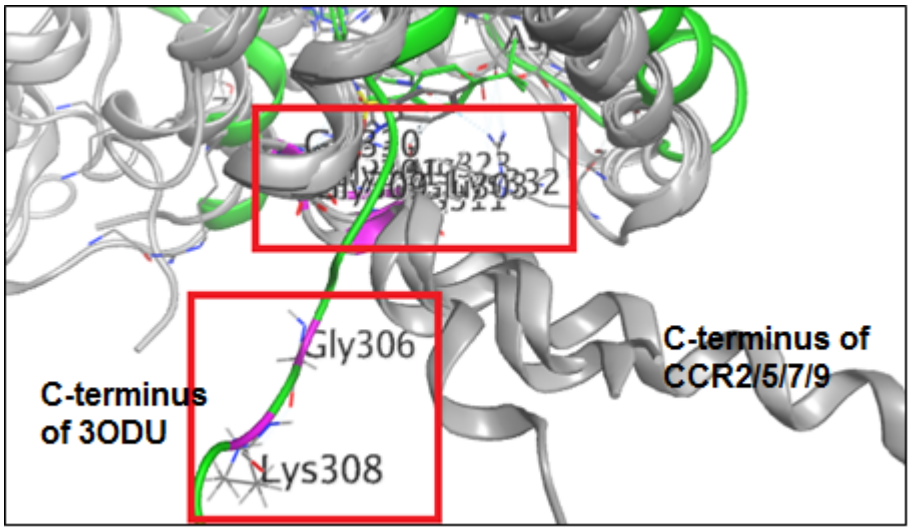


Figure 89: *Superposition of CXCR4/3ODU (Green) to the crystal structures of CCR2,5,7,9. The corresponding residues involved in the binding of Vercirnon are highlighted. The highlighted Gly306 and Lys308 belong to CXCR4.*

Of course, the reason for this may be a corollary to why the model based on crystal structure of CCR7 (with a ligand in the intracellular allosteric site) was not suitable for docking molecules in the transmembrane orthosteric site. For a similar reason, a model based on the crystal structure of CXCR4 (with a ligand in the transmembrane orthosteric site) may not be suitable for docking molecules in the intracellular allosteric site.

Therefore, we attempted to create a CXCR4 homology model for the purpose of docking the pyrrolidine compounds to the intracellular site. 6QZH and 5LWE were used for the templates. First, we used SWISS MODEL to generate the homology model, however the homology models generated from both templates lack the C-terminus, thus the important ligand interactions for intracellular binding cannot be formed and the model cannot be used for our study. The homology models generated in MOE were of poor quality. The 6QZH based model is unable to form the helix shape in one of the transmembrane helixes. The 5LWE based model has a transmembrane helix that prematurely ends and becomes part of the extracellular loop. This might be due limitations of the software. Considering the nature that it is difficult to determine whether the compounds bind at the intracellular from the docking study, and it is more likely that our compounds bind at the transmembrane site, we decided this was not a priority and did not investigate it further.

5.6.5 CCR7 binding: extracellular and transmembrane domain – homology model

Due to difficulties with using models based on CCR7 X-ray crystal structure, we decided to use homology modelling for the molecular docking of CCR7 antagonists in the transmembrane site. 5LWE has Vercirnon docked at the intracellular site. We should note that it possibly may have also changed the conformation of the orthosteric site and made it less exposed. The CCR7 homology model made from 5LWE indeed has a small binding site at the transmembrane region. The calculated binding site of CCR7 and the CCR7 homology model made from 5LWE are shown in **Figure 90A/B**. For comparison, **Figure 90C** shows the calculated binding site of CXCR4 and it has wider pocket at the transmembrane domain. This implies that although the CCR9 is closest homologue to CCR7, its use as the template will lead to a homology model of CCR7 with a concealed orthosteric site.

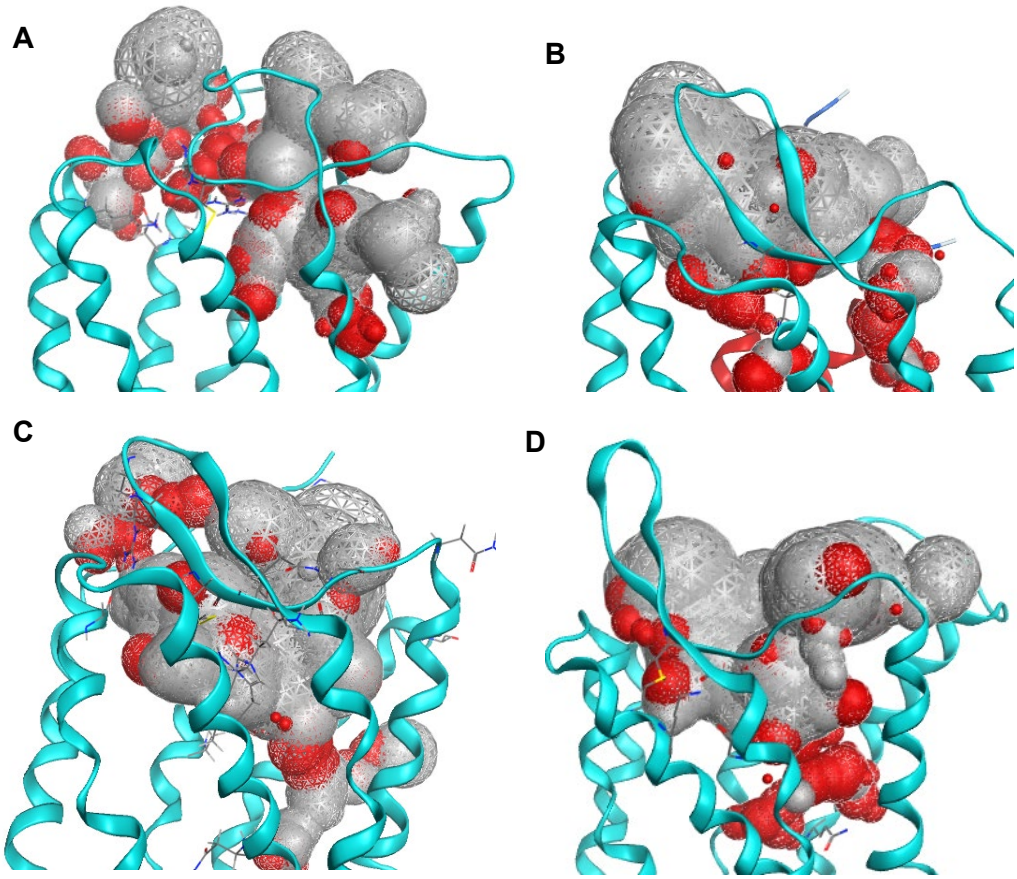


Figure 90, *3D representation of the binding site calculated by MOE. A) CCR7 homology model from 5LWE template. B) CCR7. C) CXCR4. D) CCR7 homology model made from 3ODU template.*

The next protein with the highest sequence identity for CCR7 is CXCR4 at 34%, which does not have a ligand at the allosteric site but does have a ligand at the orthosteric site. SWISS MODEL [110] were used to generate a CCR7 homology mode by using 3ODU as the template. Its calculated binding site is similar to CXCR4 (**Figure 90D**). It has a wide orthosteric site as expected.

The model generated by SWISS-MODEL [110] was assessed by Ramachandran plot and the online server tool MolProbity [111], most of the assessed properties are below the ideal values. Ramachandran plot provided by MolProbity [111] shows 2 outliers and 95.69% residues at the favoured region.

Poor rotamers	4	1.57%	Goal: <0.3%
Favored rotamers	244	95.69%	Goal: >98%
Ramachandran outliers	2	0.70%	Goal: <0.05%
Ramachandran favored	274	95.80%	Goal: >98%
Rama distribution Z-score	1.45 ± 0.49		Goal: abs(Z score) < 2
Cβ deviations >0.25Å	0	0.00%	Goal: 0
Bad bonds:	0 / 2358	0.00%	Goal: 0%
Bad angles:	13 / 3204	0.41%	Goal: <0.1%
Cis Prolines:	0 / 6	0.00%	Expected: ≤1 per chain, or ≤5%

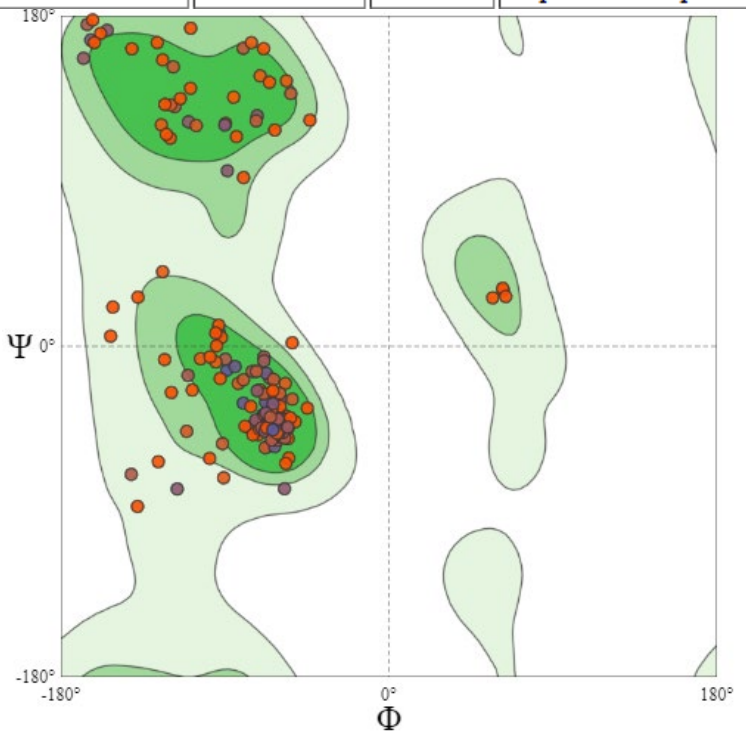


Figure 91*: The Ramachandran plot and other properties used for evaluating the CCR7 homology model.*

Docking of compounds similar to ICT13069 (Group A)

We first used the site finder application to dock compound **67** to the CCR7 homology model, after applying energy minimisation to the top scoring docking pose. The energy minimised ligand-receptor complex was used for repeating the docking of **67** and further docking studies. The newly obtained docking pose (**Figure 92**) for **67** has the GWVI/WSA ΔG score is -9.61 kcal/mol. after energy minimisation. The docking pose shows the following ligand interaction: sulfonyl forms a hydrogen bond with Leu212, the carbonyl at the polar ring forms a hydrogen bond with Ser117. The proton of the methyl amide forms a hydrogen bond with Trp114.

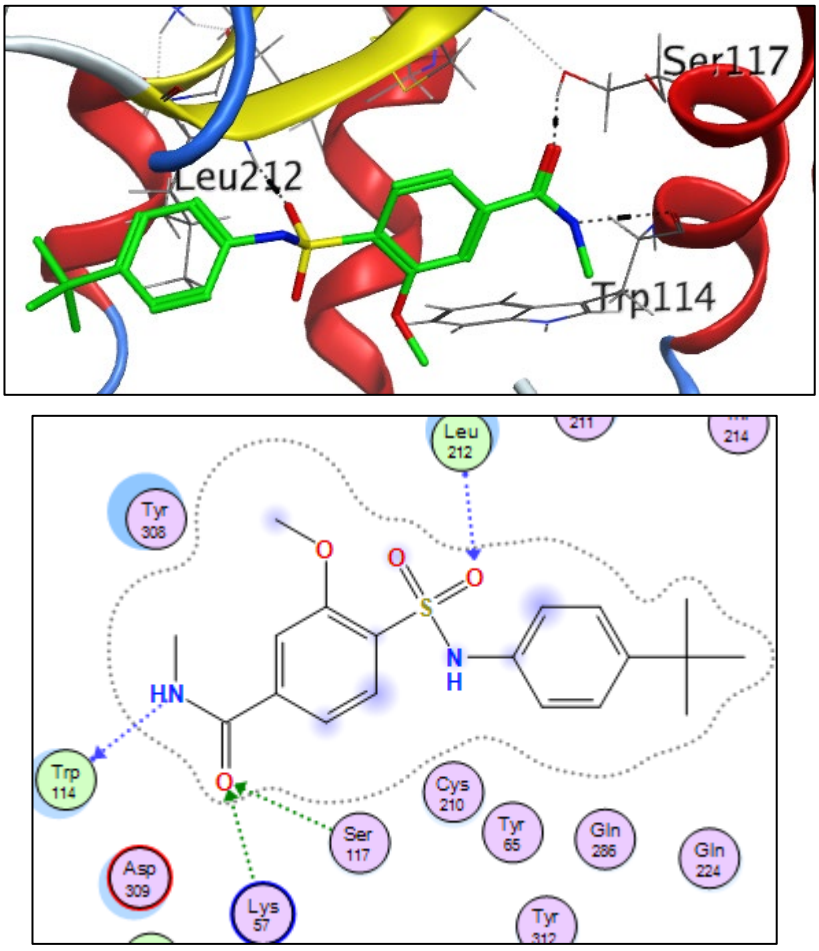


Figure 92: *Possible docking pose of compound* **67** *at the transmembrane site of the CCR7 homology model, and the ligand interaction map generated by MOE*

Compound **42** is similar to **67** but with a reversed sulfonamide: **67** is from the "reverse" sulfonamide series, **42** is from the "normal" sulfonamide series. The top scoring docking pose of **42** is very similar to that of **67**. The same ligand interactions are formed (**Figure 93A/B**), the docking scores are also similar. Superimposing the best docking pose of the two compounds show the sulfonyl interacts with Leu212 in almost same space. The same applies to the carbonyl group with Ser117 (**Figure 93C**).

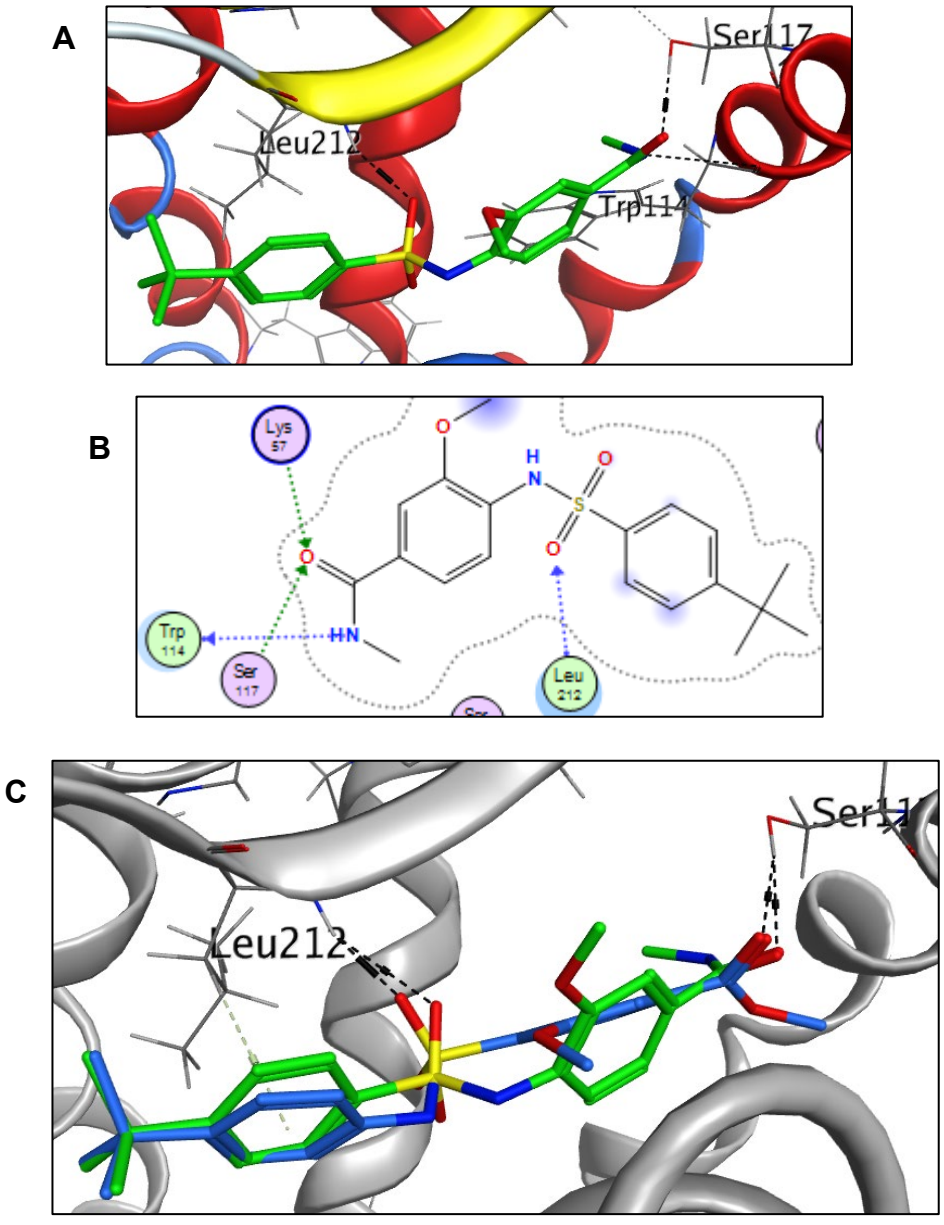


Figure 93: *A) Docking pose of compound **42** at the transmembrane site of the CCR7 homology model, B) the ligand interaction map of **42** generated by MOE C) Superposition of **42** and **67**.*

Docking of extended chain compounds

The 4-*tert*butylbenzene sulfonamide of the extended chain compounds have similar docking poses to the previous compounds (**42**, **67**), and a similar ligand interaction with Leu212. The extended chain compounds are able to form a hydrogen bond with Ser117, but the flexibility of the extended chains also allows the carboxylic acid to interact with either Lys57 or Asp309 in other docking poses. The energy difference between these poses is small. The ligand interaction maps of the carbonyl of compound **98** and **107** interacting with these 3 residues are shown below (**Figure 94**).

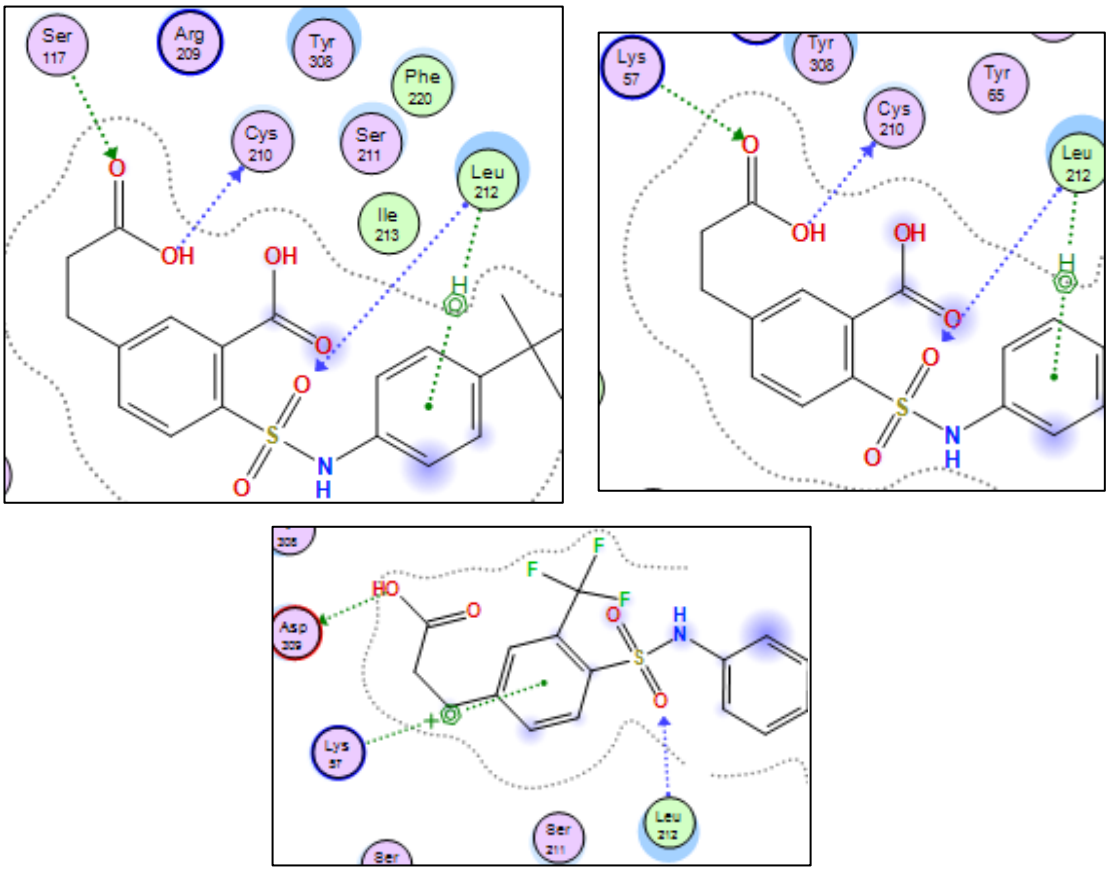


Figure 94: *Interaction maps of extended chain compounds generated by MOE. The carboxylic acid is able to form hydrogen bond with Ser117, Lys57 and Asp309 in different docking poses.*

Docking of pyrrolidine compounds

The CCR7 and CXCR4 pyrrolidine compounds have similar docking poses at the CXCR4 orthosteric site. However, the docking of compound **6** and other pyrrolidine compounds showed a different docking pose and ligand interactions at the CCR7 homology model. While the sulfonyl group interacts with Leu212, the protonated pyrrolidine forms salt bridge with Asp309 instead of Ser117 (**Figure 95**). The GWVI/WSA ΔG score after energy minimisation is -8.77 kcal/mol.

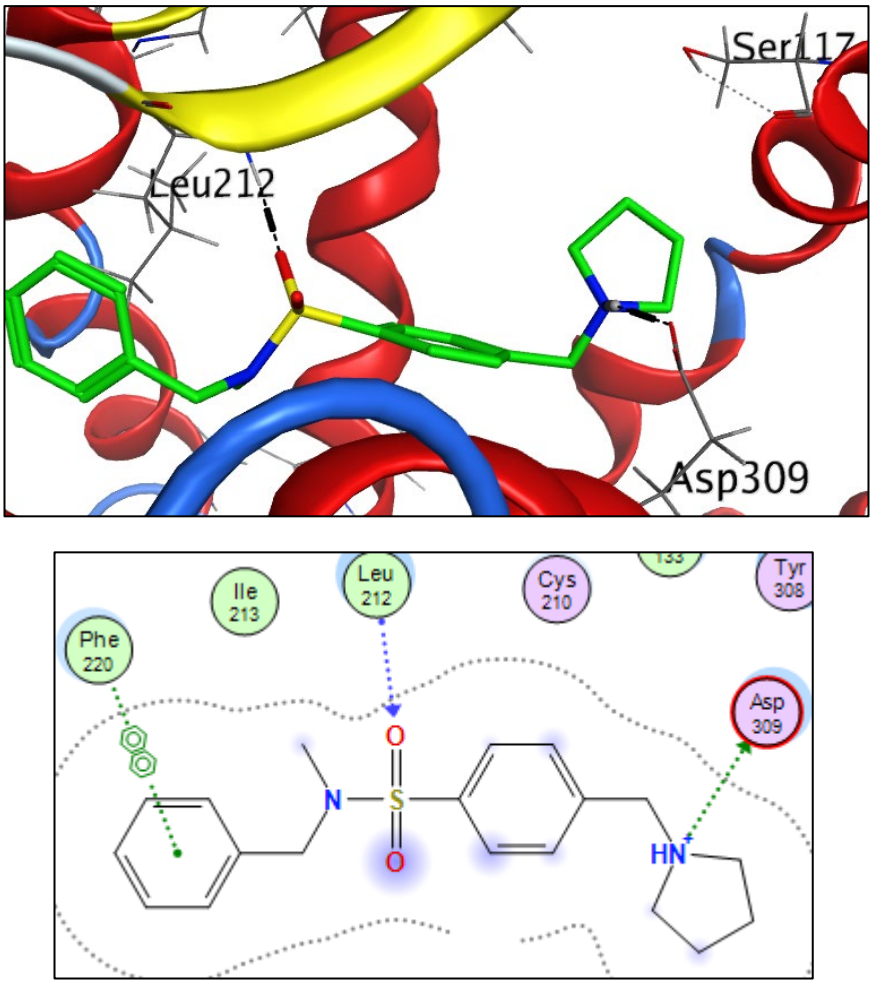


Figure 95*: Possible docking pose of compound **6** at the transmembrane site of the CCR7 homology model, and the ligand interaction map generated by MOE.*

For pyrrolidine compounds with 4-*tert*butylbenzene, some of the lower energy docking poses have a similar docking pose as the CCR7 antagonists, but the top scoring docking poses favour the docking pose observed in **figure 95**, in which the pyrrolidine fits in pocket with Asp309 and forms salt bridge (**Figure 96A/B**). The top scoring docking pose for compound **78** has the GWVI/WSA ΔG score of -

10.68 kcal/mol after energy minimisation. **Figure 96C** shows the superimposition of the docking pose of the CCR7 compounds and the pyrrolidine compounds.

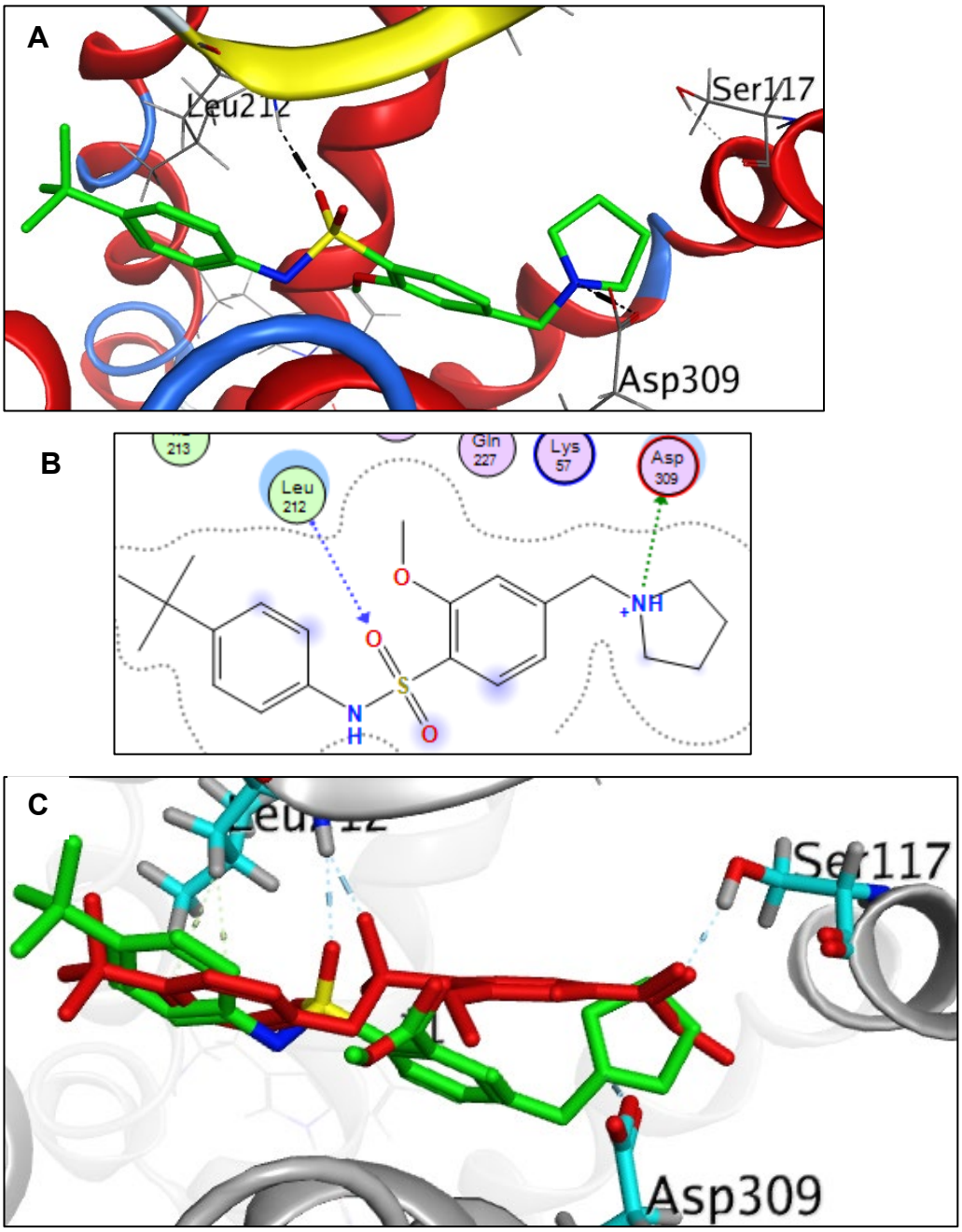


Figure 96: *A) Top docking pose of compound **78** at the transmembrane site of the CCR7 homology model. B) The ligand interaction map of **78** generated by MOE. C) Superposition of CCR7 antagonist docking pose with the pyrrolidine compound docking pose.*

The top scoring docking poses of **79** alternate between two conformations, both retain interaction with Leu212, the energy difference between the two docking poses are very small. The first one has the same ligand interaction for the pyrrolidine and Asp309 (**Figure 97**). However, the polar ring sits at the same hydrophobic ring where 4-*tert*butylbenzene would occupy. The molecule in this conformation is able to form hydrogen bond with Leu212 and Gln227. Gln227 is a residue important for CCR7 activation [118]. The GWVI/WSA ΔG score after energy minimisation is -10.15 kcal/mol.

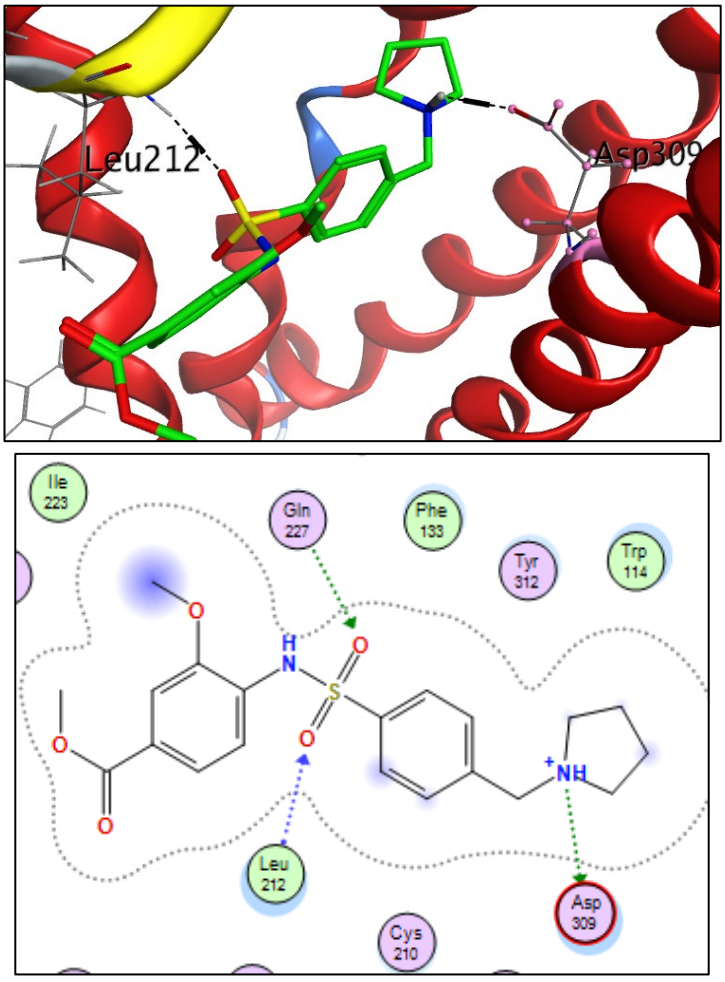


Figure 97: *Top docking pose of compound **79** at the transmembrane site of the CCR7 homology model, and the ligand interaction map generated by MOE.*

The other docking pose of **79** appears to be more favourable. Pyrrolidine doesn't dock at the pocket near Asp309. Instead it reaches further than space occupied by 4-*tert*butylbenzene in other compounds, and forms salt bridge with the carbonyl of Gln286, the carbonyl of the methyl ester is able to form hydrogen bond with Lys57 (**Figure 98**). The GWVI/WSA ΔG score after energy minimisation is -10.35 kcal/mol.

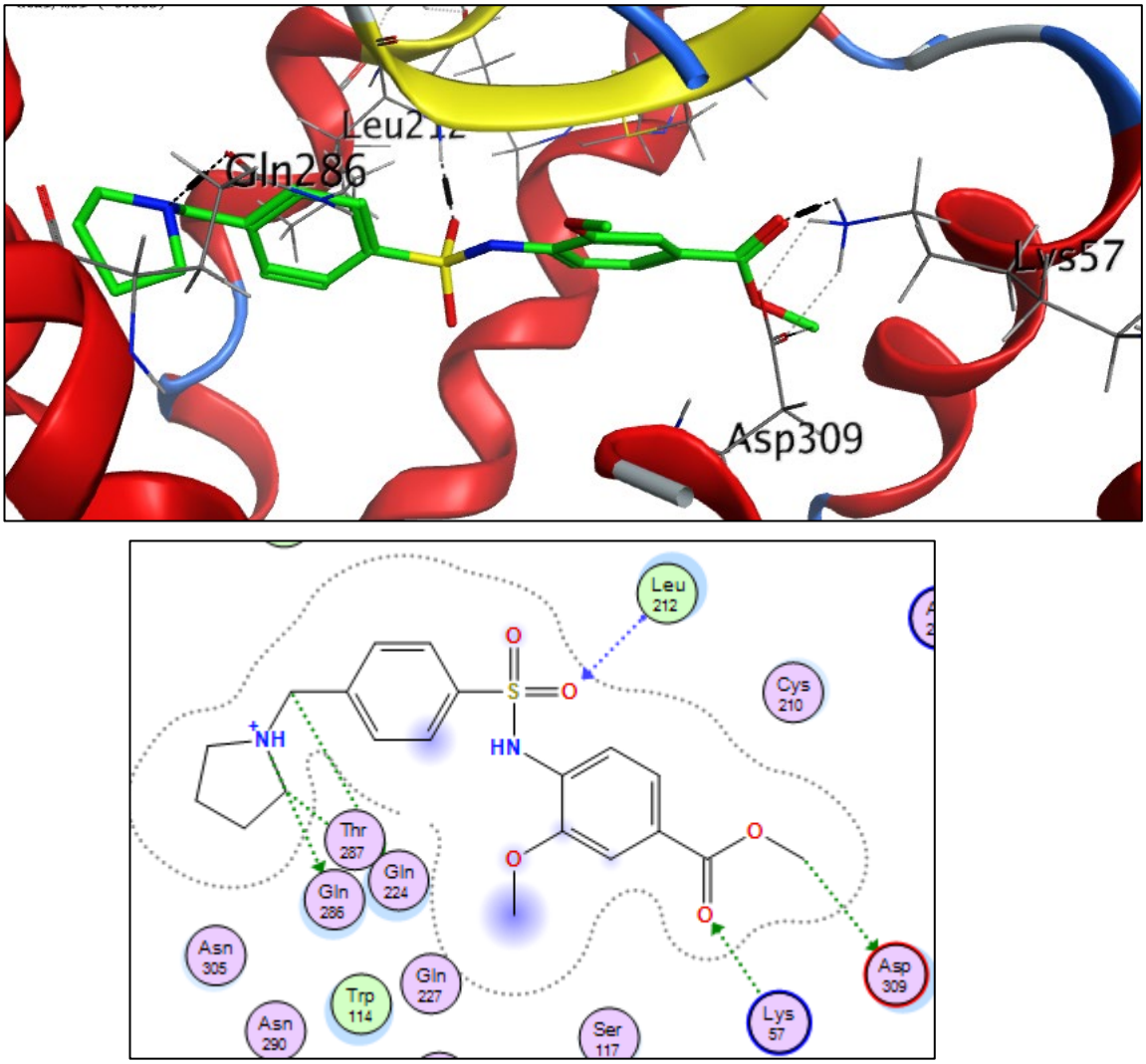


Figure 98: *Second top docking pose of compound **79** at the transmembrane site of the CCR7 homology model, and the ligand interaction map generated by MOE.*

The first docking pose of compound **79** (**Figure 97**) was used to repeat the docking of the pyrrolidine compounds to replicate the conformation that is able to form ligand interaction with Gln227 and Leu212. This was successful with other pyrrolidine compounds except for **6**. The docking pose of **87** (**Figure 99**) is representative of the pyrrolidine compounds. Their respective GWVI/WSA ΔG score after energy minimisation are above -10 kcal/mol. This might be an indication that these compounds are active in CCR7. Replicating this conformation with CCR7 compounds wasn't successful.

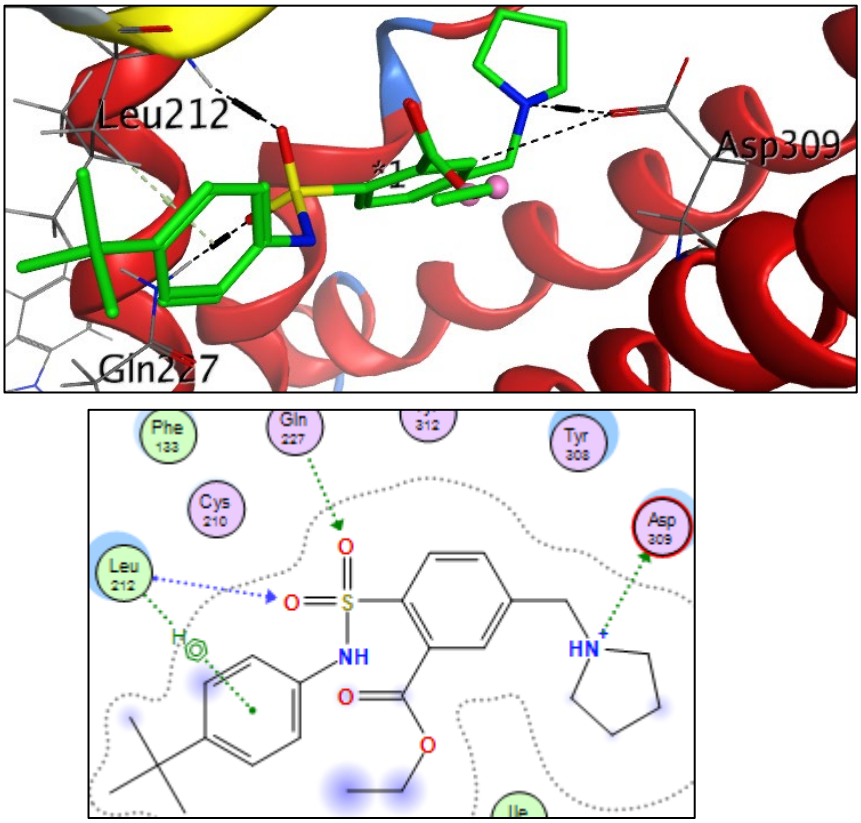


Figure 99: *Top pose of compound **87** at the transmembrane site of the CCR7 homology model, and the ligand interaction map generated by MOE.*

Summary and key findings of docking at the CCR7 homology model transmembrane site made from 3ODU

- The transmembrane binding site of CCR7/6QZH is possibly closed up due to the binding of Cmp2105 at the intracellular site, changing the conformation of CCR7.
- A new CCR7 homology model was generated from 3ODU because the homology model made from CCR9/5LWE also has a closed up transmembrane site, likely due to the binding of Vercirnon at the intracellular site.
- In general, the 4-*tert*butylbenzene moiety docks at a hydrophobic pocket consist of Phe220, Leu112, Gln286 and Tyr308.
- Sulfonyl group forms hydrogen bond with Leu212. The polar group *para* to the sulfonamide in CCR7 compounds forms hydrogen bond with Ser117
- The docking orientation and ligand interaction of CCR7 compounds in the homology model, are the same as compound **3** in CXCR4/3ODU. Leu212 and Ser117 are the corresponding residues of Arg188 and Asp97 in CXCR4. The superposition of CCR7 homology model with a CCR7 antagonists, and CXCR4 with compound 3 is shown in **Figure 100**.
- The order of the sulfonamide in the CCR7 compounds doesn't affect the ligand interaction and the orientation of the docking poses, the docking study is unable to identify whether the "normal" or the "reverse" sulfonamide have a more favourable docking pose.
- Docking poses of pyrrolidine compounds have a different conformation than the CCR7 compounds. The protonated pyrrolidine forms a salt bridge with Asp309. The rest of the molecule docks similarly to other CCR7 antagonists.
- The sulfonyl group in some docking poses of pyrrolidine compounds with 4-*tert*butylbenzene (compound **73**, **78** and **87**) are able to form hydrogen bond with Leu212 and Gln227. Gln227 is important in CCR7 activation. This indicates these pyrrolidine compounds might be active in CCR7.

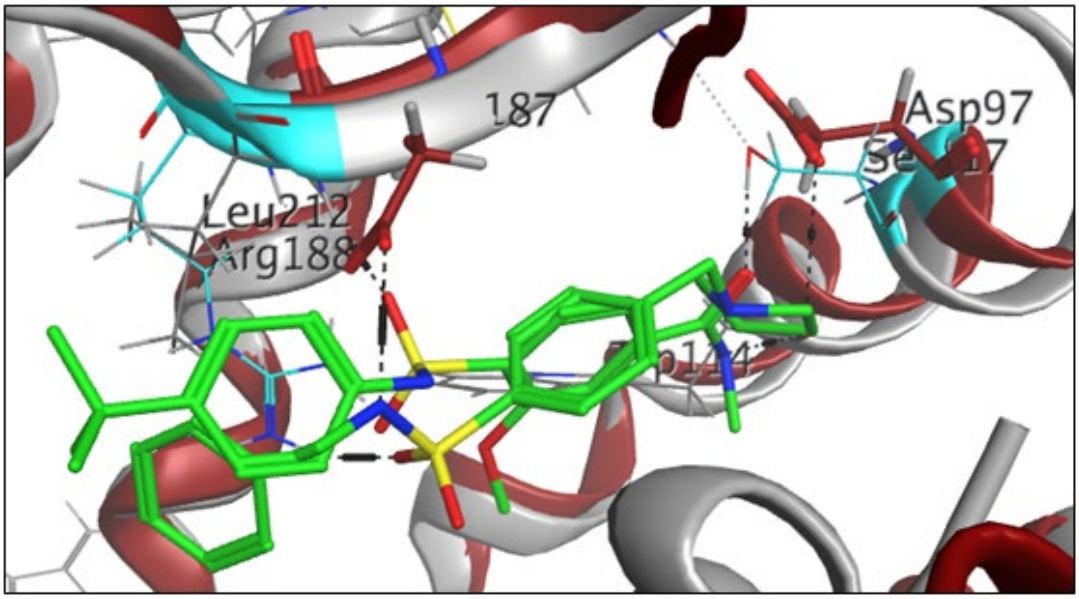


Figure 100: *The superposition of CCR7 homology model with a CCR7 antagonists, and CXCR4 with compound* ***6****.*

5.7 Correlating docking scores with IC_{50} from calcium flux assay

So far, the binding mode and binding efficacy of various actual and putative CXCR4 CCR7 antagonists are described only qualitatively. This qualitative analysis is of course important in terms of understanding how and why some molecules can bind more or less strongly. However, we also wanted to investigate if there was any quantitative correlation between binding affinities calculated using models with the experimentally derived potencies. Before we describe our results, it is worth providing a few provisos to any conclusions. Firstly, *in vitro* experimental potency/efficacy results are affected by a number of factors. For example, insoluble compounds can give unreliable results. Secondly, intracellular binding requires cell penetration, so if our antagonists do bind at the intracellular domain, the potency/efficacy results can be skewed due to lack of cell penetration. Never the less, we felt it would be interesting if we could see a correlation between experimental and computational results. 34 CCR7 compounds in group A were synthesised by other members of our group prior to this project, their structures and IC_{50} are shown in Appendix 1. We plotted the IC_{50} of these 34 compounds with their top scores for each scoring function, in both the intracellular and transmembrane site of CCR7. To identify any correlation, we used R^2 values.

Both docking scores from prior and after energy minimisation were used to generate correlation plots. Although energy minimisation generally improves the score and optimises the docking pose, in the scenario of using MOE to conduct virtual screening of a huge library of compounds, the energy minimisation step would require significantly longer time and effort. It will be helpful to find out whether the docking scores prior to energy minimisation are useful.

The initial plots revealed 7 outliers. The molecular weight affects the scoring function; compounds with higher molecular weight score higher, and smaller molecules score lower. These outliers come in three categories:

1) Potent (IC_{50}: 1 to 20 μM), but low score due to low molecular weight (below 270 Da).

2) Moderately potent (IC_{50}: 20 to 35 μM) potent but have good score due to their molecular weight (above 350 Da).

3) Least potent compounds ($IC_{50:}$ >35 μM) and low score due to low molecular weight (below 270 Da).

These outliers show compounds with molecular weight below 270 Da will have low score in our docking study regardless of the potency. Compounds that are ~350 Da or above will be prone to false postive results in screening.

These outliers were removed from the plot, the R^2 generally saw improvement. For example, the plot of GWVI/WSA ΔG vs IC_{50} at the CCR7 transmembrane increased the R_2 to 0.438 from 0.2082. The following tables show R^2 of the correlation plots prior to the removal of the outliers.

Transmembrane site	R^2 (Prior to energy minimisation)		R^2 (After energy minimisation)	
Scoring function	All 34 compounds	Outliers removed	All 34 compounds	Outliers removed
GWVI/WSA ΔG	0.2082	0.4380	0.1918	0.1874
ASE	0.2245	0.0820		
Affinity ΔG	0.2381	0.1638		
Alpha HB	0.0893	0.0611		

Table 3: R^2 of correlation plots between IC_{50} and various docking scores at the CCR7 transmembrane site.

Intracellular site	R^2 (Prior to energy minimisation)		R^2 (After energy minimisation)	
Scoring function	All 34 compounds	Outliers removed	All 34 compounds	Outliers removed
GWVI/WSA ΔG	0.1953	0.5186	0.1900	0.3569
ASE	0.3049	0.3854		
Affinity ΔG	0.2311	0.3458		
Alpha HB	0.1661	0.3896		

Table 4: R^2 of correlation plots between IC_{50} and various docking scores at the CCR7 intracellular site.

Correlation plot (before energy minimisation): CCR7 homology model transmembrane site

GWVI/WSA ΔG and ASE have the highest R^2 value of 0.438 and 0.4546 respectively, which are considered as weak correlations. R^2 value of Affinity ΔG and Alpha HB are below 0.2, showing no correlation. The GWVI/WSA ΔG and ASE plots show that their above 0.4 R^2 value are due to the 6 compounds on the upper right of the curve (**Figure 101**), the compounds with IC_{50} above 10, and docking score value higher than -7.5 kcal/mol and -27 kcal/mol respective to the scoring functions. 5 out of 6 compounds are nitro compounds, the remaining compound is another low molecular weight compound. In the scenario that we wish to use MOE to screen compounds without nitro groups and have molecular weight above 300 Da, we created a correlation plot without these 6 compounds. The R^2 is 0.0015, showing no correlation.

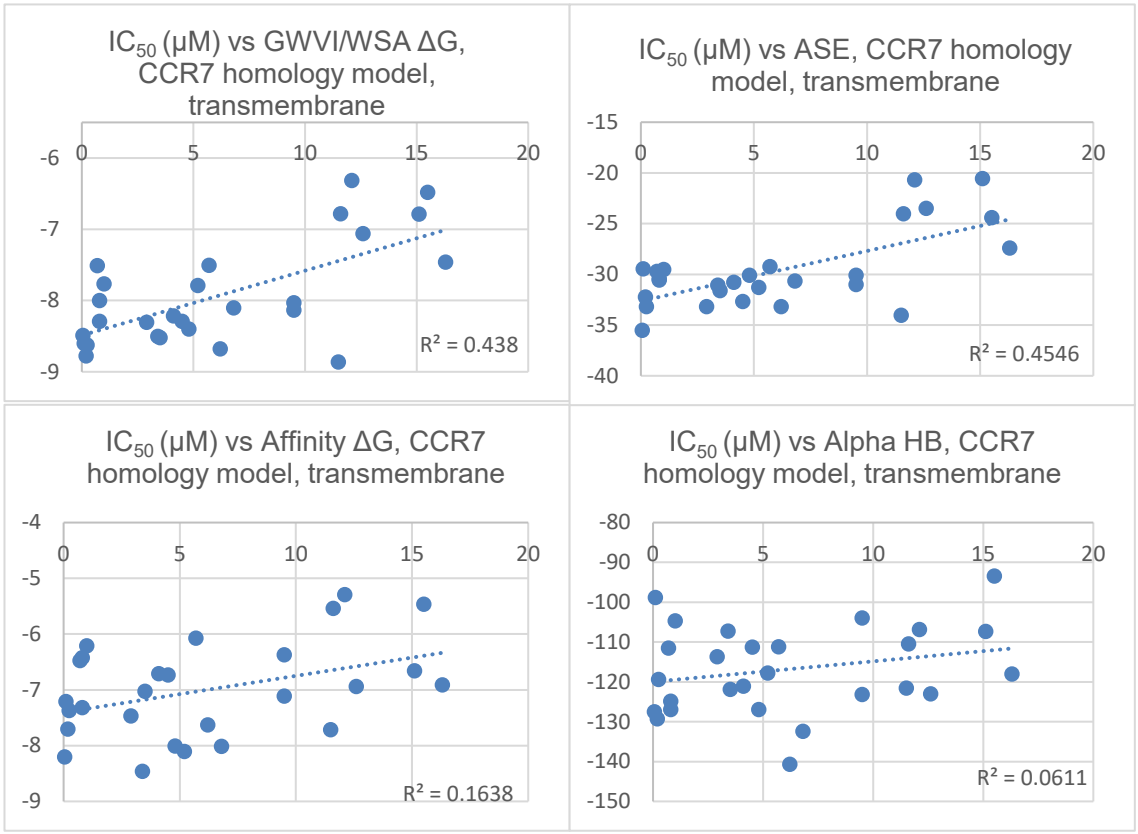


Figure 101*: Correlation plot between IC_{50} of 27 CCR7 antagonists and their top docking scores at the transmembrane site for each scoring function available to MOE.*

Correlation plot (before energy minimisation): CCR7/6QZH intracellular site

The GWVI/WSA ΔG plot has the highest R^2 of 0.5186, which is also a weak correlation. Others are below 0.4. Again, this R^2 of 0.5186 hinges on the 5 aforementioned nitro compounds and the low molecular weight compound. Removal of these 6 compounds plunges the R^2 below 0.1, showing no correlation.

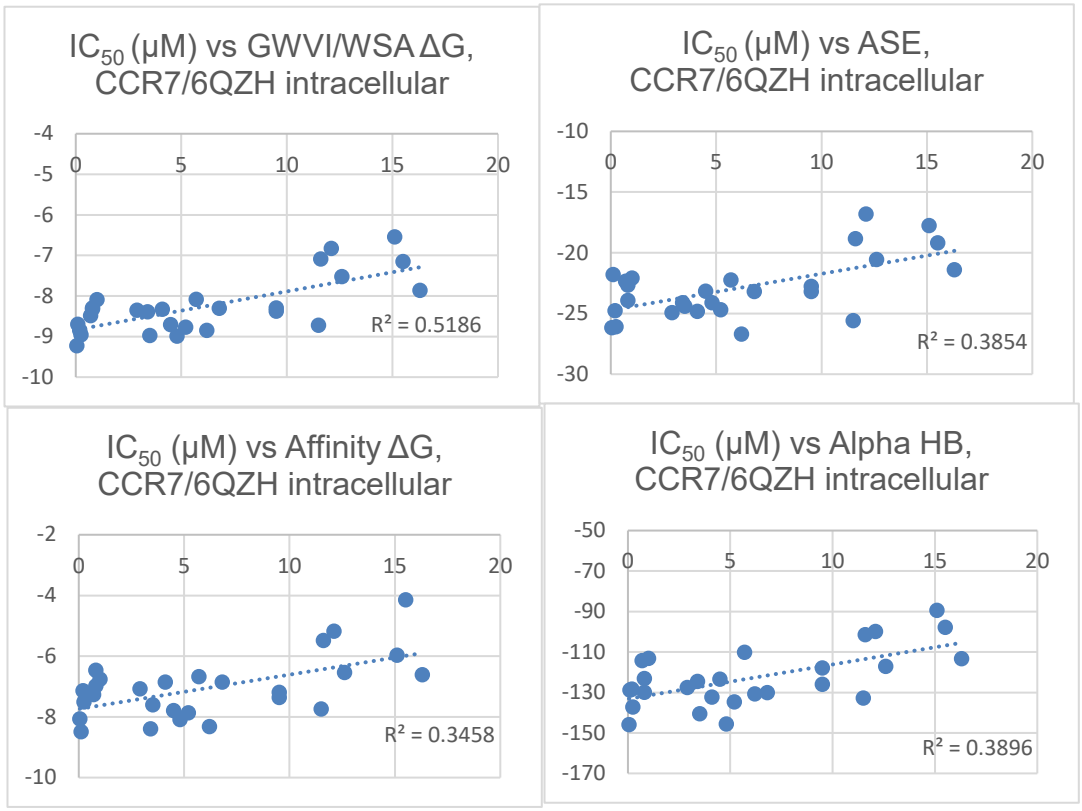


Figure 102: *Correlation plot between IC_{50} of 27 CCR7 antagonists and their top docking scores at the intracellular site for each scoring function available to MOE.*

Correlation plot (after energy minimisation): CCR7 homology model transmembrane site and CCR7/6QZH intracellular site

The energy minimisation function in MOE only calculates the new docking score in GWVI/WSA ΔG, therefore only the plots for GWVI/WSA ΔG are shown. The R^2 of the correlation plot is 0.1874 and 0.3569 for the transmembrane site and the intracellular respectively, which shows poor correlation, and they are lower than the R^2 value from the correlation plots generated from the docking scores prior energy minimisation. In the GWVI/WSA ΔG plot, four compounds with docking score above -10 kcal/mol stood out. All four compounds have higher molecular weight than other compounds, yet again highlighting the influence of molecular weight in the docking score.

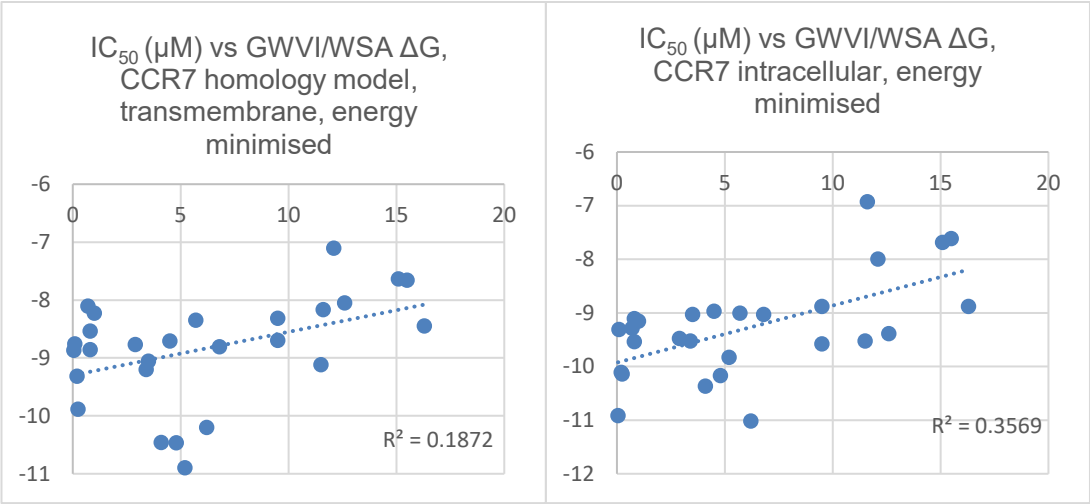


Figure 103*: Correlation plot between IC_{50} of 27 CCR7 antagonists and their post energy minimisation docking scores at the transmembrane and intracellular site.*

The plots show only a weak correlation. Therefore, we conclude that the correlation plot between the docking score and IC_{50} are not sufficient on its own to predict potency of a compound. Our plots have highlighted the problem of the molecular weight influencing the docking scores used in MOE, indicating the docking score should only be used cautiously in predicting potency.

6.0 Biological investigation of CXCR4 and dual CXCR4/CCR7 antagonists

6.1 Cytotoxicity assay

Before compounds can be assessed for activity, we needed to ensure that they were not cytotoxic at the doses used in the assay. This is to ensure that the response from the calcium flux assay can be wholly attributed to potency of the compound in blocking the receptor, rather than cell death. The cell viability at a range of concentrations was determined using an MTT assay by Dr Victoria Vinader.

In an MTT assay, cells are incubated with excess MTT reagent (3-(4,5-dimethylthiazolyl-2)-2,5-diphenyl tetrazolium bromide), a water-soluble yellow dye. The MTT penetrates by endocytosis into the cells. Cells which are viable will be metabolically active, therefore will reduce this yellow water-soluble tetrazolium salt by the mitochondrial enzyme succinate dehydrogenase (SDH) to produce formazan, a purple water-insoluble product (**Figure 104**). The amount of formazan generated is proportional to the number of viable cells present. This formazan will accumulate within the cell culture media as crystals as it is insoluble. Careful removal of cell culture medium will leave these crystals behind. The resulting purple formazan can be solubilized in DMSO and quantified using a colorimetric measurement by determining the absorbance of the resulting solution.

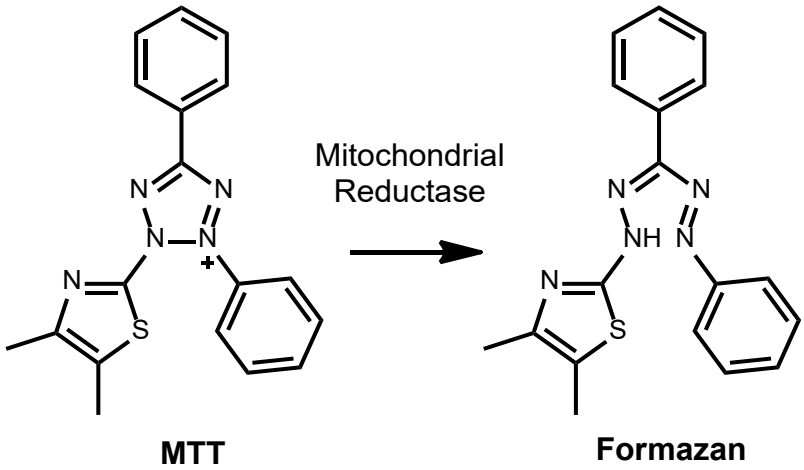


Figure 104: *The Conversion of MTT to formazan.*

The cytotoxicity of the novel antagonists synthesized was tested in OSC-19 and PANC-1 cell lines over a range of different concentrations (0.01, 0.1, 1.0, 10 and 100 μM) for 72 continuous hours of drug exposure. None of the molecules

showed cytotoxicity greater than 10% at 100 µM after 72 hours and since the calcium flux assay is performed over 3-4 hours. We concluded that this modest cytotoxicity will not affect the calcium flux assay.

6.2 Calcium flux assay

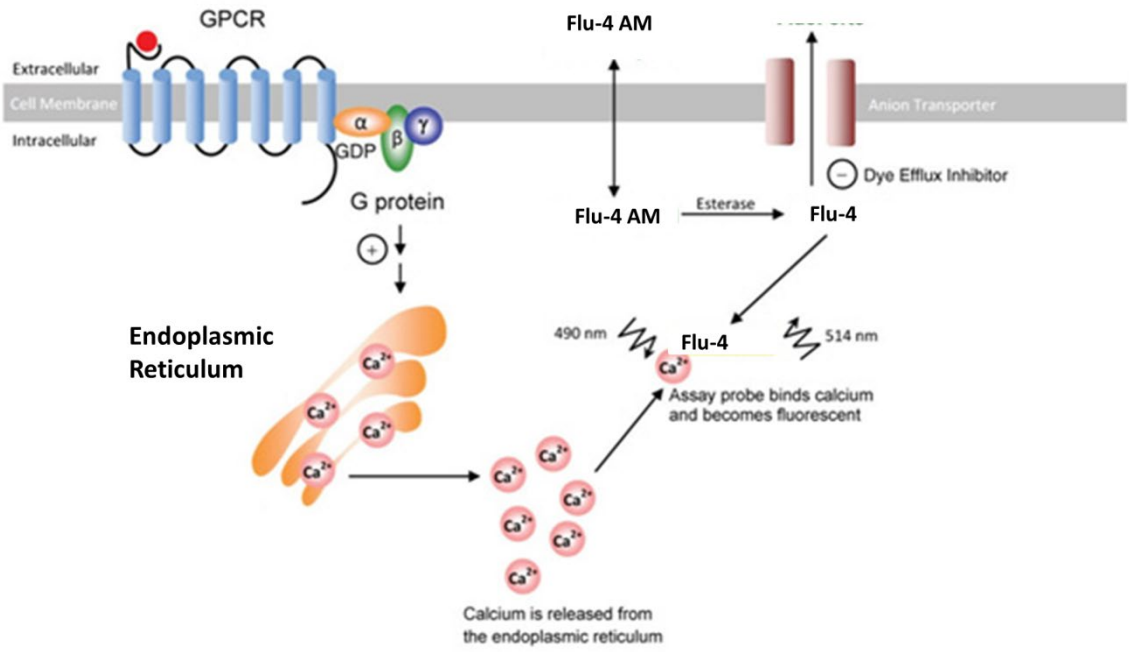


Figure 105: *The principles of the calcium flux assay.*

The calcium flux assay (**Figure 105**) was used to determine the effectiveness of the compounds as CCR7 antagonists. When a G-protein coupled receptor like CCR7 is activated, it undergoes a conformational change and activates the trimeric G protein, which in turns activates phospholipase C. Phospholipase C cleaves membrane bound PIP2 to produce IP3. IP3 diffuses across the cytoplasm and binds to the calcium channel of smooth endoplasmic reticulum. The calcium channels open and release calcium into the cytoplasm leading to further cellular responses. In the calcium flux assay the cells are treated with Flu-4-AM (AM stands for acetoxymethyl ester) which is not fluorescent in the presence of Ca^{2+} and cell-permeable. Once inside the cell, flu-4-AM hydrolyses to flu-4 which is fluorescent in the presence of Ca^{2+} and does not diffuse outside the cell. The formulation for the dye includes a dye-efflux inhibitor. It is important that flu-4 remain exclusively inside the cell as the cell media contains Ca^{2+} and leaching of the dye can result in background fluorescence and an inaccurate reading. Flu-4 binds to Ca^{2+} and this results in fluorescence which is read by an instrument. The

intensity of florescence changes depending on the concentration of the calcium in the cytoplasm. Therefore, the level of fluorescence is proportional to the activation of a GPCR. This can be used to measure the potency of an antagonist as well as an agonist. For example, in the assay already validated in our group, the addition of CCL21 to OSC-19 cells incubated in the presence of the calcium-responsive dye, will lead to an increase in fluorescence which will be used as a control. If the cells are pre-incubated with a CCR7 antagonist, activation of CCR7 by CCL21 is blocked and the fluorescent response will be less [119]. Calcium flux assays were performed by Dr Victoria Vinader and Dr Kamyar Afarinkia.

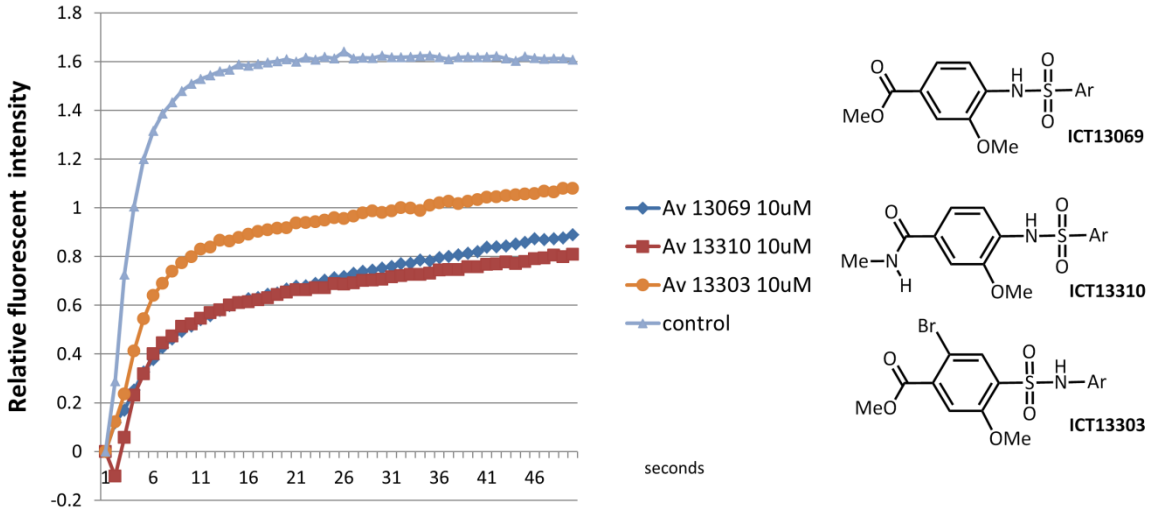
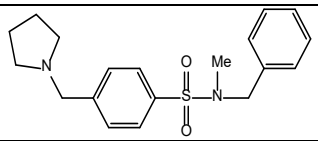
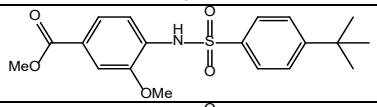
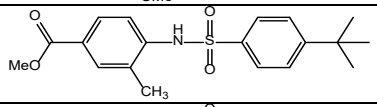
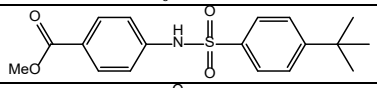
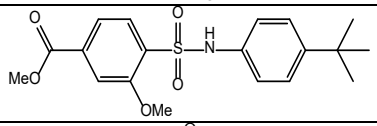
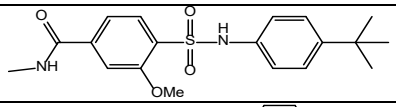
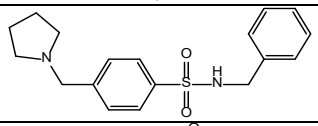
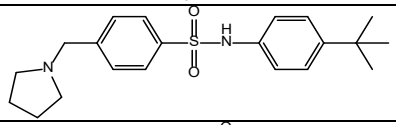
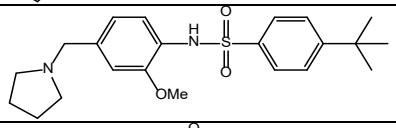
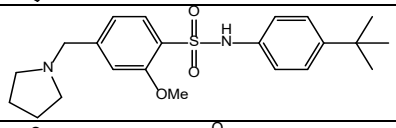
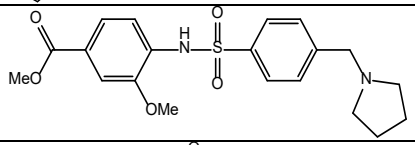
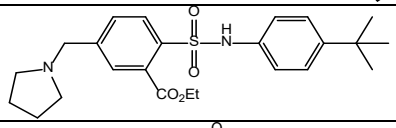
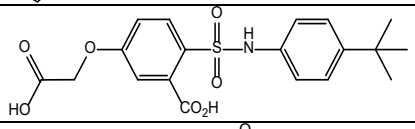
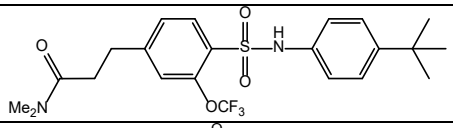
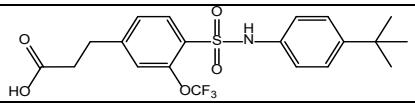


Figure 106: *Fluorescent response versus time in addition of different antagonists (OSC-19).*

In the example shown (**Figure 106**), compound **42** (ICT13310) and **64** (ICT13303) were tested against compound **30** (ICT13069) at a single concentration of 10 μM. Compound **42** (ICT13310) and **30** have very similar activity. This is likely due to the methyl ester of compound **30** and the methyl amide of compound **42** are both similar hydrogen bond acceptors. Unlike the methyl ester, the methyl amide can also act as a hydrogen bond donor, docking study shows it forms additional hydrogen bond to Trp114, but this property does not appear to make a significant difference in the activity of the compound. More importantly, since secondary amides are less prone to hydrolysis, ICT13310 is likely to be a more useful agent for *in vivo* investigation than ICT13069. Compound **64** is slightly less active compare to ICT13069. The presence of the bromine in compound **64** in the particular position appears to reduce the activity in this series of CCR7 antagonists.

Table 5: *Relative potencies of molecules using calcium flux assay (PANC-1)*

No	Structure	CCR7		CXCR4
		% inhibition at 1 μM	% inhibition at 100 μM	% inhibition at 100 μM
6		50	-97**	25
30		88	101	<10
39		72	71	<10
40		67	78	<10
66		57	-9*	<10
67		83**	94	<10
72		97	47	<10
73		88	97	<25
74		34	91	25
78		95	75	<25
79		31	58	<10
87		55	34	<10
93		79	24	<10
97		64	87	<10
98		86	70	<10
* n =1, ** n =2, all other values n =3 (error = ±10%)				

Compound **30** (ICT13069) along 14 other compounds were tested at two concentrations (1 μM and 100 μM) for their activity against CCR7 and CXCR4 in PANC-1, which is a human pancreatic cancer cell line. This cell line expresses both CCR7 [120] and CXCR4 [121], this was also shown by our group prior to this project. Any compounds with slight potency as an antagonist were expected to show a significant reduction in calcium mobilisation at 100 μM (high % inhibition). On the other hand, only potent antagonists were expected to show a significant reduction in calcium mobilisation at 1 μM (high % inhibition, but lower than that at 100 μM). Unexpectedly, compound **6**, **72**, **78**, **87** and **93** have lower inhibition when the concentration of the antagonists was increased. On the face of it, this implies the molecules are possibly inverse agonists, although further investigation is required to confirm this.

Compound **6**, the pyrrolidine CXCR4 antagonist is reported to show 100% inhibition at 10 nM in the presence of CXCL12 in a Matrigel invasion assay. The cell line MDA-MB-231 was used for that assay, which is a human breast cancer cells which shows CXCR4 expression [122]. We have previously found the levels of expression for CCR7 in MDA-MB-231 to be low, and levels of expression for CXCR4 only to be moderate compared with those in PANC-1. Compound **6** was expected to be potent in calcium flux assay, but it only showed 25% inhibition at 100 μM. Further investigation is required to verify these findings.

Effects of "reverse" sulfonamide

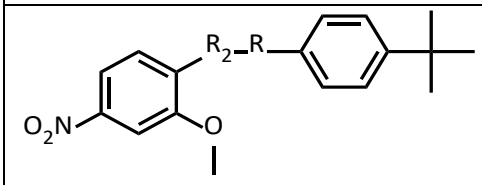
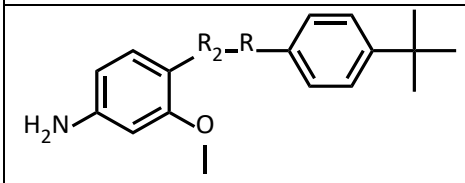
One of the main objectives of this project was to investigate whether the "reverse" sulfonamide would increase the potency of our CCR7 antagonists. Compound **30** (ICT13069) has 88±10% inhibition at 1 μM. Compound **66** is the "reverse" sulfonamide of ICT13069, and it is less potent compared to ICT13069 (57±10% inhibition at 1 μM), which was unexpected. Compound **67** is the "reverse" sulfonamide amide analogue of compound **30**, and showed similar potency (83% inhibition at 1 μM).

The extended chain compounds **93** and **98** are also "reverse" sulfonamide compounds. Even though they belong to a different chemical series due to the chain extension, they are also slightly less potent than ICT13069. From the data gathered, it is likely that the order of the sulfonamide has minor effect on the potency of the CCR7 antagonists similar to ICT13069 and the extended chain

compounds. This would be in agreement with the docking study at the CCR7 transmembrane and intracellular site, which did not show any notable difference between the docking poses of the “normal” and “reverse” sulfonamide compounds.

As for the pyrrolidine compounds, compound **78** is the reverse sulfonamide of compound **74**, and **78** is significantly more potent than **74** (95±10% and 35±10% inhibition at 10 nM). Docking studies at the CCR7 transmembrane site shows the pyrrolidine compounds have different ligand interactions and docking pose compare to the CCR7 antagonists similar to ICT13069 (**Figure 96C**). The “reverse” sulfonamide might have a different effect in pyrrolidine compounds and it might be needed for the potency of the pyrrolidine compounds. Further investigation is required to verify this.

An SAR study prior to this project has shown the two “reverse” sulfonamide compounds did indeed increased the potency of the compounds and they were validated by repeated testing (n = 3). The structures and IC_{50} in calcium flux assays of the compounds are shown in **Figure 107**. Docking study shows the nitro compounds have a similar docking pose with other CCR7 antagonists, the nitro is able to interact with Lys57 in addition to Ser117 (**Figure 107A/B**). The interaction with Lys57 might be related to “reverse” sulfonamide of the nitro compound being more potent. The aniline compounds have similar docking pose as the pyrrolidine compounds, the amine interacts with Asp309 (**Figure 107C/D**) similar to the pyrrolidine compounds, this concurred with the finding that “normal” sulfonamide is less potent in the pyrrolidine compounds. It appears the effect of the “reverse” sulfonamide depends on the substituents.

Structure	“Normal” sulfonamide: R: SO_2, R_2: NH	“Reverse” sulfonamide: R: NH, R_2: SO_2
	**ICT13070** IC_{50}: 5.7±1.3 µM (n = 3)	**ICT13124** IC_{50}: 0.10±0.3 µM (n = 3)
	**ICT13055** IC_{50}: 16.3±2.1 µM (n = 4)	**ICT13120** IC_{50}: 0.7±0.8 µM (n = 3)

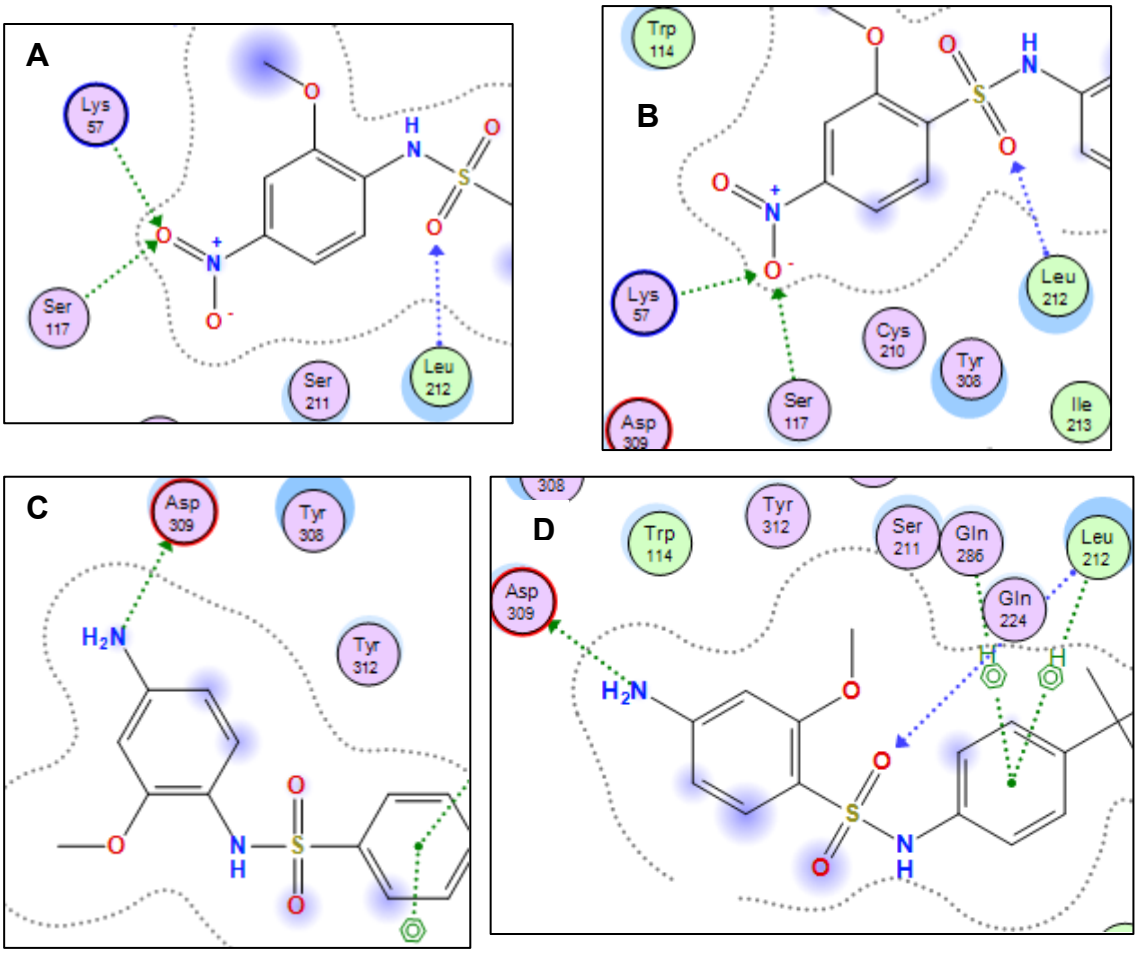
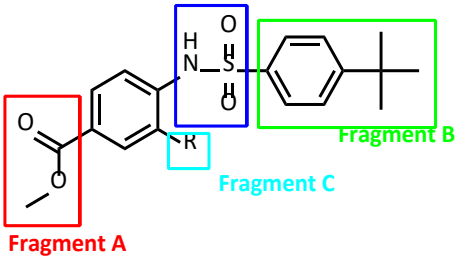


Figure 107: *Ligand interaction map generated by MOE. A) ICT13070, B) ICT13124, C) ICT13055, D) ICT13120.*

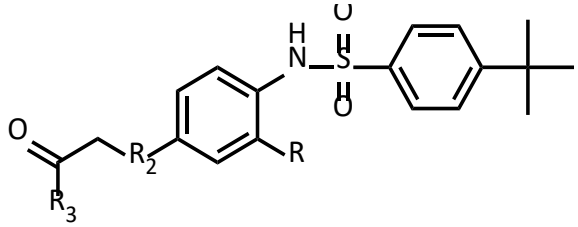
Role of Fragment C



No	R	CCR7		CXCR4
		% inhibition at 1 μM	% inhibition at 100 μM	% inhibition at 100 μM
30	OCH_3	88	101	<10
39	CH_3	72	71	<10
40	H	67	78	<10

Compound **39** and **40** were synthesised to gather more SAR data on fragment C of the CCR7 antagonists. Changing the methoxy (compound **30**) to methyl lowers the potency (compound **39**, 72±10% inhibition at 1 μM). The lack of any substituent at that position lowers the activity further (compound **40**, 67±10% inhibition at 1 μM). The docking study did not show the methoxy group forming any ligand interaction. Although the role of fragment C in the binding of the antagonists remains unclear, but the SAR established that for the CCR7 antagonists similar to ICT13069, the presence of a moiety at that position appears to improve the potency, however due to the ±10% error, further testing is required to verify this finding.

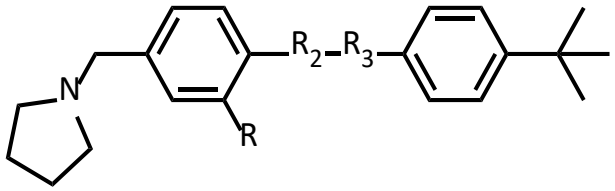
Effects of extended side chain



No	R	R_2	R_3	CCR7		CXCR4
				% inhibition at 1 μM	% inhibition at 100 μM	% inhibition at 100 μM
93	CO_2H	O	OH	79	24	<10
97	OCF_3	CH_2	$N(CH_3)_2$	67	87	<10
98	OCF_3	CH_2	OH	86	70	<10

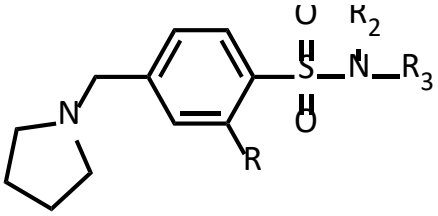
Docking study shows the extended chain compounds did not form any unique interaction at either the transmembrane nor intracellular site, and that carboxylic acid at the end of the chain (compound **93**, **98**) interacts with the same residues as ICT13069 and its analogues. The extended chain compounds are potent and shows similar potency as ICT13069. This finding validates the docking poses of the extended chain compounds, and therefore **93**, **97** and **98** can be considered as the same class of compounds as ICT13069 and related analogues. Compound **97** is the least potent extended chain compound (67±10% inhibition at 1 μM), this suggests the dimethyl amide lowers the potency of the extended chain compounds, and it is likely that the moiety and similar tertiary amide will also lower the potency of the CCR7 antagonists similar to ICT13069. Compound **93** (79±10% inhibition at 1 μM) has oxygen at the start of the extended chain, it has similar potency compare to compound **98** (86±10% inhibition at 1 μM), due to difference in moiety *para* to the sulfonamide, direct comparison cannot be made. Our docking study shows the oxygen did not have any interactions. It can be argued that the oxygen has insignificant effect on potency. This means in future SAR study on the extended chain, the synthesis route to make the ether side chain can be used instead of using Heck coupling to synthesise the alkyl side chain, which requires the costly palladium catalyst. Lastly, compound **97** and **98** show the trifluoromethyl ether in fragment C is tolerated in the CCR7 antagonists.

Pyrrolidine compounds



No	R	R_2	R_3	CCR7		CXCR4
				% inhibition at 1 μM	% inhibition at 100 μM	% inhibition at 100 μM
73	H	SO_2	NH	88	97	<25
74	OCH_3	NH	SO_2	34	91	<25
78	OCH_3	SO_2	NH	95	75	<25

Interestingly, compounds **73**, **74** and **78** show modest CXCR4 potency. Although the observed potency appears lower than what is reported by Mooring et al. [80], this observation is encouraging because it shows that a pyrrolidine methylene group does provide some CXCR4 inhibition into relatively potent CCR7 antagonists. Furthermore, compound **73** is nearly a reverse amide analogue of compound **74**. The presence of the methoxy in compound **74** makes a direct comparison not possible, but it is noteworthy that in this case, the reverse sulfonamide does appear to be more potent. Comparing compounds **73** and **78** may at first glance suggest lack of a methoxy makes the molecule more potent, however, due to the error inherent in this type of assay, that conclusion should be only cautiously considered until it is verified by further experiments.



No	R	R_2	R_3	CCR7		CXCR4
				% inhibition at 1 μM	% inhibition at 100 μM	% inhibition at 100 μM
6	H	CH_3	Benzyl	50	-97	25
72	H	H	Benzyl	97	47	<10
73	H	H	4-*tert*-buylbenzene	88	97	<25
78	OCH_3	H	4-*tert*-buylbenzene	95	75	<25
79	H	H	methyl 3-methoxybenzoate	31	58	<10
87	$CO_2C_2H_5$	H	4-*tert*-buylbenzene	55	34	<10

As described previously, the remaining analogues in this series all show good activity against CCR7 at low concentration. Compound **72** and **78** (97±10% and 95±10% inhibition at 1 μM) are more potent than ICT13069 at 1 μM. This indicates the pyrrolidine moiety can increase the potency of the CCR7 antagonists, due to the ±10% error, further testing is required to verify this finding. Our docking study shows the pyrrolidine compounds except for **6** can form additional ligand interactions with Gln227, and the CCR7 antagonists such as ICT13069 and **67** cannot. This might explain why some pyrrolidine compounds show higher potency. The CXCR4 antagonist compound **6** shows 50±10% inhibition at 1 μM. All other pyrrolidine compounds show higher potency than **6** with the exception of compound **79**.

Docking studies show that the two sulfonyl groups of the novel pyrrolidine compounds are able to form hydrogen bonds with both Gln227 and Leu212. Compound **6** was not able form a hydrogen bond with Gln227 likely to due to the methyl group on the sulfonamide. This would explain why several of the novel pyrrolidine compounds are more potent than **6** in CCR7. This suggests the docking poses are valid. Changing the methoxy group (compound **78**) to an ethyl ester significantly lowers the potency of the compound **87** (55±10% inhibition at 1 μM). Docking studies show the methoxy group and the ethyl ester did not form any ligand interaction, and the two have very similar docking poses. It is not clear why the ethyl ester reduced the potency of the compound. This finding needs to be verified.

The pyrrolidine compounds with 4-*tert*butylbenzene (**73**, **74**, **78**, **87**) have favourable docking poses at the transmembrane site from the docking study, but not at the intracellular site. Compound **73**, and **78** have over 80±10% inhibition at 1 μM. Compound **79** is the least active pyrrolidine compounds (31±10% inhibition at 1 μM). Compound **79** has favourable docking pose at the intracellular site of CCR7, but whether its docking poses are favourable at the transmembrane site is questionable. The above findings suggest the pyrrolidine compounds bind at the orthosteric transmembrane site of CCR7 and not the intracellular site. The lack of potency of compound **79** could be explained by its docking poses shown in chapter 5.6.5 (**Figure 97**), in which the pyrrolidine and sulfonyl have the same interaction as other compounds, but the benzene ring containing hydrophilic groups occupy the lipophilic pocket, making the docking pose less favourable.

The pharmacology of these compounds remains slightly ambiguous since at 100 μM, some compounds show weaker antagonism or even some agonism. Of course, 100 μM is a physiologically improbable concentration so in terms of identifying a potent dual CXCR4/CCR7, compounds **73**, **74** and **78**, remain very important lead compounds.

7.0 Summary and conclusion

This project set out to synthesise dual CCR7 and CXCR4 small molecule antagonists for the treatment of metastasis, and to generate more SAR data for the ongoing CCR7 antagonist project. Previous SAR studies potency of four biaryl sulfonamide compounds (**Figure 107**) show the order of the sulfonamide affects the potency of the compound. The more potent analogues have the sulfonamide in the same order, and they were labelled as the “reverse” sulfonamide series. This project aimed to investigate the effects of the “reverse” sulfonamide.

A series of novel sulfonamide CCR7 antagonists were synthesised, many of which belong to the “reverse” sulfonamide series. The “reverse” sulfonamide series of CCR7 antagonists were expected to increase the potency of the sulfonamide compounds, but further study and docking study suggest the effect of the “reverse” sulfonamide is related to the substituent *para* to the sulfonamide and the related binding pose. The findings of this project show the “reverse” sulfonamide increased the potency of compounds with nitro, amine and *N*-methyl pyrrolidine at that position, but it has little effect in the potency of compounds with methyl ester, methyl amide, and propanoic acid.

ICT13310 (compound **42**) and its “reverse” sulfonamide analogue compound **67**, are secondary methyl amide analogues of ICT13069 and they are nearly as potent. Since these compounds are less susceptible to hydrolysis, it represents a better CCR7 antagonist for biological investigation.

A series of novel potential dual antagonists of CCR7 and CXCR4 were synthesised. They are hybrid compounds combining the features of the sulfonamide CCR7 antagonist of this project, and the pyrrolidine of the CXCR4 antagonists of Mooring et al. [80]. A calcium flux assay has shown some of these compounds are able to inhibit GPCR activation in the presence of CCL19/21, indicating they are potent CCR7 antagonists. The most potent of the hybrid

compounds (compound **78**) shown 95±10% inhibition at 1 μM in calcium flux assay, which is more potent than ICT13069 with 88±10% inhibition at 1 μM. However, the same assay with CXCL12 shows only modest activity. This includes **6,** which has shown nanomolar potency in inhibiting CXCL12 induced migration in Matrigel invasion assay. Nevertheless, compound **72**, **73** and **78** appear to be very potent CCR7 antagonists, and modest CXCR4 antagonists, they merit further investigation as potential leads for CXCR4/CCR7 dual antagonists.

MOE 2015.10 was used to carry out a molecular docking study of the compounds of this project at the transmembrane orthosteric and intracellular allosteric site of CCR7, and at the transmembrane site of CXCR4. Cmp2105 and Vercirnon being intracellular allosteric antagonists raised the question of whether the antagonists of this project also bind at the intracellular site. The calcium flux assay results of the pyrrolidine compounds can only be rationalised if the pyrrolidine compounds bind to the transmembrane site, suggesting the pyrrolidine sulfonamide compounds are indeed orthosteric site antagonists.

Although Vercirnon has the 4-*tert*butylbenezene sulfonamide moiety similar to the compounds of this project. The docking study of this project has shown the orientation of Arg88 in CCR7 prevented Vercirnon from forming the same ligand interactions at CCR9 in CCR7, which could explain why Vercirnon is not active in CCR7. Comparing the binding mode of Cmp2105 and Vercirnon, with the docking pose of Navarixin [63], the pyrrolidine compounds and the CCR7 antagonists synthesised in this project, the CCR7 antagonists form the least ligand interactions in the intracellular site. Therefore, it is most likely these CCR7 antagonists bind to the transmembrane site, not the intracellular site. In further support of this, the docking study at the CCR7 transmembrane site was able to rationalise the findings that the “reverse” sulfonamide compounds only increased the potency of some of the compounds.

6QZH (CCR7) and 5LWE (CCR9) both have an antagonist binding to the intracellular site. Inspection of structures showed the receptors have adopted a very similar conformation that makes the orthosteric transmembrane site significantly less accessible compared to crystal structures (*CCR5/4MBS, CCR2 /5T1A and CXCR4/3ODU)* that have an antagonist bound to the transmembrane site. A CCR7 homology model based on 3ODU (CXCR4) generated by SWISS

MODEL was necessary for the docking study at the transmembrane orthosteric site of CCR7.

IC_{50} values of 34 CCR7 antagonists synthesised prior to this project were plotted against their docking score, with the aim of identifying correlations which might be useful in predicting the potency of compounds. Correlation plots of IC_{50} and docking score show poor correlation. The molecular weight was found to have significant effect on the docking score, allowing potent compounds with low molecular weight to have a low docking score. This could be due to factors such as methodology of the docking study, and the limitation of the software and docking score. MOE was able to replicate the binding pose of Vercirnon in CCR9, and it was able to generate a docking pose similar to binding pose of Cmp2105 in CCR7. The lack of correlation between the docking score and the IC_{50} does not invalidate the docking poses generated in MOE.

Future work

In terms of gathering of SAR data, a few compounds of interest still need to be tested, and some of the tested compounds require further testing to verify the findings. Compound **6** has displayed nanomolar potency in a Matrigel invasion assay, but shown poor activity in the calcium flux assay of this project. More investigation is required to confirm the findings. AMD3100 is a well-known CXCR4 antagonist, it could be used to provide a benchmark when testing these pyrrolidine compounds in the same calcium flux assay.

The discovery of Cmp2105 as an intracellular antagonist of CCR7 is a recent finding, so research of Cmp2105 in CCR7 is still limited. More research into Cmp2105 and its intracellular binding of CCR7 would be beneficial to the project. Cmp2105 might provide different avenue in developing a dual CCR7 and CXCR4 antagonists that targets the intracellular allosteric site of CCR7 and CXCR4.

This project established a framework for molecular docking at the CCR7 transmembrane/intracellular site, and at the transmembrane site of CXCR4. Some of the existing calcium flux assay results were rationalised by docking poses generated. This could be used for the rational design of CCR7 antagonists and dual CCR7 and CXCR4 antagonists in the future. The intracellular site of 3ODU/CXCR4 was unsuitable for docking studies due to the conformation of its

C-terminus. It might be beneficial to try to generate a homology model for the purpose of exploring the intracellular binding site of CXCR4, which will be important in the scenario of developing a dual CCR7 and CXCR4 intracellular allosteric antagonists.

Finally, use of compound **78** as a lead to prepare potent, dual CCR7/CXCR4 merits further investigation.

8.0 Experimental

Material and Product Information

Commercially available reagents were purchased from Sigma-Aldrich, Alfa Aesar Apollo Scientific and Fisher Scientific, and were used as received without additional purification. Dry solvents were also obtained from Sigma-Aldrich, VWR and Merck. All other solvents were of reagent grade. Petroleum ether refers to the fraction of petroleum spirit boiling in the range of 60 to 80 °C. Where stated, mixtures of solvents are referred to as percentage volume to volume (v/v) ratios.

All the compounds made were purified by flash column chromatography using silica gel high-purity Merck Grade 9385, pore size 60 Å silica gel (40-63 μm). The analytical thin layer chromatography used was based on classical Merck silica gel 60 F254 aluminium plates. Visualisation of the reaction components was accomplished by illumination under short wavelength (254 nm) ultraviolet light or using basic potassium permanganate ($KMnO_4$) stain. Proton and carbon nuclear magnetic resonance (^1H and ^{13}C NMR) spectra were recorded using Bruker AMX400 spectrometer operating at 400 MHz and 101 MHz respectively. Deuterated chloroform ($CDCl_3$) was used as the NMR solvent unless otherwise stated. Chemical shifts are reported in parts per million (δ, ppm). Coupling constants (J) are expressed in Hertz (Hz). The splitting patters in NMR spectra are reported with the following abbreviations; singlet (s), doublet (d), triplet (t), quartet (q), quintet (quint.), multiplet (m) and broad (br). ^{13}C NMR chemical shifts are reported relative to the signal of the solvent. Where necessary, phosphorus decoupled proton NMR, correlation spectroscopy (COSY), nuclear Overhauser enhanced spectroscopy (NOESY), heteronuclear multiple quantum correlation spectroscopy (HMQC) for $^{13}C/^1H$ nuclei, and distortionless enhancement by polarization transfer (DEPT) technique were employed to confirm the assignment of NMR spectra. Routine infrared (IR) spectra were recorded on a Perkin-Elmer FTIR spectrometer. Spectra are reported in wavenumbers (cm^{-1}). Routine mass spectra were carried out by Andrew Healey and were run on a Micromass Quattro Ultima spectrometer in the electron impact (EI), chemical ionisation (CI) or positive (+ve) electrospray mode (ESI) as stated.

Glassware was dried approximately >2 hours prior to the experiment in a dry oven at 140 °C. All moisture sensitive reactions were carried out under an inert atmosphere using dry argon and nitrogen. Heating was completed using an oil bath, containing silicon oil, placed on top of a hot plate stirrer. Cooling baths were prepared as described below for a range of various temperatures:

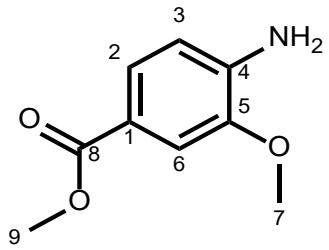
0 °C	Ice and water
-10 °C	Ice and acetone
-45 °C	Acetonitrile and dry ice
-78 °C	Acetone and dry ice

Experimental procedures

General synthesis for compounds 31-35

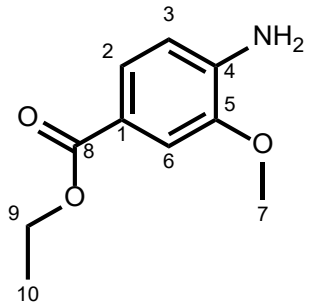
The corresponding substituted benzoic acid was dissolved in methanol (ethanol for compound **32**), and acetyl chloride (2 eqv.) was added. The reaction was stirred and heated at 65 °C for 24 hours. The volatile materials were removed at reduced pressure using a rotary evaporator to give the ester product as colourless solids which were used for next step without further purification.

Methyl 4-amino-3-methoxybenzoate (31)



Colourless solid (99% yield); m.p.: 127.6-128.2 °C (Lit., 127-129 °C)[123]; ^{1}H NMR (Acetone-d6) δ 7.33 (1H, dd, J =1.8, 8.2 Hz, H-2), 7.26 (1H, d, J = 1.8 Hz, H-6), 6.59 (1H, d, *J* = 8.2 Hz, H-3), 5.05 (2H, s, NH_2), 3.75 (3H, s, H-9), 3.68 (3H, s, H-7); ^{13}C NMR (Acetone-d6) 165.3 (C-8), 144.5 (C-5), 141.6 (C-4), 122.7 (C-2), 116.5 (C-1), 111.0 (C-3), 109.5 (C-6), 53.7 (C-9), 49.5 (C-7); LRMS (ESI) m/z: 182.8 $(M+H)^{+}$. HRMS (ESI) m/z calcd. for $C_9H_{12}NO_3$ $[M+H]^{+}$: 182.1913, found: 182.1901.

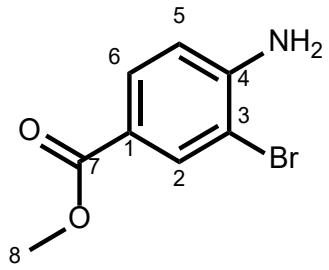
Methyl 4-amino-3-methoxybenzoate (32)



Colourless solid (99% yield); m.p.: 84.0 - 84.7 °C (Lit., 84-85 °C)[124]; ^{1}H NMR δ 7.47 (1H, dd, *J* = 8.4, 1.7 Hz, H- 2), 7.37 (1H, d, *J* = 1.7 Hz, H-6), 6.57 (1H, d, *J* = 8.4 Hz, H-3), 4.24 (2H, q, *J* = 7.1 Hz, H-9), 4.14 (2H, s, NH_2), 3.80 (3H, s, H-7), 1.28 (3H, t, *J* = 7.1 Hz, H-10); ^{13}C NMR 166.9 (C-8), 146.1 (C-5), 141.1 (C-4),

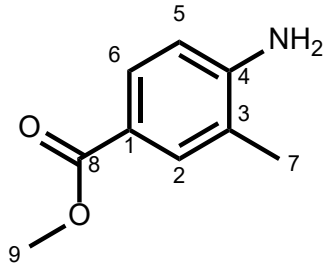
124.0 (C-2), 119.8 (C-1), 113.0 (C-3), 111.1 (C-6), 60.4 (C-9), 55.5 (C-7), 14.5 (C-10); LRMS (ESI) m/z: 196 $(M+H)^+$; HRMS (ESI) m/z calcd. for $C_{10}H_{14}NO_3$ $[M+H]^+$: 196.2112, found: 196.2105.

Methyl 4-amino-3-bromobenzoate (33)



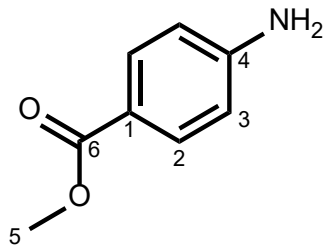
Colourless solid (99% yield); (lit., 105 – 107 °C) [125]; ^{1}H NMR: 7.48 (1H, dd, J= 1.7, 8.0 Hz, H-6), 7.42 (1H, d, J=1.7 Hz, H-2), 6.64 (1H, d, J= 8.0 Hz, H-5), 3.89 (3H, s, H-8); ^{13}C NMR: 166.2 (C-7), 146.1 (C-3), 140.6 (C-4), 123.7 (C-6), 121.6 (C-1), 113.0 (C-5), 111.1 (C-2), 55.5 (C-8); LRMS (ESI) m/z: 229.9 $(M+H)^+$.

Methyl 4-amino-3-methylbenzoate (34)



Colourless solid (99% yield). ^{1}H NMR: 7.29 (1H, dd, J= 1.7, 8.0 Hz, H-6), 7.06 (1H, d, J=1.7 Hz, H-2), 6.47 (1H, d, J = 8.0 Hz, H-5), 3.72 (3H, s, H-8), 3.29 (3H, s, H-7); ^{13}C NMR: 165.7 (C-8) 148.3 (C-4), 130.7 (C-6), 125.8 (C-2), 123.1 (C-5), 121.6 (C-3), 119.5 (C-1), 55.5 (C-9), 17.4 (C-7); LRMS (ESI) m/z: 166.8 (M+H).

Methyl 4-aminobenzoate (35)

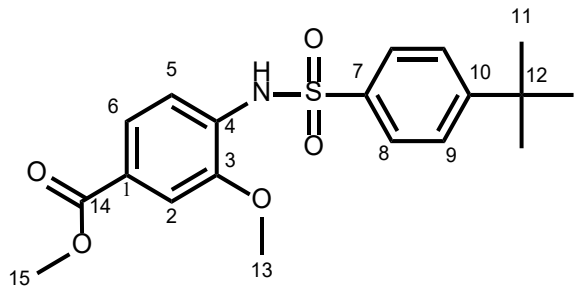


Colourless solid (99% yield). ^{1}H NMR ($CDCl_3$): δ 7.66 (2H, d, J = 8.4 Hz, H-2), 6.49 (2H, d, J = 8.4 Hz, H-3), 6.25 (2H, s, NH_2), 3.86 (3H, s, H-5); ^{13}C NMR ($CDCl_3$): δ 165.5 (C-6), 151.3 (C-4), 130.7 (C-2), 123.6 (C-1), 117.8 (C-3), 52.1 (C-5); LRMS (ESI) m/z: 151.8 (M+H).

General synthesis for compound 30, 36-40

The aniline compound (**31**-**35**) was dissolved in dichloromethane (10 mL), the appropriate sulfonyl chloride (1.05 eqv.) was added followed by pyridine (3 eqv.). The reaction was stirred at room temperature for 16 hours. The reaction mixture was diluted by addition of 1M hydrochloric acid (10 mL) and extracted with ethyl acetate (3 x 30 mL). The combined organic extracts were dried anhydrous magnesium sulfate, filtered and concentrated under reduced pressure using a rotary evaporator. The crude product was dry loaded to silica gel and purified by flash column chromatography using ethyl acetate and petroleum ether (2:8 ratio) as eluent.

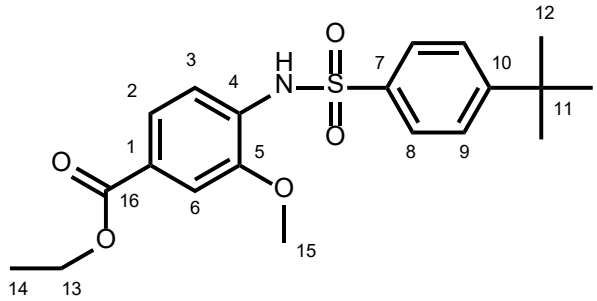
Methyl 4-(4-(*tert*-butyl)phenylsulfonamido)-3-methoxybenzoate (30, ICT13069)



Pale yellow solid (65% yield); IR: 3364, 3212, 2964, 1706, 1672, 1609, 1592, 1508, 1468, 1443 cm^{-1}; ^{1}H NMR ($CDCl_3$): δ 7.68 (2H, d, J = 8.7 Hz, H-8), 7.54

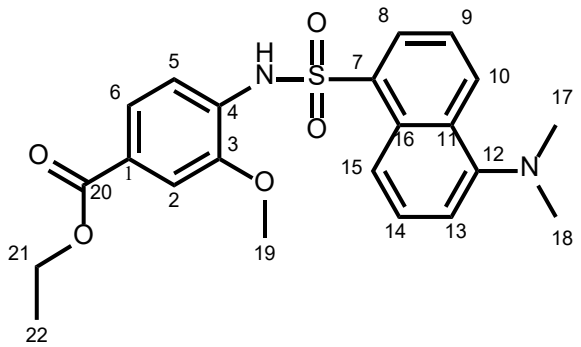
(1H, dd, J = 8.6, 1.5 Hz, H-6), 7.48 (1H, d, J = 8.6 Hz, H-5), 7.38 (2H, d, J = 8.7 Hz, H-9), 7.36 (1H, d, J = 1.5 Hz, H-2), 7.25 (1H, s, NH), 3.80 (3H, s, H-15), 3.72 (3H, s, H-13), 1.22 (9H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 166.5 (C-14), 157.1 (C-10), 147.9 (C-3), 136.0 (C- 7), 130.7 (C-4), 127.0 (C-8), 126.0 (C-9), 125.8 (C-1), 123.1 (C-6), 117.8 (C-5), 111.3 (C-2), 55.9 (C-13), 52.2 (C-15), 35.1 (C-12), 31.0 (C-11); LRMS (ESI) m/z: 378.1 $(M+H)^+$; HRMS (ESI) m/z calcd. for $C_{19}H_{24}NO_5S$ $[M+H]^+$: 378.1371, found: 378.1370.

Ethyl 4-(4-(*tert*-butyl)phenylsulfonamido)-3-methoxybenzoate (36)



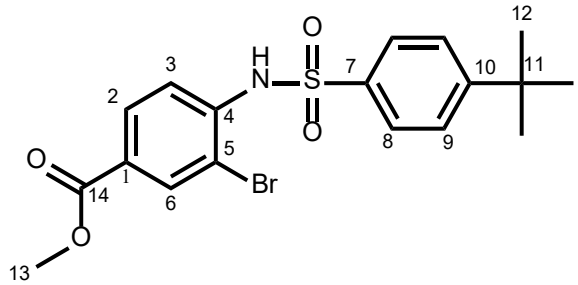
Pale yellow solid. (59% yield); m.p.: 129.3-129.8 °C; IR: 3296, 3239, 2959, 1702, 1433, 1338, 1288, 1271, 1107, 900 cm^{-1}; ^{1}H NMR ($CDCl_3$): δ 7.75 (2H, d, J= 8.6 Hz, H-8), 7.61 (1H, dd, J= 1.0 Hz, 8.1 Hz, H-2), 7.55 (1H, d, J = 8.1 Hz, H-3), 7.44 (2H, d, J = 8.62 Hz, H-9), 7.43 (1H, d, J = 1.0 Hz, H-6), 7.33 (1H, s, NH), 4.33 (2H, q, J = 7.1 Hz, H-13), 3.79 (3H, s, H-15), 1.36 (3H, t, J = 7.1 Hz, H-14) 1.28 (9H, s, H-12); ^{13}C NMR ($CDCl_3$): δ 166.1 (C-16), 157.1 (C-10), 148.0 (C-5), 136.0 (C-7), 130.6 (C-4), 127.0 (C-8), 126.3 (C-1), 126.1 (C-9), 123.1 (C-2), 117.8 (C-3), 111.3 (C-6), 61.1 (C-13), 56.0 (C-15), 35.2 (C-11), 31.0 (H-12), 14.4 (H-14); LRMS (ESI) m/z: 392.1 $(M+H)^+$; HRMS (ESI) m/z calcd. for $C_{20}H_{26}O_5NS$ $[M+H]^+$; 392.1526, found: 392.1522.

Methyl 4-(5-(dimethylamino)naphthalene-1-sulfonamido)-3-methoxybenzoate (37)



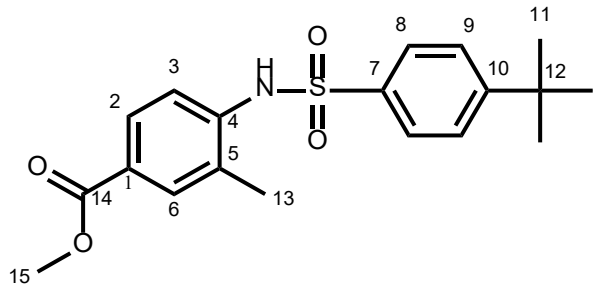
Yellow solid (63% yield); IR: 3248, 2827, 2776, 1698, 1606, 1575, 1469, 1337, 1019, 900 cm^{-1}; 1H NMR ($CDCl_3$): δ 8.42 (1H, d, J = 8.8 Hz, H-8), 8.25 (1H, d, J = 8.8 Hz, H-15), 8.17 (1H, d, J = 7.3 Hz, H-10), 7.52 (1H, dd, J = 7.3, 8.8 Hz, H-14), 7.48 (1H, s, NH), 7.44 (1H, dd, J = 1.6, 8.4 Hz, H-6), 7.41 (1H, d, J = 8.4, H-5), 7.36 (1H, dd, J = 7.3, 8.8 Hz, H-9), 7.25 (1H, d, J = 1.6 Hz, H-2), 7.10 (1H, d, J = 7.3 Hz, H-13), 4.21 (2H, q, J = 7.1 Hz, H-20), 3.56 (3H, s , H-19), 2.78 (6H, s, H-17, H-18), 1.26 (3H, t, J = 7.1 Hz, H-21); ^{13}C NMR ($CDCl_3$): δ 166.0 (C-20), 152.0 (C-12), 147.9 (C-7), 133.8 (C-3), 131.1 (C-8), 130.6 (C-4), 130.3 (C-10), 129.8 (C-16), 129.5 (C-1), 128.4 (C-14), 126.1 (C-5), 122.9 (C-6), 118.5 (C-15), 117.7 (C-9), 115.2 (C-13), 112.3 (C-11), 111.2 (C-2), 61.0 (C-21), 55.8 (C-19), 45.4 (C-17/18), 14.3 (C-22); LRMS ESI (-ve) m/z: 427.03 $(M-H)^-$.

Methyl 3-bromo-4-(4-(*tert*-butyl)phenylsulfonamido)benzoate (38)



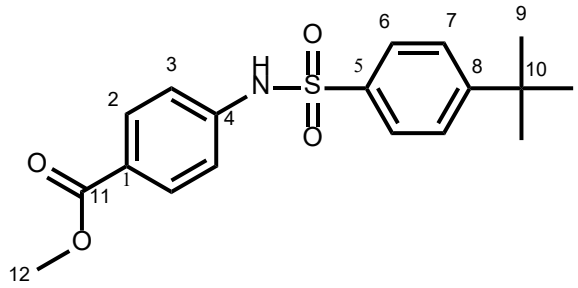
Colourless solid (52% yield); IR: 3261, 3247, 2952, 1714, 1593, 1585, 1462, 1244, 1067, 1023 cm^{-1}; 1H NMR ($CDCl_3$): δ 7.91 (1H, dd, J = 1.2, 8.4 Hz, H-2), 7.86 (1H, d, J = 1.2 Hz, H-6), 7.81 (2H, d, J = 8.4 Hz, H-8), 7.56 (1H, d. J = 8.4 Hz, H-3), 7.48 (2H, d, J = 8.6 Hz, H-9), 6.51 (3H, s, H-13), 1.43 (9H, s, H-12); ^{13}C NMR ($CDCl_3$): 165.2 (C-14), 157.7 (C-10), 138.8 (C-7), 135.6 (C-4), 134.1(C-6), 130.0 (C-2), 127.2 (C-8), 127.1 (C-5), 126.3 (C-9), 119.4 (C-3), 113.6 (C-1), 52.1 (C-13), 36.7 (C-11), 32.2 (C-12); LRMS (ESI) m/z: 426.1 $(M+H)^+$.

Methyl 4-(4-(*tert*-butyl)phenylsulfonamido)-3-methylbenzoate (39)



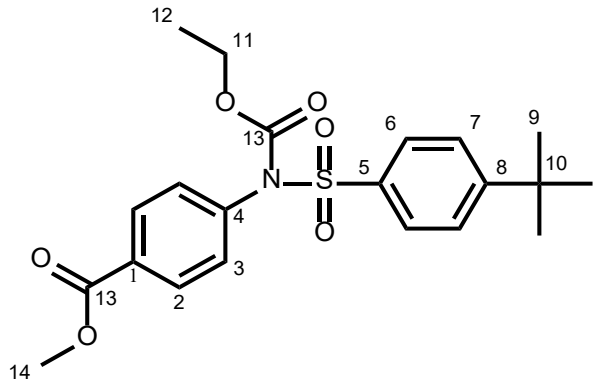
Colourless solid (64% yield); IR: 3298, 3270, 2953, 1710, 1605, 1587, 1434, 1269, 1161, 1087 cm^{-1}; ^{1}H NMR ($CDCl_3$): δ 7.82 (1H, dd, J = 1.5, 8.4 Hz, H-2), 7.78 (1H, d, J = 1.5 Hz, H-6), 7.71 (2H, d, J = 8.6 Hz, H-8), 7.48 (1H, d, J = 8.4 Hz, H-3), 7.45 (2H, d, J = 8.6 Hz, H-9), 6.51 (1H, s, NH), 3.87 (3H, s, H-15), 3.14 (3H, s, H-13), 1.31 (9H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 164.9 (C-14), 157.4 (C-10), 139.2 (C-4), 136.2 (C-7), 132.1 (C-6), 128.7 (C-2) , 128.1 (C-5), 127.0 (C-8), 126.5 (C-1), 126.2 (C-9), 120.5 (C-3), 52.1 (C-15), 31.0 (C-11), 17.4 (C-13); LRMS (ESI) m/z: 362.1 $(M+H)^{+}$.

Methyl 4-(4-(*tert*-butyl)phenylsulfonamido)benzoate (40)



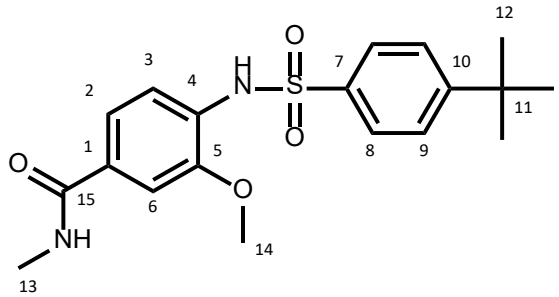
Colourless solid (61% yield); IR: 3220, 2951, 1698, 1607, 1434, 1340, 1295, 1164, 1113, 1086 cm^{-1}; ^{1}H NMR ($CDCl_3$): δ 7.92 (2H, d, J = 8.5 Hz, H-2), 7.77 (2H, d, J = 8.5 Hz, H-6), 7.46 (2H, d, J= 8.5 Hz, H-7), 7.15 (2H, d, J = 8.5 Hz, H-3), 3.87 (3H, s, H-12), 1.29 (9H, s, H-9); ^{13}C NMR ($CDCl_3$): δ 166.4 (C-11), 157.4 (C-8), 141.1 (C-4), 135.8 (C-5), 131.1 (C-2), 127.1 (C-6), 126.3 (C-7), 126.1 (C-1), 118.8 (C-3), 52.1 (C-12), 35.2 (C-10), 31.0 (C-9) ; LRMS (ESI) m/z: 348.1 $(M+H)^{+}$.

Methyl 4-((4-(*tert*-butyl)-*N*-(ethoxycarbonyl)phenyl)sulfonamido)benzoate (41)



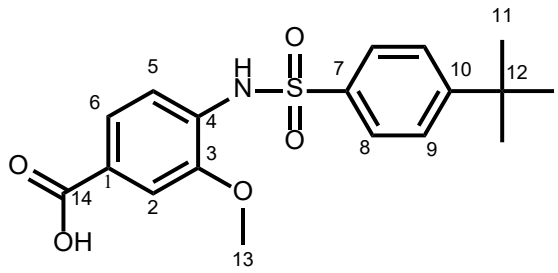
Methyl 4-(4-(*tert*-butyl)phenylsulfonamido)benzoate (50 mg, 0.14 mmol) was dissolved in tetrahydrofuran (5 mL). 60% Sodium hydride (16 mg, 0.40 mmol) was added. The reaction mixture was stirred at room temperature for 5 hours. The tetrahydrofuran was evaporated under reduced pressure. Ethyl acetate (30 ml) was added to the crude solid to form a suspension. The suspension was washed with 1M hydrochloric acid in small portions (3 x 5 mL). The organic layer was concentrated under reduced pressure to give the title compound as a colourless solid (54mg, 93%) without further purification. ^{1}H NMR ($CDCl_3$): δ 8.12 (2H, d, J = 8.7 Hz, H-2), 7.92 (2H, d, J = 8.9 Hz, H-6), 7.57 (2H, d, J = 8.9 Hz, H-7), 7.34 (2H, d, J = 8.7 Hz, H-3), 4.13 (2H, q, J = 7.14 Hz, H-11), 3.95 (3H, s, H-14), 1.37 (9H, s, H-9), 1.13 (3H, t, J = 7.14 Hz, H-12); ^{13}C NMR ($CDCl_3$): δ 166.2 (C-13), 158.1 (C-8), 140.1 (C-4), 135.9 (C-5), 130.9 (C-1), 130.6 (C-2), 129.8 (C-3), 128.9 (C-6), 126.0 (C-7), 63.7 (C-11), 52.4 (C-14), 35.4 (C-10), 31.1 (C-9), 14.0 (C-12); LRMS (ESI) m/z: 420.3 $(M+H)^+$.

4-(4-*tert*-butylbenzenesulfonamido)-3-methoxy-*N*-methylbenzamide (42)



Thionyl chloride (3 mL, 65 mmol) was added to 4-(4-*tert*-butylbenzenesulfonamido)-3-methoxybenzoic acid (100 mg, 0.275 mmol) and the reaction mixture was refluxed with stirring for 24 hours. Thionyl chloride was removed under reduced pressure as much as possible. Methylamine in ethanol (0.1 mL, 0.825mmol, ~3 equivalents) was added to the crude acid chloride at -10 °C, followed by triethylamine (0.08 mL, 0.575 mmol, ~2 eqv). The reaction mixture was left in cold storage below 0 °C for 2 days. The reaction mixture then was concentrated under reduced pressure. Ethyl acetate (40 mL) was added to the crude which was then washed with water (2 x 10mL). The organic layer was dried with anhydrous magnesium sulfate and concentrated under reduced pressure. The crude product was purified by column chromatography using petroleum ether and ethyl acetate (65:35) as the eluent to give 4-(4-*tert*-butylbenzenesulfonamido)-3-methoxy-*N*-methylbenzamide as a layer of colourless transparent solid (13 mg, 13 %). IR: 3427, 3477, 3251, 1605, 1498, 1556, 1454, 1493, 1352, 1301 cm^{-1}; ^{1}H NMR ($CDCl_3$): δ 7.70 (2H, d, J = 8.6 Hz, H-8), 7.53 (1H, d, J = 8.5 Hz, H-3), 7.42 (2H, d, J = 8.6 Hz, H-9), 7.32 (1H, d, J = 1.5 Hz, H-6), 7.20 (1H, s, SO_2NH), 7.16 (1H, dd, J = 1.5, 8.5 Hz, H-2), 6.19 (1H, d, J = 4.4 Hz, CONH), 3.74 (3H, s, H-14), 2.97 (3H, d, J = 4.4 Hz, H-13), 1.28 (9H, s, H-12). ^{13}C NMR ($CDCl_3$): δ 166.4 (C-15), 157.1 (C-10), 147.9 (C-5), 135.9 (C-7), 130.72 (C-4), 127.0 (C-8), 126.1 (C-9), 125.9 (C-1), 123.2 (C-2), 117.8 (C-3), 111.3 (C-6), 55.2 (C-14), 52.2(C-13), 36.2 (C-11), 31.0 (C-12); LRMS (ESI) m/z: 377.1 $(M+H)^+$.

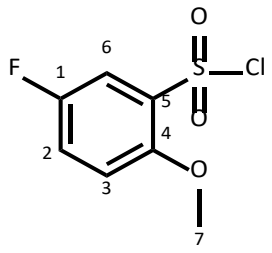
4-(4-*tert*-butylbenzenesulfonamido)-3-methoxybenzoic acid (43)



A solution of sodium hydroxide (10 mL, 3 M) was added to compound **75** (0.40 g, 1.06 mmol). Tetrahydrofuran (1 mL) was added to fully dissolve the starting material. The reaction mixture was heated at reflux with stirring for 1h. Hydrochloric acid (10 mL, 2 M) was added and reaction mixture was extracted

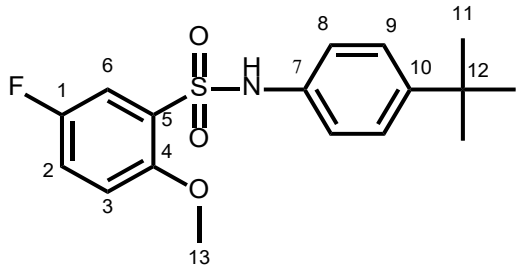
with ethyl acetate (3 x 50 mL). The combined organic extracts were dried over anhydrous magnesium sulfate, filtered and solvent evaporated under reduced pressure to give the title compound (0.28 g, 72 %) as light brown powder; m.p.: 175.1-175.6 °C; IR: 2963, 1692, 1607, 1593, 1528, 1465, 1445, 1399, 1347, 1325 cm^{-1}; ^{1}H NMR (Acetone-d6): δ 7.68 (2H, d, J = 7.4 Hz, H-8,), 7.48 (1H, dd, J = 8.8, 1.4 Hz, H-6), 7.46 (1H, d, J = 8.8 Hz, H-5), 7.43 (2H, d, J = 7.4 Hz, H-9), 7.35 (1H, d, J = 1.4 Hz, H-2), 3.64 (3H, s, H-13), 1.17 (9H, s, C-11); ^{13}C NMR (Acetone-d6): δ 170.5 (C-14), 164.3 (C-10), 152.5 (C-3), 145.0 (C-7), 137.8 (C-4), 131.6 (C-8), 127.8 (C-9), 126.8 (C-1), 123.6 (C-6), 120.2 (C-5), 112.7 (C-2), 56.3 (C-13), 55.4 (C-12), 35.7 (C-11); LRMS (ESI) m/z: 364.1 (M+H).; HRMS (ESI) m/z calcd. for $C_{18}H_{25}O_5N_2S$ $[M+NH_4]^+$: 381.1479, found: 381.1478.

***N*-(4-(*tert*-butyl)phenyl)-5-fluoro-2-methoxybenzenesulfonamide (45)**



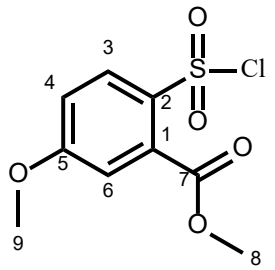
A solution of 4-fluoroanisole (0.2 mL, 1.77 mmol) dissolved in dichloromethane (5 mL) was cooled to 0 °C and kept under nitrogen. Chlorosulfonic acid (0.6 mL, 9mmol, ~5 eqv) was added under nitrogen. The reaction mixture was heated at 60 °C 16 hours. Afterward the reaction mixture was added to crushed ice and extracted with ethyl acetate (3 x 30 mL). The collective organic extracts were dried with anhydrous magnesium sulfate and concentrated using rotary evaporator. The crude product was purified using flash column chromatography using petroleum either and ethyl acetate (6:4) as the eluent to give 5-fluoro-2-methoxybenzene-1-sulfonyl chloride as a colourless solid (67 mg, 17 %). ^{1}H NMR ($CDCl_3$): δ 7.56 (1H, dd, J_H= 3.1, J_F=7.5 Hz, H-6), 7.35 (1H, ddd, J_H=3.1, 9.2, J_F=7.5 Hz, H-2), 7.06 (1H, dd, J_H=9.2, J_F=3.9 Hz, H-3), 3.96 (3H, s, H-7); ^{13}C NMR ($CDCl_3$): δ 161.8 (C-1), 155.4 (C-4), 131.5 (C-5), 123.3&123.1 (C-2), 119.8&118.0 (C-3), 112.6&112.4 (C-6), 56.9 (C-7); LRMS (ESI) m/z: 225.9 $(M+H)^+$.

***N*-(4-(*tert*-butyl)phenyl)-5-fluoro-2-methoxybenzenesulfonamide (46)**



5-Fluoro-2-methoxybenzene-1-sulfonyl chloride (67 mg, 0.3 mmol) was dissolved in dichloromethane. 4-*tert*butylaniline (0.1 mL, 0.6 mmol, ~2 eqv.) was added. The reaction mixture was stirred at room temperature for 16 hours. The reaction mixture was then poured into crushed ice and extracted with ethyl acetate (3 x 30 mL). then dried with anhydrous magnesium sulfate. The combined organic extracts were concentrated in a rotary evaporator. The crude product was purified using flash column chromatography using petroleum either and ethyl acetate (7:3) as the eluent to give *N*-(4-(*tert*-butyl)phenyl)-5-fluoro-2-methoxybenzenesulfonamide as a colourless solid (59 mg, 59 %) IR: 3298, 3264, 1711, 1605, 1588, 1509, 1289, 1270, 1161, 1018 cm^{-1}; ^{1}H NMR ($CDCl_3$): δ 7.55 (1H, dd, J_H=3.0, J_F = 7.5 Hz, H-6), 7.39 (1H, s, NH), 7.21 (2H, d, J = 8.6 Hz, H-8), 7.16 (1H, ddd, J_H= 3.0, 9.1, J_F= 7.6 Hz, H-2), 7.01 (2H, d, J = 8.6 Hz, H-9), 6.95 (1H, dd, J_H= 9.1, J_F= 3.7 Hz, H-3), 4.02 (3H, s, H-13), 1.22 (9H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 157.0 (C-1), 154.6 (C-5), 152.5 (C-4), 148.5 (C-10), 133.6 (C-7), 126.7 (C-8), 121.6 (C-9), 121.3&121.1 (C-6), 118.0&117.7 (C-3), 113.4&113.3 (C-2). 56.9 (C-13), 34.3 (C-12), 31.3 (C-11); LRMS (ESI) m/z: 338.1 $(M+H)^+$; HRMS (ESI) m/z calcd. for $C_{17}H_{21}O_3NFS$ $[M+H]^+$: 338.1215, found: 338.1219

Methyl 2-(chlorosulfonyl)-5-methoxybenzoate (49b)

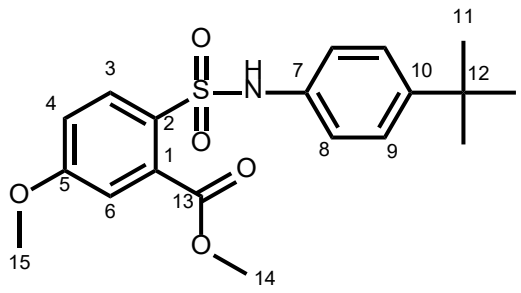


To a solution of 3-hydroxybenzoic acid (3 g, 21.7 mmol) dissolved in acetone (10 mL), potassium carbonate (15 g, 108.5 mmol, 5 eqv.) was added, followed by iodomethane (2.7 mL, 43.4 mmol, 2 eqv). The reaction mixture was heated at

reflux temperature with stirring for 24 hours. Acetone was removed at reduced pressure using a rotatory evaporator. Water (100 mL) was added the crude solid then extracted with ethyl acetate (3 x 30 mL). Collective organic layers were dried with anhydrous magnesium sulfate then concentrated at reduced pressure using a rotatory evaporator to give crude methyl 3-methoxybenzoate as a colourless solid (3.205 g, 89 %). ^{1}H NMR ($CDCl_3$): δ 7.61 (1H, dd, J = 2.3, 7.3 Hz, H-4), 7.54 (1H, dd, J = 2.3, 7.3 Hz, H-2). 7.31 (1H, t, J = 7.3 Hz, H-3), 7.07 (1H, dd, J = 2.3 Hz, H-6), 3.89 (3H, s, CO_2CH_3), 3.82 (3H, s, OCH_3).

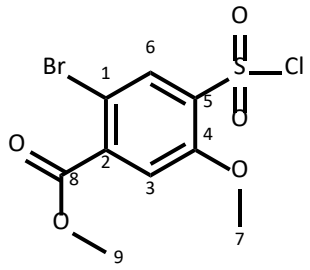
Without purification, methyl 3-methoxybenzoate (2 g, 12.048 mmol) was dissolved in dichloromethane (20 mL), cooled to 0 °C and kept under nitrogen. Chlorosulfonic acid (4 mL, 60 mmol, ~5 eqv) was added under nitrogen. The reaction mixture was left to stir at room temperature for 16 hours. Afterward the reaction mixture was added to crushed ice and extracted with ethyl acetate (3 x 30 mL). The collective organic extracts were dried with anhydrous magnesium sulfate and concentrated using rotary evaporator. The crude product was purified using flash column chromatography using petroleum either and ethyl acetate (6:4) as the eluent to methyl 2-(chlorosulfonyl)-5-methoxybenzoate as a brown oil (1.03 g, 32 %); ^{1}H NMR ($CDCl_3$): δ 8.08 (1H, d, J = 9.1 Hz, H-3), 7.16 (1H, d, J = 2.6 Hz, H-6), 7.10 (1H, dd, J = 2.6, 9.1 Hz, H-4), 3.98 (3H, s, H-8), 3.94, (3H, s, H-9); ^{13}C NMR ($CDCl_3$): δ 166.1 (C-7), 164.5 (C-5), 134.8 (C-1), 133.4 (C-2), 131.8 (C-3), 115.9 (C-4), 115.4 (C-6), 56.3 (C-9), 53.4 (C-8) LRMS (ESI) m/z: 265.7 $(M+H)^{+}$.

Methyl 2-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)-5-methoxybenzoate (50)



Methyl 2-(chlorosulfonyl)-5-methoxybenzoate (103 mg, 0.39 mmol) was dissolved in dichloromethane (5 mL), 4-*tert*butylaniline (0.13 mL, 0.8 mmol, ~2 eqv) was added. The reaction mixture was stirred at room temperature for 16 hours. The reaction mixture then was washed with brine (3 x 10 mL). The collective organic phase was dried with anhydrous magnesium sulfate then concentrated in a rotary evaporator. The crude product was purified using flash column chromatography using petroleum either and ethyl acetate (8:2) as the eluent to give the title compound as a pale-yellow solid (92 mg, 63 %). ^{1}H NMR ($CDCl_3$): δ 7.79 (1H, s, NH), 7.74 (1H, d, J = 8.8, H-3), 7.28 (1H, d, J = 2.7, H-6), 7.22 (2H, d, J = 8.7 Hz, H-8), 7.06 (2H, d, J = 8.7 Hz, H-9), 6.89 (1H, dd, J = 2.7, 8.7 Hz, H-4), 4.03 (3H, s, H-14), 3.85 (3H, s, H-15), 1.24 (9H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 168.0 (C-13), 162.4 (C-5), 148.7 (C-10), 134.1 (C-7), 132.8 (C-3), 132.3 (C-1), 132.8 (C-2), 129.9 (C-8), 126.0 (C-9), 117.0 (C-6), 115.2 (C-4), 55.8 (C-15), 55.6 (C-14), 34.4 (C-12) , 31.3 (C-11) ; LRMS (ESI) m/z: 378 $(M+H)^+$.

Methyl 2-bromo-4-(chlorosulfonyl)-5-methoxybenzoate (63)

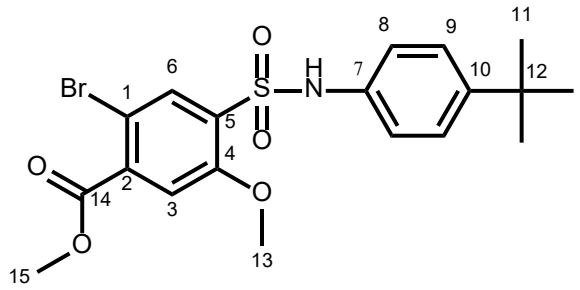


To a solution of 2-bromo-5-methoxybenzoic acid (500 mg, 2.16 mmol) dissolved in methanol (30 mL), acetyl chloride (0.3 mL, 4.2 mmol, ~2 eqv.) was added. The reaction mixture was heated under reflux with stirring for 24 hours. The solvent was removed using a rotary evaporator to give 2-bromo-5-methoxybenzoate as a colourless solid (530mg, 100%); ^{1}H NMR ($CDCl_3$): δ 7.55 (1H, d, J = 9.1 Hz), 7.33 (1H, d, J = 3.1 Hz), 6.91 (1H, dd, J = 3.1, 9.1 Hz), 3.95 (3H, s, $CO_2C\underline{H}_3$), 3.84 (3H, s, $OC\underline{H}_3$).

Without further purification, 2-bromo-5-methoxybenzoate (530 mg, 2.16 mmol) was added to a sealed tube and dissolved in chloroform (5 mL). The solution was cooled to 0-5 °C using an ice bath, chlorosulfonic acid (0.7 mL, 10.5 mmol ~5 eqv.) was added under nitrogen. The reaction mixture was heated at 70 °C with

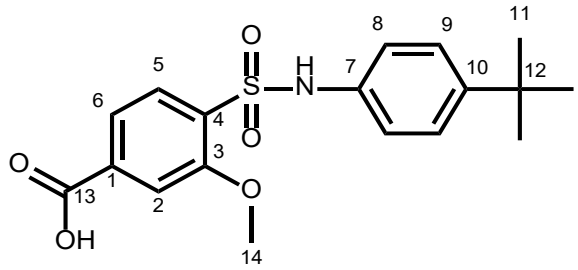
stirring for 2 days. The reaction mixture was poured into crushed ice and extracted with ethyl acetate (3 x 30 mL), then dried with anhydrous magnesium sulfate. The combined organic extracts were concentrated in rotary evaporator. The crude product was purified using flash column chromatography using petroleum either and ethyl acetate (7:3) as the eluent to give methyl 2-bromo-4-(chlorosulfonyl)-5-methoxybenzoate as a pale-yellow solid (118 mg, 16 %); ^{1}H NMR ($CDCl_3$): δ 7.93 (1H, s, H-6), 7.54 (1H, s, H-3), 4.01 (3H, s, H-9), 3.96 (3H, s, H-7); ^{13}C NMR ($CDCl_3$): δ 166.4 (C-8), 159.7 (C-4), 140.3 (C-2), 136.4 (C-5), 129.0 (C-6), 117.5 (C-3), 116.8 (C-1). 56.1 (C-7), 51.8 (C-9); LRMS (ESI) m/z: 344 $(M+H)^+$.

Methyl 2-bromo-4-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)-5-methoxybenzoate (64)



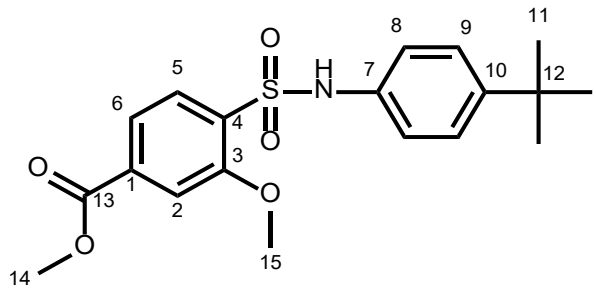
Methyl 2-bromo-4-(chlorosulfonyl)-5-methoxybenzoate (23 mg, 0.07 mmol) was dissolved in dichloromethane (5 mL), 4 *tert*butylaniline (0.03 mL, 0.19 mmol, ~2 eqv) was added. The reaction mixture was stirred at room temperature for 16 hours. The reaction mixture then was washed with brine (3 x 10 mL). The collective organic phase was dried with anhydrous magnesium sulfate then concentrated in a rotary evaporator. The crude product was purified using flash column chromatography using petroleum either and ethyl acetate (7:3) as the eluent to give the title compound as a pale-yellow solid (18 mg, 56 %). IR: 3257, 2958, 1720, 1459, 1324, 1371, 1157, 1128, 1079, 899 cm^{-1}; 1H NMR ($CDCl_3$): δ 8.06 (1H, s, H-6), 7.37 (1H, s, H-3), 7.24 (2H, d, J = 8.6 Hz, H-8), 6.97 (2H, d, J = 8.6 Hz, H-9), 6.94 (1H, s, NH), 4.06 (3H, s, H-15), 3.95 (3H, s, H-13), 1.25 (9H, s, H-11) ; ^{13}C NMR ($CDCl_3$): δ 165.6 (C-14), 154.8 (C-4), 149.1 (C-10), 137.5 (C-2), 135.8 (C-7), 133.0 (C-6), 130.6 (C-5), 126.4 (C-9), 121.7 (C-8), 114.7 (C-1), 111.9 (C-3), 57.0 (C-13), 53.0 (C-15), 34.4 (C-11), 31.3 (C-12); LRMS (ESI) m/z: 457.1 $(M+H)^+$; HRMS (ESI) m/z calcd. for $C_{19}H_{26}O_2NBr$ $[M+H]^+$: 457.0449, found: 457.0449.

4-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)-3-methoxybenzoic acid (65)



To a solution of methyl 2-bromo-4-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)-5-methoxybenzoate (546 mg, 1.2 mmol) in anhydrous butan1-ol (10 mL), palladium (II) acetate (4 mg, 0.018 mmol), triphenylphosphate (59 mg, 0.18 mmol), potassium carbonate (332 mg, 2.4 mmol) were added. The reaction mixture was stirred at 100 °C under nitrogen for 16 hours. Butanol was removed under reduced pressure. Ethyl acetate (30 mL) was added to the crude then wash with saturated sodium bicarbonate (3 x 5 mL). The organic layer was dried with anhydrous magnesium sulfate then concentrated at reduced pressure. The crude was purified using flash column chromatography (1:99, methanol and dichloromethane) to give the titled compound as a colourless solid. (278 mg, 64 %); ^{1}H NMR ($CDCl_3$): δ 7.89 (1H, d, J = 8.1 Hz, H-5), 7.66 (1H, d, J = 1.4 Hz, H-2), 7.62 (1H, dd, J = 8.1, 1.4 Hz, H-6), 7.19 (2H, d, J = 8.9 Hz, H-9), 7.00 (2H, d, J = 8.9 Hz, H-8), 4.09 (3H, s, H-14), 1.21 (9H, s, H-12).; ^{13}C NMR ($CDCl_3$): δ 165.3 (C-13), 156.2 (C-3), 148.5 (C-10), 136.2 (C-1), 133.6 (C-7), 131.0 (C-5), 130.6 (C-4), 126.2 (C-8), 121.5 (C-6), 121.3 (C-9), 112.9 (C-2), 56.6 (C-14), 34.3 (C-12), 31.2 (C-11); LRMS (ESI) m/z: 364.1 $(M+H)^+$.

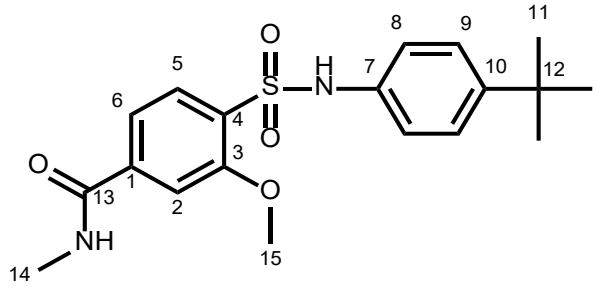
Methyl 4-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)-3-methoxybenzoate (66)



4-(*N*-(4-(*tert*-Butyl)phenyl)sulfamoyl)-3-methoxybenzoic acid (50 mg, 0.14 mmol) was dissolved in methanol (20 mL), acetyl chloride (0.03 mL, 0.42 mmol) was

slowly added on an ice bath. The reaction was reflux heated with stirring for 24 hours. The volatile was removed under reduced pressure. The crude was purified using flash column chromatography (2:9, ethyl acetate and petroleum either) to give the titled compound as a colourless solid. (49 mg, 92%); ^{1}H NMR ($CDCl_3$): δ 7.88 (1H, d, J = 7.9 Hz, H-5), 7.67 (1H, s, H-2), 7.63 (1H, d, J = 7.9 Hz, H-6), 7.20 (2H, d, J = 8.1 Hz, H-9), 7.03 (1H, s, NH), 6.98 (2H, d, J = 8.1 Hz, H-8), 4.11 (3H, s, H-15), 3.92 (3H, s, H-14), 1.22 (9H, s, H-11).; ^{13}C NMR ($CDCl_3$): δ 165.7 (C-13), 156.1 (C-3), 148.8 (C-10), 135.8 (C-1), 133.4 (C-7), 131.1 (C-5), 130.7 (C-4), 126.2 (C-8), 121.8 (C-6), 121.7 (C-9), 112.0 (C-2), 56.86 (C-15), 52.8 (C-14), 34.4 (C-12), 31.2 (C-11); LRMS (ESI) m/z: 378.1 (M+H)+; HRMS (ESI) m/z calcd. for $C_{19}H_{22}O_5NS$ [M-H]-:376.1221, found: 376.1221.

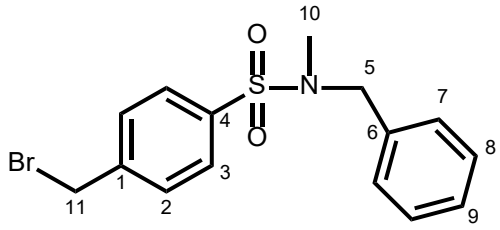
4-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)-3-methoxy-*N*-methylbenzamide (67)



Thionyl chloride (3 mL, 65 mmol) was added to 4-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)-3-methoxybenzoic acid (100 mg, 0.275 mmol) reaction mixture was refluxed with stirring for 24 hours. Thionyl chloride was removed under reduced pressure as much as possible. Methylamine in ethanol was added to the crude acid chloride at -10 °C, followed by triethylamine (0.08 mL, 0.575 mmol, ~2 eqv). The reaction mixture was left in cold storage below 0 °C for 2 days. The reaction mixture then was concentrated under reduced pressure. Ethyl acetate (40mL) was added to the crude, then the mixture was washed with water (2 x 10 mL). The organic layer was dried with anhydrous magnesium sulfate and concentrated under reduced pressure. The crude product was purified by column chromatography using petroleum ether and ethyl acetate (65:35) as the eluent to give 4-(4-*tert*-butylbenzenesulfonamido)-3-methoxy-*N*-methylbenzamide as a layer of colourless transparent solid (52 mg, 51%). ^{1}H NMR ($CDCl_3$): δ 7.72 (1H, d, J = 8.02 Hz, H-5), 7.49 (1H, d, J = 1.3 Hz, H-2), 7.18 (2H, d, J = 8.7 Hz, H-8), 7.49 (1H, dd, J = 7.14 Hz, H-6), 7.10 (1H, s, NH), 6.95 (1H, d, J = 8.73 Hz ,H-9),

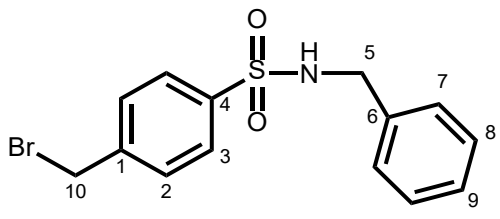
4.03 (3H, s, H-15), 2.97 (3H, d, J = 4.83 Hz, H-14), 1.20 (9H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 166.8 (C-13), 156.3 (C-3), 148.9 (C-10), 140.9 (C-1), 133.4 (C-7), 131.0 (C-5), 128.8 (C-4), 126.2 (C-8), 121.8 (C-9), 118.0 (C-6), 111.7 (C-2), 56.7 (C-15), 34.4 (C-12), 31.2 (C-11), 27.0 (C-14); LRMS (ESI) m/z: 377.1 $(M+H)^+$.

_N_-benzyl-4-(bromomethyl)-_N_-methylbenzenesulfonamide (69)



A solution of 4-(bromomethyl)benzene-1-sulfonyl chloride (500 mg, 1.86 mmol) in DCM (10mL) was added *N*-methylbenzylamine (0.5 mg, 3.88 mmol, ~2 eqv.). The reaction mixture was left to stir at room temperature for 16 hours. The reaction mixture then was washed with brine. The combined organic layers were dried over anhydrous magnesium sulfate and concentrated under reduced pressure. The crude product was purified to give the titled compound as a colourless solid (480 mg, 72%); ^{1}H NMR ($CDCl_3$): δ 7.84 (2H, d, J = 8.0 Hz, H-3), 7.59 (1H, d, J = 8.0 Hz, H-2), 7.34–7.31 (5H, m, H-7, H-8, H-9 overlap), 4.65 (2H, s, H-11), 4.16 (2H, s, H-5), 2.61 (3H, s, H-10); ^{13}C NMR ($CDCl_3$): δ 142.7 (C-1), 137.6 (C-4), 135.6 (C-6), 130.0 (C-2), 129.4 (C-3), 128.9 (C-8), 128.5 (C-7), 128.2 (C-9), 54.3 (C-5), 34.5 (C-10), 31.8 (C-11); LRMS (ESI) m/z: 339.1 $(M+H)^+$.

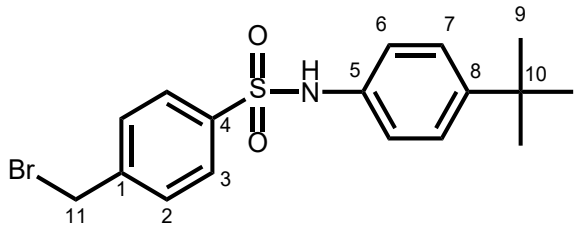
_N_-benzyl-4-(bromomethyl)benzenesulfonamide (70)



A solution of 4-(bromomethyl)benzene-1-sulfonyl chloride (1 g, 3.75 mmol) in DCM (35 mL) was added benzylamine (0.82 mL, 7.52 mmol). The reaction

mixture was left to stir at room temperature for 16 hours. The reaction mixture was washed with brine (3 x 5 mL). The combined organic layers were dried over anhydrous magnesium sulfate and concentrated under reduced pressure. The crude product was purified to give the titled compound as a colourless solid (867 mg, 68%). ^{1}H NMR ($CDCl_3$): δ 7.84 (2H, d, J = 8.0 Hz, H-3), 7.59 (2H, d, J = 8.0 Hz, H-2), 7.34–7.31 (5H, m, H-7, H-8, H-9 overlap), 4.65 (2H, s, H-11), 4.16 (2H, s, H-5), 2.61 (2H, s, H-10); 13C NMR ($CDCl_3$): δ 141.6 (C-1), 135.7 (C-4), 133.5 (C-6), 130.0 (C-2), 129.4 (C-3), 128.9 (C-8), 128.5 (C-7), 128.2 (C-9), 47.6 (C-5), 33.4 (C-10); LRMS (ESI) m/z: 339.1 $(M+H)^+$.

4-(bromomethyl)-*N*-(4-(*tert*-butyl)phenyl)benzenesulfonamide (71)

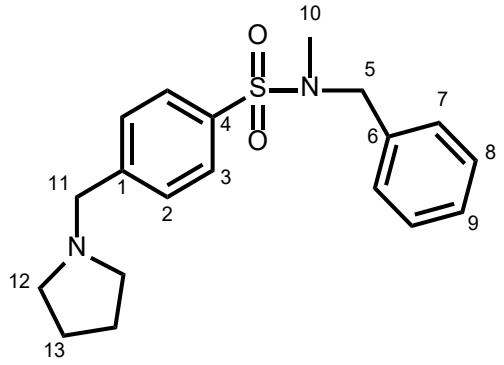


A solution of 4-(bromomethyl)benzene-1-sulfonyl chloride (1 g, 3.75 mmol) in DCM (10 mL) was added 4-*tert*butylaniline (1.2 mL, 7.55 mmol). The reaction mixture was left to stir at room temperature for 16 hours. The reaction mixture was diluted with more DCM (20 mL) and washed with brine (3 x 5 mL). The combined organic layers were dried over anhydrous magnesium sulfate and concentrated under reduced pressure. The crude was purified to give the titled compound as a colourless solid. (1.014 mg, 71%). ^{1}H NMR ($CDCl_3$): δ 7.61 (2H, t, J = 8.7 Hz, H-3), 7.24 (2H, dd, J = 6.3, 2.1 Hz, H-2), 7.05 (2H, d, J = 8.6 Hz, H-7), 6.84 (2H, d, J = 8.6 Hz, H-6), 4.36 (2H, s, H-11), 1.06 (9H, s, H-9).; ^{13}C NMR ($CDCl_3$): δ 148.6 (C-8), 142.8 (C-4), 139.2 (C-1), 133.6 (C-5), 129.6 (C-2), 127.8 (C-3), 126.3 (C-7), 121.6 (C-6), 44.9 (C-11), 34.4 (C-10), 31.6 (C-11), 31.3 (C-9); LRMS (ESI) m/z: 383.1 $(M+H)^+$.

General procedure for compound 6, 72-74)

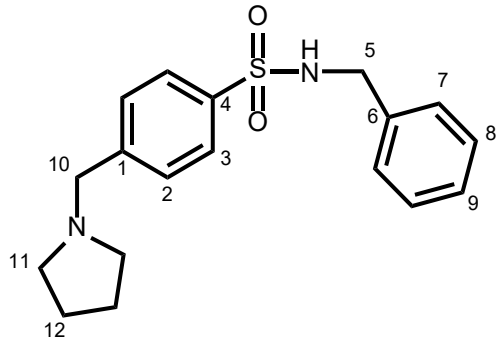
A solution of **69**/**70**/**71** in acetonitrile was added potassium carbonate (3 eqv.) and pyrrolidine (2 eqv.). The reaction mixture was left to stir at room temperature for 16 hours. Ethyl acetate was added to the reaction mixture then it was washed with saturated sodium bicarbonate. The organic phase was dried with anhydrous magnesium sulfate, then concentrated at reduced pressure. The crude product was purified using flash column chromatography (1:99 methanol: DCM).

***N*-benzyl-*N*-methyl-4-(pyrrolidin-1-ylmethyl)benzenesulfonamide (6)**



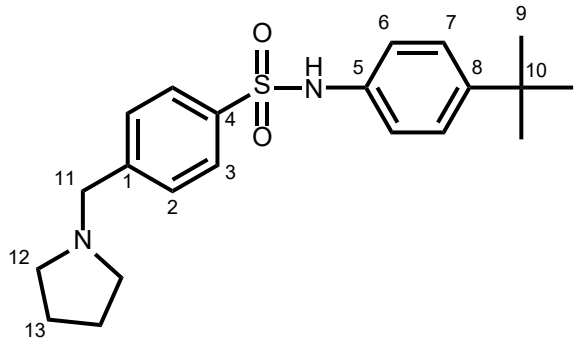
Pale yellow solid (39%). ^{1}H NMR ($CDCl_3$): δ 7.78 (2H, d, J = 8.4 Hz, H-3), 7.52 (2H, d, J = 8.4 Hz, H-2), 7.19-7.27 (5H, m, H-7, H-8, H-9 overlap), 4.13 (2H, s, H-11), 3.71 (4H, m, H-12), 3.59 (2H, s, H-5), 2.59 (3H, s, H-10), 2.46 (4H, m, J = 4.8 Hz, H-13); ^{13}C NMR ($CDCl_3$): δ 145.0 (C-1), 135.7 (C-4), 135.7 (C-6), 129.4 (C-2), 128.7 (C-8), 128.4 (C-7), 127.9 (C-9), 127.5 (C-3), 60.13 (C-11), 54.3 (C-12), 54.1 (C-5), 34.5 (C-10), 23.5 (C-13); LRMS (ESI) m/z: 345.1 $(M+H)^+$; HRMS: m/z $[M+H]^+$ calcd for $C_{19}H_{25}N_2O_2S$: 345.1631, found: 345.1629.

***N*-benzyl-4-(pyrrolidin-1-ylmethyl)benzenesulfonamide (72)**



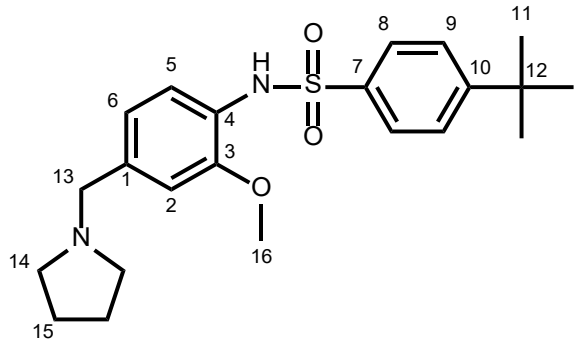
Colourless solid (35% yield). ^{1}H NMR ($CDCl_3$): δ 7.82 (2H, d, J = 8.39 Hz, H-3), 7.50 (2H, d, J = 8.39 Hz, H-2), 7.21-7.27 (m, 5H, H-7, H-8, H-9 overlap), 4.14 (2H, d, J = 5.57 Hz, H-5), 3.71 (2H, s, H-10), 2.56 (4H, m, H-11), 1.82 (4H, m, H-12); ^{13}C NMR ($CDCl_3$): δ 144.1 (C-1), 138.1 (C-4), 136.4 (C-6), 129.5 (C-2), 128.7 (C-9), 127.9 (C-7/8), 127.9 (C-7/8), 127.2 (C-3), 59.9 (C-5), 54.2 (C-11), 47.3 (C-5), 23.5 (C-12); LRMS (ESI) m/z: 331.1 $(M+H)^+$.

***N*-(4-(*tert*-butyl)phenyl)-4-(pyrrolidin-1-ylmethyl)benzenesulfonamide (73)**



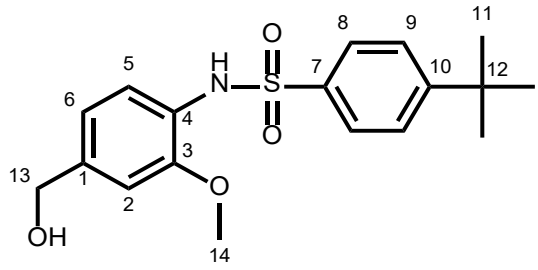
Colourless solid (42%). ^{1}H NMR ($CDCl_3$): δ 7.72 (2H, d, J = 8.4 Hz, H-3), 7.43 (2H, d, J = 8.4 Hz, H-2), 7.22 (2H, d, J = 8.5 Hz, H-7), 6.99 (2H, d, J = 8.5 Hz, H-6), 3.68 (2H, s, H-11), 2.54 (4H, m, H-12), 1.80 (4H, m, H-13), 1.24 (9H, s, H-9); ^{13}C NMR ($CDCl_3$): δ 148.5 (C-8), 144.5 (C-1), 138.12 (C-4), 133.75 (C-5), 129.3 (C-3), 127.3 (C-2), 126.1 (C-7), 121.8 (C-6), 59.9 (C-11), 54.2 (C-12), 34.4 (C-10), 31.3 (C-9), 23.4 (C-13); LRMS (ESI) m/z: 373.1 $(M+H)^+$.

4-(*tert*-butyl)-*N*-(2-methoxy-4-(pyrrolidin-1-ylmethyl)phenyl)benzenesulfonamide (74)



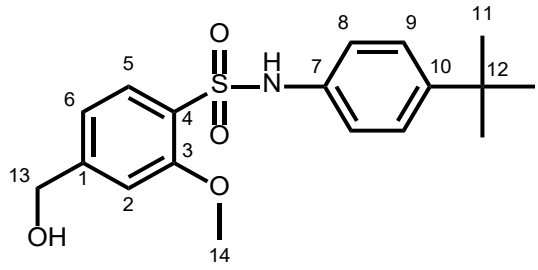
4-(*tert*-Butyl)-*N*-(4-(hydroxymethyl)-2-methoxyphenyl)benzenesulfonamide (50 mg, 0.14 mmol) was dissolved in excess thionyl chloride (3 mL) and left to stir at 60 °C for 16 hours. Thionyl chloride was removed under reduced pressure as much as possible. The crude acid chloride was dissolved in excess pyrrolidine (5 mL) and stirred at room temperature for 16 hours. The excess pyrrolidine was removed at reduced pressure. The crude was dissolved in ethyl acetate (30 mL) then washed with water (2 x 5 mL). The organic layer was dried anhydrous magnesium sulfate and concentrated at reduced pressure. The crude was purified with flash column chromatography (2:8, ethyl acetate and petroleum ether) to give the titled compound as a pale-yellow solid (35 mg, 63%). m.p.: 130.0-130.6 °C; IR: 3254, 2963, 2870, 1595, 1509, 1462, 1432, 1394, 1365, 1336 cm-1; ^{1}H NMR ($CDCl_3$): δ 7.66 (2H, d, J = 8.64 Hz, H-8), 7.43 (1H, d, J = 8.1 Hz, H-5), 7.39 (2H, d, J = 8.6 Hz, H-9), 7.03 (1H, d, J = 1.7 Hz, H-2), 6.81 (1H, dd, J = 8.1, 1.7 Hz, H-6), 3.67 (2H, s, H-13), 3.65 (3H, s, H-16), 2.65 (4H, m, H-14), 1.86 (4H, m, H-15), 1.27 (9H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 156.6 (C-10), 149.7 (C-3), 136.2 (C-7), 134.2 (C-4), 127.1 (C-8), 125.7 (C-1), 125.6 (C-9), 121.7(C-6), 120.6 (C-5), 111.7 (C-2), 60.00 (C-13), 55.9 (C-16), 53.8 (C-14), 35.1 (C-12), 31.01 (C-11), 23.3 (C-15); LRMS (ESI) m/z: 403.2 $(M+H)^+$; HRMS (ESI) m/z calcd. for $C_{22}H_{30}N_2O_3S$ $[M+H]^+$: 403.2042, found: 403.2045.

4-(*tert*-butyl)-*N*-(4-(hydroxymethyl)-2-methoxyphenyl)benzenesulfonamide (75)



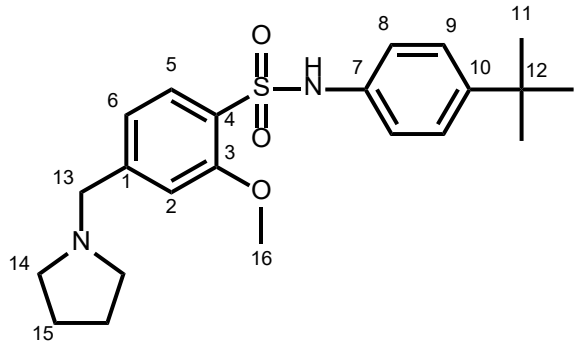
4-(4-*tert*-Butylbenzenesulfonamido)-3-methoxybenzoic acid (100 mg, 0.29 mmol) was dissolved in anhydrous THF (5 mL), lithium aluminium hydride (53 mg, 1.45 mmol) was added at -10 °C. The reaction mixture was left to stir for 16 hours at room temperature. The solvent was removed under reduced pressure. Ethyl acetate was added to the crude solid which was then washed with saturated sodium bicarbonate. The collective organic layer was dried with anhydrous magnesium sulfate and concentrated at reduced pressure. The crude product was purified with flash column chromatography (1:99, methanol and DCM) to give the titled compound as a pale-yellow powder (89 mg, 85%); m.p.: 130.0-130.6 °C; IR: 3254, 2963, 2870, 1595, 1509, 1462, 1432, 1394, 1365, 1336 cm-1; ^{1}H NMR ($CDCl_3$): δ 7.61 (2H, d, *J* = 8.7 Hz, H-8), 7.40 (1H, d, *J* = 8.3, H-5), 7.34 (2H, d, *J* = 8.7 Hz, H-9), 7.19 (s, 1H, OH), 6.94 (1H, s, NH), 6.79 (1H, dd, *J* = 8.3, 1.8 Hz, H-6), 6.73 (1H, d, *J* = 1.8 Hz, H-2), 4.54 (2H, s, H-13), 3.56 (3H, s, H-14), 1.20 (9H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 156.6 (C-10), 149.5 (C-2), 138.0 (C-7), 136.2 (C-4), 127.0 (C-8), 125.8 (C-9), 125.3 (C-1), 120.8 (C-5), 119.4 (C-6), 109.2 (C-2), 64.6 (C-13), 55.0 (C-14), 35.0 (C-12), 32.2 (C-11); LRMS (ESI) m/z: 350.1 (M+H)+; HRMS (ESI) m/z calcd. for $C_{18}H_{24}O_4NS$ $[M+H]^+$: 350.1421, found: 350.1424.

***N*-(4-(*tert*-butyl)phenyl)-4-(hydroxymethyl)-2-methoxybenzenesulfonamide (77)**



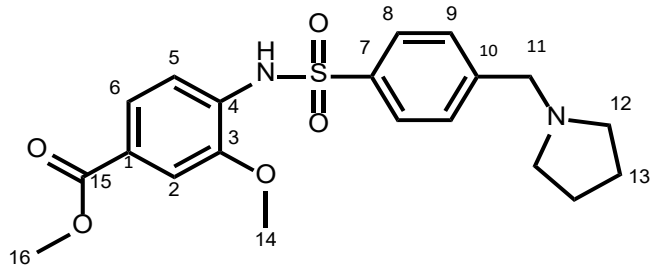
4-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)-3-methoxybenzoic acid (100 mg, 0.28 mmol) was dissolved in anhydrous THF (10 mL), lithium aluminium hydride (53 mg, 1.4 mmol) was added at -10 °C. The reaction mixture was left to stir for 16 hours at room temperature. The solvent was removed at reduced pressure. Ethyl acetate (30 ml) was added to the crude solid which was then washed with saturated sodium bicarbonate (3 x 5 ml). The collective organic layer was dried with anhydrous magnesium sulfate and concentrated under reduced pressure. The crude was purified with flash column chromatography (1:99, methanol and DCM) to give the titled compound as a pale-yellow powder (78 mg, 77%); ^{1}H NMR ($CDCl_3$): δ 7.71 (1H, d, J = 8.0 Hz, H-5), 7.18 (2H, d, J = 8.5 Hz, H-9), 7.03 (1H, s, H-2), 6.97 (2H, d, J = 8.5 Hz, H-8), 6.86 (1H, d, J = 8.0 Hz, H-6), 4.67 (2H, s, H-13), 3.49 (3H, s, H-14), 1.25 (9H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 155.4 (C-3), 147.7 (C-1), 147.3 (C-10), 132.9 (C-7), 129.8 (C-5), 125.1 (C-9), 124.2 (C-4), 120.3 (C-8), 117.2 (C-6), 108.9 (C-2), 63.4 (C-13), 55.3 (C-14), 33.3 (C-12), 30.2 (C-11); LRMS (ESI) m/z: 350.1 $(M+H)^+$; HRMS (ESI) m/z calcd. for $C_{18}H_{24}O_4NS$ $[M+H]^+$: 350.1421, found: 350.1424.

***N*-(4-(*tert*-butyl)phenyl)-2-methoxy-4-(pyrrolidin-1-ylmethyl)benzenesulfonamide (78)**



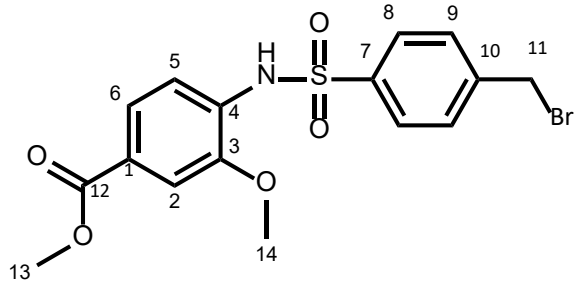
4-(*tert*-butyl)-*N*-(4-(hydroxymethyl)-2-methoxyphenyl)benzenesulfonamide (70 mg, 0.193 mmol) was dissolved in thionyl chloride (5 mL) and left to stir at 60 °C for 16 hours. Thionyl chloride was removed under reduced pressure as much as possible. The crude acid chloride was dissolved in excess pyrrolidine (6 mL) and stirred at room temperature for 16 hours. The excess pyrrolidine was removed at reduced pressure. The crude product was dissolved in ethyl acetate (30 mL) then washed with 0.5M hydrochloric acid (2 x 5 mL). The organic layer was dried with anhydrous magnesium sulfate and concentrated under reduced pressure. The crude product was purified with flash column chromatography (1:99, methanol and DCM) to give the titled compound as a pale-yellow solid (37 mg, 49%); ^{1}H NMR (MeOD): δ 7.61 (1H, d, J = 8.3 Hz, H-5). 7.08 (2H, d, J = 9.0, H-9), 7.05 (1H, d, J = 1.4 Hz, H-2), 6.91 (2H, d, J = 9.0 Hz, H-8), 6.85 (1H, dd, J = 1.4, 8.3 Hz, H-6), 3.86 (3H, s, H-16), 3.58 (2H, s, H-13), 2.47 (4H, m, H-14), 1.71 (4H, m, H-15), 1.10 (9H, s, H-11); ^{13}C NMR (MeOD): δ 156.8 (C-3), 147.1 (C-10), 145.4 (C-1), 134.9 (C-7), 130.7 (C-5), 125.5 (C-1), 125.4 (C-9), 120.5 (C-6), 120.4 (C-8), 112.7 (C-2), 59.5 (C-13), 55.2 (C-16), 53.6 (C-14), 33.7 (C-12), 30.3 (C-11), 22.7 (C-15); LRMS (ESI) m/z: 403.1 (M+H)+; HRMS (ESI) m/z calcd. for $C_{22}H_{31}N_2O_3S$ $[M+H]^+$: 403.2042, found: 403.2045.

Methyl 3-methoxy-4-(4-(pyrrolidin-1-ylmethyl)phenylsulfonamido)benzoate (79)



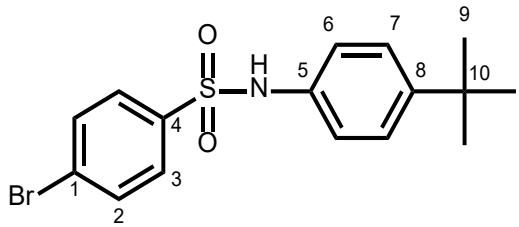
Methyl 4-(4-(bromomethyl)phenylsulfonamido)-3-methoxybenzoate (110 g, 0.27 mmol) was dissolved in acetonitrile (10 mL), followed by potassium carbonate (227mg, 1.64mmol, 3 eqv.) and pyrrolidine (0.5 mL, 0.61 mmol). The reaction mixture was stirred at room temperature for 16 hours. The crude product was dissolved in ethyl acetate (30 mL) then washed with 0.5 M hydrochloric acid (2 x 5 mL). The organic layer was dried with anhydrous magnesium sulfate and concentrated at reduced pressure. The crude product was purified with flash column chromatography (2:8, ethyl acetate and petroleum ether) to give the titled compound as a pale-yellow solid (81mg, 71%). ^{1}H NMR ($CDCl_3$): δ 7.76 (2H, d, J = 8.2 Hz, H-8), 7.59 (1H, dd, J = 8.4, 1.4 Hz, H-6), 7.54 (1H, d, J = 8.4 Hz, H-5), 7.43 (2H, d, J = 8.2 Hz, H-9), 7.42 (1H, d, J = 1.4 Hz, H-2), 3.86 (3H, s, H-16), 3.78 (3H, s, H-14), 3.65 (2H, s, H-11), 2.51 (4H, m, H-12), 1.79 (4H, m, H-13); ^{13}C NMR ($CDCl_3$): δ 166.5 (C-15), 148.1 (C-3), 144.8 (C-10), 137.7 (C-7), 130.6 (C-4), 129.3 (C-9), 127.3 (C-8), 126.1 (C-1), 123.6 (C-6), 118.1 (C-5), 111.4 (C-2), 59.8 (C-11), 56.0 (C-14), 54.2 (C-12), 52.2 (C-16), 23.47 (C-13); LRMS (ESI) m/z: 405.1 $(M+H)^+$; HRMS (ESI) m/z calcd. for $C_{20}H_{25}O_5N_2S$ $[M+H]^+$: 405.1468, found: 405.1471.

Methyl 4-(4-(bromomethyl)phenylsulfonamido)-3-methoxybenzoate (80)



To a solution of methyl 4-amino-3-methoxybenzoate (100 mg, 0.55 mmol) in DCM (10 mL) was added 4-(bromomethyl)benzene-1-sulfonyl chloride (149 mg, 0.55 mmol), followed by pyridine (0.13 mL, 1.65 mmol, ~3 eqv.). The reaction mixture was left to stir at room temperature for 16 hours. The solvent was evaporated at reduced pressure. The crude product was purified with flash column chromatography (2:8, ethyl acetate and petroleum ether) to give the titled compound as a pale-yellow solid (155 mg, 68%). ^{1}H NMR ($CDCl_3$): δ 7.80 (2H, d, J = 8.6 Hz, H-8), 7.60 (1H, dd, J = 8.4, 1.4 Hz, H-6), 7.56 (1H, d, J = 8.4 Hz, H-5), 7.44 (2H, d, J = 8.6 Hz, H-9), 7.43 (1H, d, J = 1.4 Hz, H-2), 4.55 (2H, s, H-11), 3.87 (3H, s, H-13), 3.77 (3H, s, H-14). ^{13}C NMR ($CDCl_3$): δ 166.4 (C-12), 148.3 (C-3), 142.8 (C-10), 138.9 (C-7), 130.3 (C-4), 129.7 (C-9), 127.6 (C-8), 126.5 (C-1), 118.5 (C-6), 118.0 (C-5), 111.5 (C-2), 56.0 (C-13), 52.2 (C-14), 44.8 (C-11); LRMS (ESI) m/z: 413.1 $(M+H)^+$.

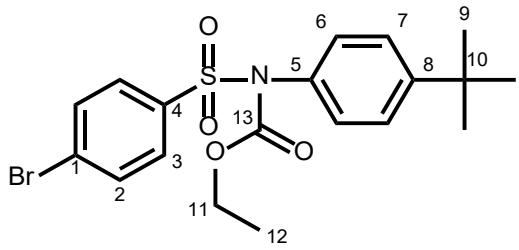
4-bromo-*N*-(4-(*tert*-butyl)phenyl)benzenesulfonamide (82)



4-Bromobenzenesulfonyl chloride (2 g, 7.84 mmol) was dissolved in pyridine (1.9 mL, 23.6 mmol, ~3 eqv.). 4-*tert*butyaniline (1.25 mL, 7.86 mmol, ~1 eqv.) was added. The reaction mixture was left to stir at room temperature for 5 hours. (2.192 mg, 76%). ^{1}H NMR ($CDCl_3$): δ 7.64 (2H, d, J = 8.8 Hz, H-3), 7.60 (2H, d, J = 8.8 Hz, H-2), 7.29 (2H, d, J = 8.6 Hz, H-7), 7.00 (2H, d, J = 8.6 Hz, H-6), 6.60, (2H, s, NH), 1.29 (9H, s, H-9); ^{13}C NMR ($CDCl_3$): δ 149.8 (C-8), 138.3 (C-1), 133.1

(C-5), 132.3 (C-3), 128.8 (C-2), 128.0 (C-4), 126.4 (C-7), 122.1 (C-6), 34.5 (C-10), 31.3 (C-9); LRMS (ESI) m/z: 369.1 $(M+H)^+$.

Ethyl ((4-bromophenyl)sulfonyl)(4-(*tert*-butyl)phenyl)carbamate (83)

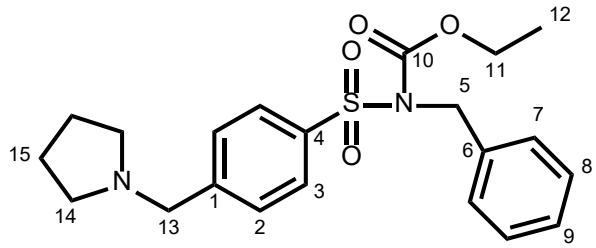


4-bromo-*N*-(4-(*tert*-butyl)phenyl)benzenesulfonamide (2.1 g, 5.71 mmol) dissolved in THF (15 mL), 60% sodium hydride (457 mg, 11.42 mmol) was added, followed by ethyl chloroformate (8.57 mmol, ~1.5 eqv). The reaction mixture was stirred at room temperature for two hours. THF was removed under reduced pressure. Ethyl acetate (40 mL) was added to the crude mixture, which was then washed with 0.5M HCl (2 x 10 mL). The collective organic layer was dried with anhydrous magnesium sulfate then concentrated under reduced pressure to give the title compound. (2.487 g, 99 %). ^{1}H NMR ($CDCl_3$): δ 7.88 (2H, d, J = 8.8 Hz, H-3), 7.70 (2H, d, J = 8.8 Hz, H-2), 7.44 (2H, d, J = 8.7 Hz, H-7), 7.44 (2H, d, J = 8.7 Hz, H-6), 4.13 (2H, q, J = 7.0 Hz, H-11), 1.34 (9H, s, H-9), 1.15 (3H, s, J = 7.03 Hz, H-12); ^{13}C NMR ($CDCl_3$): δ 152.6 (C-8), 138.3 (C-4), 132.9 (C-5), 132.1 (C-2), 130.5 (C-3), 129.0 (C-1), 128.9 (C-7), 126.4 (C-7), 63.7 (C-11), 34.8 (C-10), 31.3 (C-9), 14.1 (C-12). ; LRMS (ESI) m/z: 441.1 $(M+H)^+$.

Generation procedure for compound 88 and 89

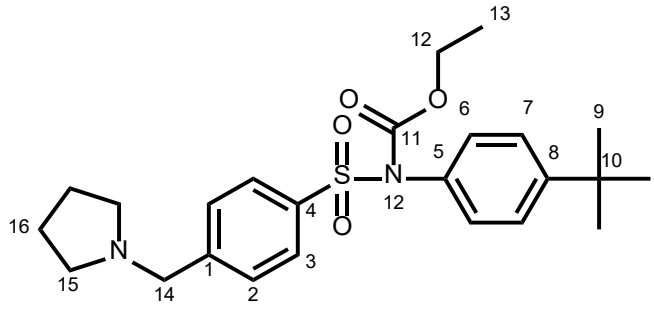
To a solution of **70**/**71** in acetonitrile was added potassium carbonate (3 eqv.), ethyl chloroformate (1 eqv.) was added slowly over an hour. The reaction mixture was left to stir room temperature for 16 hours. Ethyl acetate was added to the reaction mixture and washed with saturated sodium bicarbonate. The organic layer was dried with anhydrous magnesium sulfate and concentrated at reduced pressure. The crude product was purified using flash column chromatography (2:98 methanol and DCM) to give the product.

Ethyl benzyl((4-(pyrrolidin-1-ylmethyl)phenyl)sulfonyl)carbamate (88)



Colourless solid (32%). ^{1}H NMR ($CDCl_3$): δ 7.56 (2H, d, J = 8.5 Hz, H-3), 7.32 (2H, m, H-7), 7.30 (2H, d, J = 8.5 Hz, H-2), 7.23-7.25 (3H, m, H-8, H-9 overlap), 5.06 (2H, s, H-5), 4.13 (2H, q, J = 7.1 Hz, H-11), 3.64 (2H, s, H-13), 2.48 (4H, m, H-14), 1.79 (4H, m, H-15), 1.13 (3H, t, J = 7.1 Hz, H-12); ^{13}C NMR ($CDCl_3$): δ 152.5 (C-10), 145.9 (C-1), 137.8 (C-4), 137.0 (C-6), 128.6 (C-2), 128.5 (C-3/7/8), 128.4 (C-3/7/8), 128.4 (C-3/7/8), 127.8 (C-9), 63.6 (C-11), 60.1 (C-13), 54.2 (C-14), 49.9 (C-5), 23.5 (C-15), 14.0 (C-12); LRMS (ESI) m/z: 403.1 $(M+H)^+$.

Ethyl (4-(*tert*-butyl)phenyl)((4-(pyrrolidin-1-ylmethyl)phenyl)sulfonyl)carbamate (89)



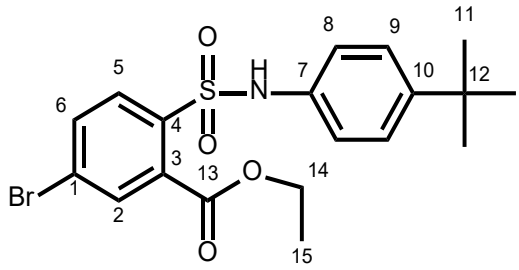
Pale-yellow solid (35%). ^{1}H NMR ($CDCl_3$): δ 7.99 (2H d, J = 8.3 Hz, H-3), 7.69 (d, 2H, J = 8.3 Hz, H-2), 7.44 (d, 2H, J = 8.7, H-7), 7.16 (d, 2H, J = 8.7 Hz, H-6), 4.12 (q, 2H, J = 7.1 Hz, H-12), 3.88 (s, 2H, H-14), 2.76 (m, 4H, H-15), 1.92 (m, 4H, H-16), 1.35 (s, 9H, H-9), 1.14 (t, 3H, J = 7.1 Hz, H-13); ^{13}C NMR ($CDCl_3$): δ 152.4 (C-11) (C-8), 142.7 (C-1), 138.3 (C-4), 133.0 (C-5), 129.6 (C-2), 129.2 (C-3), 129.0 (C-6), 126.3 (C-7), 63.6 (C-12), 59.4 (C-14), 54.0 (C-15), 34.8 (C-10), 31.3 (C-9), 23.4 (C-16), 14.1 (C-13); LRMS (ESI) m/z: 445.1 $(M+H)^+$.

General procedure for directed *ortho* metalation rearrangement reaction for 81, 86, 87

A) Preparation of Lithium diisopropylamine: A sealed flask containing solution of diisopropylamine (5 eqv.) in THF was charged with nitrogen, *N*-butyllithium (5 eqv.) was added at -78 °C. The lithium diisopropylamine solution was stirred at -78 °C under nitrogen until it is used.

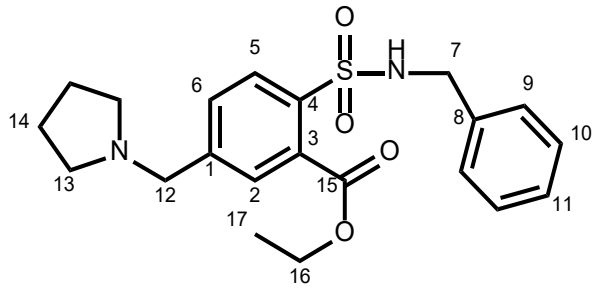
B) A sealed flask containing solution of **83**/**88**/**89** in THF was charged with nitrogen and kept at -78 °C. All of the previously prepared lithium diisopropylamine in THF solution was added through a cannula needle. The reaction mixture was stirred at -78 °C for 3 minutes. Diluted citric acid was immediately added to the reaction mixture. The mixture was extracted with ethyl acetate (3 x 15 mL). The collective organic layer was dried with anhydrous magnesium sulfate then concentrated under reduced pressure. The crude product was purified by using flash column chromatography (1:99 methanol and DCM) to give the product.

Ethyl 5-bromo-2-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)benzoate (81)



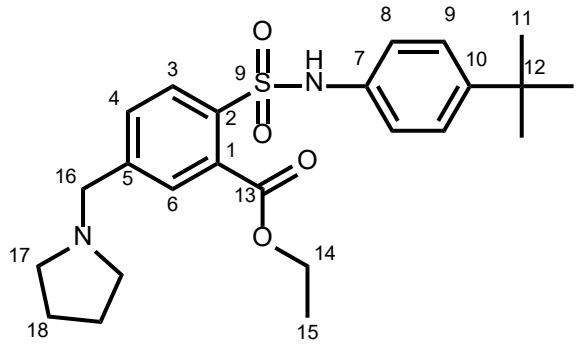
Colourless solid (68%). ^{1}H NMR ($CDCl_3$): δ 7.92 (1H, d, J = 2.0 Hz, H-2), 7.87 (1H, s, NH), 7.66 (1H, d, J = 8.4 Hz, H-5), 7.59 (1H, dd, J = 8.4, 2.0 Hz, H-6), 7.24 (2H, d, J = 8.7 Hz, H-9), 7.05 (2H, d, J = 8.7 Hz, H-8), 4.51 (2H, q, J = 7.1 Hz, H-14), 1.47 (3H, t, J = 7.1 Hz, H-15), 1.25 (9H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 166.6 (C-13), 149.1 (C-10), 137.3 (C-1), 134.3 (C-6), 133.5 (C-7), 133.5 (C-2), 132.5 (C-4), 131.8 (C-5), 127.3 (C-3), 126.2 (C-9), 122.7 (C-8), 63.3 (C-14), 34.4 (C-12), 31.3 (C-11), 14.1 (C-15); LRMS (ESI) m/z: 441.1 $(M+H)^+$.

Ethyl 2-(*N*-benzylsulfamoyl)-5-(pyrrolidin-1-ylmethyl)benzoate (86)



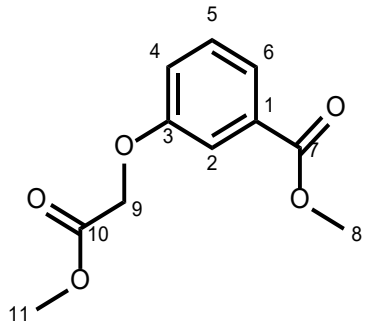
Pale-yellow solid (53%). ^{1}H NMR ($CDCl_3$): δ 7.93 (1H, d, J = 8.1 Hz, H-5), 7.76 (1H, d, J = 1.6 Hz, H-2), 7.53 (1H, dd, J = 1.6, 8.1 Hz, H-6), 7.14-7.20 (9H, m, H-9, H-10, H-11 overlap), 4.39 (2H, q, J = 5.3 Hz, H-16), 4.18 (2H, d, J = 6.2 Hz, H-7), 3.7 (2H, s, H-12), 2.54 (4H, m, H-13), 1.82 (4H, m, H-14), 1.39 (3H, t, J = 7.11 Hz, H-17); ^{13}C NMR ($CDCl_3$): δ 167.8 (C-15), 137.6 (C-1), 136.4 (C-8), 130.9 (C-3), 130.7 (C-2), 135.0 (C-4), 131.7 (C-6), 129.9 (C-5), 128.5 (C-10), 127.9 (C-9), 127.6 (C-11), 62.6 (C-16), 59.5 (C-7), 54.2 (C-13), 47.7 (C-12), 23.5 (C-14), 14.1 (C-17); LRMS (ESI) m/z: 403.1 $(M+H)^+$.

Ethyl 2-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)-5-(pyrrolidin-1-ylmethyl)benzoate (84)



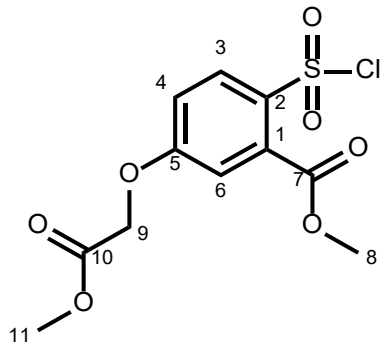
Pale-yellow solid (66%). ^{1}H NMR ($CDCl_3$): δ 7.92 (1H, s, NH), 7.77 (1H, d, J = 1.5 Hz, H-6), 7.74 (1H, d, J = 8.1 Hz, H-3), 7.43 (1H, dd, J = 1.5, 8.1 Hz, H-4), 7.21 (2H, d, J = 8.8 Hz, H-8), 7.06 (2H, d, J = 8.8 Hz, H-9), 4.49 (2H, q, J = 7.2 Hz, H-14), 3.67 (2H, s, H-16), 2.51 (4H, m, H-17), 1.80 (4H, m, H-18), 1.46 (3H, t, J = 7.20 Hz, H-15), 1.23 (9H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 168.0 (C-13), 148.7 (C-10), 144.6 (C-1), 136.7 (C-2), 134.0 (C-5), 131.2 (C-4), 131.0 (C-7), 130.7 (C-6), 130.5 (C-3), 126.0 (C-9), 122.6 (C-8), 62.8 (C-14), 59.6 (C-16), 54.2 (C-17), 34.4 (C-12), 31.2 (C-11), 23.5 (C-18), 14.1 (C-15); LRMS (ESI) m/z: 445.1 $(M+H)^+$.

Methyl 3-(2-methoxy-2-oxoethoxy)benzoate (95)



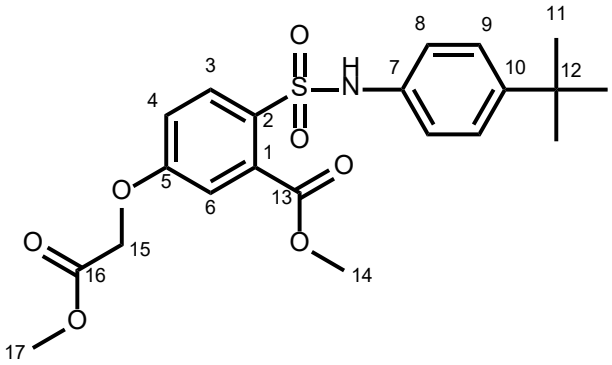
To a solution of methyl 3-hydroxybenzoate (1 g, 6.58 mmol) dissolved in tetrahydrofuran (10 mL), methyl 2-bromoacetate (1.05 mL, 6.60 mmol, ~1 eqv.) was added, followed by 60% sodium hydride (758 mg, 19.74 mmol, ~3 eqv.). The reaction mixture was stirred at 60 °C for 16 hours. Tetrahydrofuran was removed by rotary evaporation, the remaining solid was dissolved in water and diluted hydrochloric acid (1M) then extracted with ethyl acetate. The collective organic layer was dried in anhydrous magnesium sulfate. The crude product was purified using flash column chromatography with (2:8, ethyl acetate: petroleum ether) methyl 3-(2-methoxy-2-oxoethoxy)benzoate was obtained as a clear colourless oil. (1.047 g, 71%). ^{1}H NMR ($CDCl_3$): δ 7.68 (1H, dt, J = 1.1, 7.7 Hz, H-6), 7.54 (1H, dd, J = 1.1, 1.5 Hz, H-2), 7.36 (1H, t, J = 8.0 Hz, H-5), 7.14 (1H, ddd, J = 1.0, 2.7, 8.26 Hz, H-4), 4.68 (2H, s, H-9), 3.90 (3H, s, H-8), 3.81 (3H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 169.0 (C-10), 166.7 (C-7), 157.8 (C-3), 131.6 (C-1), 129.6 (C-5), 123.1 (C-6), 120.1 (C-4), 114.8 (C-2), 65.3 (C-9), 52.3 (C-11), 52.2 (C-8); LRMS (ESI) m/z: 225.9 $(M+H)^+$.

Methyl 2-(chlorosulfonyl)-5-(2-methoxy-2-oxoethoxy)benzoate (96)



To a solution of methyl 3-(2-methoxy-2-oxoethoxy)benzoate (1 g, 4.462 mmol) in dichloromethane (10 mL) charged with nitrogen, chlorosulfonic acid (0.9 mL, 13.54 mmol, ~3 eqv.) was added to solution in an ice and water bath. The reaction mixture was stirred at room temperature for 48 hours. The reaction mixture was added to crushed ice then extracted with ethyl acetate (3 x 20 mL). The collective organic layer was dried with anhydrous magnesium sulfate then concentrated under reduced pressure. The crude product was purified using flash column chromatography (2:8, ethyl acetate and petroleum either) to give the titled compound as a colourless solid (416 mg, 29 %). ^{1}H NMR ($CDCl_3$): δ 8.10 (1H, d, J = 9.1 Hz, H-3), 7.18 (1H, d, J = 2.7 Hz, H-6), 7.1 (1H, dd, J = 9.1, 2.7 Hz, H-4), 4.76 (2H, s, H-9), 3.98 (3H, s, H-8), 3.83 (3H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 167.7 (C-10), 165.8 (C-7), 162.5 (C-5), 134.8 (C-2), 134.5 (C-1), 131.9 (C-3), 116.30 (C-4), 116.2 (C-6), 65.3 (C-9), 53.5 (C-8), 52.7 (C-11); LRMS (ESI) m/z: 323.1 $(M+H)^+$.

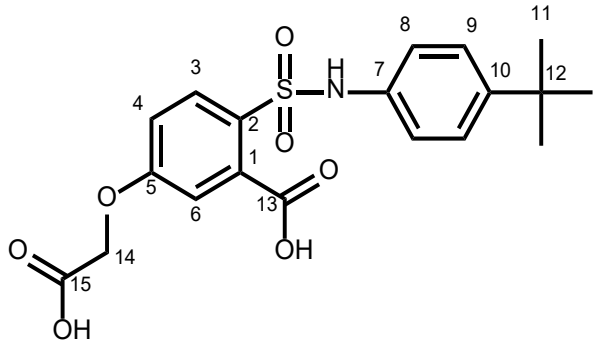
Methyl 2-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)-5-(2-methoxy-2-oxoethoxy)benzoate (92)



To a solution of methyl 2-(chlorosulfonyl)-5-(2-methoxy-2-oxoethoxy)benzoate (300 mg, 0.932 mmol) dissolved in dichloromethane (7 mL), pyridine (0.22 mL, 2.73 mmol, ~3 eqv), followed by 4-*tert*butylaniline (0.15 mL, 0.94 mmol, ~1e qv.) was added. The reaction mixture was stirred at room temperature for 5 hours. (268 mg, 66 %). ^{1}H NMR ($CDCl_3$): δ 7.81 (1H, s, NH), 7.76 (1H, d, J = 8.8 Hz, H-3), 7.33 (1H, d, J = 2.8 Hz, H-6), 7.22 (2H, d, J = 8.6 Hz, H-8) 7.06 (2H, d, J = 8.6 Hz, H-9), 6.87 (1H, dd, J = 8.8, 2.8 Hz, H-4), 4.67 (2H, s, H-15), 4.02 (3H, s, H-14), 3.80 (3H, s, H-17), 1.25 (9H, s, H-11).; ^{13}C NMR ($CDCl_3$): δ 168.1 (C-16), 167.6 (C-13), 160.4 (C-5), 148.8 (C-10), 134.0 (C-7), 132.8 (C-3), 132.5 (C-2),

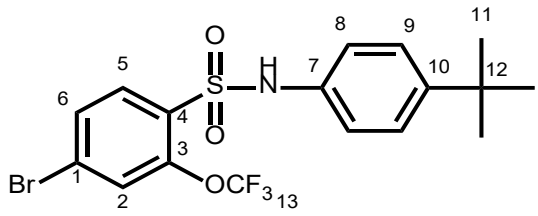
131.2 (C-1), 126.0 (C-8), 122.6 (C-9), 117.8 (C-6), 115.7 (C-4), 65.3 (C-15), 53.6 (C-14), 52.5 (C-17), 33.4 (C-12), 31.3 (C-11); LRMS (ESI) m/z: 436.1 $(M+H)^+$.

2-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)-5-(carboxymethoxy)benzoic acid (93)



Compound **89** (200 mg, 0.46 mmol) was added to 2M sodium hydroxide (10 mL). The reaction mixture was heated at 100 °C with stirring for two hours. Hydrochloric acid (3.3 mL) was added to neutralize the reaction mixture. The aqueous mixture was extracted with ethyl acetate. The collective organic layer was dried with anhydrous magnesium sulfate then concentrated under reduced pressure. The crude product was purified by using flash column chromatography (1:99 methanol and dichloromethane) to give the title compound (179 mg, 96%). ^{1}H NMR (MeOD): δ 7.59 (1H, d, J = 8.86 Hz, H-3), 7.3 (1H, d, J = 1.36 Hz, H-6), 7.21 (2H, d, J = 8.40 Hz, H-9), 7.03 (2H, d, J = 8.40 Hz, H-8), 6.89 (1H, d, J = 8.86 Hz, H-4), 4.17 (2H, s, H-14), 1.20 (9H, s, H-11); ^{13}C NMR (MeOD): δ 170.9 (C-15), 161.2 (C-5), 148.2 (C-10), 134.7 (C-7), 131.4 (C-3), 129.04, 125.5 (C-9), 122.1 (C-8), 116.6 (C-6), 114.4 (C-4), 65.1 (C-14), 33.8 (C-12), 30.3 (C-11); LRMS (ESI) m/z: 408.1 $(M+H)^+$.

4-bromo-*N*-(4-(*tert*-butyl)phenyl)-2-(trifluoromethoxy)benzenesulfonamide (100)



A solution of 4-bromo-2-(trifluoromethoxy)benzene-1-sulfonyl chloride (500 mg, 1.48 mmol) in DCM (12 mL) was added 4-*tert*butylaniline and pyridine (0.24 mL,

1.50 mmol, ~1 eqv.). The reaction mixture was stirred at room temperature for 5 hours. The reaction mixture then was washed with brine (3 x 10 mL). The collective organic phase was dried with anhydrous magnesium sulfate then concentrated in a rotary evaporator. The crude product was purified using flash column chromatography using petroleum either and ethyl acetate (7:3) as the eluent to give the title compound as a colourless solid (460 mg, 69%). ^{1}H NMR ($CDCl_3$): δ 7.73 (1H, d, J = 8.4 Hz, H-5), 7.46 (1H, d, J = 1.5 Hz, H-2), 7.39 (1H, dd, J = 8.4, 1.5 Hz, H-6), 7.17 (2H, d, J = 8.6 Hz, H-9), 6.91 (1H, d, J = 8.6 Hz, H-8), 1.18 (9H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 149.2 (C-10), 146.0 (C-3), 132.6 (C-5), 132.5 (C-7), 130.1 (C-1), 129.9 (C-6), 128.5 (C-4), 126.4 (C-8), 123.2 (C-2), 121.6 (C-9), 123.2/121.5/118.9/116.3 (C-13), 34.4 (C-12), 31.2 (C-11); LRMS (ESI) m/z: 453.1 $(M+H)^+$.

Heck coupling general procedure for compound 97, 98 and 107

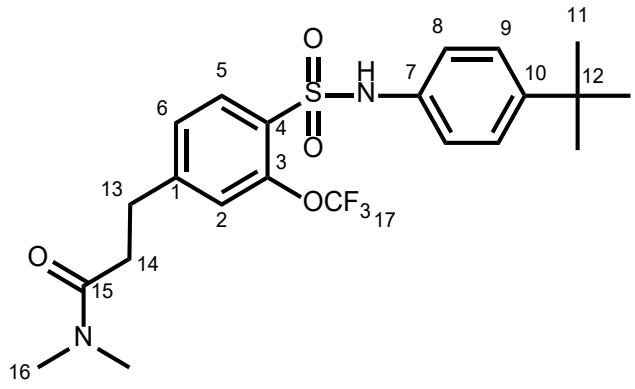
A) Compound **81**/**100** was dissolved in minimum amount of anhydrous dimethylformamide. To the solution was added palladium (II) acetate (0.004eqv) and triphenylphosphine (0.04 eqv.), methyl acrylate/ dimethylcarbamic chloride (4 eqv.) and triethylamine (4 eqv.). The reaction mixture was heated at 100 °C for 16 hours. Dimethylformamide was removed as much as possible under reduced pressure. The crude was diluted in ethyl acetate then washed with saturated sodium bicarbonate solution. The collective organic layer was dried in anhydrous magnesium sulfate then concentrated under reduced pressure to give the crude product.

B) The crude from A) was dissolved in methanol, palladium on carbon (10% (w/w), 4.00 mg was added. The flask containing the reaction mixture was charged with hydrogen and sealed. The reaction mixture was stirred at r.t. under hydrogen for 24h. The mixture was filtered through celite and concentrated under reduced pressure to give the crude product.

C) The crude from B) was added to 3M sodium hydroxide (10 mL) and heated at 100 °C for 1 hour. The reaction mixture was cooled to room temperature. 6M Hydrochloric acid (5 mL) was added to neutralize solution. The crude solid was

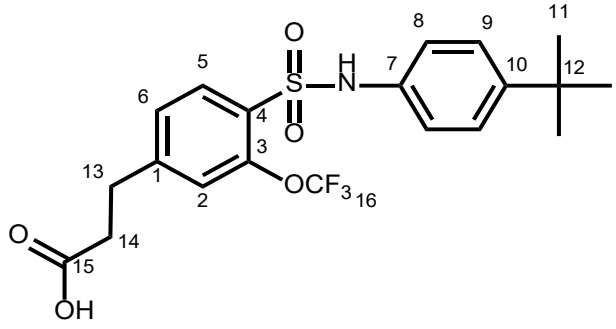
filtered, then purified by using flash column chromatography (1:99 methanol and DCM).

3-(4-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)-3-(trifluoromethoxy)phenyl)-*N*,*N*-dimethylpropanamide (97)



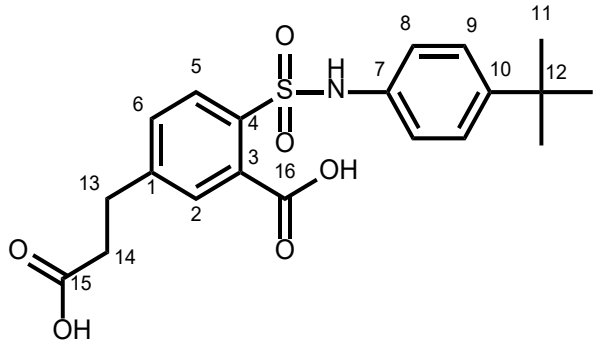
Colourless solid (48%). ^{1}H NMR ($CDCl_3$): δ 7.84 (1H, d, J = 8.1 Hz, H-5), 7.22 (2H, d, J = 8.8 Hz, H-9), 7.22 (1H, d, J = 1.5 Hz, H-2), 7.18 (1H, dd, J = 8.1, 1.5 Hz, H-6), 6.99 (1H, d, J = 8.8 Hz, H-8), 6.70 (1H, s, NH), 3.02 (2H, t, J = 7.5 Hz, H-13), 2.92 (3H, s, H-16), 2.90 (3H, s, H-16), 2.59 (2H, t, J = 7.5 Hz, H-14), 1.23 (9H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 170.9 (C-15), 149.9 (C-1), 148.7 (C-10), 145.9 (C-3), 133.1 (C-7), 131.6 (C-5), 128.7 (C-4), 126.5 (C-6), 126.2 (C-9), 121.3 (C-8), 119.9 (C-2), 124.2/121.6/119.0/116.4 (C-17), 37.1 & 35.5 (C-16), 34.4 (C-12), 34.0 (C-14), 31.2 (C-11), 31.1 (C-13); LRMS (ESI) m/z: 473.1 (M+H)+; HRMS (ESI) m/z calcd. for $C_{22}H_{28}O_4N_2F_3S$ $[M+H]^+$: 473.1712, found: 473.1712.

3-(4-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)-3-(trifluoromethoxy)phenyl)propanoic acid (98)



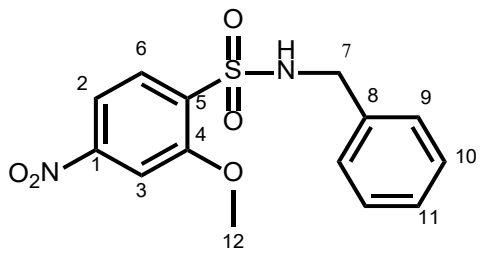
Colourless solid (52%). ^{1}H NMR ($CDCl_3$): δ 7.86 (1H, d, J = 8.1 Hz, H-5), 7.22 (1H, d, J = 1.1 Hz, H-2), 7.22 (1H, d, J = 8.8 Hz, H-9), 7.16 (1H, dd, J = 8.1, 1.1 Hz, H-6), 6.98 (1H, d, J = 8.8 Hz, H-8), 6.85 (1H, s, NH), 3.00 (2H, t, J = 7.4 Hz, H-13), 2.69 (2H, t, J = 7.4 Hz, H-14), 1.22 (9H, s, H-11); ^{13}C NMR ($CDCl_3$): δ 177.1 (C-15), 148.8 (C-10), 148.3 (C-3), 146.0 (C-1), 133.0 (C-7), 131.8 (C-5), 129.1 (C-4), 126.3 (C-6), 126.3 (C-9), 121.4 (C-8), 119.8 (C-2), 124.2/121.7/119.0/116.4 (C-16), 34.4 (C-14, C-12), 31.2 (C-11), 30.2 (C-13); LRMS (ESI) m/z: 446.1 $(M+H)^+$; HRMS (ESI) m/z calcd. for $C_{20}H_{21}O_5NF_3S$ [M-H]-: 444.1112, found: 444.1115.

2-(*N*-(4-(*tert*-butyl)phenyl)sulfamoyl)-5-(2-carboxyethyl)benzoic acid (107)



Colourless solid (51%). ^{1}H NMR (MeOD): δ 7.68 (1H, d, J = 1.7 Hz, H-2), 7.55 (1H, d, J = 8.1 Hz, H-5), 7.25 (1H, dd, J = 8.1, 1.7 Hz. H-6), 7.19 (2H, d, J = 8.8 Hz, H-9), 7.04 (2H, d, J = 8.8 Hz, H-8), 2.92 (2H, t, J = 7.6 Hz, H-13), 2.59 (2H, t, J = 7.6 Hz, H-14), 1.23 (9H, s, H-11); ^{13}C NMR (MeOD): δ 175.1 (C-15), 173.7 (C-16) 148.2 (C-10), 146.4 (C-1), 136.7 (C-4), 134.7 (C-7), 134.3 (C-3), 130.1 (C-2), 129.0 (C-6), 128.8 (C-5), 125.4 (C-9), 122.4 (C-8), 34.4 (C-14), 33.8 (C-12), 30.3 (C-11), 30.1 (C-13); LRMS (ESI) m/z: 406.1 $(M+H)^+$.

***N*-(4-(*tert*-butyl)benzyl)-2-methoxy-4-nitrobenzenesulfonamide (108)**



To a solution of 2-methoxy-4-nitrobenzene-1-sulfonyl chloride (100 mg, 0.4 mmol) dissolved in dichloromethane (6 mL), benzylamine (0.05 mL, 0.4 6mmol, ~1.1 eqv.) was added, followed by pyridine (0.1 mL, 1.2 mmol, ~3 eqv). The reaction mixture was stirred over night at room temperature. The reaction mixture was diluted with more dichloromethane then washed with brine. The collective organic phase was dried with anhydrous magnesium sulfate then concentrated in a rotary evaporator. The crude product was purified using flash column chromatography using petroleum either and ethyl acetate (7:3) as the eluent to give the title compound as pale-yellow solid (8 1mg, 63% yield); IR: 3339, 3105, 2945, 1532, 1325, 1310, 1253, 1151, 1135, 1027 cm^{-1}; ^{1}H NMR ($CDCl_3$): δ 8.06 (1H, d, J = 8.6 Hz, H-6), 7.87 (1H, dd, J = 1.7, 8.3 Hz, H-2), 7.67 (1H, d, J = 1.7 Hz, H-3), 7.12 (2H, m, H-10), 7.12 (1H, m, H-11), 7.01, (2H, m, H-9), 4.10 (2H, d, J = 6.7 Hz, H-7), 3.86 (3H, s, H-12); ^{13}C NMR ($CDCl_3$): δ 156.5 (C-3), 151.4 (C-4), 135.8 (C-8), 133.8 (C-1), 131.2 (C-5), 129.0 (C-10), 128.1 (C-11), 127.9 (C-9), 115.4 (C-6), 107.1 (C-2), 56.9 (C-12), 47.8(C-7); LRMS (ESI) m/z: 323 $(M+H)^{+}$.

Experimental procedure for calcium flux assay

Chemicals and biologicals

All general chemicals, media and media supplements were obtained from Sigma-Aldrich (Poole, UK) unless otherwise specified. Recombinant CCL19 (catalogue number 361-MI-025) and CCL21 (catalogue number 366-6C-025) were purchased from R&D Systems (Abingdon, OX14 3NB, UK). Fluo-4 NW dye mix (Catalogue number F36206), (Component A), Probenecid, water soluble (Component B) and Assay buffer (1X HBSS, 20 mM HEPES) (Component C, included in starter pack only) were purchased from life technologies, UK. 96-well black-wall, flat bottom, microtiter plates (catalogue number 27614033) were purchased from Costar, UK.

Cells and cell culture

Human prostate cancer cell line OSC-19 and PANC-1 were obtained from the American Type Culture Collection (ATCC, Middlesex, UK). The cells were cultured in 75 cm^2 cell culture flasks (T-75) as monolayers in 10% Foetal Calf Serum (FCS), sodium pyruvate (1 mM) and L-glutamine (2 mM) RPMI-1640 medium (Sigma-Aldrich), at 37 °C, 95% air and 5% CO_2 atmosphere.

Method

Cells were seeded into each well of a 96-well black-wall, flat bottom, microtiter plate at a density of 5 x 10^4 cells per well. After 24 h, the growth medium was replaced with 100 µl of the dye loading solution (Molecular Probes™ Fluo-4 NW (no wash), Invitrogen F36206). The plates were incubated at 37 °C for 30 minutes and at room temperature for an additional 30 minutes. 20 µL of a given concentration of the antagonist in medium, or plain medium as control, was added to each well and the plate was incubated at 37 °C for 15 minutes and at room temperature for an additional 30 minutes. The plate was transferred into a Fluoroskan Ascent FL instrument (ThermoScientific) and the fluorescence in response to the addition of 20 µl chemokine (to give 100 nM in-well concentration) was measured at room temperature (Ex 485 nm, Em 538 nm). IC_{50} is calculated as the concentration of the antagonist required to elicit half the maximal response to chemokine using Prism software. Data is presented as the mean ±SE of at least 3 independent experiments.

Preparation of the reagents

Following the manufacturing instructions, 1 mL of the assay buffer (component C) was added to the container labelled as component B (probenecid) and vortex mix in order to make a 250 mM stock solution of probenecid. 10 mL of the assay buffer (component C) and 100 µl of the probenecid solution (component B) was added to the component A (containing the Fluo-4 NW dye). The mixture was shaken for 2 minutes in the vortex mixer to make sure the dye dissolved properly. This dye loading solution was sufficient for one 96-well microplate as 100 µl of the dye is loaded to each well.

Preparation of ligands CCL21 and CCL19

Both ligands were prepared the same way considering the differences in molecular weight. For example, 25 µg lyophilized CCL21 (MW = 12 kDalton) was reconstituted by adding 100 µL of sterile phosphate-buffered-saline solution (PBS) containing 0.1% BSA to afford 100 µL of 2.083 x 10^{-5} M stock solution. First, both ligands were used to determine the concentration to be used in the assay, so serial dilutions were taken place. The first dilution was prepared by adding 100 µL of stock solution to 900 µL of assay buffer (without probenecid) to afford 2.083 x 10^{-6} M solution (1000 µL). If 20 µL of this solution is added to 100 µL of assay buffer (120 µL), the final concentration of CCL21 is 3.47x10^{-7}. Taking into account the 700 µl which is used to purge the system and number of wells repeated in each assay (75 µL for 3 runs and 225 µL is left for next dilutions). CCL21 showed better activity, so 100 nM of CCL21 was conducted from the EC_{50} curve to be used for the whole assay. Taking into account the dead volume of pipes in the fluoroskan, 700 µL, and number of wells used in each run, CCL21 is calculated to make 100 nM in the wells (140 µL total volume).

Preparation of antagonist solutions

A stock solution of each compound was prepared by dissolving a known amount of compound in 100 µl DMSO, to generate a 100 mM stock solution. The first top concentration was made by adding 2 µL of the stock solution to an Eppendorf tube containing 198 µL of assay buffer without probenecid to give a 1 mM solution (containing 1% DMSO v/v). When the 20 µL of first dilution solution is added directly to the well (total volume 140 µL: 100 µl loading dye mixture + 20 µl ligand

+ 20 µl antagonist) the final concentration of the molecule in cell suspension is 142 µM, containing 0.14% DMSO.

9.0 References

[1] Steeg P. S. (2016). Targeting metastasis. *Nature reviews. Cancer*, *16*(4), 201–218. https://doi.org/10.1038/nrc.2016.25

[2] Wirtz, D., Konstantopoulos, K., & Searson, P. C. (2011). The physics of cancer: the role of physical interactions and mechanical forces in metastasis. *Nature Reviews, Cancer*, *11*(7), 512–522. https://doi.org/10.1038/nrc3080

[3] Dillekås, H., Rogers, M. S., & Straume, O. (2019). Are 90% of deaths from cancer caused by metastases? *Cancer Medicine*, *8*(12), 5574–5576. https://doi.org/10.1002/cam4.2474.

[4] Sarvaiya, P. J., Guo, D., Ulasov, I., Gabikian, P. et al. (2013). Chemokines in tumor progression and metastasis. *Oncotarget*, *4*(12), 2171–2185. https://doi.org/10.18632/oncotarget.1426

[5] de Munnik, S. M., Smit, M. J., Leurs, R. et al. (2015). Modulation of cellular signalling by herpesvirus-encoded G protein-coupled receptors. *Frontiers in Pharmacology*, *6*, 40. https://doi.org/10.3389/fphar.2015.00040

[6] Balkwill F. R. (2012). The chemokine system and cancer. *The Journal of Pathology*, *226*(2), 148–157. https://doi.org/10.1002/path.3029

[7] Allen, S. J., Crown, S. E., & Handel, T. M. (2007). Chemokine: receptor structure, interactions, and antagonism. *Annual Review of Immunology*, *25*, 787–820. https://doi.org/10.1146/annurev.immunol.24.021605.090529

[8] Nimmagadda S. (2012). Differential Expression of Chemokine Receptors and their Roles in Cancer Imaging. *Frontiers in Oncology*, *2*, 46. https://doi.org/10.3389/fonc.2012.00046

[9] Bi, H., Zhang, Y., Wang, S., Fang, W. et al. (2019). Interleukin-8 promotes cell migration via CXCR1 and CXCR2 in liver cancer. *Oncology Letters*, *18*(4), 4176–4184. https://doi.org/10.3892/ol.2019.10735

[10] Ha, H., Debnath, B., & Neamati, N. (2017). Role of the CXCL8-CXCR1/2 Axis in Cancer and Inflammatory Diseases. *Theranostics*, *7*(6), 1543–1588. https://doi.org/10.7150/thno.15625

[11] Mollica Poeta, V., Massara, M., Capucetti, A. et al. (2019). Chemokines and Chemokine Receptors: New Targets for Cancer Immunotherapy. *Frontiers in Immunology*, *10*, 379. https://doi.org/10.3389/fimmu.2019.00379

[12] Zhang, Y., Xu, L., & Peng, M. (2018). CXCR3 is a prognostic marker and a potential target for patients with solid tumors: a meta-analysis. *OncoTargets and Therapy*, *11*, 1045–1054. https://doi.org/10.2147/OTT.S157421

[13] Kazanietz, M. G., Durando, M., & Cooke, M. (2019). CXCL13 and Its Receptor CXCR5 in Cancer: Inflammation, Immune Response, and Beyond. *Frontiers in Endocrinology*, *10*, 471. https://doi.org/10.3389/fendo.2019.00471

[14] Hu, W., Liu, Y., Zhou, W. et al. (2014). CXCL16 and CXCR6 are coexpressed in human lung cancer in vivo and mediate the invasion of lung cancer cell lines in vitro. *PloS One*, *9*(6), e99056. https://doi.org/10.1371/journal.pone.0099056

[15] Xiao, G., Wang, X., Wang, J. et al. (2015). CXCL16/CXCR6 chemokine signalling mediates breast cancer progression by pERK1/2-dependent mechanisms. *Oncotarget*, *6*(16), 14165–14178. https://doi.org/10.18632/oncotarget.3690

[16] Shin, S. Y., Lee, D. H., Lee, J. et al. (2017). C-C motif chemokine receptor 1 (CCR1) is a target of the EGF-AKT-mTOR-STAT3 signalling axis in breast cancer cells. *Oncotarget*, *8*(55), 94591–94605. https://doi.org/10.18632/oncotarget.21813

[17] Lim, S. Y., Yuzhalin, A. E., Gordon-Weeks, A. N. et al. (2016). Targeting the CCL2-CCR2 signalling axis in cancer metastasis. *Oncotarget*, *7*(19), 28697–28710. https://doi.org/10.18632/oncotarget.7376

[18] Aldinucci, D., Borghese, C., & Casagrande, N. (2020). The CCL5/CCR5 Axis in Cancer Progression. *Cancers*, *12*(7), 1765. https://doi.org/10.3390/cancers12071765

[19] Kadomoto, S., Izumi, K., & Mizokami, A. (2020). The CCL20-CCR6 Axis in Cancer Progression. *International Journal of Molecular Sciences*, *21*(15), 5186. https://doi.org/10.3390/ijms21155186

[20] Tu, Z., Xiao, R., Xiong, J., Tembo, K., Deng, X., & Xiong, M. et al. (2016). CCR9 in cancer: oncogenic role and therapeutic targeting. *Journal of Hematology & Oncology*, *9*(1). https://doi.org/10.1186/s13045-016-0236-7

[21] Lin, H. Y., Sun, S. M., Lu, X. F. et al. (2017). CCR10 activation stimulates the invasion and migration of breast cancer cells through the ERK1/2/MMP-7 signalling pathway. *International Immunopharmacology*, *51*, 124–130. https://doi.org/10.1016/j.intimp.2017.07.018

[22] Gunn, M. D., Tangemann, K., Tam, C. et al. (1998). A chemokine expressed in lymphoid high endothelial venules promotes the adhesion and chemotaxis of naive T lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(1), 258–263. https://doi.org/10.1073/pnas.95.1.258

[23] Otero, C., Eisele, P. S., Schaeuble, K. et al. (2008). Distinct motifs in the chemokine receptor CCR7 regulate signal transduction, receptor trafficking and chemotaxis. *Journal of Cell Science*, *121*(Pt 16), 2759–2767. https://doi.org/10.1242/jcs.029074

[24] Mukherjee, D., & Zhao, J. (2013). The Role of chemokine receptor CXCR4 in breast cancer metastasis. *American Journal of Cancer Research*, *3*(1), 46–57.

[25] Nishihira J. (2000). Macrophage migration inhibitory factor (MIF): its essential role in the immune system and cell growth. *Journal of Interferon & Cytokine Research*, *20*(9), 751–762. https://doi.org/10.1089/10799900050151012

[26] Wu, B., Chien, E. Y. T., Mol, C. D. et al. (2010). Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science*, *330*(6007), 1066-1071. https://doi.org/10.1126/science.1194396

[27] Müller, A., Homey, B., Soto, H. et al. (2001). Involvement of chemokine receptors in breast cancer metastasis. *Nature*, *410*(6824), 50–56. https://doi.org/10.1038/35065016

[28] Minn, A. J., Kang, Y., Serganova, I. et al. (2005). Distinct organ-specific metastatic potential of individual breast cancer cells and primary tumors. *The Journal of Clinical Investigation*, *115*(1), 44–55. https://doi.org/10.1172/JCI22320

[29] Günther, K., Leier, J., Henning, G. et al. (2005). Prediction of lymph node metastasis in colorectal carcinoma by expression of chemokine receptor CCR7. *International Journal of Cancer*, *116*(5), 726–733. https://doi.org/10.1002/ijc.21123

[30] Kim, J., Mori, T., Chen, S. L. et al. (2006). Chemokine receptor CXCR4 expression in patients with melanoma and colorectal cancer liver metastases and the association with disease outcome. *Annals of Surgery*, *244*(1), 113–120. https://doi.org/10.1097/01.sla.0000217690.65909.9c

[31] Ishida, K., Iwahashi, M., Nakamori, M. et al. (2009). High CCR7 mRNA expression of cancer cells is associated with lymph node involvement in patients with esophageal squamous cell carcinoma. *International Journal of Oncology*, *34*(4), 915–922. https://doi.org/10.3892/ijo_00000217

[32] Goto, M., Yoshida, T., Yamamoto, Y. et al. (2017). CXCR4 Expression is Associated with Poor Prognosis in Patients with Esophageal Squamous Cell Carcinoma. *Annals of Surgical Oncology*, *24*(3), 832–840. https://doi.org/10.1245/s10434-015-4974-5

[33] Mashino, K., Sadanaga, N., Yamaguchi, H. et al. (2002). Expression of chemokine receptor CCR7 is associated with lymph node metastasis of gastric carcinoma. *Cancer Research*, *62*(10), 2937–2941.

[34] Nikzaban, M., Hakhamaneshi, M. S., Fakhari, S. et al. (2014). The chemokine receptor CXCR4 is associated with the staging of gastric cancer. *Advanced Biomedical Research*, *3*, 16. https://doi.org/10.4103/2277-9175.124645

[35] Xu, Y., Liu, L., Qiu, X., Liu, Z. et al. (2012). CCL21/CCR7 prevents apoptosis via the ERK pathway in human non-small cell lung cancer cells. *PloS One*, *7*(3), e33262. https://doi.org/10.1371/journal.pone.0033262

[36] Dai, X., Mao, Z., Huang, J., Xie, S., & Zhang, H. (2013). The CXCL12/CXCR4 autocrine loop increases the metastatic potential of non-small cell lung cancer in vitro. *Oncology Letters*, *5*(1), 277–282. https://doi.org/10.3892/ol.2012.960

[37] Takeuchi, H., Fujimoto, A., Tanaka, M. et al. (2004). CCL21 chemokine regulates chemokine receptor CCR7 bearing malignant melanoma cells. *Clinical Cancer Research: an official journal of the American Association for Cancer Research*, *10*(7), 2351–2358. https://doi.org/10.1158/1078-0432.ccr-03-0195

[38] Somovilla-Crespo, B., Alfonso-Pérez, M., Cuesta-Mateos, C. et al. (2013). Anti-CCR7 therapy exerts a potent anti-tumor activity in a xenograft model of human mantle cell lymphoma. *Journal of Haematology & Oncology*, *6*, 89. https://doi.org/10.1186/1756-8722-6-89

[39] Kodama, J., Hasengaowa, Kusumoto, T., Seki, N. et al. (2007). Association of CXCR4 and CCR7 chemokine receptor expression and lymph node metastasis in human cervical cancer. *Annals of Oncology: official journal of the European Society for Medical Oncology*, *18*(1), 70–76. https://doi.org/10.1093/annonc/mdl342

[40] Chi, B. J., Du, C. L., Fu, Y. F., Zhang, Y. N., & Wang, R. W. (2015). Silencing of CCR7 inhibits the growth, invasion and migration of prostate cancer cells induced by VEGFC. *International Journal of Clinical and Experimental Pathology*, *8*(10), 12533–12540.

[41] Cunningham, H. D., Shannon, L. A., Calloway, P. A. et al. (2010). Expression of the C-C chemokine receptor 7 mediates metastasis of breast cancer to the lymph nodes in mice. *Translational Oncology*, *3*(6), 354–361. https://doi.org/10.1593/tlo.10178

[42] Lapteva, N., Yang, A. G., Sanders, D. E. et al. (2005). CXCR4 knockdown by small interfering RNA abrogates breast tumor growth in vivo. *Cancer Gene Therapy*, *12*(1), 84–89. https://doi.org/10.1038/sj.cgt.7700770

[43] Gan, Y., Shi, C., Inge, L., Hibner, M., Balducci, J., & Huang, Y. (2010). Differential roles of ERK and Akt pathways in regulation of EGFR-mediated signalling and motility in prostate cancer cells. *Oncogene*, *29*(35), 4947–4958. https://doi.org/10.1038/onc.2010.240

[44] Mishan, M. A., Ahmadiankia, N., & Bahrami, A. R. (2016). CXCR4 and CCR7: Two eligible targets in targeted cancer therapy. *Cell Biology International*, *40*(9), 955–967. https://doi.org/10.1002/cbin.10631

[45] Ma, S., Li, Q., & Pan, F. (2015). CXCR4 promotes GSK3β expression in pancreatic cancer cells via the Akt pathway. *International Journal of Clinical Oncology*, *20*(3), 525–530. https://doi.org/10.1007/s10147-014-0740-0

[46] Tutunea-Fatan, E., Majumder, M., Xin, X., & Lala, P. K. (2015). The role of CCL21/CCR7 chemokine axis in breast cancer-induced lymphangiogenesis. *Molecular Cancer*, *14*, 35. https://doi.org/10.1186/s12943-015-0306-4

[47] Sweat, R. S., Sloas, D. C., & Murfee, W. L. (2014). VEGF-C induces lymphangiogenesis and angiogenesis in the rat mesentery culture model. *Microcirculation (New York, N.Y.: 1994)*, *21*(6), 532–540. https://doi.org/10.1111/micc.12132

[48] Zhuo, W., Jia, L., Song, N., Lu, X. A. et al. (2012). The CXCL12-CXCR4 chemokine pathway: a novel axis regulates lymphangiogenesis. *Clinical Cancer Research: an official journal of the American Association for Cancer Research*, *18*(19), 5387–5398. https://doi.org/10.1158/1078-0432.CCR-12-0708

[49] Li, F., Zou, Z., Suo, N., Zhang, Z., Wan, F. et al. (2014). CCL21/CCR7 axis activating chemotaxis accompanied with epithelial-mesenchymal transition in human breast carcinoma. *Medical Oncology (Northwood, London, England)*, *31*(9), 180. https://doi.org/10.1007/s12032-014-0180-8

[50] Li, X., Ma, Q., Xu, Q., Liu, H., Lei, J. et al. (2012). SDF-1/CXCR4 signalling induces pancreatic cancer cell invasion and epithelial-mesenchymal transition in vitro through non-canonical activation of Hedgehog pathway. *Cancer Letters*, *322*(2), 169–176. https://doi.org/10.1016/j.canlet.2012.02.035

[51] Wheelock, M. J., Shintani, Y., Maeda, M. et al. (2008). Cadherin switching. *Journal of Cell Science*, *121*(Pt 6), 727–735. https://doi.org/10.1242/jcs.000455

[52] Bolós, V., Peinado, H., Pérez-Moreno, M. A., Fraga, M. F. et al. (2003). The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. *Journal of Cell Science*, *116*(Pt 3), 499–511. https://doi.org/10.1242/jcs.00224

[53] Hu, T. H., Yao, Y., Yu, S., Han, L. L., Wang, W. J. et al. (2014). SDF-1/CXCR4 promotes epithelial-mesenchymal transition and progression of colorectal cancer by activation of the Wnt/β-catenin signalling pathway. *Cancer Letters*, *354*(2), 417–426. https://doi.org/10.1016/j.canlet.2014.08.012

[54] Redondo-Muñoz, J., Ugarte-Berzal, E., Terol, M. et al. (2010). Matrix Metalloproteinase-9 Promotes Chronic Lymphocytic Leukemia B Cell Survival through Its Hemopexin Domain. *Cancer Cell*, *17*(2), 160-172. https://doi.org/10.1016/j.ccr.2009.12.044

[55] Lin, C. Y., Tsai, P. H., Kandaswami, C. C., Lee, P. P. et al. (2011). Matrix metalloproteinase-9 cooperates with transcription factor Snail to induce epithelial-mesenchymal transition. *Cancer Science*, *102*(4), 815–827. https://doi.org/10.1111/j.1349-7006.2011.01861.x

[56] Teng, F., Tian, W., Wang, Y., Zhang, Y., Guo, F., & Zhao, J. et al. (2016). Cancer-associated fibroblasts promote the progression of endometrial cancer via the SDF-1/CXCR4 axis. *Journal of Hematology & Oncology*, *9*(1). https://doi.org/10.1186/s13045-015-0231-4

[57] Weaver A. M. (2006). Invadopodia: specialized cell structures for cancer invasion. *Clinical & Experimental Metastasis*, *23*(2), 97–105. https://doi.org/10.1007/s10585-006-9014-1

[58] Bleul, C. C., Fuhlbrigge, R. C., Casasnovas, J. M. et al. (1996). A highly efficacious lymphocyte chemoattractant, stromal cell-derived factor 1 (SDF-1). *The Journal of Experimental Medicine*, *184*(3), 1101–1109. https://doi.org/10.1084/jem.184.3.1101

[59] Verschueren, H., Van der Taelen, I., Dewit, J. et al. (1994). Metastatic competence of BW5147 T-lymphoma cell lines is correlated with in vitro invasiveness, motility and F-actin content. *Journal of Leukocyte Biology*, *55*(4), 552–556. https://doi.org/10.1002/jlb.55.4.552

[60] Shields, J. D., Fleury, M. E., Yong, C. et al. (2007). Autologous chemotaxis as a mechanism of tumor cell homing to lymphatics via interstitial flow and autocrine CCR7 signalling. *Cancer Cell*, *11*(6), 526–538. https://doi.org/10.1016/j.ccr.2007.04.020

[61] Shah, A. D., Bouchard, M. J., & Shieh, A. C. (2015). Interstitial Fluid Flow Increases Hepatocellular Carcinoma Cell Invasion through CXCR4/CXCL12 and MEK/ERK Signalling. *PloS One*, *10*(11), e0142337. https://doi.org/10.1371/journal.pone.0142337

[62] Hull-Ryde, E. A., Porter, M. A., Fowler, K. A. et al. (2018). Identification of Cosalane as an Inhibitor of Human and Murine CC-Chemokine Receptor 7 Signaling via a High-Throughput Screen. *SLAS discovery: Advancing Life Sciences R & D*, *23*(10), 1083–1091. https://doi.org/10.1177/2472555218780917

[63] Jaeger, K., Bruenle, S., Weinert, T. et al. (2019). Structural Basis for Allosteric Ligand Recognition in the Human CC Chemokine Receptor 7. *Cell*, *178*(5), 1222-1230.e10. https://doi.org/10.1016/j.cell.2019.07.028

[64] Cushman, M., Golebiewski, W. M., McMahon, J. B. et al. (1994). Design, synthesis, and biological evaluation of cosalane, a novel anti-HIV agent which inhibits multiple features of virus reproduction. *Journal of Medicinal Chemistry*, *37*(19), 3040–3050. https://doi.org/10.1021/jm00045a008.

[65] Kuchimanchi, K. R., Udata, C., Johnston, T. P., & Mitra, A. K. (2000). Pharmacokinetics, biliary excretion, and tissue distribution of novel anti-HIV agents, cosalane and dihydrocosalane, in Sprague-Dawley rats. *Drug Metabolism and Disposition: The Biological Fate of Chemicals*, *28*(4), 403–408.

[66] Dwyer, M. P., Yu, Y., Chao, J. et al. (2006). Discovery of 2-hydroxy-N,N-dimethyl-3-{2-[[(R)-1-(5- methylfuran-2-yl)propyl]amino]-3,4-dioxocyclobut-1-enylamino}benzamide (SCH 527123): a potent, orally bioavailable CXCR2/CXCR1 receptor antagonist. *Journal of Medicinal Chemistry*, *49*(26), 7603–7606. https://doi.org/10.1021/jm0609622

[67] Alfonso-Pérez, M., López-Giral, S., Quintana, N. E. et al. (2006). Anti-CCR7 monoclonal antibodies as a novel tool for the treatment of chronic lymphocyte leukemia. *Journal of Leukocyte Biology*, *79*(6), 1157–1165. https://doi.org/10.1189/jlb.1105623

[68] Fang, L., Lee, V. C., Cha, E., Zhang, H., & Hwang, S. T. (2008). CCR7 regulates B16 murine melanoma cell tumorigenesis in skin. *Journal of Leukocyte Biology*, *84*(4), 965–972. https://doi.org/10.1189/jlb.1107776

[69] Wagstaff A. J. (2009). Plerixafor: in patients with non-Hodgkin's lymphoma or multiple myeloma. *Drugs*, *69*(3), 319–326. https://doi.org/10.2165/00003495-200969030-00007

[70] De Clercq E. (2003). The bicyclam AMD3100 story. *Nature reviews. Drug Discovery*, *2*(7), 581–587. https://doi.org/10.1038/nrd1134

[71] Smith, M. C., Luker, K. E., Garbow, J. R. et al. (2004). CXCR4 regulates growth of both primary and metastatic breast cancer. *Cancer Research*, *64*(23), 8604–8612. https://doi.org/10.1158/0008-5472.CAN-04-1844

.

[72] Fricker, S. P., Anastassov, V., Cox, J. et al. (2006). Characterization of the molecular pharmacology of AMD3100: a specific antagonist of the G-protein coupled chemokine receptor, CXCR4. *Biochemical Pharmacology*, *72*(5), 588–596. https://doi.org/10.1016/j.bcp.2006.05.010

[73] Ling, X., Spaeth, E., Chen, Y., Shi, Y., Zhang, W. et al. 2013). The CXCR4 antagonist AMD3465 regulates oncogenic signalling and invasiveness in vitro and prevents breast cancer growth and metastasis in vivo. *PloS One*, *8*(3), e58426. https://doi.org/10.1371/journal.pone.0058426

[74] Bodart, V., Anastassov, V., Darkes, M. C., Idzan, S. R. et al. (2009). Pharmacology of AMD3465: a small molecule antagonist of the chemokine receptor CXCR4. *Biochemical Pharmacology*, *78*(8), 993–1000. https://doi.org/10.1016/j.bcp.2009.06.010

[75] Debnath, B., Xu, S., Grande, F., Garofalo, A., & Neamati, N. (2013). Small molecule inhibitors of CXCR4. *Theranostics*, *3*(1), 47–75. https://doi.org/10.7150/thno.5376

[76] Zhan, W., Liang, Z., Zhu, A., Kurtkaya, S., Shim, H. et al. (2007). Discovery of small molecule CXCR4 antagonists. *Journal of Medicinal Chemistry*, *50*(23), 5655–5664. https://doi.org/10.1021/jm070679i

[77] Van Hout, A., D'huys, T., Oeyen, M. et al. (2017). Comparison of cell-based assays for the identification and evaluation of competitive CXCR4 inhibitors. *PloS One*, *12*(4), e0176057. https://doi.org/10.1371/journal.pone.0176057

[78] Li, S. H., Dong, W. C., Fan, L., & Wang, G. S. (2016). Suppression of chronic lymphocytic leukemia progression by CXCR4 inhibitor WZ811. *American Journal of Translational Research*, *8*(9), 3812–3821.

[79] Skerlj, R., Bridger, G., McEachern, E. et al. (2011). Synthesis and SAR of novel CXCR4 antagonists that are potent inhibitors of T tropic (X4) HIV-1 replication. *Bioorganic & Medicinal Chemistry Letters*, *21*(1), 262–266. https://doi.org/10.1016/j.bmcl.2010.11.023

[80] Mooring, S. R., Liu, J., Liang, Z. et al. (2013). Benzenesulfonamides: a unique class of chemokine receptor type 4 inhibitors. *ChemMedChem*, *8*(4), 622–632. https://doi.org/10.1002/cmdc.201200582

[81] Carter, P. H., Brown, G. D., Cherney, R. J. et al. (2015). Discovery of a Potent and Orally Bioavailable Dual Antagonist of CC Chemokine Receptors 2 and 5. *ACS Medicinal Chemistry Letters*, *6*(4), 439–444. https://doi.org/10.1021/ml500505q

[82] Zhao Q. (2010). Dual targeting of CCR2 and CCR5: therapeutic potential for immunologic and cardiovascular diseases. Journal of Leukocyte Biology, 88(1), 41–55. https://doi.org/10.1189/jlb.1009671

[83] Kimura, M., Kawahito, Y., Hamaguchi, M. et al. (2007). SKL-2841, a dual antagonist of MCP-1 and MIP-1 beta, prevents bleomycin-induced skin sclerosis in mice. *Biomedicine & Pharmacotherapy*, *61*(4), 222-228. https://doi.org/10.1016/j.biopha.2006.10.002

[84] Pease, J. E., & Horuk, R. (2010). Small molecule antagonists of chemokine receptors--is promiscuity a virtue? *Current Topics in Medicinal Chemistry*, *10*(13), 1351–1358. https://doi.org/10.2174/156802610791561228

[85] Stacer, A. C., Fenner, J., Cavnar, S. P. et al. (2016). Endothelial CXCR7 regulates breast cancer metastasis. *Oncogene*, *35*(13), 1716–1724. https://doi.org/10.1038/onc.2015.236

[84] Peters J. U. (2013). Polypharmacology - foe or friend? *Journal of Medicinal Chemistry*, *56*(22), 8955–8971. https://doi.org/10.1021/jm400856t

[85] Broekman, F., Giovannetti, E., & Peters, G. J. (2011). Tyrosine kinase inhibitors: Multi-targeted or single-targeted? *World Journal of Clinical Oncology*, *2*(2), 80–93. https://doi.org/10.5306/wjco.v2.i2.80

[86] Boran, A. D., & Iyengar, R. (2010). Systems approaches to polypharmacology and drug discovery. *Current Opinion in Drug Discovery & Development*, *13*(3), 297–309.

[87] Morphy, R., & Rankovic, Z. (2005). Designed multiple ligands. An emerging drug discovery paradigm. *Journal of Medicinal Chemistry*, *48*(21), 6523–6543. https://doi.org/10.1021/jm058225d

[88] Morphy, R., & Rankovic, Z. (2009). Designing multiple ligands - medicinal chemistry strategies and challenges. *Current Pharmaceutical Design*, *15*(6), 587–600. https://doi.org/10.2174/138161209787315594

[89] Wood, P. M., Woo, L. W., Thomas, M. P. et al. (2011). Aromatase and dual aromatase-steroid sulfatase inhibitors from the letrozole and vorozole templates. *ChemMedChem*, *6*(8), 1423–1438. https://doi.org/10.1002/cmdc.201100145

[90] Murugesan, N., Tellew, J. E., Gu, Z., Kunst, B. L. et al. (2002). Discovery of N-isoxazolyl biphenylsulfonamides as potent dual angiotensin II and endothelin A receptor antagonists. *Journal of Medicinal Chemistry*, *45*(18), 3829–3835. https://doi.org/10.1021/jm020138n

[91] Rees, D. C., Congreve, M., Murray, C. W., & Carr, R. (2004). Fragment-based lead discovery. *Nature reviews. Drug Discovery*, *3*(8), 660–672. https://doi.org/10.1038/nrd1467

[92] Bottegoni, G., Favia, A. D., Recanatini, M., & Cavalli, A. (2012). The role of fragment-based and computational methods in polypharmacology. *Drug Discovery today*, *17*(1-2), 23–34. https://doi.org/10.1016/j.drudis.2011.08.002

[93] Morphy, R., & Rankovic, Z. (2007). Fragments, network biology and designing multiple ligands. *Drug Discovery Today*, *12*(3-4), 156–160. https://doi.org/10.1016/j.drudis.2006.12.006

[94] Guo, C., Hou, X., Dong, L. et al. (2009). Structure-based design of novel human Pin1 inhibitors (I). *Bioorganic & Medicinal Chemistry Letters*, *19*(19), 5613–5616. https://doi.org/10.1016/j.bmcl.2009.08.034

[95] Potter, A., Oldfield, V., Nunns, C. et al. (2010). Discovery of cell-active phenyl-imidazole Pin1 inhibitors by structure-guided fragment evolution. *Bioorganic & Medicinal Chemistry Letters*, *20*(22), 6483–6488. https://doi.org/10.1016/j.bmcl.2010.09.063

[97] Frett, B., Carlomagno, F., Moccia, M. L. et al. (2015). Fragment-Based Discovery of a Dual pan-RET/VEGFR2 Kinase Inhibitor Optimized for Single-Agent Polypharmacology. *Angewandte Chemie (International ed. in English)*, *54*(30), 8717–8721. https://doi.org/10.1002/anie.201501104

[98] Klebe G. (2006). Virtual ligand screening: strategies, perspectives and limitations. *Drug Discovery today*, *11*(13-14), 580–594. https://doi.org/10.1016/j.drudis.2006.05.012

[99] Wei, D., Jiang, X., Zhou, L. et al. (2008). Discovery of multitarget inhibitors by combining molecular docking with common pharmacophore matching. *Journal of Medicinal Chemistry*, *51*(24), 7882–7888. https://doi.org/10.1021/jm8010096

[100] Bottegoni, G., Veronesi, M., Bisignano, P. et al. (2016). Development and Application of a Virtual Screening Protocol for the Identification of Multitarget Fragments. *ChemMedChem*, *11*(12), 1259–1263. https://doi.org/10.1002/cmdc.201500521

[101] Vyas, V. K., Ukawala, R. D., Ghate, M., & Chintha, C. (2012). Homology modeling a fast tool for drug discovery: current perspectives. *Indian Journal of Pharmaceutical Sciences*, *74*(1), 1–17. https://doi.org/10.4103/0250-474X.102537

[102] Liu, P., Du, Y., Song, L., Shen, J., & Li, Q. (2015). Novel, potent, selective and cellular active ABC type PTP1B inhibitors containing (methanesulfonyl-phenyl-amino)-acetic acid methyl ester phosphotyrosine mimetic. *Bioorganic & Medicinal Chemistry*, *23*(21), 7079–7088. https://doi.org/10.1016/j.bmc.2015.09.024

[103] Hogan, P., & Cox, B. (2009). Aqueous Process Chemistry: The Preparation of Aryl Sulfonyl Chlorides. *Organic Process Research & Development*, *13*(5), 875-879. https://doi.org/10.1021/op9000862

[104] Chen, J., Zhang, Y., Yang, L. et al. (2007). A practical palladium catalyzed dehalogenation of aryl halides and α-haloketones. *Tetrahedron*, *63*(20), 4266-4270. https://doi.org/10.1016/j.tet.2007.03.061

[105] Saidykhan, A., Ebert, J., Ally, H. et al. (2017). The scope and regioselectivity of intramolecular N-C rearrangements of orthogonally protected sulfonamides, including cyclization to saccharin derivatives. *Tetrahedron Letters*, *58*(32), 3089-3091. https://doi.org/10.1016/j.tetlet.2017.06.053

[106] Congreve, M., Andrews, S. P., Doré, A. S. et al. (2012). Discovery of 1,2,4-triazine derivatives as adenosine A(2A) antagonists using structure based drug design. *Journal of Medicinal Chemistry*, *55*(5), 1898–1903. https://doi.org/10.1021/jm201376w

[107] Kataria, R., & Khatkar, A. (2019). Molecular docking, synthesis, kinetics study, structure-activity relationship and ADMET analysis of morin analogous as *Helicobacter pylori* urease inhibitors. *BMC Chemistry*, *13*(1), 45. https://doi.org/10.1186/s13065-019-0562-2

[108] Camacho, C., Coulouris, G., Avagyan, V. et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics*, *10*, 421. https://doi.org/10.1186/1471-2105-10-421

[109] Oswald, C., Rappas, M., Kean, J. et al. (2016). Intracellular allosteric antagonism of the CCR9 receptor. *Nature*, *540*(7633), 462–465. https://doi.org/10.1038/nature20606

[110] Waterhouse, A., Bertoni, M., Bienert, S. et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research*, *46*(W1), W296–W303. https://doi.org/10.1093/nar/gky427

[111] Williams, C. J., Headd, J. J., Moriarty, N. W. et al. (2018). MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science: a Publication of the Protein Society*, *27*(1), 293–315. https://doi.org/10.1002/pro.3330

[112] Chan, H., Li, Y., Dahoun, T., Vogel, H., & Yuan, S. (2019). New Binding Sites, New Opportunities for GPCR Drug Discovery. *Trends in Biochemical Sciences*, *44*(4), 312–330. https://doi.org/10.1016/j.tibs.2018.11.011

[113] Hechler, B., Nonne, C., Roh, E. J. et al. (2006). MRS2500 [2-iodo-N6-methyl-(N)-methanocarba-2'-deoxyadenosine-3',5'-bisphosphate], a potent, selective, and stable antagonist of the platelet P2Y1 receptor with strong antithrombotic activity in mice. *The Journal of Pharmacology and Experimental Therapeutics*, *316*(2), 556–563. https://doi.org/10.1124/jpet.105.094037

[114] Croy, C. H., Schober, D. A., Xiao, H. et al. (2014). Characterization of the novel positive allosteric modulator, LY2119620, at the muscarinic M(2) and M(4) receptors. *Molecular Pharmacology*, *86*(1), 106–115. https://doi.org/10.1124/mol.114.091751

[115] *Molecular Operating Environment (MOE)*, 2015.10; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2015.

[116] Burg, J., Ingram, J., Venkatakrishnan, A. et al. (2015). Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor. *Science*, *347*(6226), 1113-1117. https://doi.org/10.1126/science.aaa5026

[117] Walters, M. J., Wang, Y., Lai, N. et al. (2010). Characterization of CCX282-B, an orally bioavailable antagonist of the CCR9 chemokine receptor, for treatment of inflammatory bowel disease. *The Journal of Pharmacology and Experimental Therapeutics*, *335*(1), 61–69. https://doi.org/10.1124/jpet.110.169714

[118] Ott, T. R., Pahuja, A., Nickolls, S. A. et al. (2004). Identification of CC chemokine receptor 7 residues important for receptor activation. *The Journal of Biological Chemistry*, *279*(41), 42383–42392. https://doi.org/10.1074/jbc.M401097200

[119] Hanlon, C. D., & Andrew, D. J. (2015). Outside-in signalling--a brief review of GPCR signalling with a focus on the Drosophila GPCR family. *Journal of Cell Science*, *128*(19), 3533–3542. https://doi.org/10.1242/jcs.175158

[120] Wang, L., Zhao, X. Y., Zhu, J. S. et al. (2018). CCR7 regulates ANO6 to promote migration of pancreatic ductal adenocarcinoma cells via the ERK signalling pathway. *Oncology Letters*, *16*(2), 2599–2605. https://doi.org/10.3892/ol.2018.8962

[121] Wang, Z., Ma, Q., Liu, Q. et al. (2008). Blockade of SDF-1/CXCR4 signalling inhibits pancreatic cancer progression in vitro via inactivation of canonical Wnt pathway. *British Journal of Cancer*, *99*(10), 1695–1703. https://doi.org/10.1038/sj.bjc.6604745

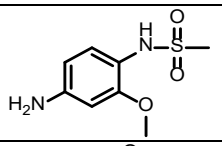
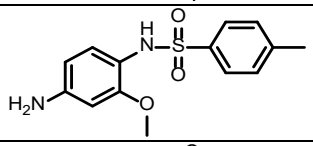
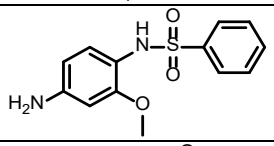
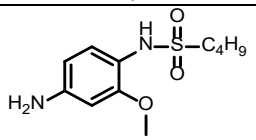
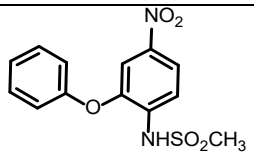
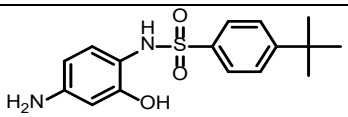
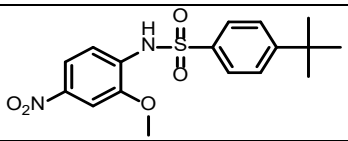
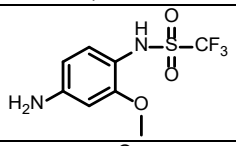
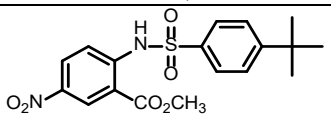
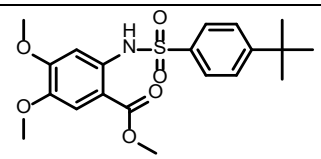
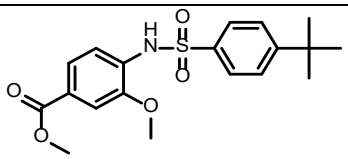
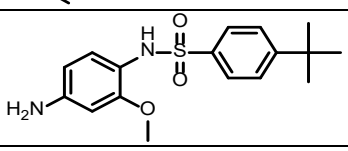
[122] Akekawatchai, C., Holland, J. D., Kochetkova, M. et al. (2005). Transactivation of CXCR4 by the insulin-like growth factor-1 receptor (IGF-1R) in human MDA-MB-231 breast cancer epithelial cells. *The Journal of Biological Chemistry*, *280*(48), 39701–39708. https://doi.org/10.1074/jbc.M509829200

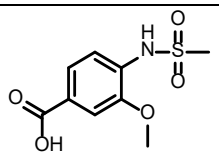
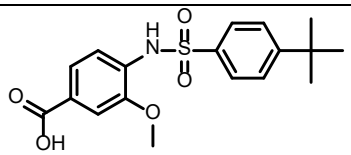
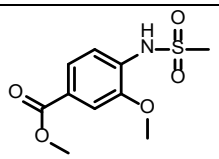
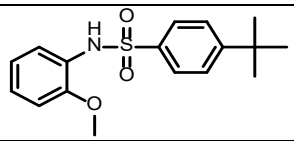
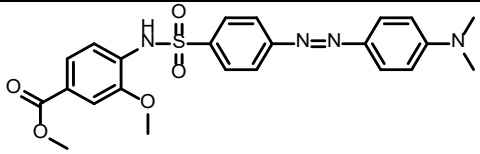
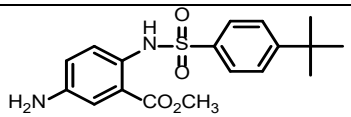
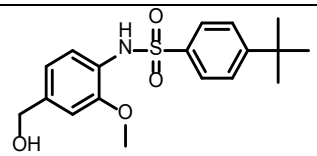
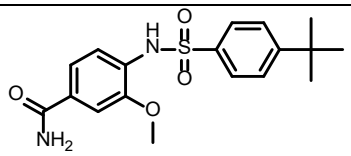
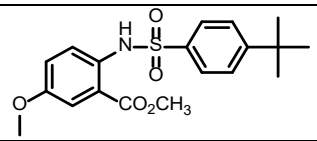
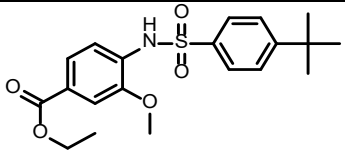
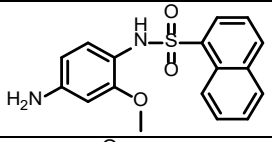
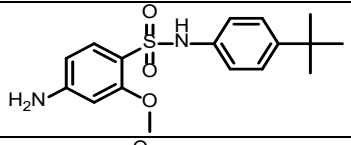
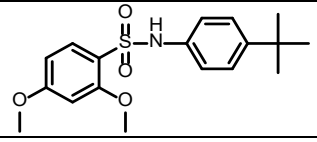
[123] Xu, R., Hong, J., Morse, C. L., & Pike, V. W. (2010). Synthesis, structure-affinity relationships, and radiolabelling of selective high-affinity 5-HT4 receptor ligands as prospective imaging probes for positron emission tomography. *Journal of Medicinal Chemistry*, *53*(19), 7035–7047. https://doi.org/10.1021/jm100668r

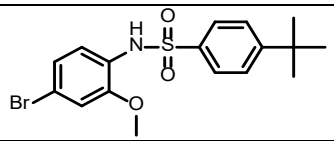
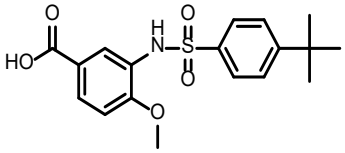
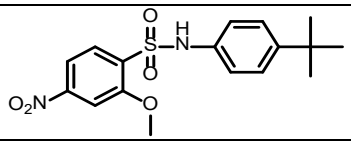
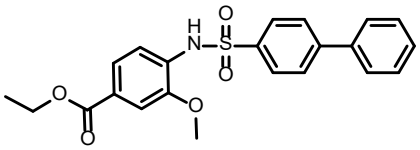
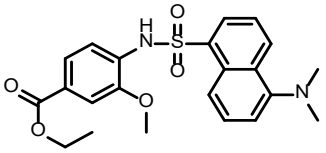
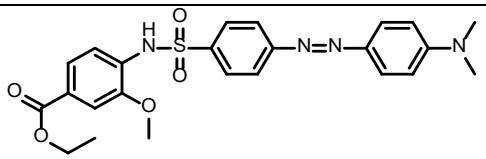
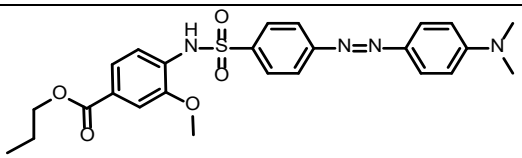
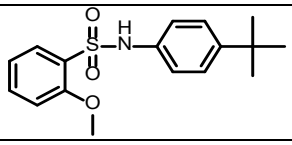
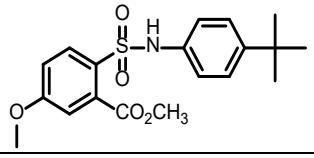
[124] Bhuiyan, M., Zhu, K., Jensen, P., & Try, A. (2010). Synthesis of Symmetric Diester-Functionalised Tröger's Base Analogues. *European Journal of Organic Chemistry*, *2010*(24), 4662-4670. https://doi.org/10.1002/ejoc.201000086

[125] Zhao, H., Yang, X., Lei, H. et al. 2019). Cu-mediated selective bromination of aniline derivatives and preliminary mechanism study. *Synthetic Communications*, *49*(11), 1406-1415. https://doi.org/10.1080/00397911.2019.1598558

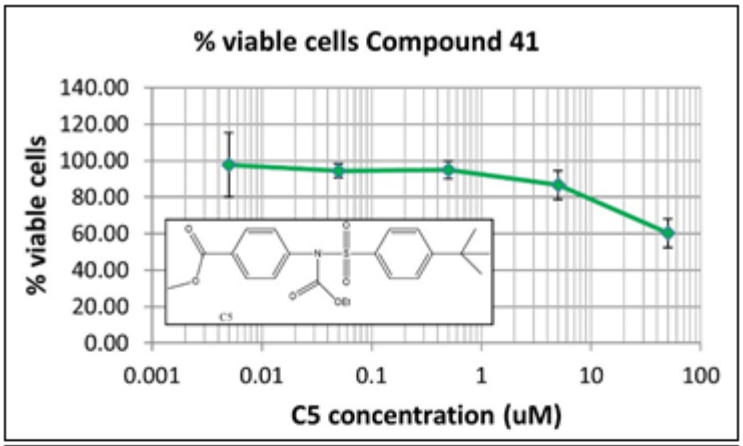
Appendix 1: Structures and IC_{50} (μM) of 34 CCR7 antagonists used in molecular docking

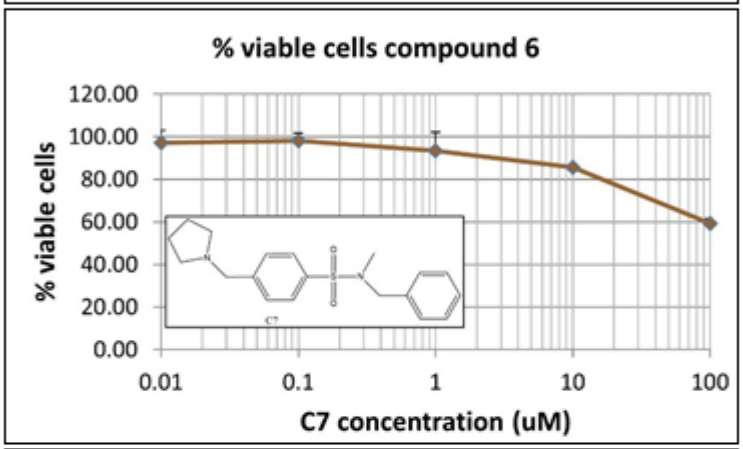
Compound name	Structure	IC_{50} (μM)
ICT13029		11.6
ICT13034		11.6
ICT13035		12.1
ICT13036		15.5
ICT13052		15.1
ICT13055		16.3
ICT13062		9.5
ICT13063		1.7
ICT13067		3.4
ICT13068		11.5
ICT13069		0.19
ICT13070		5.7

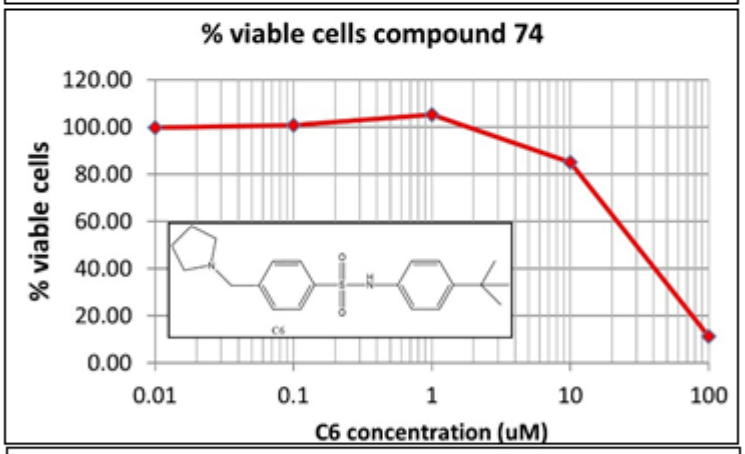
ICT13071		74.4
ICT13072		6.8
ICT13073		37.7
ICT13078		31.9
ICT13085		4.1
ICT13089		0.8
ICT13091		9.5
ICT13092		0.8
ICT13108		2.9
ICT13113		0.24
ICT13119		12.6
ICT13120		0.7
ICT13121		4.5

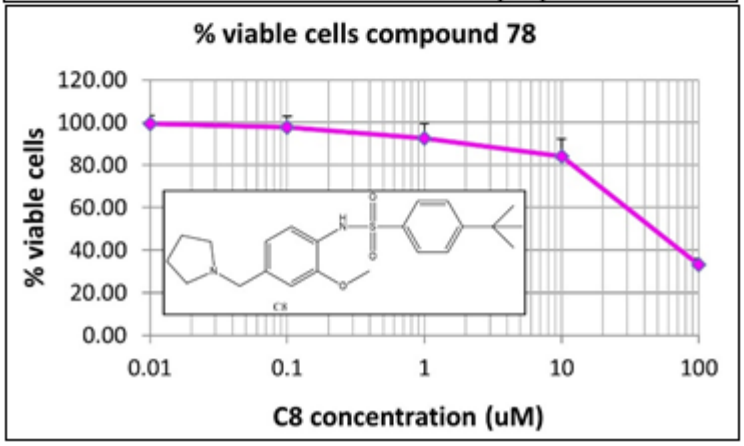
ICT13122		1
ICT13123		32.1
ICT13124		0.1
ICT13126		5.2
ICT13130		0.05
ICT13114		4.8
ICT13144		6.2
ICT13145		26.6
ICT13147		3.5

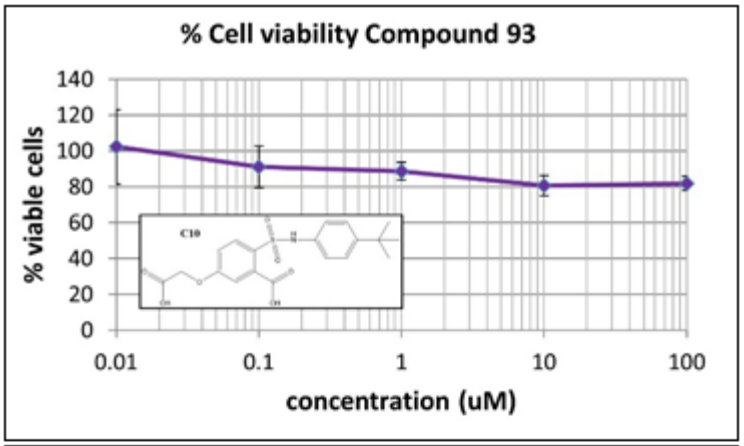
Appendix 2: MTT curves

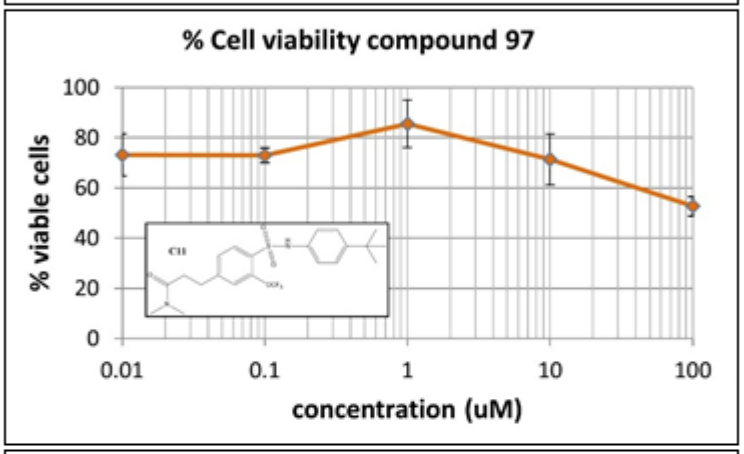

% viable cells Compound 41
% viable cells
140.00
120.00
100.00
80.00
60.00
40.00
20.00
0.00
0.001
0.01
0.1
1
10
100
C5 concentration (uM)

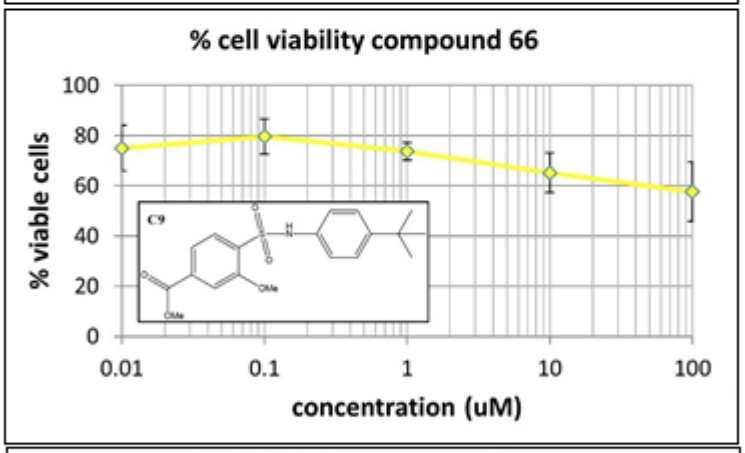

% viable cells compound 6
% viable cells
120.00
100.00
80.00
60.00
40.00
20.00
0.00
0.01
0.1
1
10
100
C7 concentration (uM)

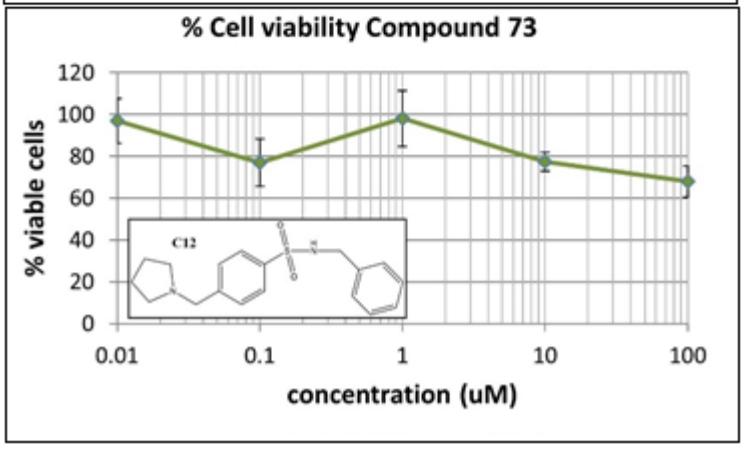

% viable cells compound 74
% viable cells
120.00
100.00
80.00
60.00
40.00
20.00
0.00
0.01
0.1
1
10
100
C6 concentration (uM)


% viable cells compound 78
% viable cells
120.00
100.00
80.00
60.00
40.00
20.00
0.00
0.01
0.1
1
10
100
C8 concentration (uM)


% Cell viability Compound 93
% viable cells
140
120
100
80
60
40
20
0
0.01
0.1
1
10
100
concentration (uM)
C10


% Cell viability compound 97
% viable cells
100
80
60
40
20
0
0.01
0.1
1
10
100
concentration (uM)
C11


% cell viability compound 66
% viable cells
100
80
60
40
20
0
0.01
0.1
1
10
100
concentration (uM)
C9


% Cell viability Compound 73
% viable cells
120
100
80
60
40
20
0
0.01
0.1
1
10
100
concentration (uM)
C12