University of Louisville

ThinkIR: The University of Louisville's Institutional Repository

Electronic Theses and Dissertations

12-2011

Structure-based design of inhibitors of CXCR4.

Jason Bradley Meier 1980-
University of Louisville

Follow this and additional works at: <https://ir.library.louisville.edu/etd>

Recommended Citation

Meier, Jason Bradley 1980-, "Structure-based design of inhibitors of CXCR4." (2011). *Electronic Theses and Dissertations*. Paper 962.
<https://doi.org/10.18297/etd/962>

This Doctoral Dissertation is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact thinkir@louisville.edu.

STRUCTURE-BASED DESIGN OF INHIBITORS OF CXCR4

By

Jason Bradley Meier

B.S., University of Louisville, 2005

A Dissertation

Submitted to the Faculty of the

School of Medicine of the University of Louisville

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

Department of Biochemistry and Molecular Biology

University of Louisville

Louisville, Kentucky

December 2011

Copyright 2010 by Jason Bradley Meier

All rights Reserved

STRUCTURE-BASED DESIGN OF INHIBITORS OF CXCR4

By

Jason Bradley Meier
B.S., University of Louisville, 2005

A Dissertation Approved on

11/17/11

DATE

By the following Dissertation Committee

John O. Trent, Dissertation Director

Thomas C. Geoghegan

Robert Mitchell

William Dean

B. Haribabu

ACKNOWLEDGEMENTS

I would like to thank my mentor, Dr. John O. Trent, for his guidance, patience and gentle prodding which helped me to bring this thesis to a close.

I would also like to thank my committee, Dr. William Dean, Dr. Thomas Geoghegan, Dr. Haribabu Bodduluri, and Dr. Robert Mitchell for their advice and assistance.

I would especially like to thank my wife, Renate, for her understanding, patience and her love and my son Noah for his unconditional love which has brought a new light into my life.

I would like to thank my lab mates, Dr. M. Clarke Miller and Dr. Indira Chaudhuri for their assistance and for creating a welcoming, family-like lab environment.

Finally, I would like to thank my mother and father; Barry and Sheryl Meier for helping me become the man that I am today.

ABSTRACT

STRUCTURE-BASED DESIGN OF INHIBITORS OF CXCR4

Jason Bradley Meier
November 17, 2011

Metastasis is a complex process requiring directed migration of metastatic cells to favorable microenvironments. Increased CXCR4 expression has been implicated in more invasive, aggressive and metastatic tumor phenotypes and poor patient survival in twenty-three forms of cancer. CXCR4 has been linked to cancer metastasis and CXCR4 expression on the cell surface of tumor cells has been linked to increased migration and homing of neoplastic cells to sites where stromal cells express the chemokine CXCL12 such as the lung and bone marrow. In this dissertation, we will utilize structure based drug design to identify inhibitors of CXCR4 targeting the extracellular surface of the receptor, as well as the intracellular interface between the GPCR and G-protein. Our screens of the extracellular surface identified one compound, ECLVS14, which inhibits chemotaxis with an IC₅₀ value of 5 μ M, and is highly selective for CXCR4 without significant cytotoxicity. Subsequent QSAR analysis of the structure of this inhibitor reveals the importance of the 1-[bis (phenylmethyl) amino] methyl moiety and the fact that electronegative modifications of the terminal benzene enhance activity. Subsequent Molecular dynamics simulations of the compound in complex with CXCR4 reveal that the compound induces significant modifications

of the receptor structure. Our intracellular screens represent a novel screening strategy targeting the intracellular region of CXCR4 interacting with Gai, which identified ten compounds selectively inhibiting CXCR4 with IC₅₀ values of 10 μ M or less. Three of the most active compounds from the extracellular and intracellular screens were tested in an *in vivo* anti-metastatic animal model, successfully demonstrating the anti-metastatic activity of these compounds. In total this work demonstrates that structure based drug design utilizing *in silico* analysis in combination with *in vitro* and *in vivo* testing can be utilized to develop novel lead compounds which can function as anti-metastatics.

TABLE OF CONTENTS

	PAGE #
ACKNOWLEDGEMENTS	iii
ABSTRACT	iv
LIST OF TABLES	vii
LIST OF FIGURES	ix
CHAPTER I: Introduction	1
CHAPTER II: Structure based drug design of extracellular loop targeted compounds by quantitative structure activity relationship analysis and mechanistic studies.	71
CHAPTER III: Discovery of Intracellular Loop Targeted Compounds by Virtual Screening	207
CHAPTER IV: Validation of Inhibitor Compounds Identified by Virtual Screening	322
CHAPTER V: Summary	465
REFERENCES:	471
CURRICULUM VITAE	490

LIST OF TABLES

Table #	Title	Page #
Table 1	GPCR Classifications.	18
Table 2	Subclassifications of Class A GPCRs.	21
Table 3	Chemokines are classified by the conformation of a conserved amino-terminal cysteine residue motif.	23
Table 4	GPCR structures serve as targets for numerous pharmaceuticals.	40
Table 5	Protomol generated for docking of ECLVS and EVS14 compounds into native structure of CXCR4.	88
Table 6	ECLVS compounds	94
Table 7	EVS14 compounds	120
Table 8	Table of ECLVS14 and EVS14 compound Chemotaxis activity.	124
Table 9	EVS14-2 derivative compounds	149
Table 10	Table of EVS14-2 and EVS14-2-AD compound Chemotaxis activity	151
Table 11	ECLVS compounds modulate the distance between the extracellular termini of the transmembrane helices of CXCR4.	196
Table 12	ECLVS compounds modulate the distance between the intracellular termini of the transmembrane helices of CXCR4.	199
Table 13	Residues of CXCR4 and Gai used in Protomol development in Surflex.	224
Table 14	Virtual screening identifies thirty-seven commercially available compounds targeting the interface of CXCR4 and Gai.	229
Table 15	Classification of compounds with anti-CXCR4 activity at 10 μ M or less.	255
Table 16	ICLGa induced conformations adopt native interhelical distances at the extracellular surface of CXCR4	297
Table 17	ICLGa induced conformations adopt native interhelical distances at the intracellular surface of CXCR4.	300
Table 18	ICLGa28 and ICLGa29 alter the position of TM7 at the intracellular surface of CXCR4.	303
Table 19	Hydrogen bonds formed by the side chains of TM7 in anti-CXCR4 compound bond molecular dynamics simulations.	306

Table 20	Hydrogen bonds formed by the side chains of TM1 in anti-CXCR4 compound bond molecular dynamics simulations.	309
Table 21	Hydrogen bonds formed by the side chains of TM3 in anti-CXCR4 compound bond molecular dynamics simulations.	311
Table 22	Hydrogen bonds formed by the side chains of TM4 in anti-CXCR4 compound bond molecular dynamics simulations.	313
Table 23	Classification of ICLGa Compounds with anti-CXCR4 activity of 10 μ M or less.	317
Table 24	Extracellular Residues of CXCR4 used in protomol development in Surflex.	336
Table 25	Residues identified for SDM.	339
Table 26	CXCR4 ECL SDM plasmids.	341
Table 27	CXCR4 ICL SDM plasmids.	343
Table 28	Interhelical distances at the extracellular surface for the ICL3 mutants.	374
Table 29	Interhelical distances at the intracellular surface for the ICL3 mutants.	376
Table 30	Interhelical distances at the extracellular surface for the helix 8 mutants.	383
Table 31	Interhelical distances at the intracellular surface for the helix 8 mutants.	386
Table 32	Interhelical distances at the extracellular surface for the distal C terminus mutants.	392
Table 33	Interhelical distances at the intracellular surface for the distal C terminus mutants.	395
Table 34	Interhelical distances at the extracellular surface for the polar extracellular mutants.	403
Table 35	Interhelical distances at the intracellular surface for the polar extracellular mutants.	406
Table 36	Interhelical distances at the extracellular surface for the non-polar extracellular mutants.	414
Table 37	Interhelical distances at the intracellular surface for the non-polar extracellular mutants.	417
Table 38	Mutation of extracellular residues of CXCR4 alters chemotaxis.	421
Table 39	Mutation of intracellular residues of CXCR4 alters chemotaxis.	425
Table 40	Mutation of extracellular residues of CXCR4 alters CXCL12 –induced calcium mobilization.	428
Table 41	Mutation of intracellular residues of CXCR4 alters calcium mobilization.	431
Table 42	Comparison of Mutant Molecular Dynamics and Experimental data.	450

LIST OF FIGURES

Figure	#	Title	Page
Figure	1	The process of metastasis of neoplastic cells from primary tumors to target organs.	2
Figure	2	Metastatic breast cancer cells migrate to regions of elevated CXCL12 concentration.	5
Figure	3	Complex interactions between CXCR4, gp120, CD4 and the V3 loop of HIV-1 guide fusion of the HIV-capsule to the host cell membrane.	11
Figure	4	Linear structure of CXCR4.	16
Figure	5	Receptor structures and ligand binding sites for GPCR crystal structures.	28
Figure	6	Two-site / two-step interaction between CXCL12 and CXCR4.	33
Figure	7	Anti-CXCR4 compounds	43
Figure	8	Bicyclic aromatic compounds inhibit GPCR activity	45
Figure	9	Polyphemusin derivative structures.	50
Figure	10	There are 5 potential structural targets for inhibiting CXCR4 signaling	58
Figure	11	Structures for G-protein inhibitors	60
Figure	12	Comparison of the CXCR4 x-ray crystal determined structure of CXCR4 and the CXCR4 homology model.	65
Figure	13	Anti-CXCR4 compounds.	73
Figure	14	Structure-based drug design is an iterative and cyclical process.	78
Figure	15	ECLVS compounds	92
Figure	16	ECLVS Compounds Inhibit CXCR4 Mediated Chemotaxis.	96
Figure	17	IC50 of Active Inhibitors of CXCR4 Mediated Chemotaxis.	98
Figure	18	ECLVS14 selectively inhibits CXCR4 Mediated Chemotaxis.	101
Figure	19	ECLVS compounds inhibit CXCR4 mediated Intracellular Calcium Mobilization.	103
Figure	20	IC50 of ECLVS compound mediated inhibition of calcium mobilization.	106

Figure 21	ECLVS59 selectively inhibits CXCR4 Mediated Calcium Mobilization.	108
Figure 22	ECLVS compounds produce relatively little non-specific cytotoxicity.	111
Figure 23	Modifications of ECLVS14 for QSAR analysis.	113
Figure 24	EVS14 compounds	116
Figure 25	ECLVS14 Derivative Compounds Inhibit CXCR4 Mediated Chemotaxis.	126
Figure 26	IC50 of Active Inhibitors of CXCR4 Mediated Chemotaxis.	128
Figure 27	EVS14 compounds selectively inhibit CXCR4 Mediated Chemotaxis.	130
Figure 28	EVS14 compounds do not inhibit CXCR4 mediated calcium mobilization.	133
Figure 29	Increasing length of the hydrazine moiety decreases activity.	135
Figure 30	Effect of terminal benzene modifications on anti-CXCR4 activity.	137
Figure 31	Absence of the 1-BPMAM moiety alone decreases anti-chemotaxis activity.	140
Figure 32	Terminal benzene electronegativity confers greater influence over anti-chemotaxis activity than the presence or absence of the 1-BPMAM moiety.	143
Figure 33	EVS14-2 derivative compounds	146
Figure 34	EVS14-2 derivative Compounds Inhibit CXCR4 Mediated Chemotaxis	153
Figure 35	IC50 of Active Inhibitors of CXCR4 Mediated Chemotaxis.	155
Figure 36	EVS14-2 derivatives selectively inhibit CXCR4 Mediated Chemotaxis.	157
Figure 37	EVS14-2 derivatives do not inhibit CXCR4 mediated calcium mobilization.	160
Figure 38	EVS14-2 derivatives modify or replace the 1-BPMAM moiety.	162
Figure 39	Acetamide ECLVS14 derivatives produce elevated non-specific cytotoxicity.	165
Figure 40	Final averaged structure of Native CXCR4 dynamically bound to compounds.	169
Figure 41	Energy plot of ECLVS compounds bound to native CXCR4 versus time.	171
Figure 42	RMSD values of CXCR4 bound to ECLVS compounds across 5 ns of the production trajectory.	174
Figure 43	Dynamic binding site of ECLVS14 to native CXCR4.	176
Figure 44	Dynamic binding site of EVS14-2-AB to native CXCR4.	178

Figure 45	Dynamic binding site of ECLVS58 to native CXCR4.	181
Figure 46	Dynamic binding site of ECLVS59 to native CXCR4.	183
Figure 47	Dynamic binding site of ECLVS82 to native CXCR4.	186
Figure 48	Transmembrane regions of the ECLVS compounds complexed with CXCR4 adopt an inactive conformation.	192
Figure 49	Structural modifications of ECLVS14 and effect on activity.	203
Figure 50	There are 5 potential structural targets for inhibiting CXCR4 signaling	208
Figure 51	Chalcones inhibit CXCR4 signaling by binding CXCL12.	211
Figure 52	Anti-CXCR4 compounds	213
Figure 53	SBDD of inhibitors of the intracellular surface of CXCR4.	221
Figure 54	ICLGA Compounds	233
Figure 55	ICLGA compounds inhibit CXCR4 mediated Chemotaxis.	240
Figure 56	IC50 of ICLGA compound mediated chemotaxis inhibition.	242
Figure 57	ICLGA compounds selectively inhibit CXCR4 Mediated Chemotaxis.	245
Figure 58	ICLGA compounds inhibit CXCR4 mediated Intracellular Calcium Mobilization.	248
Figure 59	IC50 of ICLGA compound mediated inhibition of calcium mobilization.	250
Figure 60	ICLGA compounds selectively inhibit CXCR4 Mediated Calcium Mobilization.	252
Figure 61	ICLGA compounds produce little cytotoxicity.	258
Figure 62	Toxic concentrations of ICLGA compounds	260
Figure 63	Inhibition of cell proliferation by ICLGA compounds.	263
Figure 64	Final averaged structure of Native CXCR4 dynamically bound to ICLGA Compounds.	268
Figure 65	Energy plot of ICLGA compounds bound to CXCR4.	271
Figure 66	RMSD values of CXCR4 bound to ICLGA Compounds across 5 ns of the production trajectory.	273
Figure 67	Dynamic binding site of ICLGa8 to native CXCR4.	275
Figure 68	Dynamic binding site of ICLGa12 to native CXCR4.	278
Figure 69	Dynamic binding site of ICLGa25 to native CXCR4	280
Figure 70	Dynamic binding site of ICLGa28 to native CXCR4.	282
Figure 71	Dynamic binding site of ICLGa29 to native CXCR4.	285

Figure 72	Dynamic binding site of ICLGa37 to native CXCR4.	287
Figure 73	Transmembrane regions of ICLGa compound-CXCR4 complexes adopt an inactive conformation.	294
Figure 74	Progression in the SBDD protocol for developing inhibitors of CXCR4.	323
Figure 75	Virtual screening can be used to assess the interaction of multiple conformations of a single compound within an active site.	348
Figure 76	Virtual rescreening of extracellular loop targeted compounds identifies specific contact residues.	350
Figure 77	Virtual rescreening of intracellular loop targeted compounds identifies specific contact residues.	353
Figure 78	Final averaged structures of molecular dynamics of mutated CXCR4.	357
Figure 79	Transmembrane regions of the ICL3 mutants adopt an inactive conformation.	371
Figure 80	Transmembrane regions of the proximal C-terminal mutants adopt an inactive conformation.	380
Figure 81	Transmembrane regions of the C-terminal mutants adopt an inactive conformation.	390
Figure 82	Transmembrane regions of the polar/charged ECL mutants adopt an inactive conformation.	399
Figure 83	Transmembrane regions of the non-polar ECL mutants adopt an inactive conformation.	410
Figure 84	ECLVS14 and EVS14-2-AB produce greater calcium inhibition in V177A-F276A mutant CXCR4.	433
Figure 85	EVS14-21 produces greater calcium mobilization in V177A-F189A-L194A mutant CXCR4.	436
Figure 86	ECLVS14 and EVS14-2-AB produce greater inhibition of calcium mobilization in N192A mutant CXCR4.	439
Figure 87	ECLVS14, EVS14-21, EVS14-22, and EVS14-2-AB produce greater inhibition of calcium mobilization in S178A-E179A-N192A-H281A mutant CXCR4.	441
Figure 88	ECLVS14 and EVS14-2-AB produce greater inhibition of calcium mobilization in F276A mutant CXCR4.	443
Figure 89	EVS14-2-AB produces greater inhibition of calcium mobilization in H281A mutant CXCR4.	445
Figure 90	ICLGa12 produces greater inhibition of calcium mobilization in K239A mutant CXCR4	448
Figure 91	TC-1 cells express increased cell surface CXCR4.	453
Figure 92	Lewis lung cancer cell metastasis is not altered by anti-CXCR4 compounds.	456
Figure 93	Treatment with ECLVS14, ICLGa28 and ICLGa36 each reduce tumor incidence and tumor burden.	458

CHAPTER I: INTRODUCTION

Metastasis presents an unresolved and difficult clinical problem

Tumor metastasis represents a significant portion of the morbidity and mortality attributable to neoplasia [1]. Axillary lymph node micrometastasis or occult metastasis of 2mm or less correlates poorly with disease-free and overall survival [2]. Bone metastasis is present in 30-50% of patients with highly metastatic cancers such as breast, prostate, and non-small cell lung cancer [3].

Development of metastases requires a complex process of cellular evolution including several distinct transitions from primary carcinoma *in situ* to cells capable of establishing new metastatic foci, (see Figure 1) [4, 5].

Neoplastic cells must proliferate and invade at the primary site. These cells must then leave the primary tumor and enter the blood stream, in a process known as intravasation. Neoplastic cells extravate from the blood stream and attach to surfaces at foci of metastasis [6]. Established by Paget in 1889, the “seed and soil hypothesis” stipulates that neoplastic cells only establish metastatic foci in favorable microenvironments [6, 7].

The diffuse and systemic nature of metastatic disease renders many conventional treatment modalities ineffective, presenting a particularly perplexing predicament for treatment. Prior to metastasis, neoplasia can be effectively treated by surgical excision with clear margins and localized radiation. Once the neoplastic cell's metastatic capability is disabled, the disease reverts to

Figure 1: The process of metastasis of neoplastic cells from primary tumors to target organs. Neoplastic cells proliferate and invade at their primary locus. Selected cells gain the capability of leaving the primary tumor and entering the systemic circulation. Tumor cells extravate in target organs such as the lung, liver, bone marrow and brain. Once adherent, selected neoplastic cells will survive in favorable microenvironments and proliferate to form metastatic foci. Reproduced from [6].

Figure 1: The process of metastasis of neoplastic cells from primary tumors to target organs.

an earlier stage, and becomes a cancer that is once again isolated to its primary site and any metastatic foci. Neoplastic growth would continue in a linear fashion, but without systemic spread, conventional surgery/chemoradiation could have a greater probability of success. An anti-metastatic approach largely mirrors the clinical reality. For example, 84% of newly diagnosed lung cancer has already progressed past early stage disease upon time of diagnosis [8], making treating and inhibiting the metastasis of the primary tumor a greater concern than inhibition of the growth of the primary tumor. Unfortunately there are no primary anti-metastatic treatments approved by the FDA or in clinical trials.

CXCR4 plays a key role in Tumor metastasis.

A wide number of molecules such as cytokines, chemokines and their receptors and growth factors have been implicated in the metastatic properties of cancer cells [9-15]. In 2001 the involvement of chemokine receptors in breast cancer metastasis was reported [9]. In a seminal Nature paper, Müller *et al* delineated a theory of metastasis which centered on chemoattraction, a process illustrated in Figure 2. Metastatic breast cancer cells displayed increased cell surface expression of CXCR4 and CXCR4's ligand CXCL12 was highly expressed in lung, lymph node, bone marrow and liver tissues. Müller found that CXCL12 induced more malignant and metastatic phenotypes with increased actin modification and pseudopodia formation in cells expressing CXCR4. CXCL12 and CCL21 also were found to induce metastatic cell migration. Protein extracts from several organs were found to induce chemotactic migration by the cancer cells, with lung, lymph node, bone marrow and liver tissues inducing the

Figure 2: Metastatic breast cancer cells migrate to regions of elevated CXCL12 concentration. Reproduced from [16]. Müller et al. found that metastatic breast cancer cells (patient samples and in the MDA-MB-231 cell line) express elevated levels of cell surface CXCR4. CXCR4 ligand CXCL12 expression was increased in sites of metastatic foci, including lymph nodes, lung, liver and bone marrow. Metastatic cells migrated to regions of elevated CXCL12 concentration, and cells with elevated CXCR4 surface expression formed stable metastatic foci.

Figure 2: Metastatic breast cancer cells migrate to regions of elevated CXCL12 concentration.

most migration. Finally, metastatic models, using mammary fat pad or tail vein injection of metastatic MDA-MB-231 cells, CXCR4 neutralization by antibody was found to reduce lung metastasis with a rate of relative suppression of 61-68%.

CXCR4 has been demonstrated to play a critical role in angiogenesis [17] and metastasis [18] of several tumors, including basal cell carcinoma [17], thyroid cancer [18], squamous cell carcinoma [19], neuroblastoma [20], ovarian [21], melanoma [22], renal cell carcinoma [23], hepatocellular carcinoma [24], breast [25], colon [26], lung [25], pancreatic [27] and prostate cancers [25]. Chinni *et al* expanded on the Müller theory by demonstrating that bone metastatic prostate cancer cells express cell surface CXCR4, and the CXCR4/CXCL12 interaction was essential to migration and bone marrow invasion [28]. Additional research has determined that knockdown or inhibition of CXCR4 impedes growth [29] and metastasis [9, 30, 31] of neoplastic breast cancer cells. Phillips *et al* [32] demonstrated that non-small cell lung cancer (NSCLC) tumor specimens and NSCLC derived cell lines, (such as A549) express CXCR4 and that these cells migrate towards CXCL12, with increased intracellular signaling in the form of increased intracellular calcium mobilization and MAPK activation. They also demonstrated that NSCLC metastasis could be prevented in nude mice through the application of a CXCR4-specific antibody, with rate of relative suppression of around 50% for bone marrow and brain metastasis.

A recent publication by Karnoub *et al* [33] suggested an alternative mechanism for metastasis. In this report, they demonstrated that mesenchymal stem cells (MSCs) within the tumor stroma can induce increased metastasis in

primary tumors. Their mechanism for this increased metastasis centered on increased CCL5 production by MSCs inducing enhanced AKT phosphorylation which resulted in increased tumor motility and invasiveness. While this paper would seem to cast some doubt on CXCR4's role in metastasis, several factors speak to the importance of CXCR4 in this process. Active MSCs have been characterized as retaining enhanced expression of CXCR4 and CXCL12 [34]. Also, a more recent study using CD90 high expressing cells (a common marker for MSCs) found that co-culture or incubation in conditioned medium resulted in increased CXCR4 expression in BPH-1 cells [35]. Muelberg *et al* [36] recently demonstrated increased motility, invasion and metastasis in breast cancer cells using Adipose tissue derived stem cells. They found this increased metastasis was largely effected by BCCs inducing increased secretion of CXCL12 by the Adipose tissue derived stem cells, and that the enhanced- metastasis phenotype could be abolished by CXCR4 knockdown in tumor cells.

CXCR4 expression by the tumor cells allows for tumor cell migration, and homing of the neoplastic cells to sites where non-malignant stromal cells express CXCL12 [37]. This implies that tumor cell metastasis is not a random event, but directed by the expression of chemokine receptors and adhesion molecules on the neoplastic cells, and their respective ligands in the target organs [9, 16]. Tumor cells utilize this mechanism to access microenvironments, such as the bone marrow, that provide factors that favor their growth. For example, the importance of CXCR4 for breast cancer metastasis to different target organs has been demonstrated *in vitro* and *in vivo* [9]. Leukemia cells from patients with

chronic lymphocytic leukemia (CLL) and acute lymphoblastic leukemia utilize CXCR4 for homing to marrow stromal cells *in vitro* [38] and *in vivo* [39].

Moreover, CXCL12 has a direct growth and survival-promoting effect for various cancer cells, such as breast cancer cells [40] or CLL cells [41]. CXCL12 promotes tumor progression by recruiting endothelial progenitor cells to tumors for angiogenesis [40].

CXCR4 was originally characterized as a co-receptor for T-tropic strains of HIV-1.

Originally dubbed “LESTR”, or “leukocyte-derived seven-transmembrane domain receptor” [42], CXCR4 was an orphan G-protein coupled receptor (GPCR) when it was identified as the co-receptor for T-tropic strains of the human immunodeficiency virus-1 (HIV-1) [43]. CXCR4 was officially named following the identification of its cognate ligand CXCL12, (previously identified as stromal derived factor-1) through experiments demonstrating that CXCR4 interaction with CXCL12 inhibited HIV-1 infection [44, 45]. It was found that the N-terminal region of CXCR4 was essential to HIV-1 entry [46] and that complex interactions between the N-terminus and second extracellular loop (ECL2) of CXCR4, GP120 and CD4 and the V3 loop of HIV-1 are important for HIV-1 infection, as illustrated in Figure 3 [47-50]. Interaction between gp120 and CD4 results in a conformational change of gp120, allowing binding of the co-receptor, in this case, CXCR4. Interactions between gp120 and CXCR4 induce further conformational changes in the complex, resulting in insertion of the fusion peptide of gp41 into the cell membrane of the host T-cell membrane [51].

CXCR4 and CXCL12 play essential roles in the physiology of the developing embryo.

CXCR4 is constitutively expressed in many tissues including stomach, small intestine, brain, lymphatics, thymus and spleen and serves to direct migration for numerous cell types [52]. CXCR4 mediated migration serves a variety of functions, from facilitating repair of damaged liver tissue [53] to endochondral bone formation [54].

CXCR4 also plays an important role in development as CXCR4 deletions are lethal *in utero*, with embryos displaying defects in neuronal and cardiac development [55-57] as well as defects in neovascularization [24]. CXCR4 regulates the migration of primordial germ cells in the formation of the nervous system [58]. CXCR4 and CXCL12 are also essential to the proper migration of cells associated with the cerebellum, dentate gyrus and dorsal root ganglia [59-61]. Finally, CXCR4/CXCL12 interaction controls cortical interneuron migration and positioning [62].

CXCR4 mediated signaling is also essential to the formation of the primordial gut, known as the cloaca, [63] as well as the formation of the vascular

Figure 3: Complex interactions between CXCR4, gp120, CD4 and the V3 loop of HIV-1 guide fusion of the HIV-capsule to the host cell membrane.

Reproduced from Cohen et al [64]. Interaction between gp120 and CD4 results alters the conformation of gp120 resulting in CXCR4 binding of gp120. CXCR4: gp120 interaction induces insertion of the fusion peptide of gp41 into the cell membrane of the host T-cell membrane [51].

Figure 3: Complex interactions between CXCR4, gp120, CD4 and the V3 loop of HIV-1 guide fusion of the HIV-capsule to the host cell membrane.

system of the GI tract [65, 66]. The CXCL12/CXCR4 axis is also responsible for maintaining mucosal barrier integrity in the gut by regulating epithelial cell migration, barrier maturation, and reconstitution [67].

CXCR4 heterozygous mutation can lead to WHIM (warts, hypogammaglobinemia, infections, and myelokathexis) syndrome [68]. The structural basis of WHIM syndrome pathology stems from a C-terminus deletion of CXCR4, most commonly at R334 [69]. This C-terminal deletion produces a two-fold increase in signaling, while rendering the GPCR strongly resistant to ligand-mediated downregulation [69].

Essential role of CXCR4 signaling in cell mediated immunity

CXCR4 and CXCL12 signaling is essential to immune physiology. CXCR4 is widely expressed on leukocytes, and serves to regulate leukocyte hematopoiesis and trafficking [68] as well as anchoring developing cells to specific regions of the bone marrow [70]. Naïve T-lymphocytes express CXCR4, and this expression is essential in directing T-cell migration to the thymus for maturation. Once in the thymus, the interaction between CXCR4 and CXCL12 is critical to the intra-thymic migration of immature T-cells [71]. In the spleen, lymph nodes, and tonsils, CXCR4 is essential to the formation of properly structured lymphoid follicles and population of germinal centers by controlling the trafficking of B-cells, T-cells and plasma cells [72, 73]. In mature neutrophilic leukocytes, increased CXCR4 expression has been associated with cell aging, and sequestration in the liver, spleen and bone marrow before apoptosis [74]. While increased CXCR4 expression and activation has been associated with increased

migration in the leukocyte compartment, decreased CXCR4 expression and activation has been associated with increased circulating active leukocytes [75], and increased activation of T-cell receptors on T-lymphocytes [76].

One of the more interesting roles of CXCR4 in the immune compartment involves its role in regulation of neutrophil mobilization. Petit *et al* demonstrated that neutrophils are maintained in the bone marrow by CXCR4-CXCL12 signaling. Specifically, it is the production of CXCL12 by bone marrow stromal cells which maintains the interaction between CXCL12 and CXCR4, and when factors such as granulocyte colony stimulating factor (G-CSF) decrease stromal CXCL12 production by inhibiting CXCL12 mRNA expression, stem cells are mobilized [77, 78]. This effect is seen in a variety of pathological situations, including sepsis where depressed marrow CXCL12 expression combined with enhanced CXCL12 expression in the spleen has been demonstrated to drive neutrophil mobilization and leukocytosis in response to overwhelming infection [79]. CXCR4's role in sequestration of hematological stem cells in the bone marrow is also highlighted by CXCR4 inhibitor AMD3100. Inhibition of CXCR4 decreases CXCR4-CXCL12 signaling, producing rapid mobilization of hematopoietic stem cells [80-82], a finding that resulted in FDA approval of AMD3100 as a mobilization agent for hematopoietic stem cells for autologous transplantation in non-Hodgkin's lymphoma and multiple myeloma [80].

Because of CXCR4's critical role in immunity some may question the wisdom of using it as a chemotherapeutic agent. They may suggest that inhibition of immunity at a time when patients are already immunocompromised

due to elevated tumor burden would be unwise. However, this really isn't the case with cancer chemotherapy. One of the most common side effects of common chemotherapeutic agents such as cisplatin, adriamycin, cyclophosphamide and mercaptopurine is myelosuppression, which results in a nadir 21-28 days after treatment is begun when white blood count can drop to almost zero. Several chemotherapy regimens include prednisone, a corticosteroid which suppresses the immune system. The counter to symptomatic myelosuppression has been treatment with G-CSF, which acts by inhibiting CXCR4 signaling.

G-Protein Coupled Receptors: sensors of external and internal stimuli

CXCR4 is a seven-transmembrane-helix G-protein coupled receptor (GPCR) which binds CXCL12, a CXC chemokine. Structurally, CXCR4 and other GPCRs are characterized by seven membrane spanning helical domains, an extracellular amino terminus, and an intracellular carboxyl terminus. The seven transmembrane (TM) domains are joined by three extracellular (ECL) and three intracellular (ICL) loops as illustrated in Figure 4. Humans express 750 GPCRs which serve as sensors of external and internal stimuli by transmitting signals through the cell membrane and activation of cytoplasmic signaling cascades [83]. As seen in Table 1, GPCRs are divided into 6 classes, based on the Vriend classification system for GPCRs [84]. Class A GPCRs, which are rhodopsin-like, encompass 80% of GPCRs [85]. CXCR4 is a Class A GPCR. Class B, secretin-like, class C, metabotropic glutamates, class D, pheromones, class E, cAMP

Figure 4: Linear structure of CXCR4. G-Protein coupled receptors are characterized by seven transmembrane receptors joined by three extracellular and three intracellular loops, with an extracellular N terminus and an intracellular C terminus. Adapted from [86].

Figure 4: Linear structure of CXCR4.

Table 1: GPCR Classifications. Under the Vriend classification system [84] for GPCRs, class A GPCRs, which are rhodopsin-like, encompass 80% of GPCRs. CXCR4 is a Class A GPCR. Class B, secretin-like, class C, metabotropic glutamates, class D, pheromones, class E, cAMP receptors, and the class F frizzled/smoothed family encompass the other 20% of these receptors.

Table 1: GPCR Classifications.

GPCR Family	Family Description
Class A	Rhodopsin-Like
Class B	Secretin-Like
Class C	Metabotropic glutamates/pheromones
Class D	Fungal Pheromone
Class E	cAMP receptors
Class F	Frizzled/Smoothed

receptors, and the class F frizzled/smoothed family encompass the other 20% of these receptors. Phylogenetic analysis has allowed further subclassification of the Class A GPCRs into distinct sub-families as depicted in Table 2 [87]. CXCR4 belongs to subfamily A2 in this classification system, which encompasses the CC and CXC chemokines, as well as adrenomedullin and the Duffy antigen.

GPCR activation triggers coupling with heterotrimeric G-proteins consisting of α , β , and γ subunits. Upon interaction with an activated GPCR, GDP bound to $G\alpha$ will be exchanged for GTP, facilitating a separation of the heterotrimer into $G\alpha$ and $G\beta\gamma$ subunits. These subunits in turn activate downstream signaling pathways until intrinsic GTP-ases in $G\alpha$ convert GTP to GDP, facilitating the reformation of the G-protein heterotrimer.

Chemokines: small protein regulators of a variety of physiological processes

Chemokines are a superfamily of small molecule chemoattractive cytokines that mediate several cellular functions. Chemokines are classified by the conformation of a conserved amino-terminal cysteine residue motif, as depicted in Table 3 [88, 89]. Chemokines and their receptors regulate chemotaxis and the transendothelial migration of leukocytes were recently observed to play an important role in the metastasis of various cancers [90], as well as homeostasis, cell proliferation, hematopoiesis, angiogenesis and neovascularization [88]. Chemokines also have important roles outside the hematopoietic compartment, for example, $GRO\alpha$ has been shown to be an

Table 2: Subclassifications of Class A GPCRs. CXCR4 belongs to subfamily A2 in this classification system, which encompasses the CC and CXC chemokines, as well as adrenomedullin and the Duffy antigen. The BLT/LTB4 signaling pathway belongs to the A5 class of receptors. Rhodopsin is classified as a class A16 receptor with the rest of the Opsin receptors.

Table 2: Subclassifications of Class A GPCRs.

Subfamily	Representative Receptors
A1	CC and CX3C Chemokines
A2	CC, CXC, and IL-8
A3	Angiotensin and Bradykinin
A4	Opioid, Somatostatin, Neuropeptides B/W
A5	Leukotriene B4 receptors, Relaxin, Galanin
A6	Cholecystokinin, Vasopressin, Orexin
A7	Bombesin, Endothelin, TRHR
A8	Anaphylatoxin receptor, Formyl peptide receptors
A9	Melatonin, Neurokinin, Neuropeptide Y
A10	FSH receptor, LH Receptor, Thyrotropin receptor
A11	GPR40 and P2 Purinoreceptor
A12	P2 Purinoreceptor, Platelet-activating factor
A13	Cannabinoid receptor, Lysophosphatidic acid receptor, Melanocortin/ACTH receptor
A14	Eicosanoid receptor
A15	Protease-activated receptor, Proton-sensing GPCRs
A16	Opsins
A17	5-Hydroxytryptamine (5-HT) receptor, Adrenergic receptor, Dopamine receptor, Histamine H2 receptor
A18	Histamine Receptor, Adenosine receptor, Muscarinic acetylcholine receptor
A19	5-Hydroxytryptamine (5-HT) receptor

Table 3: Chemokines are classified by the conformation of a conserved amino-terminal cysteine residue motif. Chemokines serve a variety of functions in the regulation of the immune system as well as in control of homeostasis, cell proliferation, haematopoiesis, angiogenesis and neovascularization in neoplasia. Adapted from [88].

Table 3A: CXC, C, and CX3C Chemokines, and their receptors.

Chemokine	Human Ligand	Receptor(s)
CXC Chemokine		
CXCL1	GRO α /MGS α	CXCR1, CXCR2
CXCL2	GRO β /MGS β	CXCR2
CXCL3	GRO γ	CXCR2
CXCL4	PF4	CXCR3B
CXCL5	ENA-78	CXCR2
CXCL6	GCP-2	CXCR1, CXCR2
CXCL7	NAP-2	CXCR2
CXCL8	IL-8	CXCR1, CXCR2
CXCL9	MIG	CXCR3, CXCR3B
CXCL10	IP-10	CXCR3, CXCR3B
CXCL11	I-TAC	CXCR3, CXCR3B, CXCR7
CXCL12	SDF-1 α/β	CXCR4, CXCR7
CXCL13	BLC/BCA-1	CXCR5
CXCL14	BRAK, Bolekine	Unknown
CXCL15	Unknown	Unknown
CXCL16	SICB6	CXCR6
CXCL17	DMC	Unknown
C chemokines		
XCL1	Lymphotactin/ATAC/SCM-1 α	XCR1
XCL2	SCM-1 β	XCR1
CX3C Chemokines		
CX3CL1	Fractalkine	CX3CR1

Table 3B: CC Chemokines, and their receptors.

Chemokine	Human Ligand	Receptor(s)
CC Chemokine		
CCL1	I-309	CCR8
CCL2	MCP-1/MCAF/TDCF	CCR2
CCL3	MIP-1 α /LD78 α	CCR1, CCR5
CCL3L1	LD78 β	Unknown
CCL3L3	LD78 β	Unknown
CCL4	MIP-1 β	CCR5
CCL4L1	AT744.2	Unknown
CCL5	RANTES	CCR1, CCR3, CCR5
CCL7	MCP-3	CCR1, CCR2, CCR3
CCL8	MCP-2	CCR1, CCR2, CCR3, CCR5
CCL11	Eotaxin	CCR3
CCL13	MCP-4	CCR1, CCR2, CCR3
CCL14	HCC-1	CCR1
CCL15	HCC-1/LKN1/MIP-1 γ	CCR1, CCR3
CCL16	HCC-4/LEC/LCC-1	CCR1, CCR2, CCR5, HRH4
CCL17	TARC	CCR4
CCL18	DC-CK1/PARC/AMAC-1	Unknown
CCL19	MIP3 β /ELC/exodus-3	CCR7
CCL20	MIP3 α /LARC/exodus-1	CCR6
CCL21	SLC/6Ckine/exodus-2	CCR7
CCL22	MDC/STCP-1	CCR4
CCL23	MPIF/CK β 8/CK β 8-1	CCR1
CCL24	Eotaxin-2/MPIF-2	CCR3
CCL25	TECK	CCR9
CCL26	Eotaxin-3	CCR3
CCL27	CTACK/ILC	CCR10
CCL28	MEC	CCR3, CCR10

autocrine growth factor in melanoma [91]. In addition, decoy chemokine receptors bind ligands with high affinity but negligible activity, in order to control the availability of chemokine for receptor and fine tune directionality of chemical gradients [88, 89, 92].

GPCR signal transduction is based on movement of transmembrane helixes

The theory of G-protein function is largely based on the opsin model, wherein interaction between ligand and receptor disrupts an ionic lock between TM3 and TM6. This in turn disrupts several stabilizing interactions between the TM regions allowing the GPCR to shift from an inactive to an active state capable of interacting with a G-protein [93-96]. GPCRs are also capable of forming homo and hetero dimers and oligomers, allowing protein function to vary with quaternary structure [83, 97-103].

The lack of insight into GPCR structure is largely due to the difficulty in expressing, purifying, and solubilizing the large amounts of protein required for X-ray crystallography. The original mammalian crystal structure, bovine rhodopsin [104] was generated by micelle suspension of rhodopsin, which required harvesting the GPCR from a large volume of bovine retinas, a process which benefitted from the relatively undynamic structure of rhodopsin when compared to other GPCRs [104-107]. Similar large volume purification protocols were used in the crystallization of opsin [95].

Developing the capability of producing massive volumes of functional receptor from baculovirus infected insect cells, coupled with antibody-mediated

stabilization methods to stabilize the third intracellular loop (ICL3) region of the GPCR allowed initial crystallization of the β 2-adrenergic receptor [108]. Higher resolution structures of the β 2-adrenergic receptor were generated by replacement of the ICL3 region with T4 Lysozyme (T4L) a soluble and stable protein [109, 110] which exchanges a mobile and flexible unit for a solid and immobile element which is a desirable characteristic in the development of well-ordered three dimensional crystals. Utilization of the T4L replacement technique along with replacement of the C-terminal tail with a histidine purification epitope allowed successful crystallization of the α 2A adenosine receptor [111].

Comparison of the rhodopsin, β 2-adrenergic receptors and the α 2A adenosine receptor structures offers some key insights into the functional differences between the GPCRs, as depicted in Figure 5. At the extracellular surface of the receptors, there are significant differences in the placement of ECL2 and the N-terminus. In Rhodopsin, the ECL2 and the N-terminus the GPCR combine to form a β -sheet which covers the binding pocket of the protein and prevents cis-retinal hydrolysis. ECL2 in the β 2-adrenergic receptor forms an α helix which is displaced away from the ligand binding site by two disulfide bonds. ECL2 on the α 2A adenosine receptor is characterized by formation of an α helix restrained by disulfide binding as in the case of the β 2-adrenergic receptor however in the case of α 2A adenosine, there are 2 additional disulfide bonds on the helix, and the surface of the binding site displays significantly lower electron density than that of the β 2-adrenergic receptor.

Figure 5: Receptor structures and ligand binding sites for GPCR crystal structures. Structures are presented in ribbon format for A) bovine rhodopsin/retinal (PDBID: 1F88), B) β 2-adrenergic receptor/carazolol (PDBID: 1F88), C) β 2-adrenergic receptor/timolol (PDBID: 1F88), and D) α 2A-adrenergic receptor/ZM241385. Bound ligands are depicted as spheres in the ribbon helix format and as stick figures in the ligand binding site images. Modified from [112].

Figure 5A: Receptor structures for GPCR crystal structures.

Figure 5B: Receptor ligand binding sites for GPCR crystal structures.

On the intracellular surface of the rhodopsin crystal structure, the second intracellular loop (ICL2) is displaced away from the central TM helices. In the crystal structures of β 2-adrenergic receptor and α 2A adenosine receptor, ICL2 interacts with the intracellular surfaces of TM2, TM3, TM5 and TM6, specifically with the thirtieth residue of TM6, disrupting the formation of the ionic lock formed by this glutamate residue as seen in rhodopsin. In addition, ICL2 of the α 2A adenosine receptor forms a short helical motif not seen in the other receptor structures. In the cases of the β 2-adrenergic receptor, the α 2A adenosine receptor replacement of the ICL3 region and removal of the C-terminal tail could also explain these differences with the rhodopsin structure. (Reviewed in [93, 107, 112-115])

CXCL12 activity is governed by structure and posttranslational modifications

The structure of CXCL12 has been identified by crystal structure determination. CXCL12 is a 67 residue polypeptide with an n-terminal loop, followed by three anti-parallel β -strands, and a c-terminal α -helix packed against the β -sheet [116]. CXCL12 exists in equilibrium between monomeric and dimeric structures and the chemokine aggregates at nonacidic pH and in response to multivalent anions like phosphate, sulfate, citrate, or heparin [117, 118]. CXCL12 activity is regulated by the protein carboxypeptidase N, which cleaves CXCL12's C-terminal lysine, disabling the chemokine's proliferative and chemoattractant abilities [119]. A CXCL12 ligand lacking the C-terminal α -helix can still bind and

induce internalization of CXCR4, but does not induce calcium mobilization, ERK phosphorylation, or chemotaxis [120].

CXCL12 signaling occurs in a two-site / two-step reaction as depicted in Figure 6. In this reaction, the N-terminus of CXCR4 interacts with the R-F-F-E-S-H motif of the N-terminal loop of CXCL12 [121-123]. This triggers a conformational change in the CXCR4 N-terminus bringing CXCL12 into contact with the extracellular portions of TM2 and TM7 as well as ECL1 and ECL2 [122]. The extracellular binding interface of CXCL12 was described by Brelot *et al* as composed of residues D97 on TMII, D187 on ECL2, and E288 on TMVII [124]. Ligand binding at these points is largely driven by hydrogen bond interactions between the highly basic surface of CXCL12 and the acidic residues of CXCR4. Interaction between the N-terminus of CXCL12 and extracellular surfaces of the TM helixes produces signal transduction [121, 125, 126].

CXCR4 signal transduction is determined by CXCR4 structure.

CXCR4 is characterized structurally by seven membrane spanning helical domains, an extracellular amino terminus, and an intracellular carboxyl terminus. The seven TM domains are joined by three ECL and three ICL loops. The extracellular loop region consists of 4 sub-regions: the N-terminal strand from residues 1 to 39, the ECL1 encompassing residues 98 to 109, the ECL2 extending from residue 176 to 202, and the ECL3 from 262 to 282. Initial studies into the interaction of CXCR4/CXCL12 revealed that the N-terminal strand is essential to the interaction of chemokine and receptor, while reduced binding and signaling was seen in mutations to the extracellular portions of TM2 and TM7 as

Figure 6: Two-site / two-step interaction between CXCL12 and CXCR4. In the two-site / two-step reaction, the RFFESH loop of the N-terminus of CXCL12 binds to the N-terminus of CXCR4, and the N-terminus of CXCL12 subsequently interacts with the extracellular surfaces of the transmembrane helixes. Adapted from [121].

Figure 6: Two-site / two-step interaction between CXCL12 and CXCR4.

well as ECL1 and ECL2. The N-terminus was essential to receptor transport to the cell surface [124]. N-terminal deletions of CXCR4 impair CXCL12 binding and signaling. Specifically, acidic (Glu14 and Glu15) and tyrosine (Tyr21) residues of the n-terminus are required for CXCL12 binding [124].

The intracellular loop region consists of 4 sub-regions: the first intracellular loop encompassing residues 64-76, the second intracellular loop extending from residue 132 to 155, ECL3 from 223 to 241, and the carboxyl-terminus of the protein extending from residue 308 to 352. This region governs signal transduction of the receptor, and has multiple interactions to stimulate signaling cascades. Mutational analysis of the intracellular surface has defined several essential elements to signal transduction, including the DRY motif of the second intracellular loop, which will inactivate receptor signaling if disrupted [124], and the N-terminal region of ICL3, disruption of which eliminates CXCL12 mediated signaling [124].

The interhelical segments are also essential for CXCR4 mediated signal transduction. Zhang *et al* developed a series of mutations focused on the TM3 residue Asn 119 (N119). N119 mutation to Ser (N119S) or Ala (N119A) produced a constitutively active conformation of CXCR4, while mutation to Lys (N119K) produced a constitutively inactive conformation of the receptor [127]. The N119S conformation produced an elevated signaling level which could be further stimulated by CXCL12 binding, in contrast to the N119A mutation which although constitutively signaling at a level greater than the N119S mutant, could not be further stimulated. The constitutively active mutants also both bound G-proteins

to a greater extent than the native state of the receptor. In contrast to both of these strains, N119K signaled poorly in response to CXCL12 binding, and did not bind G-proteins to any extent greater than the unstimulated native receptor.

CXCR4/CXCL12 signaling occurs by a variety of pathways

CXCR4 signaling occurs primarily via the pertussis toxin-sensitive Gi pathway, composed of G α i, G β and G γ [128]. CXCL12 interaction with CXCR4 induces dissociation of G α i from the trimeric G-proteins, and G $\beta\gamma$ activates several phosphorylation pathways, including ERK1/2, JNK, MAPK, and GSK 3 α/β [129]. Activation of the G $\beta\gamma$ subunit also results in signaling by the Phospholipase C pathway to enhance calcium mobilization as well as signaling by the Phosphoinositide-3-kinase pathway to activate AKT. CXCR4 binding of CXCL12 also serves to augment intracellular cAMP levels [130] and activate cAMP-dependant signaling pathways, such as PK α or CREB [131, 132]. The combined activation of these factors results in cell migration [133-135].

Activation of these signaling pathways allows CXCR4 to induce several changes in the cell phenotype. CXCR4 enhances cell adhesion by enhancing binding between integrin couples such as VLA-4 and VCAM-1 [136]. CXCR4 stimulation produces a phenotype with enhanced invasiveness by increasing AKT mediated MMP9 expression. Finally, CXCR4 facilitates cytoskeletal rearrangement by stimulating actin polymerization by formation of F-actin [137]. CXCR4 expression in cancer cells can be enhanced by several factors including demethylation [138], or increased expression of hypoxia inducible factor (HIF) 1 α [139]. CXCR4 activity can also be down-regulated in several fashions, such as

internalization on lipid rafts and inactivation by β -arrestin binding due to excess CXCL12 [139] or repression of CXCR4 DNA by G-CSF [140].

CXCR4 internalization is in fact essential to CXCR4 function. Interactions between β -arrestin and the C-terminus of CXCR4 govern receptor internalization [141], but coexpression of CXCR4 with overexpression of β -arrestin has been demonstrated to produce a 3-6 fold increase in CXCL12-induced ERK phosphorylation [142], as well as increased sensitivity to CXCL12 to induce chemotaxis [143], effects that could be blocked by knockdown of β -arrestin expression [143]. In addition, once β -arrestin has internalized CXCR4 it serves to prevent degradation of the receptor [144] and modulates endosomal sorting to recycle inactivated receptors back to the cell surface [145].

This produces a signaling model of CXCR4 based on receptor recycling. Interactions between CXCR4 and CXCL12 produce an active receptor, which initiates multiple G-protein mediated signaling cascades. The active receptor then interacts with β -arrestin, which silences CXCR4 by internalization. Internalized receptors are then protected from degradation, inactivated and recycled to the cell surface. The necessary balance between activation and inactivation of CXCR4 is reflected in the activity of the WHIM variant of CXCR4. In the CXCR4-WHIM variant, absence of several C-terminus residues impairs β -arrestin-2 recruitment and binding, which produces a delay in internalization of the receptor. Cells expressing the CXCR4-WHIM variant also displayed delayed phosphorylation of Erk 1/2 [146]. This suggests that rapid cycling of CXCR4 augments the ability of the receptor to detect changes over time in the availability

of ligand, producing a mechanism to fine tune the directionality of chemotaxis. This explains the poor immune response of WHIM syndrome immune cells, in spite of enhanced CXCL12 mediated chemotaxis in WHIM patient leukocytes [147] and lymphocytes [148].

There are multiple assays available to assess CXCR4-mediated signaling. Some of the most common methods include assessment of CXCL12-induced calcium mobilization [149-151], phosphorylation of secondary signaling molecules such as ERK1/2, JNK, MAPK, and GSK 3 α / β [151-153], and internalization assays to assess the activity of β -arrestin [142-144]. Assessment of CXCL12-induced calcium mobilization involves pretreating cells with Fura-2 or Indo-1, fluorescent dyes which alter the ratio of their emissions in response to exposure to calcium. Treatment with Triton X-100 releases mitochondrial calcium stores providing a consistent measure by which assay results can be normalized, thus this method provides a rapid, real time, consistent measure of CXCL12-induced signaling which is very amenable to usage in high-throughput processes.

Western blot analysis can be used to determine the quantity of phosphorylated ERK1/2 and PI3K/AKT in a given sample. However, assessment of protein phosphorylation requires lysis and purification of the protein samples, preventing any assessment of signaling in real time. In addition, variability in protein expression and signaling makes translation of this measurement to a high-throughput process difficult.

Internalization assays to assess the activity of β -arrestin can be performed utilizing fluorescently tagged receptor proteins, allowing cell surface expression

of the receptor to be detected by fluorescent microscopy. Although receptor internalization would be a very qualitative measurement of β -arrestin activity, the results would be difficult to quantify, making comparisons between samples different, and making it difficult to translate this method to high throughput protocols.

GPCRs are common targets for pharmaceuticals

GPCRs account for the majority of the top selling pharmaceuticals and roughly 40% of all licensed pharmaceuticals, a partial listing of which is presented in Table 4. The serotonin receptor serves as the target for several medications such as Resperdal, Zyprexa, Imitrex, Buspar, Prozac, Paxil, Zoloft, and Effexor. The histamine receptors serve as targets for many pharmaceuticals as well, including Claratin, Allegra, Zantac, Pepcid, and Gaster. Cardura, Flomax, Hytrin, and Wellbutrin all function through inhibition of the alpha-adrenergic receptor. Tenormin, Ventolin, Toprol-XL, and Serevent all inhibit the beta-adrenergic receptor. Diovan, Vasotec, and Cozaar all inhibit the angiotensin II receptor and the angiotensin-converting enzyme. Several other GPCR receptors serve as targets for pharmaceuticals including the GABA (Diprivan), GNRH (Lupron), histone deacetylase (Depakote), Leukotriene (Singulair), LHRH (Zoladex), Muscarinic acetylcholine (Atrovent), and the prostanoid selective FP receptor (Xalatan).

Several anti-CXCR4 compounds have been developed

Despite significant research, there is currently only one FDA-approved CXCR4 inhibitor on the market, AMD3100, and that compound is designated

Table 4: GPCR structures serve as targets for numerous pharmaceuticals.

Anti-GPCR compounds serve as targets for more than 40% of all licensed pharmaceuticals for a wide variety of indications ranging from hypertension, gastric esophageal reflux disease, allergy, to psychosis.

Table 4: GPCR structures serve as targets for numerous pharmaceuticals.

Pharmaceutical	Receptor target	Indication
Cardura	α adrenergic	Prostatic Hypertrophy
Flomax	α adrenergic	Prostatic Hypertrophy
Hytrin	α adrenergic	Hypertension
Wellbutrin	α adrenergic	Depression
Diovan	Angiotensin II	Hypertension
Vasotec	Angiotensin Converting Enzyme	Hypertension
Cozaar	Angiotensin Converting Enzyme	Hypertension
Tenormin	β adrenergic	Angina
Ventolin	β adrenergic	Bronchospasm
Toprol-XL	β adrenergic	Hypertension
Serevent	β adrenergic	Asthma
Diprivan	GABA	Anesthesia
Lupron	GnRH	Prostate cancer
Claratin	Histamine	Allergy
Allegra	Histamine	Allergy
Zantac	Histamine	GERD
Pepcid	Histamine	GERD
Gaster	Histamine	GERD
Depakote	Histone Deacetylase	Epilepsy
Singulair	Leukotriene	Asthma
Zoladex	LHRH	Prostate cancer
Atrovent	Muscarinic Acetylcholine	Bronchospasm
Xalatan	Prostanoid Selective FP	Glaucoma
Risperdal	Serotonin	Psychosis
Zyprexa	Serotonin	Psychosis
Imitrex	Serotonin	Migraine
BuSpar	Serotonin	Anxiety
Prozac	Serotonin	Depression
Paxil	Serotonin	Depression
Zoloft	Serotonin	Depression
Effexor	Serotonin	Depression

solely for stem-cell mobilization and not as adjuvant chemotherapy for cancer or HIV-1 infection. AMD3100, as depicted in Figure 7, is a bicyclam used as a mobilization agent for hematopoietic stem cells [80]. AMD3100 exhibited antagonistic activity to proliferation and chemotaxis of acute lymphoblastic leukemia [154], acute myelogenous leukemia [155], ovarian cancer [156], primary brain tumors [157], anaplastic thyroid carcinoma [158], and colorectal cancer [159]. However, cardiotoxic side effects limit its utility as a chemotherapeutic agent [160].

There is a long history of bicyclic amine compounds being used to modulate the activity of GPCRs, several of which are depicted in Figure 8. Zimelidine was one of the earliest selective serotonin antagonists (SSRI) prescribed as anti-depressants [161] although it has been withdrawn from the market for inducing Guillain-Barré syndrome, a form of peripheral neuropathy [162]. Zimelidine, along with citalopram [163-165], represents some of the earliest use of bicyclic amines as chemotherapy, and they continued to play an important role in the pharmacotherapy of depression through the development of subsequent bicyclic SSRIs such as venlafaxine [166-168].

There has been limited success with bicyclam derivatives such as AMD-070 (Figure 7), which recently completed Phase I clinical trials as anti-HIV-1 chemotherapy [169]. There has also been successful development of *in vitro* CXCR4 inhibitors in the benzimidazole [170] and imidazopiperazine [171] analogs of the AMD bicyclams, but these compounds are heavily protein bound and suffer from severe protein shift in antiviral assays [170]. These bicyclam

Figure 7: Anti-CXCR4 compounds. Several compounds have been demonstrated to selectively inhibit CXCR4, targeting the extracellular surface of the receptor and the interaction between CXCR4 and either its ligand CXCL12, or the interaction between CXCR4, gp120, and the V3 loop of HIV-1.

Figure 7: Anti-CXCR4 compounds.
AMD3100

KRH-3955

AMD-070

MSX-122

AMD-070 benzamidazole derivative

T140

KRH-1636

CTCE-9908

Figure 8: Bicyclic aromatic compounds inhibit GPCR activity. Bicyclic compounds are some of the earliest characterized anti-GPCR compounds. These compounds function by inhibiting serotonin receptors.

Figure 8: Bicyclic aromatic compounds inhibit GPCR activity.

derivatives strongly interact with two residues, Asp171 and Asp262, located on the extracellular side of TM4 and TM6, on either end of the CXCL12 binding pocket [172].

Other small molecule derivatives, such as the KRH series, depicted in Figure 7, can inhibit CXCR4 mediated chemotaxis and intracellular calcium mobilization [173], but none of these compounds have entered clinical trials. These compounds share a binding site with the AMD bicyclams, targeting the aspartate residues in the CXCL12 binding pocket [174].

Peptide compounds represent another family of CXCR4 inhibitors with polyphemusin derivatives, such as T140 [175], (outlined below) and CXCL12 peptidomimetics such as CTCE-9908 (Figure 7) [176] which generate significant anti-CXCR4 activity. CTCE-9908 is a 17-amino acid palindromic dimer of the N-terminal region of CXCL12, which has been demonstrated to function as a CXCR4 antagonist and to decrease metastasis of breast cancer cells [177, 178], osteosarcoma cells [176] and melanoma cells [176]. CTCE-9908 has also been found to induce mitotic catastrophe in ovarian cancer cells [179], and induce apoptosis and inhibit tumor proliferation in prostate tumor cells [180]. CTCE-9908 entered phase Ib/II clinical testing and showed efficacy with stable disease in six of twenty patients and no serious adverse effects associated with the drug, however Chemokine Therapeutics Corp, the parent corporation of CTCE-9908 filed for bankruptcy in 2008, leaving any further development in limbo. All of these compounds have been designed targeting the extracellular surface of CXCR4 and their primary objective has been inhibition of interaction with HIV-1.

There are no FDA approved chemotherapeutic or anti-metastatic anti-CXCR4 agents

AMD3100 is the only anti-CXCR4 agent that has been FDA approved, however it is not approved for HIV-1 or cancer chemotherapy, but as a mobilization agent for hematopoietic stem cells. There are currently forty clinical trials involving CXCR4: seventeen of these are studies observing CXCR4 expression in a variety of clinical scenarios, such as rheumatoid arthritis or acute myocardial infarction. CXCR4 expression in response to pharmaceuticals such as atorvastatin or an anti-herpes simplex gel was also assessed. Another ten of these studies are exploring the role of anti-CXCR4 compounds AMD3100 and TG-0054 in stem cell mobilization. Two of the studies involve anti-CXCR4 compound AMD070 and its role as an anti-HIV-1 agent. Two other studies are preliminary safety studies of anti-CXCR4 compounds MSX-122 and ALX-0651. The remaining nine studies concern CXCR4 and anti-CXCR4 compounds AMD3100, BKT140, and BMS-936564 as chemotherapy or as neoadjuvant therapy for leukemia, lymphoma, myelodysplasia, or multiple myeloma. All of the anti-CXCR4 compounds currently in clinical trials target the extracellular surface of the receptor.

Structure based drug design is an effective technique for identifying new pharmaceuticals.

Structure based drug design (SBDD) has had substantial success in the field of drug discovery, with over 50 novel compounds entering clinical trials and multiple FDA approvals [181]. SBDD is an empirical method based on developing

lead compounds either targeting specific structures of a target molecule, mimicking structural features or ligands, or generating new structures derived from established active compounds.

One particularly illuminating example of structure based drug design highlights the development of T140 and its subsequent derivatives by SBDD based on a successful lead compound. T140 structurally originates from a series of antimicrobial 19-mer polypeptide compounds derived from Atlantic horseshoe crab hemocytes called polyphemusins [182]. Initial SBDD work generated a synthetic 19-mer polypeptide based on the polyphemusin sequence which specifically prevents HIV-1 binding, T22, depicted in Figure 9 [183]. Extensive study went into identifying the interaction of T22 with CXCR4 on the cell surface [184-189], and structure-activity relationships of the polypeptide [190, 191]. The N-terminal and C-terminal arginine residues as well as the disulfide rings were identified as essential elements of anti-HIV-1 activity. Aromatic substitution of Trp3 was found to maintain activity of the compound, while the positive charge of the beta turn side chain was found to be essential to strong anti-HIV-1 activity. Through quantitative structure activity relationship (QSAR) analysis, pharmacophores or regions and properties of the T22 molecular structure which were essential to its interaction with CXCR4 were identified and investigated through a series of analog peptide structures [192-194], allowing eventual downsizing of the original 19-mer polypeptide to a smaller 14-mer peptide, T140 [195]. Subsequent structure-activity relationship analysis of T140 [127, 175, 196-200], allowed additional pharmacophore development based on this structure

Figure 9: Polyphemusin derivative structures. Polyphemusins act as specific inhibitors of CXCR4. Structure based drug design was used to refine the structure of the compounds from a 19-mer peptide compound to a 14 residue peptide. Further refinement of the T140 structure produced size reductions in the form of cyclic pentapeptide compounds, and finally small molecule inhibitors of CXCR4.

Figure 9: Polyphemusin derivative structures.

T22

T140

TE14011

Cyclic Pentapeptide T140 Derivative

Tetrapeptide T140 Derivative

Small Molecule T140 Derivatives

[30, 201-204]. Basic amino acid residues of T140 were found to produce non-specific binding and cytotoxicity, and replacement of these residues with Glu and L-citrulline was found to produce selective anti-CXCR4 activity with low cytotoxicity and high anti-HIV-1 activity [201]. Pharmacophore analysis of T140 illustrated that four of the fourteen residues (Arg2, Nal3, Tyr5, and Arg14) were essential to CXCR4 function [196], allowing the 14-mer T140 peptide to be reduced to a cyclic pentapeptide with strong anti-CXCR4 activity. A series of analogs based on this cyclic pentapeptide, TE14011, identified a 4-F-benzoyl derivative with enhanced anti-HIV-1 activity, identifying this group as a novel pharmacophore for CXCR4 antagonism [202]. Further size reduction of the T140 pharmacophore was produced by replacing the Nal-Gly subunit of the cyclic pentapeptide with a γ -amino acid or pseudopeptides cyclized by disulfide and olefin bridges to develop novel cyclic tetrapeptide derivatives of T140 [205]. Subsequent QSAR analysis of the cyclic pentapeptide derivatives of T140 [206], in conjunction with the identification of an enhanced pharmacophore in conjunction with identification of an electro-deficient N-terminal aromatic moiety helped to develop small-molecule inhibitor based on the cyclic pentapeptide derivatives of T140 [204].

Lipid bilayer simulations suggest that T140 binding to CXCR4 is mediated by interaction with a series of residues, Asp171 on TM4, Arg188 and Tyr190 on ECL2, Gly207 on TM5, and Asp262 on ECL3 [200]. Photolabeling studies of radiolabeled T140 identified TM4 as the most important segment of CXCR4 involved in binding [207].

Those of us not fortunate enough to have unlimited access to Atlantic horseshoe crabs need to develop our own lead compounds, and rational methods for identifying new lead compounds and drug development scaffolds are another extension of SBDD. *In silico* screening methodologies such as virtual screening are some of the most promising of these new rational methods for lead compound identification. Virtual screening prescreens vast databases, such as the Zinc compound library a compound virtual database [208], the 2010 version of which contains the refined three dimensional structures of thirteen million commercially available compounds. This prescreening is accomplished by targeting a three dimensional structure to determine those compound which best fit, or dock, into a chosen allosteric binding site or ligand binding site. The theory is that by prescreening and identifying compounds with structurally favorable features, the total volume of compounds needed to identify a new lead inhibitor can be reduced by several orders of magnitude. This database enrichment would allow reduced *in vitro* or *in vivo* screening of compounds.

Virtual screening has been successful in determining new lead compounds for many different targets such as: carbonic anhydrase II [209], Ptp1b [210], estrogen receptor [211], thrombin [211], factor Xa [211], thymidylate synthase [212], retinoic acid receptor [213], farnesyl transferase [214], kinesin [215], hypoxanthine-guanine-xanthine phosphoribosyl transferase [216], DNA gyrase [217], HIV-1 RNA transactivation response element [218], aldose reductase [219], rac GTPase [220], checkpoint kinase-1 [221], alpha amylase [222], BCR-ABL tyrosine kinase [223], protein kinase CK2 [224], and alpha-

reductase [225]. In one documented case virtual screening targeting DNA gyrase identified lead compounds where experimental high throughput screening of libraries was unsuccessful [181, 217].

Docking of compounds into the target structure can occur using a pre-positioned, or rigid, chemical conformation for the screened compounds or by docking compound fragments and constructing the full compound based of energetically favorable positions of the initial fragment placements. Docked compound conformations are then evaluated based on the electrostatic and hydrophobic interactions they would form between the compounds and the target site, which in combination with any identified steric conflicts formed can be used to derive a ranking of the interaction between the compound and the target, or score. There are numerous docking methods employed in virtual screening, the most well reviewed of which include DOCK [226], Autodock [227], GOLD [228], eHiTS [229], Glide [230, 231], Cdocker [232], MCDock [233] and Surflex [234], and while these docking methods employ their own scoring methodologies, several outside scoring programs such as PLP1 [235, 236], PLP2 [237], Ludi [238], PMF [239], FlexX [240], SCORE [226], and Ligandfit [241] also exist. Unfortunately, no docking or scoring algorithm has been determined to be superior than others, so common practice has been to determine consensus scoring for docked ligands using two or more docking programs to identify bound molecules identified by one than one scoring methodology [181, 242, 243]. The DOCK [226], Autodock [227], and Surflex [234], have the advantage of being designed for performance with large database virtual screening, allowing efficient

flexible docking and evaluation of docked conformations at high volumes. One important facet of this process is the basis of the virtual screening target itself: the target of the virtual screening protocol. In our experience with virtual screening we have found that the strongest determinant of identification of active compounds instead of false positives has been a well defined target structure, and strong experimental and *in vivo* evidence governing the development of the targeted site for the docking.

Our prior experience with the Surflex-Dock [234, 244] docking program has demonstrated that this program is highly amenable to high-throughput virtual screening. In unpublished research, benchmarking Surflex-Dock [234, 244] with three other common molecular docking programs, MolDock [245], DOCK [226], and Autodock [227], we found that Surflex-Dock consistently reproduced predicted compound binding modes as illustrated by compound/receptor crystal structures to a greater degree than the other docking programs. The Surflex-Dock [234, 244] method is centered on development of a protomol, a pseudo-molecule used to virtually define the binding site targeted on the protein structure. This protomol can be generated based on available ligands, inhibitors or partner proteins or development can be designed based on a defined series of target residues composing an active site. Once the protomol is developed, Surflex-Dock conducts flexible docking by fragmenting the compounds into components and docking the fragments into the defined site, constructing the compound in the active site based on favorable positions for each of the compound fragments. Multiple poses of each compound are returned to the user

and assigned a score based on binding affinity, degree of inappropriate penetration into the protein, and internal structural conflicts.

Receptor inhibition can be in the form of agonists, antagonists, and inverse agonists.

Compounds targeting a receptor can alter signaling in four primary patterns: agonism, partial agonism, antagonism, and inverse agonism. In contrast to agonists and superagonists, which bind and activate an inhibitor producing equivalent or greater signaling responses, partial agonists bind the receptor and result in signaling, but the amplitude of the response is much less than that of the full agonist. Effectively, partial agonists cause the receptor to shift into a partially active conformation, incapable of full signal transduction, but also not in fully inactive states. The degree of signaling can vary with the partial agonist, and very weak partial agonists, such as AMD3100 [127], can almost completely abolish receptor signal transduction by inducing alterations of the intracellular TM regions and the intracellular loops [200]. In contrast to partial agonists, inverse agonists do not induce the active conformation of the receptor or result in signal transduction. These compounds interact with receptors with some basal level of activity, and upon binding, reduce the constitutive signaling of a receptor. These compounds, such as T140 [127], preferentially stabilize the inactive, or native in the case of CXCR4, conformation of the receptor [200].

CXCR4 as a Virtual Screening Target.

CXCR4 represents an interesting potential target for virtual screening. From our prior descriptions of CXCR4 signaling there are five potential

mechanisms for inhibiting signal transduction. The natural ligand could be intercepted or manipulated in such a manner that it does not interact with the extracellular surface of CXCR4 as illustrated by Target 1 in Figure 10. Conversely, the extracellular surface where CXCL12 is bound could be manipulated to prevent ligand binding as illustrated by Target 2 as per Figure 10. This has been the approach of prior attempts at developing anti-CXCR4 agents, and has been the binding site of AMD3100 and T140. Next, the TM helix could be restrained preventing movement necessary to transmit the signal from the extracellular surface to the intracellular surface as illustrated by Target 3 in Figure 10. The intracellular surface itself could be manipulated to either restrain the protein's intracellular surface in an inactive conformation, or prevent G-proteins from interacting with the surface as illustrated by Target 4 in Figure 10. Finally, the G-protein heterotrimer could be manipulated in such a manner that it does not interact with the intracellular surface of CXCR4 as illustrated by Target 5 in Figure 10. Although the majority of anti-CXCR4 inhibition efforts have been directed at just the extracellular surface of the receptor, there is a precedent for targeting the intracellular components of the signaling process. YM-254890 (Figure 11) is a specific G α q/11 inhibitor designed to prevent acute thrombosis and neointima formation by preventing signal transduction at the G-protein level rather than at the GPCR level like the P2Y₁₂ receptor antagonists ticlopidine and clopidogrel [246-251]. M119 (Figure 11) is a G β γ inhibitor which modulates the μ -opioid receptor to act as an anti-nociceptive.[252-254] This compound has also been found to inhibit metastasis of MDA-MB-231 cells [255]. Covalent

Figure 10: There are 5 potential structural targets for inhibiting CXCR4

signaling: 1) Targeting the ligand CXCL12 to prevent ligand/receptor interaction, 2) Targeting the Extracellular ligand binding site as in AMD3100 and T140, 3) Stabilizing the transmembrane helices to prevent ligand induced conformation changes, 4) Targeting the Intracellular region of CXCR4 at the G-Protein interface, 5) Targeting the G-protein to prevent GPCR-G-protein interaction.

Figure 10: There are 5 potential structural targets for inhibiting CXCR4 signaling

Figure 11: Structures for G-protein inhibitors. Specific inhibitors of G-proteins have been designed. YM-25489 is an anti-thrombotic agent that targets the Gαq/11 subunit. M119 is an anti-nociceptive agent targeting the Gβγ subunit. Covalent modification of Gβγ inhibits Gβγ function.

Figure 11: Structures for G-protein inhibitors

M119K

YM-254890

Selenocystamine

modification of the G β γ subunit by selenocystamine (Figure 11) has also been found to selectively inhibit G β γ function [256].

A novel crystal structure for CXCR4.

Five novel crystal structures for CXCR4 were released on November 19, 2010 [257]. These structures were developed in a manner similar to the crystal structures for the β 2-adrenergic receptor and the α 2A adenosine receptor, with replacement of ICL3 with T4L, as well as a stabilizing L125W mutation, and a C terminal tail truncation. Testing of the crystallization protein demonstrated that the modified CXCR4 was capable of binding a bis (imidazolylmethyl) amine (BIMA) ligand, although at a lower binding affinity than wild-type CXCR4. Calcium flux assays using the crystallization protein were ineffective, which the authors suggested was due to ineffective G-protein binding due to C-terminal truncation and replacement of ICL3. Structures were generated as co-crystals with either IT1t, an isothioureia derivative which has been demonstrated to inhibit interaction between CXCR4 and HIV-1 [258], or CVX15, a 16-mer peptide derivative of T140 with demonstrated anti-metastatic activity by CXCR4 inhibition [30]. Three constructs of CXCR4, CXCR4-1, CXCR4-2, and CXCR4-3, were cocrystallized with IT1t with additional cocrystals of the CXCR4-2/IT1t complex and CXCR4-2/CVX15.

The primary description of the structure was for the highest accuracy (2.5 angstrom, monomer, and bound to IT1t) structure, with the other four conformations being described in relation to it. Significant differences between the primary CXCR4 structure and other GPCRs involved significant structural

deviations by TM1 and TM4 from their expected positions in the β 2-adrenergic receptor and the α 2A adenosine receptor. A tighter helical turn at Pro 92 produced an additional 120 degree rotation in TM2, which produces a different series of residues facing the binding pocket. TM6 is characterized by a sharp kink at residue 254, producing a deviation of the extracellular end of the helix. Characterization of the extracellular region links the base of the N-terminal segment to the cytoplasmic side of TM7 by a disulfide bond and an additional disulfide linkage between TM3 and ECL2. The ligand binding pocket is larger, more open, and more superficial than in comparable GPCRs. Characterization of the intracellular surface indicates that positioning of the intracellular side of the TM helices is conserved between the CXCR4, β 2-adrenergic receptor, α 2A adenosine receptor, and rhodopsin structures, while TM7 was one turn shorter than expected.

Differences were noted between the five crystal structures. Construct CXCR4-3 contained a T240P mutation which abolished signal transduction capability, which was structurally characterized to contain a disruption of the intracellular side of TM6 corresponding to a region between Gly231 and Pro240. Constructs bound to inhibitors were found to have a larger, more open and more shallow cavity on the extracellular surface compared to other GPCRs, and constructs were found to have alteration of the position of TM5 by 0.5 angstroms from the center of the receptor, TM6 1.5 angstroms towards the center of the receptor and TM7 by 1.5 tangential to the receptor.

Subsequent investigation of these crystal structures by Indira Chaudhuri, a post doctoral fellow in the Trent laboratory, has identified several interesting factors concerning the conformation of the receptors and their utility for virtual screening. Comparison of the crystal structure to the lipid bilayer structure of CXCR4 in Figure 12 reveals that there is a significant truncation of the N and C terminal regions, as well as disruption of ICL3 in the crystal structure of CXCR4. In addition, an extreme kink in TM6 corresponding to residue 254 producing a sharp deviation of TM6 and TM7. Further, unlike in the homology model and in the other available GPCR structures, ECL2 is displaced upwards and outwards, extending between TM1 and TM2. Differences are also noted in the N119S structures, in that the N terminal region, rather than being more compact in the N119S structure seems to be almost completely disordered in the homology model based on the crystal structure.

Several features of the CXCR4 crystal structure represent causes for concern if it is to be used as a basis for SBDD and virtual screening in these studies. The structure was released on November 19, 2010 [257], well after the experiments involved in this structure were completed. Truncations of the C-terminal segment and replacement of ICL3 with T4L mean that essential intracellular features of the receptor would be absent, precluding investigation of the intracellular surface of the protein. Further, the absence of these structures could artificially influence movement or rigidity of the TM helix regions especially TM5, TM6 and TM7. Truncation of the C-terminal tail could cause the remaining portion of the sequence to adopt an unfavorable position, such as the novel kink

Figure 12: Comparison of the CXCR4 x-ray crystal determined structure of CXCR4 and the CXCR4 homology model. A) On the left, the CXCR4 crystal structure, (PDB ID: 3OE0) is presented in blue ribbon. ICL3 is replaced with T4L protein on the cytoplasmic surface of the receptor. Homology model of CXCR4 illustrated on right in green ribbon presentation. B) CXCR4 crystal structure, (PDB ID: 3OE0) is presented in yellow ribbon for comparison with the homology model of CXCR4 on the right in green ribbon. Truncation of the N and C termini as well as loss of the ICL3 loop region are visible in the CXCR4 crystal structure. Also, note the displacement of TM6, TM7 and the altered conformation of the ECL2 loop.

Figure 12A: Comparison of the CXCR4 x-ray crystal determined structure of CXCR4 and the CXCR4 homology model.

Figure 12B: Comparison of the CXCR4 x-ray crystal determined structure of CXCR4 and the CXCR4 homology model.

at residue 254, by failure to effectively anchor TM7 on the intracellular side of the membrane. This could allow TM7 to unnaturally shift towards the extracellular side of the membrane distorting ECL3 and TM6. In addition, the author's own testing indicated that there were ligand binding deficiencies by the crystallization protein, suggesting irregularities in the structure of the ligand binding pocket and by extension the extracellular loop structure. Thus, perhaps rhodopsin based homology modeling of this receptor would be a more effective in virtual screening until more representative crystal structures are completed.

Dissertation Overview

This dissertation investigates several potential inhibitors of CXCR4 identified by virtual screening targeting the ECL and ICL regions of this receptor. We hypothesize that we will be able to effectively inhibit CXCR4 mediated metastasis by targeting multiple sites of CXCR4. Compounds targeting the extracellular portion of CXCR4 will be investigated in the second chapter. Based on research by Dr. Kelby Napier, we evaluate compounds targeting the T140 binding site of CXCR4 for their capacity to inhibit CXCL12-induced chemotaxis. We successfully identify one compound, ECLVS14, which inhibits chemotaxis with an IC₅₀ value of 5 μ M, and is highly selective for CXCR4 without significant cytotoxicity. Subsequent QSAR analysis of the structure of this inhibitor reveals the importance of the 1-[bis (phenylmethyl) amino] methyl moiety and the fact that electronegative modifications of the terminal benzene enhance activity. Subsequent Molecular dynamics simulations of the compound in complex with

CXCR4 reveal that the compound induces significant modifications of the receptor structure.

The third chapter of this dissertation will investigate novel compounds targeting the intracellular region of CXCR4 responsible for interaction with Gai. Based on *in silico* studies by Dr. Kelby Napier, we will utilize virtual screening to identify compounds which will selectively inhibit CXCR4. Our screens will identify thirty-seven compounds, twenty four of which will inhibit CXCL12-mediated chemotaxis at 50 μ , for an exceptionally high initial hit rate of 65%. Ten of these compounds will inhibit chemotaxis at IC50 values of 10 μ M or less and will be highly selective for CXCR4. These compounds displayed relatively no cytotoxicity, and 3 of the compounds significantly inhibited tumor cell proliferation. Molecular dynamics revealed that the compounds acted by a variety of mechanisms with some compounds restricting the receptor to the native conformation, while others induced an intermediary configuration, and still others were incapable of interacting with the native conformation of the receptor.

The fourth chapter of this dissertation involves validation work for the compounds identified in the previous two chapters. We will use *in silico* docking methods to predict the binding mode of the compounds with the receptor. Based on these docking simulations, we will conduct additional molecular dynamics simulations to examine the effect of mutating these residues on receptor structure. Mutagenic analysis will also investigate the effect of mutation on receptor function. Finally, three of the most active compounds from the previous chapters will be selected and tested in an *in vivo* anti-metastatic animal

model, successfully demonstrating the anti-metastatic potential of these compounds.

In total this work demonstrates that structure based drug design utilizing *in silico* analysis in combination with *in vitro* and *in vivo* testing can be utilized to develop novel lead compounds which can function as anti-metastatics.

CHAPTER II

**STRUCTURE BASED DRUG DESIGN OF EXTRACELLULAR LOOP
TARGETED COMPOUNDS BY QUANTITATIVE STRUCTURE ACTIVITY
RELATIONSHIP ANALYSIS AND MECHANISTIC STUDIES.**

Introduction

CXCR4 has been extensively linked to cancer metastasis (see recent reviews [259-263]). CXCR4 expression by the tumor cell allows for tumor cell migration, and homing of neoplastic cells to sites of where non-malignant stromal cells express CXCL12 [37]. This implies that tumor cell metastasis is not a random event, but directed by the expression of chemokine receptors and adhesion molecules on the neoplastic cells, and their respective ligands in the target organs [9, 16]. Elevated CXCR4 expression is also associated with increased circulating tumor cells in non-small cell lung cancer [264], as well as poor outcomes in advanced cases of this disease [265].

It is also important to note that while CXCR4 expression has been correlated with increased metastasis of neoplastic cells [9], reductions in CXCR4 expression, or knockdown, can produce the opposite effect. Knockdown of CXCR4 expression has been demonstrated to reduce proliferation and invasion of ovarian cancer cells [266], and metastasis in breast cancer cells [267] and prostate cancer cells [268]. AMD3100, an anti-CXCR4 inhibitor, has consistently inhibited CXCR4 signaling and migration in *in vivo* experiments in a variety of cell

lines [269-271]. It has also exhibited antagonistic activity to proliferation and chemotaxis of acute lymphoblastic leukemia [154], acute myelogenous leukemia [155], ovarian cancer [156], primary brain tumors [157], Anaplastic thyroid carcinoma [158], and colorectal cancer [159].

AMD3100 is a bicyclam compound as depicted in **Figure 13**. Originally identified as an impurity on a series of anti-HIV-1 compounds, AMD3100 was found to selectively antagonize viral entry [272] by T-tropic strains of the HIV-1 virus through inhibition of CXCR4 [273-275]. Although AMD3100, successfully completed initial phase I clinical trials [276], poor *in vivo* Anti-HIV-1 activity and cardiotoxic side effects halted phase II trials and limit AMD3100's utility as an anti-HIV-1 therapy [160].

In spite of its cardiotoxic side effects, AMD3100 is currently the only FDA-approved CXCR4 inhibitor on the market, and is only indicated for use in stem-cell mobilization and not as adjuvant chemotherapy for cancer or as therapy for HIV-1 infection. There has been limited success in the development of bicyclam derivatives such as AMD-070 (**Figure 13**), which recently completed Phase I clinical trials as anti-HIV-1 chemotherapy [169]. There has also been successful development of *in vitro* CXCR4 inhibitors in the benzimidazole (**Figure 13**) [170] and imidazopiperazine [171] analogs of the AMD bicyclams as anti-HIV-1 agents, but these compounds are heavily protein bound and suffer from severe protein shift in antiviral assays [170] suggesting that the drug concentration available in the systemic circulation would be greatly reduced by circulating albumin proteins. These bicyclam derivatives strongly interact with two residues, Asp171 and

Figure 13: Anti-CXCR4 compounds. Illustrated compounds specifically interfere with CXCR4 activity and are targeted at the extracellular surface of the CXCR4 receptor.

Figure 13: Anti-CXCR4 compounds.
AMD3100

KRH-1636

AMD-070

KRH-3955

AMD-070 benzimidazole derivative

CTCE-9908

T140

Asp262, located on the extracellular side of transmembrane helix 4 and 6, on either end of the CXCL12 binding pocket [172]. In reference to the potential mechanisms of inhibiting CXCR4 depicted in Figure 10, these compounds target Target 2, the extracellular interface of the receptor.

Other small molecule derivatives, such as the KRH compounds (Figure 13) KRH-1636 and KRH-3955 are selective CXCR4 inhibitors. KRH-1636 was extensively investigated as an anti-HIV-1 entry inhibitor, and was identified as selectively binding CXCR4 [277], but development was halted due to poor oral bioavailability. KRH-3955 produced greater oral bioavailability, while inhibiting HIV-1 entry as well as CXCL12 binding and intracellular calcium mobilization [173]. The compound was found to bind in the extracellular region of the receptor and interact with all three extracellular loops [278], sharing a binding site with the AMD bicyclams, targeting the aspartate residues in the CXCL12 binding pocket [174]. Like the bicyclam derivatives, these compounds also target the extracellular surface of the receptor, Target 2 in Figure 10. None of the KRH compounds have entered clinical trials.

Peptide compounds such as polyphemusin derivatives T140 (Figure 13) [175] and CXCL12 peptidomimetics such as CTCE-9908 (Figure 13) [176] generate significant anti-CXCR4 activity. CTCE-9908 has been demonstrated to function as a CXCR4 antagonist and decrease metastasis of breast cancer cells [177, 178], osteosarcoma cells [176] and melanoma cells [176] *in vitro* and in tail vein injection studies of mouse lung metastasis. CTCE-9908 presents an especially compelling proof of principle for anti-CXCR4 agents in that it

demonstrates that inhibition of CXCR4 can effect a reduction in tumor cell metastasis.

The peptide CXCR4 inhibitors have been designed targeting the extracellular surface of CXCR4. CTCE-9908 mimics the ligand binding mechanism of CXCL12, primarily interacting with the N-terminus of CXCR4. Lipid bilayer simulations suggest that T140 binding to CXCR4 is mediated by interaction with a series of residues, Asp171 on transmembrane helix 4, Arg188 and Tyr190 on extracellular loop 2, Gly207 on transmembrane helix 5, and Asp262 on extracellular loop 3 [200]. Photolabeling studies of radiolabeled T140 identified transmembrane helix 4 as the most important segment of CXCR4 involved in binding [207].

SBDD is an empirical method of drug design based on developing lead compounds either targeting specific structures of a target molecule, mimicking structural features or ligands, or generating new structures derived from established active compounds. One of the earliest examples of utilization of this method for targeting CXCR4 activity utilized synthetic peptides based on the transmembrane domains of CXCR4 [279]. Peptides correlating to TM2 were capable of completely abolishing CXCL12-induced calcium mobilization at 0.2 μ M. Elimination of the acidic apex and basal amino acids was demonstrated to improve potency 10-fold, while replacement with positively charged lysine residues improved potency 100-fold.

As seen in the description of the evolution of the polyphemusin compounds in Chapter 1, drug discovery and development is an iterative,

cyclical, and evolutionary process, illustrated in **Figure 14**. In the first step of this process, initial lead compounds are identified through a screening assay. Compounds which demonstrate activity in the screening assay are then further characterized through *in vitro* biological assays to identify their impact on cell signal transduction, homeostasis, cytotoxicity, and possible identification of their molecular mechanism. Active compounds are then subsequently modeled, and derivatives are generated to test a hypothesis of interaction based on the model. The activity of initial derivatives in *in vitro* assays is used to perform quantitative structure activity relationship analysis and determine the essential elements of the active compound and methods by which activity could be increased. Compounds are then tested in animal models for toxicity and efficacy, and subsequent generations of derivatives are generated to test hypotheses of interaction.

We will be applying this model of SBDD to designing inhibitors targeting the extracellular surface of CXCR4. The initial step of identifying a lead compound was carried out in previous research by Dr. Napier. His work used a N119S structure of XCR4 derived from homology modeling and developed through solvated lipid bilayer molecular dynamics, as a target for identifying anti-CXCR4 compounds targeting the extracellular surface of the receptor, at the binding site for T140 [200]. N119S was chosen as the original virtual screening target because the inverse agonist T140 was one of the docking targets, and this inverse agonist could not bind to the inactive native conformation of the receptor.

Figure 14: Structure-based drug design is an iterative and cyclical process.

After initial identification of a structural basis for pathology and identification of an active site, initial compounds are screened for toxicity and functionality. The activity of functional compounds is then rationalized by *in silico* methods.

Compounds will then be tested in *in vivo* models. Successful lead compounds will then be subjected to QSAR analysis and derivation and testing of derivatives in functional screens will identify active pharmacophores for further testing rationalization and derivation.

Figure 14: Structure-based drug design is an iterative and cyclical process.

An initial rigid docking protocol, DOCK [226], was used to prescreen the Zinc library database, and the FlexX (Tripos) program was used to flexibly dock the top 2100 compounds from the rigid body analysis, with the top 100 compounds from the FlexX analysis scored and ranked by energy (DOCK) and Cscore (FlexX). Dr. Napier was then able to use chemotaxis experiments to identify 10 compounds which inhibited CXCR4-mediated chemotaxis. One of these compounds, designated compound 14, selectively inhibited CXCR4-mediated chemotaxis, but not CCR5-mediated chemotaxis. Because of the selective anti-CXCR4 activity of compound 14, it was chosen as the lead compound to be used as a scaffold for development of subsequent derivatives.

In this chapter, we will describe the continuation of the SBDD process by further characterizing the anti-CXCR4 activity of compound 14 as well as the other identified CXCR4 inhibitors. We will further characterize the nature of compound 14's activity on CXCR4 by designing a series of inhibitors to investigate important regions of the compound 14 structure. Finally, the interaction between compounds with anti-CXCR4 activity will be investigated through molecular dynamics simulations to analyze the interaction between identified inhibitors and CXCR4.

This chapter will highlight the following results: Compound 14 (renamed ECLVS14) selectively inhibits CXCR4-mediated chemotaxis, but not intracellular calcium flux without significant cytotoxicity. QSAR will be used to identify terminal benzene modifications in conjunction with acetamide derivation of the ECLVS14 structure as a potential mechanism for increasing anti-CXCR4 activity. Molecular

dynamics simulations demonstrated that ECLVS14 induces a more closed extracellular position of the extracellular loops (ECL) but an intermediately open position for the intracellular loops, suggesting that ECLVS14 could act as a weak-partial agonist. In total, this work will provide an extensive characterization of ECLVS14 through *in vitro* analysis and investigation of multiple derivative structures, as well as dynamic simulation of compound binding, producing significant progress in the SBDD process.

Methods

Compound Selection

The prior virtual screening studies by Napier *et al* identified ten compounds as both specific and non-specific inhibitors of CXCR4 mediated chemotaxis, which will be referred to as ECLVS compounds after “ExtraCellular Loop Virtual Screen”. ECLVS14 was the most promising identified compound with specific inhibitory activity of CXCR4-mediated chemotaxis. The other nine ECLVS compounds also exhibited significant inhibition of CCR5-mediated chemotaxis. Of the ten compounds displaying activity towards CXCR4, all were selected for testing except ECLVS48 and ECLVS88. As these compounds displayed significant cytotoxicity in the original research, they were excluded from these studies.

As ECLVS14 displayed significant specific inhibition of CXCR4 activity, it was selected to perform QSAR analysis in order to identify modifications that could alter the activity of the compound in respect to CXCR4. Twenty three ECLVS14 derivatives were synthesized by our organic chemistry collaborators led by G. B. Hammond, Dept. of Chemistry, University of Louisville, and prepped for analysis in human CXCR4 expressing cell lines, which will be referred to with the prefix EVS14 and a numerical designation. These derivatives explored the role of the bis (phenylmethyl) amino methyl moiety and effects of modification of this group, as well as exploring the effect of lengthening the hydrazine-linked

alkyl moiety or converting it to an acetamide and modifying the terminal benzyl group.

Fourteen compound EVS14-2 derivatives were also synthesized for testing, with an EVS14-2 prefix and an alphanumeric designation. These compounds retained ECLVS14-2's conversion of the hydrazine-linked alkyl to a propylhydrazine and demonstrated modification or replacement of the bis (phenylmethyl) amino methyl moiety.

Transfection and generation of RBL-2H3 and 300.19 stable cell lines expressing hCXCR4-GFP

To fully evaluate the effects of the potential inhibitors we stably transfected overexpressing CXCR4 cell lines in collaboration with Dr. Haribabu Bodduluri. CXCR4-GFP was generated in Dr. Bodduluri's lab by in frame cloning of the entire coding region of CXCR4 ahead of the GFP coding regions. RBL-2H3 and 300.19 cells were obtained from our co-investigator, Dr. Haribabu Bodduluri. RBL-2H3 cells are maintained as monolayer cultures in DMEM supplemented with 15% FBS, 2 mM L-glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin. 300.19 cells are maintained as suspension cultures in RPMI supplemented with 10% FBS, 100 U/ml penicillin and 100 µg/ml streptomycin, and 55 µM of β-mercaptoethanol. Transfections are performed by electroporation. The RBL-2H3 and 300.19 stable single cell lines expressing GFP were derived with pEGFPN2 vectors. The stable cell lines are maintained in the presence of 1 mg/ml of G418, an aminoglycoside antibiotic which inhibits polypeptide synthesis [280] in the absence of a resistance gene present on the

pEGFPN2 vector. Cells are cultured in complete medium for 24 hrs after transfection and used for microscopy. The cell surface expression of hCXCR4 is also determined by incubating receptor or parental cells with 12G5 Ab followed by PE labeled secondary Ab, and analyzed by FACS (BD Biosciences).

300.19 lymphocyte derived cells will be used for the majority of described experiments. This line was selected because as a non-adherent cell line, greater cell populations could be acquired quickly, allowing higher-throughput screening of the investigated compounds. Also, as a lymphocyte derived cell line, these cells retained stronger basal chemotactic capabilities than even highly metastatic tumor derived cell lines. In a chemotaxis experiment this would make the difference between normalizing migration to a control value of 150-300 cells per counted field in the 300.19 cell line, or 150-300 migrated cells total.

Chemotaxis analysis

Chemotaxis analysis for suspension 300.19 cells expressing hCXCR4-GFP or BLT1 was performed as previously described [151, 281]. In brief, 24 well Transwell chambers (Corning Costar, Cambridge, MA) with pore diameter 8.0 μm were used. CXCL12 (Peprotech) was added to chemotaxis buffer (RPMI media + 10% FBS + 1 mM Hepes) and added to the lower chamber. Cell variability required that CXCL12 concentration be optimized for each run, and results were normalized to the CXCL12 alone control. Inhibitor was then added to the CXCL12 + Chemotaxis buffer solution. 1,000,000 Cells were resuspended in Chemotaxis buffer and added to the upper chamber. Cells were then allowed to migrate for 3 hours at 37 °C. Migrated cells in the bottom chamber were then

counted via hemocytometer. Each run featured positive and negative controls run in duplicate and each compound was run in duplicate. Negative controls lacked CXCL12 in the bottom chamber, while positive controls included both CXCL12 alone chambers and CXCL12 + DMSO to account for the suspension medium of the inhibitors.

Cell signaling analysis

Cell signaling by CXCR4 was investigated by assessment of intracellular calcium mobilization. Intracellular calcium mobilization was performed using both 300.19 cells as described previously [149-151]. Briefly, cells were suspended in loading buffer (1xPBS +1% FBS) and loaded with INDO-1 for 30 minutes at 37 °C. Cells were then washed with washing buffer (1 x PBS +0.1% FBS) and resuspended in assay buffer (1XHBSS +1%FBS). Cells were then loaded into fluorimeter cuvettes (Sigma-Aldrich) and assessed for fluorescence by FACScaliber (BD Biosciences). Cells were allowed to equilibrate in the chamber for 60 seconds, after which compound was added. Cells were allowed to equilibrate for another 60 seconds, after which response to CXCL12 addition was assessed. Cell variability required that CXCL12 or other chemokine concentration be optimized for each run, and results were normalized to the CXCL12 or other chemokine alone control. Positive and negative controls and compounds were assessed in duplicate for each run. Negative controls were assessment of untreated cells, while positive controls included CXCL12 alone and CXCL12 + DMSO samples.

Class A GPCR selectivity for non-family member BLT1

The selectivity of the compounds for CXCR4 was assessed by investigating their effect on the Class A GPCR BLT1. Assays for chemotaxis and intracellular calcium mobilization, as described earlier, were used to assess compound activity in 300.19 cells expressing BLT1 [149] grown in Dr. Bodduluri's laboratory.

Determination of nonspecific cytotoxicity

The potential cytotoxic effects of the compounds were investigated in our own hCXCR4-GFP expressing 300.19 cells. Cytotoxicity was investigated using a modified cytotoxicity protocol [282]. In brief: 300.19 cells are plated into a 24 well plate at 100000 cells/ml in a total volume of 600 μ l and exposed for 3 hours to DMSO, inhibitor compounds or camptothecin (positive control). Cells were then collected by centrifugation at 8000xg, washed with PBS, and resuspended in 0.2 ml PBS + 0.1% BSA. Cell suspension was mixed with 2 μ l 100 μ g/ml PI (Sigma) incubated for five minutes at room temperature in the dark and then analyzed by FACScaliber (BD Biosciences). PI emission is detected through a 600 nm long-pass filter and a 615 nm short-pass filter on FL2 channel, counting 10,000 events per sample.

Compound Screening And Dose Response Analysis

Compounds identified in the Napier virtual screens (ECLVS) and the subsequent ECLVS14 derived compounds were tested *in vitro* using the above assays. CXCR4 response to the compounds was initially evaluated at a screening concentration of 50 μ M. 50 μ M is a high initial screening concentration

for a drug discovery effort when concentrations of 10 to 20 μM are typically used [173, 283]. However, for a conventional compound screening protocol, the compounds are usually derived from structurally similar lead compounds, for which there is established EC₅₀ data. Screening of compounds exhibiting activity at 50 μM or derivatives at 10 μM underwent dose-response assays to eliminate any potential false-positives produced by the elevated screening concentration.

Molecular dynamics modeling of CXCR4 and ECLVS complexes

Complexes of active ECLVS compounds and CXCR4 were investigated by molecular dynamics. The highest scoring conformations of the ECLVS compounds were utilized as a guide to initial placement of the inhibitor. As the initial virtual screens for the active compounds were targeted at the N119S conformation of CXCR4 and the interface surface of Gai, the compounds were re-docked into the native structure using Surflex-Dock [234, 244]. A novel protocol was generated for the native CXCR4 structure using the residues identified in Table 5, which adequately provided docking coverage to the N-terminus and extracellular loops two and three. Surflex-Dock [234, 244] was used to dock ECLVS compounds with anti-CXCR4 activity at 10 μM or less into the defined extracellular surface of the native structure of CXCR4. All compounds were successfully docked into the structure. The top scoring conformation for each compound was utilized in generating the CXCR4 compound complexes.

Fully solvated lipid bilayer molecular dynamics simulations of native CXCR4 bound with active ECLVS compounds were performed using the AMBER suite of programs using the parm99 force field [284]. Restraints were initially

Table 5: Protomol generated for docking of ECLVS and EVS14 compounds into native structure of CXCR4. Surflex-Dock [234, 244] will be used to generate a novel residue based protomol, based on the amino acids within 8 angstroms of the T140 binding site residues, Asp171 on TM4, Arg188 and Tyr190 on ECL2, Gly207 on TM5, and Asp262 on ECL3 identified in lipid bilayer simulations [200].

Table 5: Protomol generated for docking of ECLVS and EVS14 compounds into native structure of CXCR4.

Amino Acid	Residue
Asn	176
Val	177
Ser	178
Glu	179
Ala	180
Cys	186
Asp	187
Arg	188
Phe	189
Tyr	190
Pro	191
Asn	192
Leu	194

used to guide conformation of the receptors as previously described in [285]. In brief, an initial 5000 step steepest decent minimization was conducted with Protein and compound restrained ($100 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$), followed by another 5000 step steepest decent minimization with no protein or compound restraints. This was followed by a 25 ps MD ($T=100 \text{ K}$) holding protein fixed ($100 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$), and another unrestrained 5000 step steepest decent minimization without restraints. The temperature of the system was then increased from an initial point of 100K to 300K while maintaining $100 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ restraints. An additional three 25 ps steps of minimization were then conducted to gradually decrease the restraints from $100 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ to $10 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ and finally to $1 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$. Ten production runs of 500 ps were then conducted free of restraints for a 5 ns total production run.

Results

ECLVS compounds significantly inhibit CXCR4 mediated chemotaxis.

Our goal was to assess the identified ECLVS compounds targeting the extracellular surface of CXCR4 in order to discover novel small molecules which could selectively inhibit receptor function and be used as scaffolds for development of clinically active anti-metastatic compounds. The ECLVs compounds are depicted in **Figure 15**, and listed in **Table 6**. We sought to evaluate the compounds by chemotaxis assays which quantitatively examined the capability of cells to migrate towards a ligand gradient. Initial screening of the eight ECLVS compounds at 50 μ M identified two compounds which significantly inhibited chemotaxis as depicted in **Figure 16**. Results were normalized to the CXCL12 alone control. In contrast to treatment with CXCL12 alone treatment with DMSO, the vehicle used to suspend the compounds, induced no chemotaxis. At the initial screening concentration of 50 μ M, ECLVS14 reduced CXCL12-induced chemotaxis by 79% while ECLVS58 reduced CXCL12-induced chemotaxis by 94%. Dose response analysis was used to determine the concentration of the active compounds ECLVS14 and ECLVS58 where chemotaxis was inhibited by 50% in reference to the CXCL12 alone control (IC₅₀). Results of the dose response analysis are depicted in **Figure 17**. ECLVS14 inhibited chemotaxis with an IC₅₀ value of 5 μ M, while ECLVS58 had an IC₅₀ value of 20 μ M.

Figure 15: ECLVS compounds. Compounds identified by virtual screening using Ludi [238] molecular docking software targeting the T140 binding site of the extracellular surface of CXCR4.

Figure 15: ECLVS compounds.

ECLVS14

ECLVS58

ECLVS59

ECLVS43

ECLVS44

ECLVS72

ECLVS55

ECLVS82

Table 6: ECLVS compounds. Compounds from the ECLVS series are listed below with chemical name, supplier, and molecular weight.

Table 6: ECLVS compounds.

Compound	NAME	Supplier	MW
ECLVS14	(Z)-2-(((E)-benzylideneamino)oxy)-N'-(1-((dibenzylamino)methyl)-2-oxoindolin-3-ylidene)acetohydrazide	TimTec Stock Library	531.6
ECLVS43	methyl 4-(2,3-di(pyridin-2-yl)quinoxaline-6-carboxamido)benzoate	ChemBridge:	461.47
ECLVS44	N-((3-(oxazolo[4,5-b]pyridin-2-yl)phenyl)carbamothioyl)cinnamamide	ChemBridge:	40.45
ECLVS55	N-(4-(2,5-dioxo-3-((pyridin-3-ylmethyl)amino)pyrrolidin-1-yl)phenyl)acetamide	ChemBridge:	339.36
ECLVS58	(3E,5E)-1-methyl-3,5-bis(pyridin-3-ylmethylene)piperidin-4-one	ChemBridge:	291.35
ECLVS59	(E)-3-amino-N'-cyano-4-(4-methoxyphenyl)-6-phenylbenzo[b]thiophene-2-carboximidamide	ChemBridge:	398.48
ECLVS72	4-((5,7-bis(trifluoromethyl)-1,8-naphthyridin-2-yl)oxy)-N-(pyridin-3-ylmethyl)benzenesulfonamide	Bionet Research	528.43
ECLVS82	N,N'-(1,10-phenanthroline-2,9-diyl)dibenzamide	Specs	418.45

Figure 16: ECLVS Compounds Inhibit CXCR4 Mediated Chemotaxis.

Compounds ECLVS14 and ECLVS58 Significantly inhibit CXCR4-mediated chemotaxis in response to CXCL12. DMSO alone did not significantly induce chemotaxis in hCXCR4-GFP expressing 300.19 cells.

Figure 16: ECLVS Compounds Inhibit of CXCR4 Mediated Chemotaxis.

Figure 17: IC50 of Active Inhibitors of CXCR4 Mediated Chemotaxis.

Concentrations which inhibited 50% of chemotaxis in relation to CXCL12 alone signaling (IC50) were determined for active ECLVS compounds. Vertical axis is presented in terms of IC50 of chemotaxis in μM . Of the original ECLVS compounds, only ECLVS14 had an IC50 value less than 10 μM .

Figure 17: IC50 of Active Inhibitors of CXCR4 Mediated Chemotaxis.

We then sought to determine the specificity of the ECLVS compounds for CXCL12-induced chemotaxis. To do this, chemotaxis assays were carried out using 300.19 cells expressing the class A GPCR BLT1 which mediates chemotaxis in response to the chemokine LTB₄. Initial screening of ECLVS14 and ECLVS58 revealed that both compounds significantly inhibited LTB₄-mediated chemotaxis in cells expressing BLT1 at 50 μ M. However, as can be seen in **Figure 18**, dose response analysis of the ECLVS compounds revealed that while ECLVS58 inhibited BLT1 chemotaxis to a greater degree than CXCR4 mediated chemotaxis (IC₅₀ of 20 μ M vs. 5 μ M), ECLVS14 produced significantly greater anti-CXCR4 activity at 5 μ M than anti-BLT1 at 25 μ M. This suggests that the inhibitory effects produced by ECLVS14 are somewhat specific for CXCR4.

Division of ECLVS compounds into inhibitors and non-inhibitors of CXCL12 induced calcium mobilization

Intracellular calcium mobilization is a common method for assessing GPCR receptor activation and signaling. Thus, we investigated the impact of ECLVS compounds on CXCL12-induced calcium mobilization. CXCL12-induced signaling was assessed by the change in amplitude of the calcium signal and deviation from baseline produced by CXCL12 administration, and results were normalized to the CXCL12-alone control. Initial screening of the eight ECLVS targeted compounds at 50 μ M for inhibition of calcium mobilization identified five ECLVS compounds which produced significant reductions in calcium release, as depicted in **Figure 19**. Interestingly, the two compounds which produced significant reduction in CXCR4 mediated chemotaxis had no effect on calcium

Figure 18: ECLVS14 selectively inhibits CXCR4 Mediated Chemotaxis.

Specificity screening of active ECLVS compounds in BLT1 expressing cells revealed that both compounds inhibited chemotaxis in both systems. ECLVS58 inhibited BLT1 chemotaxis to a greater degree than CXCR4 mediated chemotaxis. ECLVS14 produced significantly greater anti-CXCR4 activity at 5 μ M than anti-BLT1 at 25 μ M.

Figure 18: ECLVS14 selectively inhibits CXCR4 Mediated Chemotaxis.

Figure 19: ECLVS compounds inhibit CXCR4 mediated Intracellular Calcium Mobilization. Calcium Mobilization assays performed with ECLVS compounds at a concentration of 50 μ M. Vertical axis is presented in terms of normalized calcium signal, wherein the amplitude of the calcium response to the CXCL12-alone control corresponds to 1. Decreased signal amplitude correlates with inhibition of calcium mobilization. Five of eight tested ECLVS compounds significantly inhibited calcium mobilization.

Figure 19: ECLVS compounds inhibit CXCR4 mediated Intracellular Calcium Mobilization.

mobilization. The final compound ECLVS72 significantly enhanced calcium release (data not shown). Subsequent dose response analysis conducted to determine the IC₅₀ of the active compounds revealed that two of the five active compounds, ECLVS59 (5 μ M) and ECLVS82 (2.5 μ M) had calcium mobilization IC₅₀ values of 10 μ M or less as seen in [Figure 20](#).

We then sought to determine the specificity of the ECLVS compounds for CXCL12-induced intracellular calcium mobilization. To do this, calcium mobilization assays were carried out using 300.19 cells expressing BLT1 and signaling was assessed in response to LTB₄. Initial screening of the five ECLVS compounds which inhibited CXCL12-induced calcium mobilization revealed that all five compounds significantly inhibited LTB₄-induced intracellular calcium mobilization. Subsequent dose response analysis, depicted in [Figure 21](#), revealed that the IC₅₀ of four of the five ECLVS compounds were comparable between CXCR4 and BLT1 mediated calcium mobilization: ECLVS43 (42 μ M vs. 50 μ M), ECLVS44 (47.5 μ M vs. 50 μ M), ECLVS55 (47.5 μ M vs. 50 μ M) and ECLVS82 (2.5 μ M vs. 3.5 μ M). In contrast ECLVS59 produced very specific inhibition of CXCR4 mediated calcium mobilization, with anti-CXCR4 activity of 5 μ M compared to 40 μ M for inhibition of BLT1-mediated calcium mobilization, suggesting that anti-calcium mobilization activity in ECLVS59 is selective for CXCR4.

As inhibition of chemotaxis and calcium mobilization by the ECLVS compounds could be due to nonspecific toxic side effects, the ECLVS compounds were screened for cytotoxicity. Mimicking the conditions in the

Figure 20: IC₅₀ of ECLVS compound mediated inhibition of calcium mobilization. Concentrations which inhibited 50% of calcium mobilization in relation to CXCL12 alone signaling (IC₅₀) were determined for active ECLVS compounds. Vertical axis is presented in terms of IC₅₀ of calcium mobilization in μM . Two of the five active compounds exhibited significant inhibition of calcium mobilization at 10 μM or less.

Figure 20: IC₅₀ of ECLVS compound mediated inhibition of calcium mobilization.

Figure 21: ECLVS59 selectively inhibits CXCR4 Mediated Calcium

Mobilization. The five ECLVS compounds which inhibited CXCL12-induced calcium mobilization were tested for specificity in BLT1 expressing cells, revealing that all five compounds inhibited calcium mobilization in both systems at a 50 μ M screening concentration. Four of the five compounds produced BLT1 mediated calcium mobilization equivalent to CXCR4 mediated calcium mobilization. ECLVS59 was selective for CXCR4 mediated calcium mobilization.

Figure 21: ECLVS59 selectively inhibits CXCR4 Mediated Calcium

Mobilization

chemotaxis assay, 300.19 cells were cultured in the presence of ECLVS compounds at the initial screening concentration of 50 μ M and cytotoxicity was assessed using propidium iodide (PI) based flow cytometry. PI is unable to penetrate the cell membrane in viable cells, so cells staining positive for PI do not have intact membranes. Cytotoxicity values are presented as percentage of cells staining positive for PI in **Figure 22**. Of the nine ECLVS compounds only two produced displayed elevated cytotoxicity when compared to untreated cells or cells treated with DMSO only control. Compounds ECLVS44 (88.3%) and ECLVS59 (99.8%) displayed significant increases in cytotoxicity above that of untreated (3.5%) and DMSO-treated (3.7%) controls.

Three of the ECLVS compounds possess anti-CXCR4 activity at less than 10 μ M, ECLVS14 with a chemotaxis IC₅₀ of 5 μ M, and ECLVS59 and ECLVS82 with calcium mobilization IC₅₀s of 5 μ M and 2.5 μ M respectively. Both ECLVS14 and ECLVS59 were demonstrated to selectively inhibit CXCR4, while ECLVS82 was non-selective. In addition, while ECLVS14 produced relatively little cytotoxicity, ECLVS59 produced almost 100% cytotoxicity after 3 hours. Thus, because ECLVS14 selectively inhibited CXCR4-mediated chemotaxis and produced no non-specific cytotoxicity, this compound was chosen as a parent compound for subsequent derivation and for the development of QSAR analysis.

Development of ECLVS14 derivatives for QSAR analysis.

ECLVS14 derivatives were developed targeting specific modifications of the compound structure as depicted in **Figure 23**. Modifications of the central chain could introduce greater flexibility between the terminal benzylidenemethanamine

Figure 22: ECLVS compounds produce relatively little non-specific cytotoxicity. Vertical axis is expressed in terms of cells staining positive for PI, a marker for membrane disruption and non-viable cells. Only two of the eight ECLVS compounds with exhibited significant cytotoxicity. These compounds produced greater than 80% cytotoxicity at 50 μ M.

Figure 22: ECLVS compounds produce relatively little non-specific cytotoxicity.

Figure 23: Modifications of ECLVS14 for QSAR analysis. ECLVS14

derivatives will be designed with several modifications in mind. Acetamide derivatives will be developed. The central chain will be lengthened. The benzylidenemethanamine moiety will also be modified to include additional electronegative or polar substituents. The necessity of the 1-[bis(phenylmethyl)amino]methyl (1-BPMAM) moiety will be assessed in compounds lacking this moiety. The phenylmethanamine moiety will also be assessed by addition of electronegative or polar substituents or replacement with a less aromatic substituent.

Figure 23: Modifications of ECLVS14 for QSAR analysis.

moiety and the imidazole, thus the effect of increasing chain length on this group will be assessed. Loss of the acetamide moiety could result in loss of hydrogen bonding between the compound and the receptor thus derivatives based on this moiety will be developed. Increasingly electronegative or polar substituents targeting the benzylidenemethanamine moiety could facilitate increasing ionic or hydrogen bonding interactions between compound and receptor, while the addition of aliphatic elements or decreasing the aromatic characteristics of the group could modulate binding as well. Modifications of the 1-[bis(phenylmethyl)amino]methyl (1-BPMAM) moiety could facilitate compound binding in a manner similar to that of modifications of the benzylidenemethanamine moiety, while removal of the entire group could facilitate greater and deeper binding of the compound in the binding site.

EVS14 compounds significantly inhibit CXCR4 mediated chemotaxis.

Twenty-three ECLVS14 derivatives assigned the designation EVS14 (illustrated in **Figure 24** and listed in **Table 7**) were synthesized by our organic chemist collaborators led by G. B. Hammond, Dept. of Chemistry, University of Louisville and prepped for analysis in human CXCR4 expressing cell lines by suspending the compounds in DMSO to a concentration of 20 mM. These derivatives explored the role of the bis (phenylmethyl) amino-methyl moiety and effects of modification of this group, as well as exploring the effect of lengthening the hydrazine-linked alkyl moiety or converting it to an acetamide and modifying the terminal benzyl group.

Figure 24: EVS14 compounds. Derivatives of ECLVS14 were generated focusing on conversion of the hydrazine alkyl linker moiety to an acetamide alkyl linker moiety, as well as investigating the effects of lengthening the hydrazine alkyl linker, and assessing the effect of presence or absence of the 1-BPMAM moiety.

Figure 24: EVS14 compounds

EVS14-1

EVS14-5

EVS14-2

EVS14-6

EVS14-3

EVS14-7

EVS14-4

EVS14-8

EVS14-17 **EVS14-21** **EVS14-18** **EVS14-22** **EVS14-19** **EVS14-23** **EVS14-20**

Table 7: EVS14 compounds. Compounds from the EVS14 series are listed below with chemical name, supplier, and molecular weight.

Table 7: EVS14 compounds.

Compound	NAME	MW
EVS14-1	(Z)-3-(((E)-benzylideneamino)oxy)-N'-(1-((dibenzylamino)methyl)-2-oxoindolin-3-ylidene)propanehydrazide	545.63
EVS14-2	(Z)-4-(((E)-benzylideneamino)oxy)-N'-(1-((dibenzylamino)methyl)-2-oxoindolin-3-ylidene)butanehydrazide	559.66
EVS14-3	(Z)-5-(((E)-benzylideneamino)oxy)-N'-(1-((dibenzylamino)methyl)-2-oxoindolin-3-ylidene)pentanehydrazide	573.68
EVS14-4	(Z)-2-(((E)-(4-fluorobenzylidene)amino)oxy)-N-(2-oxoindolin-3-ylidene)acetamide	325.29
EVS14-5	(Z)-2-(((E)-(4-chlorobenzylidene)amino)oxy)-N-(2-oxoindolin-3-ylidene)acetamide	341.75
EVS14-6	(Z)-N-(2-oxoindolin-3-ylidene)-2-(((E)-(4-(trifluoromethyl)benzylidene)amino)oxy)acetamide	375.3
EVS14-7	(Z)-2-((cyclohexylideneamino)oxy)-N-(2-oxoindolin-3-ylidene)acetamide	299.32
EVS14-8	(Z)-2-(((E)-(4-methoxybenzylidene)amino)oxy)-N-(2-oxoindolin-3-ylidene)acetamide	337.33
EVS14-9	(Z)-2-(((E)-(4-methylbenzylidene)amino)oxy)-N-(2-oxoindolin-3-ylidene)acetamide	321.33
EVS14-10	(Z)-2-(((E)-(4-nitrobenzylidene)amino)oxy)-N-(2-oxoindolin-3-ylidene)acetamide	352.3
EVS14-11	(Z)-2-(((E)-(2,4-dimethoxybenzylidene)amino)oxy)-N-(2-oxoindolin-3-ylidene)acetamide	367.36

Table 7, continued: ECLVS Compounds

Compound	NAME	MW
EVS14-12	(Z)-N'-(1-((dibenzylamino)methyl)-2-oxoindolin-3-ylidene)-2-(((E)-(4-(trifluoromethyl)benzylidene)amino)oxy) acetohydrazide	599.6
EVS14-13	(Z)-N'-(1-((dibenzylamino)methyl)-2-oxoindolin-3-ylidene)-2-(((E)-(4-methylbenzylidene)amino)oxy)acetohydrazide	545.63
EVS14-14	(Z)-N'-(1-((dibenzylamino)methyl)-2-oxoindolin-3-ylidene)-2-(((E)-(4-fluorobenzylidene)amino)oxy)acetohydrazide	549.59
EVS14-15	(Z)-2-(((E)-(4-chlorobenzylidene)amino)oxy)-N'-(1-((dibenzylamino)methyl)-2-oxoindolin-3-ylidene)acetohydrazide	566.05
EVS14-16	(Z)-2-((cyclohexylideneamino)oxy)-N'-(1-((dibenzylamino)methyl)-2-oxoindolin-3-ylidene)acetohydrazide	523.63
EVS14-17	(Z)-N'-(1-((dibenzylamino)methyl)-2-oxoindolin-3-ylidene)-2-(((E)-(2,4-dimethoxybenzylidene)amino)oxy)acetohydrazide	591.66
EVS14-18	(Z)-N'-(1-((dibenzylamino)methyl)-2-oxoindolin-3-ylidene)-2-(((E)-(4-nitrobenzylidene)amino)oxy)acetohydrazide	576.6
EVS14-19	(Z)-N'-(2-oxoindolin-3-ylidene)-4-(((E)-(4-(trifluoromethyl)benzylidene)amino)oxy)butanehydrazide	418.37
EVS14-20	(Z)-2-(((E)-benzylideneamino)oxy)-N'-(2-oxoindolin-3-ylidene)acetohydrazide	322.32
EVS14-21	(Z)-N'-(2-oxoindolin-3-ylidene)-2-(((E)-(2-(trifluoromethyl)benzylidene)amino)oxy)acetohydrazide	390.32
EVS14-22	(Z)-N'-(2-oxoindolin-3-ylidene)-2-(((E)-(3-(trifluoromethyl)benzylidene)amino)oxy)acetohydrazide	390.32
EVS14-23	(Z)-N'-(1-((dibenzylamino)methyl)-2-oxoindolin-3-ylidene)-4-(((E)-(4-(trifluoromethyl)benzylidene)amino)oxy) butanehydrazide	627.66

Chemotaxis analysis was pursued using the EVS14 compounds to identify structural elements which could enhance anti-CXCR4 activity. The chemotaxis values of ECLVS14 and the EVS14 compounds are presented in Table 8. Of the twenty-three EVS14 compounds tested, ten significantly inhibited CXCR4 mediated chemotaxis. Initial screening chemotaxis at 50 μ M is depicted in Figure 25. Dose response analysis depicted in Figure 26, was pursued to identify the IC₅₀ values of the nine EVS14 compounds which inhibited chemotaxis. Two of the EVS14 compounds produced IC₅₀ values of less than 10 μ M; EVS14-21 and EVS14-22 exhibited IC₅₀ values of 5 μ M and 7.5 μ M respectively.

Specificity screening of active EVS compounds in BLT1 expressing cells revealed that eight of the ten compounds inhibited chemotaxis in both systems. Dose response analysis revealed that of the eight compounds with anti-BLT1 activity, only EVS14-4 and EVS14-5 produced inhibition of chemotaxis at greater than the screening concentrations for BLT1, as depicted in Figure 27. All of the compounds other than EVS14-5 produced significantly greater anti-CXCR4 activity than anti-BLT1 activity. EVS14-5 produced significantly greater reduction in BLT1 mediated chemotaxis than CXCR4-mediated chemotaxis. EVS14 compounds were further evaluated for their capability of inhibiting CXCR4 mediated intracellular calcium mobilization. The EVS14 compounds frequently precipitated at higher concentrations during the intracellular calcium mobilization experiments. As such, screening of these compounds was carried out at 10 μ M. At this concentration precipitation of the compound was not noted in the

Table 8: Table of ECLVS14 and EVS14 compound chemotaxis activity.

Chemical characteristics of parent compound ECLVS14 and the EVS14 derivatives are listed below along with chemotaxis index listings and the IC₅₀ of active compounds.

Table 8: Table of ECLVS14 and EVS14 compound Chemotaxis activity.

Compound	Linker length	Terminal Phenyl	1-BPMAM Moeity	Chemotaxis index	Compound IC50
ECLVS14	1	phenyl	1-BPMAM	0.21 ± 0.16	5 µM
EVS14-1	2	phenyl	1-BPMAM	0.24 ± 0.11	40 µM
EVS14-2	3	phenyl	1-BPMAM	0.46 ± 0.23	Not Active
EVS14-3	4	phenyl	1-BPMAM	0.39 ± 0.20	Not Active
EVS14-4	acetamide	4-fluorophenyl	H	0.10 ± 0.12	15 µM
EVS14-5	acetamide	4-chlorophenyl	H	0.18 ± 0.24	20 µM
EVS14-6	acetamide	4- trifluroro methylphenyl	H	0.15 ± 0.24	15 µM
EVS14-7	acetamide	Cyclohexylidene amino	H	0.53 ± 0.26	Not Active
EVS14-8	acetamide	4-methoxy phenyl	H	0.38 ± 0.21	Not Active
EVS14-9	acetamide	4-methylphenyl	H	0.23 ± 0.22	35 µM
EVS14-10	acetamide	4-nitrophenyl	H	0.64 ± 0.08	Not Active
EVS14-11	acetamide	2,4-dimethoxy phenyl	H	0.72 ± 0.14	Not Active
EVS14-12	1	4-trifluroro methylphenyl	1-BPMAM	0.39 ± 0.24	Not Active
EVS14-13	1	4-methylphenyl	1-BPMAM	0.73 ± 0.16	Not Active
EVS14-14	1	4-fluorophenyl	1-BPMAM	0.73 ± 0.24	Not Active
EVS14-15	1	4-chlorophenyl	1-BPMAM	0.59 ± 0.24	Not Active
EVS14-16	1	Cyclohexylidene amino	1-BPMAM	0.82 ± 0.03	Not Active
EVS14-17	1	2,4-dimethoxy phenyl	1-BPMAM	0.53 ± 0.13	Not Active
EVS14-18	1	4-nitrophenyl	1-BPMAM	0.57 ± 0.19	Not Active
EVS14-19	3	4-trifluroro methylphenyl	H	0.28 ± 0.11	25 µM
EVS14-20	1	phenyl	H	0.09 ± 0.07	25 µM
EVS14-21	1	2-trifluroro methylphenyl	H	0.28 ± 0.10	5 µM
EVS14-22	1	3-trifluroro methylphenyl	H	0.02 ± 0.02	7.5 µM
EVS14-23	1	4-trifluroro methylphenyl	1-BPMAM	0.21 ± 0.08	27.5 µM

Figure 25: ECLVS14 Derivative Compounds Inhibit CXCR4 Mediated Chemotaxis. Ten of the twenty three (43.5%) ECLVS14 derived compounds significantly inhibited CXCR4-mediated chemotaxis. Chemotaxis assays performed at a concentration of 50 μ M.

Figure 25: ECLVS14 Derivative Compounds Inhibit CXCR4 Mediated Chemotaxis.

Figure 26: IC₅₀ of Active Inhibitors of CXCR4 Mediated Chemotaxis.

Concentrations which inhibited 50% of chemotaxis in relation to CXCL12 alone signaling (IC₅₀) were determined for active EVS14 compounds. Vertical axis is presented in terms of IC₅₀ of chemotaxis in μM . Of the EVS compounds, EVS14-21 and EVS14-22 had IC₅₀ values less than 10 μM .

Figure 26: IC₅₀ of Active Inhibitors of CXCR4 Mediated Chemotaxis.

Figure 27: EVS14 compounds selectively inhibit CXCR4 Mediated

Chemotaxis. Specificity screening of active EVS compounds in BLT1 expressing cells revealed that eight of the ten compounds inhibited chemotaxis in both systems. Only EVS14-4 and EVS14-5 produced inhibition of chemotaxis at greater than the screening concentrations for BLT1. All of the compounds other than EVS14-5 produced significantly greater anti-CXCR4 activity than anti-BLT1 activity. EVS14-5 produced significantly greater reduction in BLT1 mediated chemotaxis than CXCR4-mediated chemotaxis.

Figure 27: EVS14 compounds selectively inhibit CXCR4 Mediated Chemotaxis.

spectrophotometer cuvette. None of the compounds significantly inhibited CXCR4 mediated intracellular calcium mobilization, as can be seen in **Figure 28**. In all, the EVS14 compounds maintain the characteristics of their parent compound ECLVS14. Several of the derivatives selectively inhibit CXCR4-mediated chemotaxis, while maintaining ECLVS14's inability to modulate CXCR4-mediated calcium mobilization.

QSAR analysis reveals essential pharmacophores of ECLVS14

Specific analysis of the EVS14 compounds reveals several specific factors of ECLVS14 involved in its anti-CXCR4 activity. One of the most obvious modifications of the parent compound involved extension of the hydrazine-linked alkyl moiety. Modifications of the hydrazine-linked alkyl moiety to ethyl (EVS14-1), propyl (EVS14-2), and butyl (EVS14-3) derivatives revealed that as the hydrazine-linked alkyl group was lengthened, activity decreased, as can be seen in **Figure 29**. When the hydrazine-linked alkyl moiety is the only modification from ECLVS14, neither the propyl or butyl derivatives significantly inhibited chemotaxis. The ethyl derivative significantly inhibited chemotaxis, but the IC₅₀ was much lower than the hydrazine-linked alkyl parent compound (40 μ M vs. 5 μ M).

Several of the ECLVS14-derived compounds are comparable, in that they were modified in the same manner at the terminal benzene as depicted in **Figure 30**. These compounds differ in that some retain the 1-BPMAM moiety and the hydrazine-linked alkyl of the parent compound, while their counterparts lack this moiety and have converted the hydrazine-linked alkyl into an acetamide. When

Figure 28: EVS14 compounds do not inhibit CXCR4 mediated calcium mobilization. Calcium Mobilization assays performed with EVS14 compounds at a concentration of 10 μ M. Vertical axis is presented in terms of normalized calcium signal, wherein the amplitude of the calcium response to the CXCL12-alone control corresponds to 1. Decreased signal amplitude correlates with inhibition of calcium mobilization. None of the tested EVS14 compounds significantly inhibited calcium mobilization.

Figure 28: EVS14 compounds do not inhibit CXCR4 mediated calcium mobilization.

Figure 29: Increasing length of the hydrazine-linked alkyl moiety decreases activity. Sequential modifications of the hydrazine-linked alkyl moiety of ECLVS14, depicted as X in the figure, increased the length from the hydrazine (ECLVS14) to ethyl (EVS14-1), propyl (EVS14-2) and butyl (EVS14-3) derivatives, with decreasing chemotaxis activity with increasing chain length.

Figure 29: Increasing length of the hydrazine-linked alkyl moiety decreases activity.

Figure 30: Effect of terminal benzene modifications on anti-CXCR4 activity.

Acetamide derivatives lacking the 1-BPMAM moiety with terminal benzene modifications with greater electronegativity have greater anti-chemotaxis activity than their hydrazine-linked alkyl counterparts which retain the 1-BPMAM moiety.

Figure 30: Effect of terminal benzene modifications on anti-CXCR4 activity.

comparing these two groups, we find that the absence of the 1-BPMAM moiety and the conversion to an acetamide confers a significant increase in chemotaxis index. None of the compounds which retained the 1-BPMAM moiety accompanied by terminal benzene modifications significantly inhibited chemotaxis.

Identifying which alteration was responsible for this change is difficult to tease out with the available compounds. The structure of compound EVS14-20 mimics parent compound ECLVS14 except that it lacks the 1-BPMAM moiety as can be seen in [Figure 31](#). Although both compounds significantly inhibit chemotaxis, the IC₅₀ of the parent compound retaining the group is significantly less than the derivative lacking the group (5 μ M vs. 25 μ M). Compounds EVS14-6, EVS14-12, EVS14-21, EVS14-22, all are modified at the terminal benzene with a trifluoromethyl group, but are variably altered at the hydrazine-linked alkyl and 1-BPMAM moiety. All three compounds lacking the 1-BPMAM moiety significantly inhibited chemotaxis while the compound retaining the moiety did not. However, among the compounds lacking the moiety, the two compounds which retained the hydrazine-linked alkyl group both had IC₅₀ values which were less than the compound with the acetamide group (5 μ M and 7.5 μ M vs. 15 μ M). This comparison must be qualified in that the three position of the trifluoromethyl substituent differs in ortho/meta/para position in all three compounds. Unfortunately there was no compound which was modified solely in terms of the hydrazine-linked alkyl group, so a direct comparison of this modification to the parent compound is not available from this data.

Figure 31: Absence of the 1-BPMAM moiety alone decreases anti-chemotaxis activity. A comparison of ECLVS14 which retains the 1-BPMAM moiety and EVS14-20 which lacks it reveals that the presence of this moiety confers greater anti-chemotaxis inhibitory activity.

Figure 31: Absence of the 1-BPMAM moiety alone decreases anti-chemotaxis activity.

Further modification of the compounds lacking the 1-BPMAM moiety and converting to the acetamide group reveals that activity of the compounds is related to the electronegativity of the terminal benzyl group modifications. A parabolic relationship governs chemotaxis activity. There was no significant inhibition in the least electronegative methoxy and dimethoxy derivatives, and the most electronegative nitro derivatives. However, as electronegativity increased from the methyl derivatives to the chloro, trifluoromethyl and fluoro derivatives, the IC₅₀ of the compounds decreased, suggesting that electronegativity of the terminal group modifications significantly influences the inhibition of chemotaxis. The influence of terminal benzene electronegativity can be seen in the relationship between compounds EVS14-2, EVS14-19, and EVS14-23. All three of these compounds retain the propylhydrazine modification, however, where the terminal benzene on EVS14-2 is unmodified, that of EVS14-19 and EVS14-23 is modified to a trifluoromethyl derivative, as illustrated in **Figure 32**. Whereas EVS14-2 did not significantly inhibit chemotaxis, both EVS14-19 and EVS14-23 did significantly inhibit this activity. And while the IC₅₀ values of the compounds are very similar, (25 μ M and 27.5 μ M respectively) they have a significant structural difference in that EVS14-19 lacks the 1-BPMAM moiety. This suggests that terminal benzene electronegativity confers a greater influence over activity than the presence or absence of the 1-BPMAM moiety.

Development of EVS14-2 derivatives for QSAR analysis

Based on preliminary chemotaxis results and a desire to further evaluate the effects of extension of the hydrazine-linked alkyl moiety, Dr. Hammond

Figure 32: Terminal benzene electronegativity confers greater influence over anti-chemotaxis activity than the presence or absence of the 1-BPMAM moiety. Comparison of EVS14-9 with EVS14-23 reveals that both compounds with modification of the terminal benzene have roughly equivalent inhibition of chemotaxis despite differing in the presence or absence of the 1-BPMAM moiety.

Figure 32: Terminal benzene electronegativity confers greater influence over anti-chemotaxis activity than the presence or absence of the 1-BPMAM moiety.

Compound	X	Y	Activity
EVS14-2	H	1-BPMAM	No significant inhibition
EVS14-19	Trifluoromethylphenyl	H	25 μ M
EVS14-23	Trifluoromethylphenyl	1-BPMAM	27.5 μ M

developed an additional series of ECLVS14 derivatives based on the propylhydrazine derivative EVS14-2. Fourteen compound EVS14-2 derivatives were synthesized for testing. These compounds, illustrated in [Figure 33](#) and listed in [Table 9](#), retained EVS14-2's conversion of the hydrazine-linked alkyl to a propylhydrazine and demonstrated modification or replacement of the 1-BPMAM moiety.

Chemotaxis analysis was pursued using the EVS14-2 derivative compounds to identify structural elements which could enhance anti-CXCR4 activity. The chemotaxis values of the EVS14-2 derivatives are presented in [Table 10](#). Of the fifteen compounds tested, ten significantly inhibited CXCR4 mediated chemotaxis. Initial screening chemotaxis at 50 μ M is depicted in [Figure 34](#). Dose response analysis depicted in [Figure 35](#), was pursued to identify the IC₅₀ values of the five EVS14-2 derivative compounds which inhibited chemotaxis. Compound EVS14-2-AB was the sole compound to exhibit an IC₅₀ less than 10 μ M

Specificity screening of active EVS14-2 derivatives, depicted in [Figure 36](#), in BLT1 expressing cells revealed that five of the seven compounds inhibited chemotaxis in both systems. All of the compounds produced significantly greater anti-CXCR4 activity than anti-BLT1 activity.

EVS14-2 derivatives were further evaluated for their capacity for inhibiting CXCR4 mediated intracellular calcium mobilization. Much like their parent compound, the EVS14-2 derivatives frequently precipitated at higher concentrations during the intracellular calcium mobilization experiments and

Figure 33: EVS14 -2 derivative compounds. Derivatives of EVS14-2 were generated focusing on modification of the benzene groups of the 1-BPMAM moiety while maintaining the propylhydrazine modifications of EVS14-2.

Figure 33: EVS14-2 derivative compounds

EVS14-2-B

EVS14-2-J

EVS14-2-C

EVS14-2-K

EVS14-2-D

EVS14-2-L

EVS14-2-E

EVS14-2-N

EVS14-2-P

EVS14-2-AB

EVS14-2-T

EVS14-2-AD

EVS14-2-AA

EVS14-2-AE

Table 9: EVS14-2 derivative compounds. Derivatives of EVS14-2 were generated focusing on modification of the benzene groups of the 1-BPMAM moiety while maintaining the propylhydrazine modifications of EVS14-2. The following table lists compound name and molecular weight.

Table 9: EVS14-2 derivative compounds

Compound	NAME	MW
EVS14-2-B	(Z)-N'-(1-((benzyl(4-methylbenzyl)amino)methyl)-2-oxoindolin-3-ylidene)-4-(((E)-benzylideneamino)oxy)butanehydrazide	573.7
EVS14-2-C	(Z)-N'-(1-((benzyl(3-methoxybenzyl)amino)methyl)-2-oxoindolin-3-ylidene)-4-(((E)-benzylideneamino)oxy)butanehydrazide	589.7
EVS14-2-D	(Z)-N'-(1-((benzyl(4-methoxybenzyl)amino)methyl)-2-oxoindolin-3-ylidene)-4-(((E)-benzylideneamino)oxy)butanehydrazide	589.7
EVS14-2-E	(Z)-N'-(1-((benzyl(2,4-dimethoxybenzyl)amino)methyl)-2-oxoindolin-3-ylidene)-4-(((E)-benzylideneamino)oxy)butanehydrazide	619.7
EVS14-2-J	(Z)-N'-(1-((benzyl(4-chlorobenzyl)amino)methyl)-2-oxoindolin-3-ylidene)-4-(((E)-benzylideneamino)oxy)butanehydrazide	594.1
EVS14-2-K	(Z)-N'-(1-((benzyl(2-fluorobenzyl)amino)methyl)-2-oxoindolin-3-ylidene)-4-(((E)-benzylideneamino)oxy)butanehydrazide	577.7
EVS14-2-L	(Z)-N'-(1-((benzyl(4-fluorobenzyl)amino)methyl)-2-oxoindolin-3-ylidene)-4-(((E)-benzylideneamino)oxy)butanehydrazide	577.7
EVS14-2-N	(Z)-N'-(1-((benzyl(4-(trifluoromethyl)benzyl)amino)methyl)-2-oxoindolin-3-ylidene)-4-(((E)-benzylideneamino)oxy)butanehydrazide	627.7
EVS14-2-P	(Z)-N'-(1-((benzyl(4-nitrobenzyl)amino)methyl)-2-oxoindolin-3-ylidene)-4-(((E)-benzylideneamino)oxy)butanehydrazide	604.7
EVS14-2-T	(Z)-N'-(1-((benzyl(pyridin-4-ylmethyl)amino)methyl)-2-oxoindolin-3-ylidene)-4-(((E)-benzylideneamino)oxy)butanehydrazide	560.7
EVS14-2-AA	(Z)-N'-(1-((benzyl(ethyl)amino)methyl)-2-oxoindolin-3-ylidene)-4-(((E)-benzylideneamino)oxy)butanehydrazide	497.6
EVS14-2-AB	(Z)-N'-(1-((benzyl(vinyl)amino)methyl)-2-oxoindolin-3-ylidene)-4-(((E)-benzylideneamino)oxy)butanehydrazide	495.6
EVS14-2-AD	(Z)-4-(((E)-benzylideneamino)oxy)-N'-(2-oxo-1-(piperidin-1-ylmethyl)indolin-3-ylidene)butanehydrazide	447.5
EVS14-2-AE	(Z)-4-(((E)-benzylideneamino)oxy)-N'-(1-(morpholinomethyl)-2-oxoindolin-3-ylidene)butanehydrazide	449.5

Table 10: Table of EVS14-2 compound Chemotaxis activity Chemical characteristics of the EVS14-2 derivatives are listed below along with chemotaxis index listings and the IC₅₀ of active compounds.

Table 10: Table of EVS14-2 and EVS14-2-AD compound Chemotaxis activity

Compound	Terminal Phenyl	1-BPMAM Moeity	Chemotaxis index	Compound IC50
EVS14-2-B	phenyl	N-benzyl-N-(4-methylbenzyl)ethanamine	1.59 ± 0.30	Not Active
EVS14-2-C	phenyl	N-benzyl-N-(3-methoxybenzyl)ethanamine	1.09 ± 0.29	Not Active
EVS14-2-D	phenyl	N-benzyl-N-(4-methoxybenzyl)ethanamine	0.20 ± 0.05	17.5 µM
EVS14-2-E	phenyl	N-benzyl-N-(2,4-dimethoxybenzyl)ethanamine	1.34 ± 0.37	Not Active
EVS14-2-J	phenyl	N-benzyl-N-(4-chlorobenzyl)ethanamine	0.44 ± 0.13	Not Active
EVS14-2-K	phenyl	N-benzyl-N-(2-fluorobenzyl)ethanamine	1.43 ± 0.18	Not Active
EVS14-2-L	phenyl	N-benzyl-N-(4-fluorobenzyl)ethanamine	0.23 ± 0.05	22.5 µM
EVS14-2-N	phenyl	N-benzyl-N-(4-(trifluoromethyl)benzyl)ethanamine	1.45 ± 0.19	Not Active
EVS14-2-P	phenyl	N-benzyl-N-(4-nitrobenzyl)ethanamine	0.66 ± 0.20	Not Active
EVS14-2-T	phenyl	N-benzyl-N-(pyridin-4-ylmethyl)ethanamine	0.17 ± 0.01	40 µM
EVS14-2-AA	phenyl	N-benzyl-N-ethylethanamine	0.22 ± 0.09	20 µM
EVS14-2-AB	phenyl	N-benzyl-N-ethylethenamine	0.36 ± 0.22	Not Active
EVS14-2-AD	phenyl	1-ethylpiperidine	0.03 ± 0.01	Not Active
EVS14-2-AE	phenyl	4-ethylmorpholine	0.21 ± 0.07	22.5 µM

Figure 34: EVS14-2 derivative compounds Inhibit CXCR4 Mediated Chemotaxis. Seven of the fifteen EVS14-2 derived compounds significantly inhibited CXCR4-mediated chemotaxis. Chemotaxis assays performed at a concentration of 50 μ M.

Figure 34. EVS14-2 derivative Compounds Inhibit CXCR4 Mediated Chemotaxis

Figure 35: IC₅₀ of Active Inhibitors of CXCR4 Mediated Chemotaxis.

Concentrations which inhibited 50% of chemotaxis in relation to CXCL12 alone signaling (IC₅₀) were determined for active EVS14-2 derivatives. Vertical axis is presented in terms of IC₅₀ of chemotaxis in μM . Of the EVS14-2 compounds, EVS14-2-AB had an IC₅₀ value of less than 10 μM .

Figure 35: IC₅₀ of Active Inhibitors of CXCR4 Mediated Chemotaxis.

Figure 36: EVS14-2 derivatives selectively inhibit CXCR4 Mediated Chemotaxis. Specificity screening of active EVS14-2 derivatives in BLT1 expressing cells revealed that five of the seven compounds inhibited chemotaxis in both systems. All of the compounds other than EVS14-5 produced greater anti-CXCR4 activity than anti-BLT1 activity.

Figure 36: EVS14-2 derivatives selectively inhibit CXCR4 Mediated Chemotaxis.

screening was carried out at 10 μ M. At this concentration precipitation of the compound was not noted in the spectrophotometer cuvette. None of the compounds significantly inhibited CXCR4 mediated intracellular calcium mobilization, as can be seen in **Figure 37**.

In all, some of the EVS14-2 derivatives improve on their predecessor. Several of the derivatives selectively inhibit CXCR4-mediated chemotaxis, while maintaining ECLVS14's inability to modulate CXCR4-mediated calcium mobilization.

QSAR analysis of EVS14-2 derivatives illustrates the effects of alteration of the 1-BPMAM moiety.

The majority of the EVS14-2 derived compounds deal with replacement or modification of the 1-BPMAM moiety, as illustrated in **Figure 38**. It is difficult to identify relationships between different modifications of the benzene elements of the 1-BPMAM moiety. Only two of the modifications significantly inhibited chemotaxis, the methoxy and fluoro derivatives. Replacement of the 1-BPMAM moiety seemed to offer greater alteration of chemotaxis and seemed to correlate with size. Replacement of one benzene substituent with a pyridine significantly inhibited chemotaxis, and inhibition increased as the substituents shrank in size, from pyridine to ethyl to ethenyl (40 μ M vs. 20 μ M vs. 5 μ M). Wholesale replacement of the 1-BPMAM moiety with a piperidine or morpholine substituent conferred a similar enhancement of inhibition compared to the EVS14-2 parent compound (20 μ M and 22.5 μ M).

Figure 37: EVS14-2 derivatives do not inhibit CXCR4 mediated calcium mobilization. Calcium Mobilization assays performed with EVS14-2 derivatives at a concentration of 10 μ M. Vertical axis is presented in terms of normalized calcium signal, wherein the amplitude of the calcium response to the CXCL12-alone control corresponds to 1. Decreased signal amplitude correlates with inhibition of calcium mobilization. None of the tested EVS14-2 derivatives significantly inhibited calcium mobilization.

Figure 37. EVS14-2 derivatives do not inhibit CXCR4 mediated calcium mobilization.

Figure 38: EVS14-2 derivatives modify or replace the 1-BPMAM moiety. The EVS14-2 derived compounds deal with replacement or modification of the 1-BPMAM moiety. Seven of the fourteen modifications significantly inhibited chemotaxis. Replacement of one benzene substituent with a pyridine significantly inhibited chemotaxis, and inhibition increased as the substituents shrank in size, from pyridine to ethyl to ethenyl (40 μ M vs. 20 μ M vs. 5 μ M). Wholesale replacement of the 1-BPMAM moiety with a piperidine or morpholine substituent conferred a similar enhancement of inhibition compared to the EVS14-2 parent compound (20 μ M and 22.5 μ M).

Figure 38: EVS14-2 derivatives modify or replace the 1-BPMAM moiety.

Compound	Substitution	IC ₅₀	X
EVS14-2-D	<i>N</i> -benzyl-1-(4-methoxyphenyl)- <i>N</i> -methylmethanamine	17.5 μ M	
EVS14-2-L	<i>N</i> -benzyl-1-(4-fluorophenyl)- <i>N</i> -methylmethanamine	22.5 μ M	
EVS14-2-T	<i>N</i> -benzyl- <i>N</i> -methyl-1-(pyridin-4-yl)methanamine	40 μ M	
EVS14-2-AA	<i>N</i> -benzyl- <i>N</i> -methylethanamine	20 μ M	
EVS14-2-AB	<i>N</i> -benzyl- <i>N</i> -methylethenamine	5 μ M	
EVS14-2-AD	1-methylpiperidine	20 μ M	
EVS14-2-AE	4-methylmorpholine	22.5 μ M	

Acetamide derivatives of ECLVS14 are associated with elevated levels of non-specific cytotoxicity.

As inhibition of chemotaxis by the ECLVS14 derivative compounds could be due to nonspecific toxic side effects, the ECLVS compounds were screened for cytotoxicity. Mimicking the conditions in the chemotaxis assay, 300.19 cells were cultured in the presence of ICLGa compounds at the initial screening concentration of 50 μ M and cytotoxicity was assessed using propidium iodide (PI) based flow cytometry. PI is unable to penetrate the cell membrane in viable cells, so cells staining positive for PI do not have intact membranes. Cytotoxicity values are presented as percentage of cells staining positive for PI in Figure 39. Of the thirty seven EVS14 and EVS14-2 compounds only six produced elevated cytotoxicity when compared to untreated cells or cells treated with DMSO only control. Interestingly, four of the eight acetamide derivatives of ECLVS14 displayed elevated cytotoxicity, suggesting that modification of this group adversely affects the capability of CXCR4 to exhibit specific anti-CXCR4 activity.

Figure 39: Acetamide ECLVS14 derivatives produce elevated non-specific cytotoxicity. Vertical axis is expressed in terms of cells staining positive for PI, a marker for membrane disruption and non-viable cells. Four of the eight acetamide derivatives of ECLVS14 exhibited significant cytotoxicity.

Figure 39: Acetamide ECLVS14 derivatives produce elevated non-specific cytotoxicity.

Analysis of anti-CXCR4 compound mechanism through molecular dynamics simulations

To test the hypothesis that compound binding will result in structural changes associated with inactivation of CXCR4, the conformational differences between native CXCR4, the constitutively active N119S mutant [286, 287], and bound CXCR4 inhibitors were investigated.

The initial native CXCR4 and N119S mutant structures were developed by Dr. Kelby Napier via homology modeling utilizing the high resolution rhodopsin crystal structure [288]. When this work was conducted, the x-ray crystal structure for CXCR4 had not yet been released, so this structure was not a viable option for these studies. Also, the rhodopsin structure was chosen rather than the new β 1-adrenergic, β 2-adrenergic, or α 2a-adrenergic receptors, because these structures were generated with significant modifications of the intracellular loop 3 and the C-terminal region. Homologies developed based on these structures would have required rhodopsin based modeling of intracellular loop 3 and the C-terminal region, making this method redundant.

Molecular modeling of CXCR4

Fully solvated lipid bilayer molecular dynamics simulations of native CXCR4 bound with active ECLVS compounds were performed using the AMBER suite of programs using the parm99 force field [284]. Anti-CXCR4 compounds were docked into the T140 binding site [200, 207] of CXCR4 using Surflex-Dock [234]. Restraints were initially used to guide conformation of the receptors as previously described in [285]. In brief, an initial 5000 step steepest decent

minimization was conducted with protein and compound restrained ($100 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$), followed by another 5000 step steepest decent minimization with no protein or compound restraints. This was followed by a 25 ps MD ($T=100 \text{ K}$) holding protein fixed ($100 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$), and another unrestrained 5000 step steepest decent minimization without restraints. The temperature of the system will be increased from an initial point of 100K to 300K while maintaining $100 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ restraints. An additional three 25 ps steps of minimization will be conducted gradually decreasing the restraints from $100 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ to $10 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ and finally to $1 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$. Ten production runs of 500 ps were then conducted. The impact of the ECLVS and EVS14 compounds on CXCR4 structure has not yet been determined.

Some of the ECLVS and EVS14 compounds do not bind the native conformation of CXCR4 in molecular dynamics simulations.

The final averaged structures of native CXCR4 in complex with each of the anti-CXCR4 compounds from the 5 ns molecular dynamics simulation are shown in **Figure 40**. Of the seven docked compounds, stable CXCR4: compound complexes were generated for compounds ECLVS14, ECLVS58, ECLVS59, ECLVS82 and EVS14-2-AB. Compounds EVS14-21 and EVS14-22 did not form stable complexes with the native conformation of CXCR4. Neither of these compounds possessed the 1-BPMAM moiety, suggesting that this element is essential for interacting with the extracellular surface of the receptor.

The molecular dynamics simulation maintained equilibrium throughout the production run, as can be seen in **Figure 41**. All energy values were noted to be

Figure 40: Final averaged structure of Native CXCR4 dynamically bound to compounds. Average structure of CXCR4 presented in ribbon format.

Compounds are bound to dynamically determined binding position. Interacting residues are highlighted in stick conformations.

Figure 40: Final averaged structure of Native CXCR4 dynamically bound to compounds.

Figure 41: Energy plot of ECLVS compounds bound to native CXCR4 versus time. Energy values across the trajectory for total energy (blue), kinetic energy (red) and potential energy (green) are presented. CXCR4 was found to be energetically stable and maintained equilibrium across the production run. A) ECLVS14, B) EVS14-2-AB, C) ECLVS58, D) ECLVS59, E) ECLVS82

Figure 41: Energy plot of compounds bound to native CXCR4 versus time.

A) ECLVS14, B) EVS14-2-AB, C) ECLVS58, D) ECLVS59, E) ECLVS82

stable, indicating that the simulation was well equilibrated prior to production analysis. Conformational stability of the CXCR4-compound complexes was investigated using root RMSD analysis, presented in **Figures 42**. RMSD values remained stable for the majority of the production runs, with the final average structure representative of the trajectory after approximately 1.5 ns. The helical domains were highly stable throughout the production run, while the majority of the structural variation was seen in the loop regions. The RMSD of the ECLVS14 complex from the native conformation of CXCR4 was 3.1, compared to RMSD values of 3.0 angstroms for EVS14-2-AB, 3.8 angstroms for ECLVS58, 3.7 angstroms for ECLVS59 and 3.6 angstroms for ECLVS82.

The dynamic binding site of ECLVS14, shown in **Figure 43** consists of the N-terminus and ECL2 posterior to the compound. One of the benzene substituents of the 1-BPMAM moiety of ECLVS14 interacts with the aliphatic Ile6 side chain as well as the backbone of the Ser5, Tyr7, and Thr8 residues. The terminal benzene of the compound interacts with Val197 and Val198. The indolin-2-one moiety interacts with the side chain of Trp195 through π - π stacking interactions as well as with the Leu194 side chain.

The dynamic binding site of EVS14-2-AB, shown in **Figure 44** consists anteriorly of the C-terminus and ICL2, and superiorly and posteriorly of ICL2. The side chain of Ser144 interacts with the ethenamine moiety of the compound, while the indoleine moiety interacts with the guanidine of Arg148 by π - π stacking. The hydrazine linker region interacts with the backbone of Pro147, as well as the side chain of Asn143. The terminal benzene is bound anterior and

Figure 42: RMSD values of CXCR4 bound to ECLVS compounds across 5 ns of the production trajectory. Total RMSD calculated at the C α position presented in green. Helical domains are presented in blue, and loop regions are presented in red. The CXCR4 structure was noted to be stable, and the average structure was noted to represent the conformation present in the trajectory after 0.5 ns of simulation time. A) ECLVS14, B) EVS14-2-AB, C) ECLVS58, D) ECLVS59, E) ECLVS82

Figure 42: RMSD values of CXCR4 bound to ECLVS14 across 5 ns of the production trajectory. A) ECLVS14, B) EVS14-2-AB, C) ECLVS58, D) ECLVS59, E) ECLVS82

Figure 43: Dynamic binding site of ECLVS14 to native CXCR4. Critical residues for ECLVS14 binding to CXCR4 are highlighted. ECLVS14 is depicted in red ball-stick representation, while CXCR4 residues within 3 angstroms of the compound are presented in stick representation and superstructure of CXCR4 is presented as cyan ribbon. ECLVS14 interacts with the N-terminus and ECL2 regions of the extracellular surface.

Figure 43: Dynamic binding site of ECLVS14 to native CXCR4.

Figure 44: Dynamic binding site of EVS14-2-AB to native CXCR4. Critical residues for EVS14-2-AB binding to CXCR4 are highlighted. EVS14-2-AB is depicted in red ball-stick representation, while CXCR4 residues within 3 angstroms of the compound are presented in stick representation and superstructure of CXCR4 is presented as magenta ribbon. EVS14-2-AB interacts with the C-terminus and ICL2 regions of the Intracellular surface.

Figure 44: Dynamic binding site of EVS14-2-AB to native CXCR4.

posterior by side chain of Ser352 to the anterior surface of the aromatic ring, and by the guanidine of Arg146 by π - π stacking posterior to the compound. As the original docked conformation placement of EVS14-2-AB was on the extracellular surface of the receptor, the fact that the compound stably and dynamically bound to the intracellular surface of the receptor came as a surprise. However, the compound structure of EVS14-2-AB was sufficiently different from the parent compound to suggest that it may bind in a separate conformation than ECLVS14.

The dynamic binding site of ECLVS58, shown in **Figure 45** consists of the N-terminus posterior to the compound, ECL2 inferior to the compounds and ECL3 anterior to the anterior of the compound. The methyl end of the methyl piperidin-1-one moiety is stabilized by interaction with the aliphatic elements of the side chain of Glu32 and the backbone of Glu31. One of the pyridine moieties interacts with the aliphatic portion of Tyr12 as well as with anterior and posterior cysteine residues 274 and 28. The carboxyl end of the methyl piperidin-1-one moiety interacts with the guanidine elements of the Arg183 side chain. The other pyridine moiety interacts with the aliphatic elements of the Arg183 side chain as well as the side chain of his 281.

The dynamic binding site of ECLVS59, shown in **Figure 46** consists of ECL3 posterior to the compound, ECL2 inferior to the compound and the N-terminus anterior to the compound. The methoxybenzene moiety is stabilized by interaction with the aliphatic elements of the phe276 and Val280 side chains. The cyanoacetimidamide is stabilized by multiple hydrogen bonds, between the backbones of Cys274, Thr279 and His281 as well as the guanidine of Arg 183.

Figure 45: Dynamic binding site of ECLVS58 to native CXCR4. Critical residues for ECLVS58 binding to CXCR4 are highlighted. ECLVS58 is depicted in red ball-stick representation, while CXCR4 residues within 3 angstroms of the compound are presented in stick representation and superstructure of CXCR4 is presented as yellow ribbon. ECLVS58 interacts with the N-terminus, ECL2 and ECL3 regions of the extracellular surface.

Figure 45: Dynamic binding site of ECLVS58 to native CXCR4.

Figure 46: Dynamic binding site of ECLVS59 to native CXCR4. Critical residues for ECLVS59 binding to CXCR4 are highlighted. ECLVS59 is depicted in red ball-stick representation, while CXCR4 residues within 3 angstroms of the compound are presented in stick representation and superstructure of CXCR4 is presented as orange ribbon. ECLVS59 interacts with the N-terminus and ECL2 regions of the extracellular surface.

Figure 46: Dynamic binding site of ECLVS59 to native CXCR4.

The dynamic binding site of ECLVS82, shown in **Figure 47** consists of ECL3 and the proximal portion of the N-terminus lateral to the compound, and the distal portion of the N-terminus posterior to the compound. One of the benzamide elements is stabilized by interaction with the aliphatic elements of the Phe276 side chain, as well as the backbones of the Glu275 and Cys274 residues inferior to the compound, and the backbones of the Glu31 and Glu32 residues. The phenanthroline element of the compound interacts with the side chain of Glu32 to the anterior of the compound, and with aliphatic elements of the Met6 and Thr13 side chains posterior to the compound, and the Asn11 and Tyr12 side chains lateral to the compound.

These dynamics determined structures shed light on the potential mechanism of the activity of the ECLVS compounds. The compounds primarily interacted with the N-terminal and the second extracellular loop regions. Stabilization of these regions could allosterically disrupt the conformation of the CXCL12 binding site, or prevent movement of TM7 necessary for effective signal transduction. Interaction between the compounds and ECL2 could also directly interfere with CXCL12 ligand binding. Interaction with the intracellular surface of the receptor could produce effects similar to those seen with the ICLVS compounds (described in Chapter III). Interaction by the compounds with the C-terminal region could potentially stabilize TM7, or interfere with the interface between the C-terminus and Gai. Interposition of the compound between the C-terminus and ICL2 could prevent effective formation of the cleft which would be occupied by bound Gai.

Figure 47: Dynamic binding site of ECLVS82 to native CXCR4. Critical residues for ECLVS82 binding to CXCR4 are highlighted. ECLVS82 is depicted in red ball-stick representation, while CXCR4 residues within 3 angstroms of the compound are presented in stick representation and superstructure of CXCR4 is presented as orange ribbon. ECLVS82 interacts with the N-terminus and ECL3 regions of the extracellular surface.

Figure 47: Dynamic binding site of ECLVS82 to native CXCR4.

ECLVS compound binding alters the conformation of the extracellular and intracellular surfaces of CXCR4.

Given the interactions of the ECLVS compounds with the extracellular surface of CXCR4, we sought to investigate what effect inhibitor binding could have on overall receptor conformation. To do this, we sought to compare the structure of the ECLVS complexed conformations of CXCR4 to the inactive native conformation and the constitutively active N119S conformations of the protein.

The comparison between the structures of the ECLVS14-bound, native CXCR4, and the N119S conformation is characterized by differences in loop positioning. Although the structure of the loops is conserved between the three conformations, there are several differences in positioning. In the N-terminal region, the anti-parallel β -sheet is more compact and the distance between ECL1 and the N-terminal region is decreased, mimicking the N119S conformation by movement of ECL. The entire region N-terminal is displaced away from the extracellular surface of the TM regions and towards ECL3. ECL3 is outwardly displaced from the center of the receptor. At the intracellular surface of CXCR4, ICL1 and the C-terminus of the protein retain the native conformation and position, while ICL2 is shifted opposite that of the N119S conformation and towards the center of the molecule and ICL1. ICL3 adopts the shape of the native conformation, while the apex of the loop is shifted towards the C-terminus. The comparison between the structures of the EVS14-2-AB, native CXCR4, and the N119S conformation is characterized by differences in loop positioning.

Although the structure of the loops is conserved between the three conformations, there are several differences in positioning. In the N-terminal region, the anti-parallel β -sheet is preserved, but more linear than in either the native or N119S conformations. The region is also shifted away from ECL2 and towards the ECL3. ECL1 and ECL largely adopt the native conformation, while ECL3 is outwardly displaced from the center of the receptor. At the intracellular surface of CXCR4, ICL1, ICL2 and the C-terminus essentially adopt the native conformation of the receptor. The conformation of ICL3 is shifted such that the apex of the loop is shifted towards the C-terminus and the center of the receptor. The comparison between the structures of the ECLVS58, native CXCR4, and the N119S conformation is characterized by differences in loop positioning. In the N-terminal region, the anti-parallel β -sheet is more compact and the distance between ECL1 and the proximal portion of the N-terminal region is decreased, mimicking the N119S conformation by movement of ECL1. The anti-parallel β -sheet is also shifted away from ECL2, resulting in a conformation significantly different from either the native or N119S conformations. At the intracellular surface of CXCR4, ICL1 retains the native conformation and position, while ICL2 is shifted opposite that of the N119S conformation and towards the center of the molecule and ICL1. ICL3 adopts the shape of the native conformation, but a mild flexion of TM6 shifts the region towards the center of the molecule and ICL3.

More significant differences characterize the comparison of the structures of the ECLVS59, native CXCR4, and the N119S conformation. In the N-terminal region, the anti-parallel β -sheet is preserved, but more compact than the native

conformation. In addition, the proximal portion of the region is much more disordered adopting a more extended loop structure than in either the native or N119S conformations. ECL1 is shifted towards the N-terminal region, and the ECL2 region is shifted away from the extracellular surface of the TM helices. The apex of the ECL3 loop is shifted away from the N-terminus. At the intracellular surface of CXCR4, ICL1 and the C-terminus essentially adopt the native conformation of the receptor. ICL2 is shifted opposite that of the N119S conformation and towards the center of the molecule and ICL1 while the conformation of ICL3 is shifted such that the apex of the loop is shifted towards the C-terminus and the center of the receptor.

The comparison between the structures of the ECLVS82, native CXCR4, and the N119S conformation is characterized by conservation of the native CXCR4 structure. In the N-terminal region, each of the ECL loops and the N-terminal region adopt the native conformation of the receptor. At the intracellular surface of CXCR4, ICL1 is shifted slightly downward away from the surface of the TM regions, while ICL2 is shifted opposite that of the N119S conformation and towards the center of the molecule and ICL1. The apex of ICL3 is displaced away from that of the native conformation and away from the C-terminus and the TM surface. The C-terminus is displaced slightly downward from the conformation, while the distal portion of the loop is more extended and disordered than in the native structure.

These compound: CXCR4 complex conformations suggest that the receptors adopt intermediary conformations of the native and N119S

conformation, characterized by movement of the N-terminal regions and movement of the ICL loops towards the center of the receptor.

ECLVS compound binding alters the conformation of the transmembrane helices of CXCR4.

Given the alterations of the ECLVS compounds of the surface of CXCR4, we sought to investigate what effect inhibitor binding could have on the structure of the transmembrane helices. We compared the reduced representations of the transmembrane helices for each of the ECLVS complexed conformations of CXCR4 to the inactive native conformation and the constitutively active N119S conformation of the protein. Multiple receptor conformations were compared using the helixanal program developed by Dr. Trent.

The structure of the ECLVS14 transmembrane helices is presented in **Figure 48a**. TM1, TM2 and TM3 structure and positioning is conserved between the complexed conformation and the native form. The TM4 kink at Pro170 and flexion at Ile173 aligned with the native conformation of the receptor, although the helix is shifted towards TM5. The TM5 bulge present in the native structure between Met205 and Ile213 is present in the complexed structure. TM6 helix is more compact in conformation than native, and the apical turn of the helix is displaced towards the plasma membrane. The apex of TM7 is less flexed at Glu288 than in the native conformation, displacing the apex of the helix towards an intermediate position between N119S and native.

The structure of the EVS14-2-AB transmembrane helices is presented in **Figure 48b**. TM1 adopts the conformation of the native helix other than a flexion

Figure 48: Transmembrane regions of the ECLVS compounds complexed with CXCR4 adopt a similar conformation to native CXCR4. Reduced representations of receptor transmembrane helices are presented: native CXCR4 in green, N119S in red, ECLVS14 conformation in teal, EVS14-2-AB conformation in violet, ECLVS58 conformation in yellow, ECLVS59 conformation in orange, ECLVS82 conformation in blue.

Figure 48: Transmembrane regions of ECLVS compounds complexed with CXCR4 adopt a similar conformation to native CXCR4.

of Ile44 displacing the apex of the helix towards TM2. TM2, TM3, TM4, TM5 and TM6 structure and positioning is conserved between the complexed conformation and the native form. Angularity at the apex of TM7 align with native structure, while the base of the helix is more linear than in the native conformation, lacking the kink at Leu301, which results in an outward displacement of the base of the helix.

The structure of the ECLVS58 transmembrane helices is presented in **Figure 48c**. TM1 is more linear in conformation than native and lacks a kink at Val54 present in the N119S conformation. TM2, TM3 and TM4 adopt the structure and positioning of the native conformation of the receptor. The although the apical position of the TM5 helix matches the native conformation, the apical portion of the bulge more closely resembles the more extended structure of the N119S mutant. Torsion of the helix shifts the apex of the bulge outward from the center of the receptor, while a decreased flexion at the base of the bulge shifts the base of the helix away from the center of the receptor. TM6 helix is more linear than native, without bowing of N119S, apical turn is intact. TM7 apical flexions align with native structure, while the base of the helix is more linear than in the native conformation, lacking the kink at Leu301, which results in an outward displacement of the base of the helix.

The structure of the ECLVS59 transmembrane helices is presented in **Figure 48d**. TM1 is more linear in conformation than native and lacks a kink at Val54 present in the N119S conformation, producing an outward displacement of the apex of the receptor. TM2, TM3, TM4 and TM6 structure and positioning is

conserved between the complexed conformation and the native form. The TM5 bulge present in the native structure between Met205 and Ile213 is present in the complexed structure, but the bulge is more extended than in the native structure. Angularity at the apex of TM7 aligns with native structure, while the base of the helix is more linear than in the native conformation, lacking the kink at Leu301, which results in an outward displacement of the base of the helix.

The structure of the ECLVS82 transmembrane helices is presented in **Figure 48e**. TM1 is more linear in conformation than native and lacks a kink at Val54 present in the N119S conformation, resulting in displacement of the apex and base of the helix. The remaining helices adopt the structure and positioning of the native conformation.

Further examination of the transmembrane helices involved assessment of the distances between the intracellular and extracellular termini of the interhelical segments. At the extracellular surface, as seen in **Table 11**, the compounds adopted several specific conformations. ECLVS14 produced similar interhelical distances to the native conformation of the receptor aside from the TM1-2, TM2-3, TM1-6, and TM2-6 distances, which were all increased compared to the native conformation for the receptor. EVS14-2-AB produced increases in the TM1-7 and TM1-6 interhelical distances when compared to the native conformation, while the TM-15 distance was comparable to the N119S distance. ECLVS58 produced increased interhelical distances between TM1-2 and TM1-7, while the interhelical distance between TM2-3 was decreased, and the TM5-6 distance was intermediate between native and N119S. ECLVS59 produced

Table 11: ECLVS compounds modulate the distance between the extracellular termini of the transmembrane helices of CXCR4. Interhelical distances measured at extracellular termini of transmembrane domains of native CXCR4, constitutively active N119s, and compound docked structures. Distances measured at the extracellular termini of reduced representations of the averaged structures. Illustrated interhelical distances demonstrate 1.5 angstrom or greater deviation from either the Native or N119S structure.

Table 11: ECLVS compounds modulate the distance between the extracellular termini of the transmembrane helixes of CXCR4.

TMA	TMB	Native	N119S	ECLVS 14	EVS14- 2-AB	ECLVS 58	ECLVS 59	ECLVS 82
1	2	6.5	7.6	7.3	7.0	7.9	9.4	5.9
5	6	11.2	13.1	11.3	11.1	12.3	10.5	11.2
6	7	12.1	13.1	12.4	12.2	12.0	13.2	11.5
1	7	15.8	14.6	16.2	16.9	16.7	18.2	11.5
2	3	18.3	18.6	19.4	18.2	17.3	18.1	17.5
2	7	19.1	19.3	19.9	19.5	20.5	19.4	18.8
1	3	20.9	22.2	21.9	21.6	20.5	22.9	19.6
1	4	21.9	23.3	22.4	22.9	21.2	24.9	20.9
1	6	26.0	25.4	27.0	27.8	26.3	27.8	23.3
2	6	28.7	28.0	30.3	29.6	29.2	29.5	27.1
1	5	33.6	34.8	34.3	35.3	33.6	35.2	27.9

increases in the interhelical distances of TM1-2, TM1-4, and TM1-6, while decreasing the interhelical distance between TM5-6. TM6-7 and 1-5 adopted an interhelical distance comparable to the N119S conformation. Finally, ECLVS82 decreased the interhelical distances between TM1-2, TM1-7, TM1-6, Tm 1-5 and TM2-6.

At the intracellular surface, as seen in Table 12, the compounds adopted several specific conformations. ECLVS14 produced an increased distance between TM6-7 when compared to native, while the distances between TM2-6, TM3-4, and TM4-6 were intermediate between the native and N119S conformations. EVS14-2-AB produced an increased distance between TM2-7, Tm3-7 and TM6-7 when compared to native, while the distances between TM3-4 and TM5-7 were intermediate between the native and N119S conformations. ECLVS58 produced an increased distance between TM3-6 when compared to native, while the distances between TM2-6 were intermediate between the native and N119S conformations. The interhelical distances between TM1-3, TM1-5, TM2-5, and TM2-7 were decreased when compared to the native conformation. ECLVS59 produced an increased distance between TM2-7, TM3-7, TM4-5, Tm4-7 and TM6-7 when compared to native, while the distances between TM2-5, TM3-4, TM-3-5 and TM5-7 were intermediate between the native and N119S conformations. The interhelical distances between TM1-3 were decreased when compared to the native conformation. ECLVS82 produced an increased distance between TM6-7 when compared to native, while the distance between TM3-4 was intermediate between the native and N119S conformations.

Table 12: ECLVS compounds modulate the distance between the intracellular termini of the transmembrane helices of CXCR4. Interhelical distances measured at intracellular termini of transmembrane domains of native CXCR4, constitutively active N119S, and compound docked structures. Distances measured at the intracellular termini of reduced representations of the averaged structures. Illustrated interhelical distances demonstrate 1.5 angstrom or greater deviation from either the Native or N119S structure.

Table 12: ECLVS compounds modulate the distance between the intracellular termini of the transmembrane helices of CXCR4.

TMA	TMB	Native	N119S	ECLVS 14	EVS14 -2-AB	ECLVS 58	ECLVS 59	ECLVS 82
2	6	10.7	12.9	11.1	10.8	10.3	11.1	10.6
3	5	11.5	16.5	10.9	11.3	11.9	13.2	11.6
2	7	11.7	12.3	11.9	14.4	11.0	13.7	11.7
3	6	12.0	13.4	11.5	11.5	12.5	11.2	11.0
6	7	12.3	11.8	13.4	16.5	11.5	15.5	13.9
3	4	16.0	12.6	15.6	15.2	15.6	15.5	16.4
4	6	19.5	17.6	18.8	18.7	19.1	19.5	19.0
3	7	21.5	22.2	21.9	25.4	21.0	23.8	22.0
4	7	22.3	21.0	22.2	24.9	21.6	24.2	22.1
2	5	22.5	25.3	22.6	22.3	21.6	23.7	22.6
1	3	24.6	25.7	24.3	24.6	24.6	22.9	24.1
5	7	24.8	23.7	25.3	27.8	21.8	27.7	25.4
4	5	26.6	26.8	25.5	25.7	26.3	28.0	27.1
1	5	32.2	33.9	31.7	31.5	29.8	32.2	31.6

The overall representation of these interhelical distance measurements and the reduced representation analysis speaks to a difference in activity between the anti-CXCR4 compounds. ECLVS14 binding seemed to alter the interhelical distances by modulating TM2 and TM6 position at the extracellular surface producing a more open conformation of the receptor on this surface, and seemed to modulate the interaction between TM6 and the other helices at the intracellular surface to produce a similarly open conformation at this surface. This suggests that ECLVS14 may drive the receptor towards an intermediate conformation, more open than the native, but not as open as the constitutively active N119S conformation.

EVS14-2-AB binding alters TM1 interaction with the other helices at the extracellular surface, and TM7 at the intracellular surface, primarily through its interactions with the C-terminus. ECLVS58 and ECLVS59 primarily increased intrahelical distances at the extracellular surface and decreased intrahelical distances at the intracellular surface, suggesting that binding of these compounds could drive the intracellular surface of the receptor towards a more closed conformation which would be less receptive to G-protein binding. ECLVS82 on the other hand, decreased intrahelical distances at the extracellular surface, suggesting that it was driving the receptor to a more closed conformation at this surface, which could stabilize the receptor in an inactive conformation.

Discussion

There is been a dearth of successful compounds which are capable of acting as neoadjuvant chemotherapies for anti-metastasis. We have utilized SBDD to target CXCR4 with homology modeling and virtual screening to develop novel inhibitors of CXCR4. In this work, non-peptide CXCR4 antagonists were structurally characterized through QSAR analysis of structural derivatives.

These compounds represent an investigation into the structural determinants of activity and the role that the subgroups of ECLVS14 play in determining its anti-CXCR4 activity. Retesting of the original ECLVS compounds identified ECLVS14 as a selective inhibitor of CXCR4-mediated chemotaxis and ECLVS58 as a non-selective inhibitor of CXCR4-mediated chemotaxis, neither of which inhibited CXCR4-mediated intracellular calcium flux. This is an important distinction to make, as there are multiple signaling pathways that could potentially influence CXCR4-mediated chemotaxis. Cell migration due to CXCR4 activation is the result of activation of multiple signaling pathways[133-135], including the ERK1/2, JNK, MAPK, and GSK 3 α / β phosphorylation systems, [129] activation of phospholipase C to enhance calcium mobilization [130], activation of the phosphoinositide-3-kinase pathway to activate AKT [130], and activation of PKA or CREB [131, 132]. Thus anti-chemotaxis activity of ECLVS14 and ECLVS58 could be due to inactivation of multiple signaling pathways outside of calcium mobilization.

Derivatives of the parent ECLVS14 structure revealed several important structural-activity relationships governing ECLVS14 activity, seen in Figure 49.

Figure 49: Structural modifications of ECLVS14 and their effects on activity. Data from chemotaxis and cytotoxicity assays were analyzed to correlate the chemical structure of available analogs to inhibitory activity.

Figure 49: Structural modifications of ECLVS14 and their effects on activity

Analysis of the 1-[bis (phenylmethyl) amino] methyl moiety revealed that absence of this group diminished anti-CXCR4 activity as confirmed by EVS14-20.

However, comparison between compounds EVS14-19 and EVS14-2-1 which differed solely in the presence or absence of the 1-BPMAM moiety revealed no significant difference in anti-CXCR4 activity.

Modifications of the hydrazine-alkyl linker moiety to ethyl, propyl, and butyl derivatives revealed that as the hydrazine-alkyl linker is lengthened, activity decreases, as seen with compounds EVS14-1, EVS14-2, and EVS14-3. There was no direct comparison available between the parent compound with a hydrazine-alkyl linker moiety and one which was altered to the shortened acetamide moiety. However, several compounds had modified 1-BPMAM moieties and acetamide along with comparable terminal benzene modifications. In each case the compound with the modified 1-BPMAM moiety and acetamide demonstrated increased anti-CXCR4 activity. And as the absence of the 1-BPMAM moiety has been shown to have either a negligible or negative effect on anti-CXCR4 activity it would stand to reason that shortening the hydrazine-alkyl linker moiety of the parent compound to an acetamide would increase anti-CXCR4 activity. The benefits of this alteration would have to be balanced by the increased potential for cytotoxic side effects of conversion of the compound to an acetamide.

Modification of the terminal benzene by halogenation, methylation, nitroxylation or conversion to a cyclhexylideneamine disrupted the activity of the parent compound, as seen in compounds EVS14-12, EVS14-13, EVS14-14,

EVS14-15, EVS14-16, EVS14-17, and EVS14-18. However, increases in anti-CXCR4 activity were seen in compounds with increasing electronegativity as well as acetamide replacement and removal of the 1-BPMAM moiety.

It is important to note that ECLVS14 did not significantly impair intracellular calcium mobilization. In fact, the vast majority of the ECLVS14 derivatives were incapable of inhibiting this signal transduction mechanism. The relative insensitivity of calcium mobilization to compound ECLVS14 suggests that this compound may allow selective inhibition of CXCR4-mediated chemotaxis without modification of calcium signaling. ECLVS14 is very unique in its inability to modify calcium signaling, a characteristic that is not shared by any other CXCR4 inhibitors [289-291], including T140 [195] and AMD3100 [275].

Our 5 ns molecular dynamics simulations have allowed us to gain a finer mechanistic understanding of the interaction between CXCR4 and ECLVS14. We have found that ECLVS14 drives CXCR4 into a unique conformation. It seems to be a more open extracellular surface than seen in the native conformation, but less open than seen in the N119S conformation.

This work represents a significant examination of the activity and development of an anti-CXCR4 lead compound with direct application to the inhibition of cancer metastasis as a neoadjuvant chemotherapeutic agent. Investigation of multiple generations of ECLVS14 derivatives as well as extensive molecular dynamics simulations have allowed us to carry forward the structure based drug design program and gain greater insight into the function of CXCR4 and its interaction with a potent inhibitor.

CHAPTER III

DISCOVERY OF INTRACELLULAR LOOP TARGETED COMPOUNDS BY VIRTUAL SCREENING

Introduction

As depicted in **Figure 50**, there are 5 sites where signal transduction by CXCR4 can be disrupted. One could attempt to bind or sequester CXCL12 to prevent CXCR4 from interacting with its cognate ligand. One could attempt to inhibit the interaction of CXCL12 and the extracellular surface of the receptor by blocking the interaction points between CXCL12 and CXCR4 on the extracellular surface of the receptor. One could attempt to stabilize the transmembrane helices in an effort to prevent extracellular to intracellular signal transduction. One could attempt to prevent the interaction of G-proteins with the intracellular surface of CXCR4 by attempting to block the interaction points between CXCR4 and the G proteins. Finally, one could attempt to bind and sequester the G protein in such a manner that it will not interact with CXCR4.

Although CXCL12 knockdown has been demonstrated to inhibit tumor cell proliferation and migration [292, 293], these genetic therapies are not yet applicable to the clinical setting. In addition, in the face of rising concentrations of CXCL12, as would be expected in favorable microenvironments by CXCR4 expressing neoplastic cells [34-36], inhibition by binding and sequestering

Figure 50: There are 5 potential structural targets for inhibiting CXCR4

signaling: 1) Targeting the ligand CXCL12 to prevent ligand/receptor interaction,
2) Targeting the Extracellular ligand binding site as in AMD3100 and T140,
3) Stabilizing the transmembrane helices to prevent ligand induced conformation changes, 4) Targeting the Intracellular region of CXCR4 at the G-Protein interface, 5) Targeting the G-protein to prevent GPCR-G-protein interaction.

Figure 50: There are 5 potential structural targets for inhibiting CXCR4 signaling

CXCL12 would become increasingly difficult. In spite of this, a series of compounds has been designed which selectively bind CXCL12. These compounds, chalcones (Figure 51), inhibit intracellular calcium mobilization, ligand binding and chemotaxis mediated by CXCL12 at an IC₅₀ of 100 nM [294].

The majority of anti-CXCR4 drug discovery efforts have been targeted at the extracellular surface of the receptor, or Target 2 in Figure 50. CTCE-9908 is a palindromic 17-amino acid cyclic peptide analog of CXCL12 as depicted in Figure 52, which features dual 8 amino acid repeats which mimic the 8 amino acid N-terminus of CXCL12 [176]. Binding of CTCE-9908 is purported to mimic the CXCL12 binding model, primarily interacting with the N-terminal region of CXCR4. Identified initially as anti-HIV agents, many other anti-CXCR4 compounds act at the extracellular surface of CXCR4 including AMD3100 and T140 which are depicted in Figure 52. AMD3100 strongly interacts with residues Asp171, Asp262 and Glu288, located on the extracellular side of TM4, TM6 and TM7 [295-297]. T140 utilizes a similar binding mechanism interacting with residues Asp171, Arg188, Tyr190, Gly207 and Asp262 [200]. The net effect of these interactions is that although some of the specific residues may be different, both AMD3100 and T140 interact at the CXCL12 binding site, and inhibit signal transduction, CXCL12 –CXCR4 interactions, and stabilizing the conformation of the native receptor or driving the conformation into an intermediate position which does not effectively transmit signal [200].

Direct targeting of the TM region of CXCR4 would be a difficult task. Unlike many GPCRs, CXCR4 lacks a central channel leading to an

Figure 51: Chalcones inhibit CXCR4 signaling by selectively binding CXCL12. Chalcones are small molecules which neutralize CXCL12. These compounds effectively and selectively inhibit CXCL12 mediated CXCR4 internalization, chemotaxis, and calcium mobilization in addition to CXCR4:CXCL12 ligand binding.

Figure 51: Chalcones inhibit CXCR4 signaling by selectively binding CXCL12.

Figure 52: Anti-GPCR compounds. Two different classes of compounds are illustrated in Figure 69. Those in Column A interact with CXCR4 at the extracellular surface of the GPCR. Those in Column B interact with GPCRs at the G-protein level, inhibiting $G\alpha$ and $G\beta\gamma$.

Figure 52: Anti-GPCR compounds

Column A
CTCE-9908

Column B
YM-254890

AMD3100

BIM-46174

T140

Selenocytamine

M119K

intramembranous rhodopsin-like binding pocket, rather the pocket is larger, more open and located more superficially in the extracellular surface of the receptor than in other GPCRs [257]. Structural studies of CXCR4 have also suggested that movement of TM7 compared to the rhodopsin structure reduces the size of the intramembrane region which would correspond to a retinal binding site [200].

Drug development targeting G-proteins

The intracellular surface of CXCR4 is largely undiscovered country for drug development, in that there have been no prior attempts to inhibit CXCR4 function by targeting the intracellular surface of the receptor, and only a handful of attempts to target G-proteins as a method for inhibiting GPCRs. G-proteins transmit GPCR mediated signals through a series of reactions. G-proteins are heterotrimers composed of $G\alpha$, $G\beta$ and $G\gamma$. Upon interaction with an activated GPCR, GDP bound to $G\alpha$ will be exchanged for GTP, facilitating a separation of the heterotrimer into $G\alpha$ and $G\beta\gamma$ subunits. These subunits in turn activate downstream signaling pathways until intrinsic GTP-ases in $G\alpha$ convert GTP to GDP, facilitating the reformation of the G-protein heterotrimer.

Antagonism of the interaction between $G\alpha$ and GPCRs has been demonstrated to modulate receptor function. Peptides based on the C-terminal residues of $G\alpha_i$ and $G\alpha_o1$ have been found to inhibit agonist binding and signal transduction in A1 adenosine receptors [298]. YM-254890 depicted in Figure 52 is a specific $G\alpha_q/11$ inhibitor [299] designed to prevent acute thrombosis and neointima formation. YM-254890 was purified from *Chromobacterium* sp. QS36666 [251], and inhibits $G\alpha_q/11$ function by preventing GDP/GTP exchange

by the $G\alpha$ subunit [300]. YM-254890 mediated effects on thrombosis and response to vascular injury were compared to other antithrombotic compounds: P2Y₁₂ receptor antagonists ticlopidine and clopidogrel [246-251] and was found to have greater inhibitory potency than the anti-GPCR compounds. This makes for an interesting proof of principle in that it demonstrates that inhibition of GPCR function can be as effective at the G-protein level as it is at the GPCR level.

More substantial efforts have targeted $G\beta\gamma$ as a method of inhibiting GPCR signal transduction. BIM-46174 (Figure 52) has been developed as an inhibitor of the heterotrimeric G-protein complex, and effectively inhibits GPCR signaling in these cells while inhibiting the growth and proliferation of several cancer cell lines, although this activity also extended to non-tumor cell lines such as HEK-293 [301]. The $G\beta\gamma$ subunit has been successfully targeted by virtual screening for the development of anti- $G\beta\gamma$ compounds, which actively inhibit G-protein mediated signaling including calcium mobilization [253]. Covalent modification of the $G\beta\gamma$ subunit by selenocystamine (Figure 52) has also been found to selectively inhibit $G\beta\gamma$ function [256]. Unfortunately, non-specific modulation of $G\beta\gamma$ function would produce widespread failures in GPCR function, with potentially devastating outcomes in the immune compartment such as myelosuppression.

However, a selective $G\beta\gamma$ inhibitor could potentially have great therapeutic utility. M119K, depicted in Figure 52, is one such compound which inhibits $G\beta\gamma$ and acts as an anti-pain medication by modulating the μ -opioid receptor [252-

254]. Additional use of this compound has expanded its scope to potentially include treatment of metastatic disease. M119K was found to inhibit migration and invasion of MDA-MB-231 and MDA-MB-436 cells, but perhaps most importantly, M119K was found to selectively inhibit CXCR4 mediated migration and chemotaxis by selectively inhibiting G $\beta\gamma$ functions, while G α_i functions remained intact [255]. This compound demonstrates that CXCR4 function can be successfully inhibited by targeting the intracellular portion of the signaling apparatus.

Development efforts targeting the interface between GPCRs and G-proteins has been especially sparse, with the only efforts tied to development of small molecule inhibitors of the interaction between rhodopsin and its G-protein transducin [302, 303]. The first of these studies utilized a reduced intracellular loop representation of the rhodopsin crystal structure to identify a series of compounds which inhibited transducin binding to rhodopsin [303]. In the second study, a portion of the transducin α subunit was used as a pharmacophore model to identify compounds which stabilized the active state of Rhodopsin [302].

Heterotrimeric G-proteins only interact with active CXCR4 in metastatic breast cancer cell lines.

One of the origins of this work was in a finding by Dr. Kelby Napier. Dr. Napier had generated a series of molecular dynamics structures featuring mutants of the TM3 residue asparagine 119 (N119). This series of mutants was based on findings by Zhang *et al* which found that N119 mutation to Ser (N119S) or Ala (N119A) produced a constitutively active conformation of CXCR4, while

mutation to Lys (N119K) produced a constitutively inactive conformation of the receptor [127]. One of Dr. Napier's objectives was to attempt to recreate the interaction between CXCR4 and a heterotrimeric G-Protein, but he found that he was unable to generate a structure with a viable interaction between the native CXCR4 conformation and the heterotrimeric G-Proteins. While he was also unable to produce a viable interaction structure using the N119K conformation of CXCR4, CXCR4 did interact with the G-protein when the N119S and N119A structures were used. This suggested that the heterotrimeric G-proteins only interacted with activated CXCR4 conformations. This finding is supported by clinical data. Holland *et al* demonstrated that although the G α and G $\beta\gamma$ subunits could interact individually with CXCR4 in both metastatic and non-metastatic breast cancer cell lines, the heterotrimer G $\alpha\beta\gamma$ only interacted with CXCR4 in the meta static cell lines [137]. This data in combination with the proof of principle studies demonstrating that the G $\beta\gamma$ inhibitor M119 can inhibit migration and invasion of MDA-MB-231 and MDA-MB-436 breast cancer cells [255], suggests that inhibition of CXCR4 targeting the intracellular surface of the receptor is a viable target to inhibit CXCR4 mediated metastasis.

Virtual screening as a method for drug discovery

Identification of new small molecule inhibitors begins with identification of novel lead compounds. Virtual screening presents the capability to prescreen vast databases, such as the Zinc compound library, a 13 million compound virtual database [208], by targeting a three dimensional structure to determine those compound which best fit, or dock, into a chosen allosteric or ligand binding

site. This method has proven successful in the identification of a wide variety of targets [209-225].

This method is dependent on the development of an accurate three-dimensional structure. GPCR crystal structures other than bovine rhodopsin [104] have only recently become available so the majority of virtual screening targeting GPCRs has been accomplished using homology modeling, in spite of the fact that there is often low sequence similarity between rhodopsin and the target GPCR. Although new high resolution crystal structures of the β 2-adrenergic receptor [108] and the α 2A adenosine receptor [111] have been released, in each of these structures intracellular loop 3 has been replaced with T4 lysozyme and the C-terminus has been truncated and replaced with a histidine purification epitope. Thus, use of these structures would have required homology modeling based on rhodopsin to generate structures for these two regions. Novel crystal structures for CXCR4 were released on November 19, 2010 [257], too late to be used in any of these studies. These structures carried the same T4 Lysozyme and C-terminal substitutions as seen in the β 2-adrenergic and α 2A adenosine receptors. Thus modeling of CXCR4 focusing on the intracellular surface of the receptor based on these crystal structures would require extensive homology modeling to develop a workable model. Further, such drastic modifications of ICL3 and the C-terminus could have dramatic implications on the positioning and structure of their attached transmembrane helixes, TM5, TM6 and TM7.

One method of performing virtual screening is Surflex-Dock [234, 244]. This docking program utilizes protomols, or pseudo-molecules, to virtually define

binding sites targeting the protein structure. Protomol development can be based on ligands, inhibitors or partner proteins acting at the targeted site, or residues of the protein itself which comprise or surround an active site can be used to define the protomol. The target protein can be screened by attempting to dock compounds into the defined protomol site. Surflex-Dock conducts flexible docking by fragmenting the compounds into components and docking the fragments into the defined site, constructing the compound in the active site based on favorable positions for each of the compound fragments. Multiple poses of each compound are returned to the user and assigned a score based on binding affinity, degree of inappropriate penetration into the protein, and internal structural conflicts.

In this chapter, we have used SBDD to identify inhibitors of the CXCR4 by targeting the intracellular surface of the receptor. As illustrated in Figure 53, we have used Surflex-Dock to virtually screen the interface between constitutively active N119S CXCR4 and the heterotrimeric G-protein. We identified compounds which interacted with either the CXCR4 or the G α side of the interface, and several of these compounds significantly inhibited chemotaxis for an exceptionally high initial hit rate of sixty-two percent. Two of these compounds, ICLGa28 and ICLGa36, produced IC₅₀ values of 1 μ M. Two other compounds, ICLGa9 and ICLGa25, specifically inhibited CXCL12-induced calcium mobilization with an IC₅₀ value of 1 μ M. These compounds exhibited almost no cytotoxicity, and several were also capable of inhibiting tumor cell proliferation. These compounds represent the first successful attempt to target the intracellular interface of a non-rhodopsin GPCR for drug development and drug discovery.

Figure 53: SBDD of inhibitors of the intracellular surface of CXCR4. Surflex-Dock was used to virtually screen the interface between constitutively active N119S CXCR4 and the heterotrimeric G-protein. We identified thirty-seven compounds which displayed interacted with either the CXCR4 or the G α side of the interface, and eleven of these compounds significantly and specifically inhibited CXCR4 with an IC₅₀ value of 10 μ M in either chemotaxis or calcium mobilization. The interaction between the compounds and CXCR4 was further rationalized by molecular dynamics.

```

graph TD
    subgraph Pathology [Identification of Pathology]
        A[Metastatic tumor cells are CXCR4+ and migrate in towards CXCL12]
        B[Metastasis]
    end

    subgraph QSAR [QSAR of active inhibitors]
        C[Derivations of inhibitor structures]
        D[Pharmacophore identification]
    end

    subgraph ActiveSite [Identification of active site]
        E[Extracellular CXCR4 Ligand binding site]
        F[Intracellular G-protein binding site]
        G[Virtual screening]
    end

    subgraph Toxicity [General Toxicity Screening]
        H[CXCR4+ 300.19  
CXCR4- 300.19  
HS27, HMEC]
        I[Nonspecific and specific toxicity]
    end

    subgraph FunctionalScreening [CXCR4 Functional Screening]
        J[CXCR4+ 300.19 and TC-1  
CXCR4- 300.19  
Signaling Ca++ Flux  
Migration: Chemotaxis  
μ25-, SDF competition]
        K[CXCR4 Specific]
    end

    subgraph GPCRScreening [General GPCR Screening]
        L[BLT1+ 300.19  
Signaling Ca++ Flux  
Migration: Chemotaxis]
        M[Nonspecific G-protein or GPCR binding]
    end

    subgraph Rationalization [Rationalization of inhibitor binding site]
        N[Virtual re-docking  
Molecular dynamics simulations]
        O[In silico rationalization]
    end

    subgraph InVivo [In vivo testing of active compounds]
        P[Tail vein injection model of lung metastasis  
Intraventricular injection model of bone metastasis]
    end

    subgraph Analysis [In vivo analysis]
        Q[In vivo analysis]
    end

    A --> E
    B --> E
    C --> E
    D --> E
    E --> G
    G --> I
    I --> J
    J --> K
    K --> L
    L --> M
    M --> O
    O --> N
    N --> P
    P --> Q
  
```

Methods

Virtual Screening

The lipid-bilayer simulation of the CXCR4 N119S mutant docked with the G-protein heterotrimer was used to identify residues on the interaction surfaces of both proteins. The interface between two proteins occurred primarily between CXCR4 and the Gai subunit of the heterotrimer. Specific structural considerations came into play in designing the protomol files. CXCR4 residues within 8 Å of Gai and vice-versa were identified using a distance assessment PERL script, and these residues (Table 13) were used to produce residue based protomol files targeting both proteins. In addition, an α -helix of Gai came into close contact with the intracellular surface of CXCR4. This structure was used to develop a ligand-based protomol targeting CXCR4. A portion of the third intracellular loop of CXCR4 composed of residues 234-239 came into close contact with the surface of Gai. This structure was used to define a ligand-based protomol targeting Gai.

In evaluating the protomols targeting CXCR4, the ligand-based protomol best defined the interface surface on intracellular loops 1-3 and the c-terminus without positioning itself toward the transmembrane helices and away from Gai, as was seen with the residue based protomol. The ligand and residue based protomols targeting Gai both identified portions of the interaction surface with CXCR4, but neither completely covered the interface. As there was relatively little overlap between the protomols, they were combined into a single file and this combined protomol provided excellent coverage of the entire interface surface. The virtual screen was set up using the ZINC compound library, which in 2006

Table 13: Residues of CXCR4 and Gai used in Protomol development in

Surflex. Residues on CXCR4 within 8 Å of Gai and residues on Gai within 8 Å of CXCR4 were identified using a distance assessment PERL script to evaluate the interaction structure of CXCR4 and Gai. These residues were used to produce residue based protomol files targeting both proteins.

Table 13: Residues of CXCR4 and Gai used in Protomol development in Surflex.

CXCR4		Gai	
Residue	AA	Residue	AA
71	Ser	670	Glu
72	Met	671	Ile
140	His	691	Val
141	Ala	693	Asp
143	Asn	694	Val
144	Ser	696	Ile
234	Lys	697	Lys
235	Arg	698	asn
236	Lys	700	Leu
237	Ala	701	Lys
238	Leu	706	Phe
239	Lys		
240	Thr		
305	Leu		
306	Gly		
307	Ala		
308	Lys		
309	Phe		

comprised a data set of 2.5 million compound structures. These were individually docked, scored and ranked by Surflex-Dock using the Dataseam GRID distributed computing system (Dataseam [304]). The top 500 compounds were re-run as a separate subset using higher accuracy parameters (including randomized conformations in multiple starting positions) in Surflex-Dock.

Transfection and generation of RBL-2H3 and 300.19 stable cell lines expressing hCXCR4 and hCXCR4-GFP

To fully evaluate the effects of the potential inhibitors we have stably transfected overexpressing CXCR4 cell lines in collaboration with Dr. Haribabu Bodduluri, as described in chapter 2.

Chemotaxis analysis

Chemotaxis analysis for suspension 300.19 cells expressing hCXCR4-GFP or BLT1 was performed as previously described in Chapter 2 [151, 281] .

Cell signaling analysis

Cell signaling by CXCR4 was investigated by assessment of intracellular calcium mobilization as described in Chapter 2.

Class A GPCR selectivity for non-family member BLT1

The selectivity of the compounds for CXCR4 was assessed by investigating their effect on the Class A GPCR BLT1 as described in chapter 2.

Determination of nonspecific cytotoxicity

The potential cytotoxic effects of the compounds were investigated in our own hCXCR4-GFP expressing 300.19 cells as described in chapter 2.

Cell proliferation MTT assay

Anti-proliferative effects of the potential inhibitors were investigated using tissue derived cell lines. Control non-neoplastic human skin fibroblast (HS27) cells as well as lung (A549) and breast (MDA-MB-231 and MCF7) carcinoma cell lines were used in a 3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyl tetrazolium bromide (MTT) assay [305] with our protocols [306-309]. In brief, 10^2 cell per well were plated in 96 well plates, allowed to grow for one day prior to the addition of small molecules. Concentrations of the small molecules at 100 μ M, 10 μ M, 1 μ M, and 100nM were screened for activity. Four days after treatment, 125 μ g MTT (Sigma) was added to each well and incubated for 4 h at 37 C. Cells were then lysed with 100 μ l lysis buffer (20% sodium dodecyl sulfate, 50% N,N-dimethylformamide, and 0.8% acetic acid, pH 4.7) incubated for another 22 h and finally absorbance at 490 nm was determined. Samples were run in triplicate with untreated and DMSO-alone controls. Samples involving ICL compounds were normalized to average absorbance of the untreated samples [310-312].

Molecular modeling of CXCR4 and ICLGa complexes

Complexes of active ICLGa compounds and CXCR4 were investigated by molecular dynamics as described in chapter 2.

Results

Virtual screening identifies thirty-seven commercially available compounds targeting the CXCR4:Gai interface

Virtual screening targeting the CXCR4:Gai interaction site was conducted utilizing structures of CXCR4 and Gai initially developed by Dr. Napier. These structures were derived from molecular dynamics simulations of CXCR4 in a fully solvated lipid bilayer. As constitutively active mutants of CXCR4 were the only configurations of this receptor that were capable of interacting with the G-protein heterotrimer and interaction between the heterotrimer and CXCR4 primarily occurred at the Gai subunit, the dynamically-refined models of the N119S-mutant CXCR4 and Gai were used in virtual screening to identify small molecule antagonists targeting the CXCR4:Gai interface. Use of an active conformation of a receptor for virtual screening has been demonstrated to be beneficial for drug development, with active conformation of the β 2-adrenergic receptor enriching screening studies by 38% when compared to similar studies using an inactive conformation of the receptor [313].

Surflex-Dock [234, 244] was conducted in a two stage process. First, the ZINC compound library [208] was individually docked, scored and ranked by Surflex-Dock using GRID (Dataseam [304]). The top 5000 compounds were re-run as a separate subset using five starting positions and ten randomized initial compound conformations. The top 37 commercially-available compounds, (Table 14 and Figure 54) were assayed using human CXCR4 expressing cell lines. This is the first virtual screen to target the intracellular surface of CXCR4.

Table 14: Virtual screening identifies thirty-seven commercially available compounds targeting the interface of CXCR4 and Gαi. In one of the first virtual screens to target the GPCR:G-protein interface, 37 commercially available compounds were selected based on virtual screening scores and assessment by “Lipinski’s rule of five” for drug-like characteristics.

Table 14: Virtual screening identifies thirty-seven commercially available compounds targeting the interface of CXCR4 and Gai.

	Target	NAME	Supplier	MW
ICLGA1	CXCR4	N-(benzo[d][1,3]dioxol-5-ylmethyl)-2-hydroxy-3-(4-nitrophenoxy)propan-1-aminium	Maybridge	347.3
ICLGA2	Gai	(E)-1-(naphthalen-2-yl)-5-(2-(pyrrolidin-1-yl)vinyl)-1H-tetrazole	Maybridge	291.4
ICLGA3	CXCR4	1-(4-nitrophenyl)-2-((5-(4-pentylphenyl)-4H-1,2,4-triazol-3-yl)thio)ethanone	Maybridge	410.5
ICLGA4	Gai	N-(2-(4-ethylphenoxy)ethyl)-5-(3-methylpiperazin-1-yl)-2-nitroaniline	Chemical Block	384.5
ICLGA5	Gai	1-(4-chlorophenoxy)-3-(2-imino-3-(2-(piperidin-1-yl) ethyl) -2,3-dihydro-1H-benzo[d]imidazol-1-yl)propan-2-ol	Chem Bridge	429.0
ICLGA6	Gai	N1-(3-(dimethylamino)propyl)-N3,N3-dimethyl-N1-((4-(prop-1-en-2-yl)cyclohex-1-en-1-yl)methyl)propane-1,3-diamine	Chem Bridge	321.5
ICLGA7	CXCR4	4-methoxy-N-((1-(2-(o-tolyloxy)ethyl)-1H-benzo[d]imidazol-2-yl)methyl)aniline	Specs	387.5
ICLGA8	CXCR4	4-ethoxy-N-((1-(2-phenoxyethyl)-1H-benzo[d]imidazol-2-yl)methyl)aniline	Specs	387.5
ICLGA9	CXCR4	N4-(benzo[g]quinolin-4-yl)-N1-(6-methoxyquinolin-8-yl)pentane-1,4-diamine	Specs	436.6
ICLGA10	Gai	N-((1-methyl-1H-benzo[d]imidazol-2-yl)methyl)-1-propyl-1H-benzo[d]imidazol-5-amine	Specs	319.4
ICLGA11	Gai	1-(((4-(4-((2,5-dioxopyrrolidin-1-yl)(methyl)amino)benzyl)phenyl)amino)methyl)pyrrolidine-2,5-dione	Asinex	420.5
ICLGA12	CXCR4	2-((3-cyano-5,7-dimethylquinolin-2-yl)amino)ethyl 2-methoxybenzoate	Asinex	375.4
ICLGA13	CXCR4	N-(2-((3-cyano-5,7-dimethylquinolin-2-yl)amino)ethyl)propionamide	Asinex	296.5

Table 14, continued: Virtual screening identifies thirty-seven commercially available compounds targeting the interface of CXCR4 and Gai.

	Target	NAME	Supplier	MW
ICLGA14	CXCR4	1-(3,4,5-trimethoxybenzyl)-5'H-spiro[piperidine-,4'-pyrrolo[1,2-a]quinoxaline]	Asinex	419.5
ICLGA15	CXCR4	5-(3-((4-ethoxyphenyl)amino)-2-hydroxypropyl)-3-methyl-1-oxo-1,5-dihydrobenzo[4,5]imidazo[1,2-a]pyridine-4-carbonitrile	Asinex	416.5
ICLGA16	CXCR4	1-(9H-carbazol-9-yl)-3-((4-ethoxyphenyl)amino)propan-2-ol	Asinex	360.5
ICLGA17	CXCR4	2-(3-(3-(dimethylamino)propyl)-2-imino-2,3-dihydro-1H-benzo[d]imidazol-1-yl)-1-phenylethanol	Asinex	338.5
ICLGA18	CXCR4	N-((2,3-dihydrobenzo[b][1,4]dioxin-2-yl)methyl)-2-hydroxy-3-(4-nitrophenoxy)propan-1-aminium	Asinex	361.4
ICLGA19	CXCR4	1-methyl-2-(piperidin-1-ylmethyl)-N-(pyridin-2-ylmethyl)-1H-benzo[d]imidazol-5-amine	ChemDiv	335.5
ICLGA20	CXCR4	2-(benzo[d][1,3]dioxol-5-yl)-9-(4-methylbenzyl)-9H-benzo[d]imidazo[1,2-a]imidazole	IBScreen	381.4
ICLGA21	CXCR4	N-(1-(2,3-dihydrobenzo[b][1,4]dioxin-2-yl)ethyl)-2-hydroxy-3-(4-nitrophenoxy)propan-1-aminium	IBScreen	375.4
ICLGA22	Gai	N-((1-benzyl-1H-benzo[d]imidazol-2-yl)methyl)-1,2-dimethyl-1H-benzo[d]imidazol-5-amine	IBScreen	381.5
ICLGA23	CXCR4	2-(((1,1'-biphenyl)-4-yloxy)methyl)-5-((2-morpholinoethyl)amino)oxazole-4-carbonitrile	IBScreen	404.5
ICLGA24	CXCR4	N-(3,4-dimethylisoxazol-5-yl)-4-((3-(phenylsulfonamido)quinoxalin-2-yl)amino)benzenesulfonamide	Enamine	550.6
ICLGA25	CXCR4	N-(3-((4-(dimethylamino)phenyl)amino)quinoxalin-2-yl)-4-nitrobenzenesulfonamide	Enamine	464.5
ICLGA26	Gai	1-(1-(4-fluorophenyl)-2,5-dimethyl-1H-pyrrol-3-yl)-2-(furan-2-yl(methyl)amino)ethanone	Enamine	326.4

Table 14, continued: Virtual screening identifies thirty-seven commercially available compounds targeting the interface of CXCR4 and Gai.

	Target	NAME	Supplier	MW
ICLGA27	CXCR4	ethyl 2-(2-(benzo[d]thiazol-2-ylthio)-2-phenylacetamido)-6-methyl-4,5,6,7-tetrahydrobenzo[b]thiophene-3-carboxylate	Enamine	522.7
ICLGA28	CXCR4	(E)-4-(((2-(3,4-dimethylphenyl)-1,3-dioxo-2,3-dihydroisoquinolin-4(1H)-ylidene)methyl)amino)-N-(4-hydroxyphenyl)benzenesulfonamide	Enamine	539.6
ICLGA29	CXCR4	2-(1-(1,5-dimethyl-3-oxo-2-phenyl-2,3-dihydro-1H-pyrazol-4-yl)-2,5-dimethyl-1H-pyrrol-3-yl)-2-oxoethyl 2-(3,5-dimethyl-4-nitro-1H-pyrazol-1-yl)acetate	Enamine	520.54
ICLGA30	CXCR4	1-(3-(3-((3,4-dimethoxyphenethyl)amino)-2-hydroxypropoxy)phenyl)ethanone	Enamine	373.44
ICLGA31	CXCR4	N-(3,4-dimethylphenyl)-2-(((5-((4-methoxyphenyl)amino)-1,3,4-thiadiazol-2-yl)thio)acetamide	Enamine-REAL	400.52
ICLGA32	CXCR4	2-(4-(ethoxycarbonyl)-3,5-dimethyl-1H-pyrrol-2-yl)-2-oxoethyl 5-methyl-3-phenylisoxazole-4-carboxylate	Enamine-REAL	410.42
ICLGA33	CXCR4	2-(2-chlorophenyl)-N'-(4-methylcyclohex-1-en-1-yl)quinoline-4-carbohydrazide	Enamine	391.89
ICLGA34	CXCR4	1-(naphthalen-1-ylamino)-1-oxopropan-2-yl 2-(3,5-dimethyl-4-nitro-1H-pyrazol-1-yl)acetate	Enamine-REAL	396.4
ICLGA35	CXCR4	1-(4-(methoxycarbonyl)phenyl)-2-(methylamino)-2-oxoethyl 5-methyl-3-phenylisoxazole-4-carboxylate	Enamine	408.4
ICLGA36	CXCR4	N-(3-((9-ethyl-9H-carbazol-3-yl)amino)quinoxalin-2-yl)-4-methylbenzenesulfonamide	Otava	507.61
ICLGA37	Gai	2-(((3,4-dimethylphenyl)amino)-N-(5-(3-(trifluoromethyl)benzyl)thiazol-2-yl)acetamide	Otava	419.46

Figure 54: ICLGA Compounds. Structures of thirty-seven compounds identified in virtual screens targeting the interface between the intracellular surface of CXCR4 and Gai are illustrated below.

Figure 54: ICLGA Compounds

ICLGA1

ICLGA5

ICLGA2

ICLGA6

ICLGA3

ICLGA7

ICLGA4

ICLGA8

ICLGa9 **ICLGa13** **ICLGa10** **ICLGa14** **ICLGa11** **ICLGa15** **ICLGa12** **ICLGa16**

ICLGa17

ICLGa20

ICLGa21

ICLGa18

ICLGa19

ICLGa22

ICLGa23

ICL Ga24 **ICL Ga28** **ICL Ga25** **ICL Ga29** **ICL Ga26** **ICL Ga30** **ICL Ga27** **ICL Ga31**

ICLGa32

ICLGa33

ICLGa34

ICLGa36

ICLGa35

ICLGa37

Compounds targeting the CXCR4:G α i interface specifically inhibit chemotaxis

Our goal was to target the CXCR4:G-protein interface to discover novel small molecules which could selectively inhibit receptor function and be used as scaffolds for development of clinically active anti-metastatic compounds. We sought to evaluate the compounds by chemotaxis assays which quantitatively examined the capability of cells to migrate towards a ligand gradient. Initial screening of the thirty-seven ICL compounds at 50 μ M identified twenty-four compounds which significantly inhibited chemotaxis as depicted in **Figure 55**. Dose response analysis was used to determine the concentration of the active compounds where chemotaxis was inhibited by 50% in reference to the CXCL12 alone control (IC₅₀). Ten of the twenty-four active compounds had chemotaxis IC₅₀ values of 10 μ M or less (**Figure 56**), for an exceptionally high initial hit rate of 27%. Compounds ICLGa28 and ICLGa36 were two of the most promising compounds identified, with effective inhibition of chemotaxis at 1 μ M. Other notable compounds included ICLGa 8 with effective inhibition at 3 μ M and ICLGa compounds 5, and 37 with effective inhibition at 5 μ M.

It is important to note that the hit rate of our virtual screening studies is significantly higher than what is seen in similar published studies. Our initial hit rate of 64.9% at 50 μ M for the initial screening of our thirty seven ICL compounds represents a great increase over contemporary studies. For comparison, recent investigations using virtual screening of glycogen synthase kinase-3 β (GSK-3 β) had an initial hit rate of 12.5% for 162 compounds [314]. A recent investigation of

Figure 55: ICLGa compounds inhibit CXCR4 mediated Chemotaxis.

Chemotaxis assays performed with ICLGa compounds at a concentration of 50 μ M. Vertical axis is presented in terms of chemotaxis index wherein sample chemotaxis is normalized to chemotaxis of the CXCL12 alone control. Successful inhibition of chemotaxis reduced chemotaxis of samples compared to CXCL12 alone. 24 of 37 tested compounds significantly inhibited chemotaxis at the initial screening concentration of 50 μ M.

Figure 55: Inhibitors of CXCR4 Mediated Chemotaxis

Figure 56: IC50 of ICLGa compound mediated chemotaxis inhibition.

Concentrations which inhibited 50% of chemotaxis in relation to CXCL12 alone signaling (IC50) were determined for active ICLGa compounds. Vertical axis is presented in terms of IC50 of chemotaxis in μM . 10 of the 26 active compounds exhibited significant inhibition of chemotaxis at 10 μM or less.

Figure 56: IC50 of Inhibitors of CXCR4 Mediated Chemotaxis

integrin $\alpha\beta3$ using virtual screening of a large compound library had an initial hit rate of 14% for 50 compounds, with IC₅₀ values ranging from 30 to 200 μM [315]. Another study targeting bacterial NAD synthetase had an initial hit rate of 20% at 300 μM , but only 2 of the 40 (5%) active compounds had IC₅₀s less than 50 μM [316].

We then sought to determine the specificity of the ICLGa compounds for CXCL12-induced chemotaxis. Chemotaxis assays were carried out using 300.19 cells expressing the class A GPCR BLT1 mediating chemotaxis in response to LTB₄. Initial screening of the twenty-four active ICLGa compounds revealed that fifteen of the compounds also inhibited LTB₄-mediated chemotaxis at 50 μM . However, as can be seen in **Figure 57**, dose response analysis of the ICLGa compounds with anti-BLT1 activity revealed that only four of the compounds (ICLGa6, ICLGa12, ICLGa14 and ICLGa16) had IC₅₀ values that were lower for BLT1 than for CXCR4. There was an average difference of 18.5 μM in the IC₅₀ values of the compounds in reference to CXCR4 versus BLT1. It is important to note that of the ten ICLGa compounds which inhibited CXCR4-mediated chemotaxis at 10 μM or less, only five exhibited any anti-BLT1 activity, and none of these possessed an anti-BLT1 IC₅₀ of less than 25 μM . This suggests that the inhibitory effects produced by these compounds are specific for CXCR4.

Division of ICLGa compounds into calcium and non-calcium inhibitors of CXCR4

Intracellular calcium mobilization is a common method for assessing GPCR receptor activation and signaling. CXCL12-induced calcium mobilization is

Figure 57: ICLGa compounds selectively inhibit CXCR4 Mediated Chemotaxis. Specificity screening of active ICLGa compounds in BLT1 expressing cells revealed that 15 compounds inhibited chemotaxis in both systems. Only 3 of the 15 compounds with BLT1 activity inhibited BLT1 chemotaxis to a greater degree than CXCR4 mediated chemotaxis. Only 5 of the ICLGa compounds with anti-CXCR4 activity at 10 μ M or less exhibited anti-BLT1 activity, and none had anti-BLT1 IC₅₀ values of less than 25 μ M.

Figure 57: Selective IC50 of Inhibitors of CXCR4 Mediated Chemotaxis

a pertussis toxin sensitive [317] process, suggesting that this process is dependent on G α signaling. Thus, we investigated the impact of ICLGa compounds on CXCL12-induced calcium mobilization. CXCL12-induced signaling was assessed by the change in amplitude of the calcium signal and results were normalized to the CXCL12-alone control. Initial screening of the ICLGa targeted compounds at 50 μ M identified twelve compounds which significantly reduced calcium release, seen in [Figure 58](#). Dose response analysis determined that six of the twelve active compounds, ICL9 (1 μ M), ICL12 (5 μ M), ICL13 (5 μ M), ICLGA25 (1 μ M), ICL28 (10 μ M) and ICL36 (6 μ M) had calcium mobilization IC₅₀ values of 10 μ M or less as seen in [Figure 59](#).

We then sought to determine the specificity of the ICLGa compounds for CXCL12-induced intracellular calcium mobilization. Calcium mobilization assays were carried out using 300.19 cells expressing BLT1 and signaling was assessed in response to LTB₄. Initial screening of the twelve active ICLGa compounds revealed that nine of the twelve significantly inhibited LTB₄-induced intracellular calcium mobilization. Subsequent dose response analysis, depicted in [Figure 60](#), revealed that six of the nine ICLGa compounds with anti-BLT1 activity had IC₅₀ values that were lower for BLT1 than for CXCR4. Differences in calcium mobilization were relatively small for ICLGa2 (45 μ M vs. 40 μ M), ICLGa4 (35 μ M vs. 27.5 μ M), ICLGa12 (5 μ M vs. 2.5 μ M), ICLGa13 (5 μ M vs. 2.5 μ M), and ICLGa16 (35 μ M vs. 30 μ M), suggesting that inhibition of calcium signaling by these five compounds is not selective. In contrast the IC₅₀ of calcium mobilization was significantly greater for CXCR4 than BLT1 for ICLGa9 (1 μ M vs.

Figure 58: ICLGa compounds inhibit CXCR4 mediated Intracellular Calcium Mobilization. Calcium Mobilization assays performed with ICLGa compounds at a concentration of 50 μ M. Vertical axis is presented in terms of normalized calcium signal, wherein the amplitude of the calcium response to the CXCL12-alone control corresponds to 1. Decreased signal amplitude correlates with inhibition of calcium mobilization. Eleven of thirty seven tested ICLGa compounds significantly inhibited calcium mobilization.

Figure 58: ICLGa compounds inhibit CXCR4 mediated Intracellular Calcium Mobilization.

Figure 59: IC50 of ICLGa compound mediated inhibition of calcium mobilization. Concentrations which inhibited 50% of calcium mobilization in relation to CXCL12 alone signaling (IC50) were determined for active ICLGa compounds. Vertical axis is presented in terms of IC50 of calcium mobilization in μM . Six of the twelve active compounds exhibited significant inhibition of calcium mobilization at 10 μM or less.

Figure 59: IC50 of ICLGa compound mediated inhibition of calcium mobilization.

Figure 60: ICLGa compounds selectively inhibit CXCR4 Mediated Calcium Mobilization. The twelve ICLGa compounds which inhibited CXCL12-induced calcium mobilization were tested for specificity in BLT1 expressing cells, revealing that nine compounds inhibited calcium mobilization in both systems at a 50 μ M screening concentration. Five of the nine compounds produced BLT1 mediated calcium mobilization equivalent to CXCR4 mediated calcium mobilization. Three of the compounds were selective for CXCR4 mediated calcium mobilization with IC₅₀ values less than 10 μ M.

Figure 60: ICLGa compounds selectively inhibit CXCR4 Mediated Calcium Mobilization.

20 μ M), ICLGa25 (1 μ M vs. 40 μ M), and ICLGa28 (10 μ M vs. 22.5 μ M) suggesting that anti-calcium mobilization activity in these compounds is selective for CXCR4.

The thirteen ICLGa anti-CXCR4 compounds with activity at 10 μ M or less are depicted in **Table 15**. Seven compounds exhibit CXCR4-specific and exclusive anti-chemotaxis activity. Three compounds exhibit CXCR4-specific anti-chemotaxis and anti-calcium mobilization activity. Two compounds exhibit non-specific exclusive anti-calcium mobilization activity. The remaining compound exhibits CXCR4-specific and exclusive anti-calcium mobilization activity. Given that CXCL12-induced calcium modulation is dependent on G α signaling, one would expect that compounds targeting G α i would preferentially inhibit intracellular calcium mobilization; however that was not the case for our compounds. Of the nine compounds targeting the G α i side of the interface, only three of these (ICLGa2, ICLGa4, and ICLGa5) significantly inhibited calcium mobilization, and the IC₅₀ values of these compounds were much greater than for many of the CXCR4 targeted compounds which inhibited this signaling. One explanation for this discrepancy lies in the signaling pathway. CXCL12-induced calcium mobilization is a pertussis toxin sensitive process [317], but also requires activation of phospholipase C β [318]. G α i, the G α subunit which interacts with CXCR4, is incapable of activating phospholipase C β , and the only G α subunit which activates phospholipase C β , G α q, is insensitive to pertussis toxin [319, 320]. The G $\beta\gamma$ subunit is also capable of activating phospholipase C β [321]. Thus, CXCL12-induced calcium mobilization could be dependent on interactions

Table 15: Classification of compounds with anti-CXCR4 activity at 10 μ M or less. ICLGa compounds with anti-CXCR4 activity with IC50 values of less than 10 μ M in either chemotaxis or calcium mobilization are listed in the table below. Three of the compounds ICLGa25, ICLGa28 and ICLGa36 produced inhibition of CXCR4 below 10 μ M for both activities.

Table 15: Classification of compounds with anti-CXCR4 activity at 10 μ M or less.

Compound	Chemotaxis IC50 (μM)	Calcium IC50 (μM)	Activity	Target
ICL Ga4	10	35	Chemotaxis	Gai
ICL Ga5	5	45	Chemotaxis	Gai
ICL Ga8	3	--	Chemotaxis	CXCR4
ICL Ga9	45	1	Calcium	CXCR4
ICL Ga12	12.5	5	Calcium	CXCR4
ICL Ga13	12.5	5	Calcium	CXCR4
ICL Ga17	10	--	Chemotaxis	CXCR4
ICL Ga22	10	--	Chemotaxis	Gai
ICL Ga25	10	1	Dual	CXCR4
ICL Ga28	1	10	Dual	CXCR4
ICL Ga29	10	--	Chemotaxis	CXCR4
ICL Ga36	1	6	Dual	CXCR4
ICL Ga37	5	--	Chemotaxis	Gai

between Gai and either Gaq or Gβγ. In the case of the ICLGa compounds, some compounds could interfere with the interaction between CXCR4 and Gai, completely preventing signal transmission through the receptor inhibiting both chemotaxis and calcium mobilization as is the case with ICLGa25, ICLGa28 and ICLGa36.

Compounds targeting the CXCR4:Gai interface have relatively little non-specific cytotoxicity, but three active ICLGa compounds selectively inhibit tumor cell proliferation.

Inhibition of chemotaxis and calcium mobilization by ICLGa compounds could be due to nonspecific toxic side effects, the active ICLGa compounds were screened for cytotoxicity. Mimicking the conditions in the chemotaxis assay, 300.19 cells were cultured in the presence of ICLGa compounds at the initial screening concentration of 50 μM and cytotoxicity was assessed using propidium iodide (PI) based flow cytometry. PI is unable to penetrate the cell membrane in viable cells, so cells staining positive for PI do not have intact membranes. Cytotoxicity values are presented as percentage of cells staining positive for PI in Figure 61. Of the twenty-six ICLGa compounds producing significant inhibition of chemotaxis and calcium mobilization, only seven produced displayed elevated cytotoxicity when compared to untreated cells or cells treated with DMSO only control. To further investigate the cytotoxicity of these compounds, dose response assays were conducted at concentrations of 10, 5, and 1 μM, seen in Figure 62. Compounds ICLGa9, ICLGa25, and ICLGa36 displayed no cytotoxicity beyond the initial 50 μM screening concentration. Compounds

Figure 61: ICLGa compounds produce relatively little non-specific cytotoxicity. Vertical axis is expressed in terms of cells staining positive for PI, a marker for membrane disruption and non-viable cells. Only seven of the twenty-six ICLGa compounds with anti-CXCR4 activity compounds exhibited significant cytotoxicity. Only five of the compounds produced greater than 50% cytotoxicity at 50 μ M. Two additional compounds, ICLGa25 and ICLGa36, displayed elevated cytotoxicity compared to untreated and DMSO-alone controls, but to a lesser extent than 50% cytotoxicity.

Figure 61: ICLGa compounds produce relatively little non-specific cytotoxicity

Figure 62: Toxic concentrations of ICLGa compounds. Only one of the 7 ICLGa compounds exhibiting cytotoxicity at 50 μ M exhibited cytotoxicity at 5 μ M or less. The three compounds with significant anti-CXCR4 activity below 10 μ M exhibited no toxic effects below 50 μ M.

Figure 62: Toxic concentrations of ICLGa compounds.

ICLGa2, ICLGa7 and ICLGa16 all retained significant cytotoxicity at 10 μ M. Only compound ICLGa4 maintained significant cytotoxicity at 5 μ M. There were no structural similarities between the compounds with cytotoxicity beyond 10 μ M.

ICLGa9, ICLGa25 and ICLGa36 all selectively inhibited CXCR4 activity at 1 μ M in either chemotaxis or calcium mobilization. All three of these compounds also displayed significant cytotoxicity at 50 μ M, but no cytotoxic effects at lower concentrations. As BIM-46174 had been demonstrated to inhibit tumor proliferation and induce apoptosis [301] we sought to investigate the impact of ICLGa on tumor cell proliferation. To determine the potential for the ICLGa compounds to impact the proliferation of tumor cells, we assessed cell growth capacity using the 3-(4,5-dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide (MTT) assay. We investigated the effect of these compounds on four cell lines: HS27, a “normal” fibroblast cell line; MCF7, an estrogen-dependant non-metastatic breast cancer cell line; MDA-MB-231, an estrogen-independent metastatic breast cancer cell line; and A549, a lung cancer cell line. Cells were left untreated, treated with 12.5 mM DMSO, or treated with varying concentrations of three ICLGa compounds, ICLGa9, ICLGa25 and ICLGa36, then incubated for four days before analysis. In our analysis, the absorbance values were normalized to untreated samples.

The three ICLGa compounds showed particularly interesting effects on cell proliferation. In the normal cell line HS27 cells as depicted in Figure 63a, both ICLGa25 and ICLGa36 displayed marginal proliferation inhibition at higher concentrations, but these effects abated as concentration dropped below 20 μ M.

Figure 63: Inhibition of cell proliferation by ICLGa compounds. Cell proliferation was assessed by MTT assay. Vertical axis represents MTT signal detected in each sample normalized to the untreated control. Horizontal axis represents concentration of ICLGa sample.

Figure 63: Inhibition of cell proliferation by ICLGa compounds.

ICL Ga9 consistently inhibited cell proliferation even at concentrations nearing 1 μM . (Fig. III9)

In MCF7 cells as depicted in [Figure 63b](#), ICL Ga9 inhibited cell proliferation with an IC_{50} value of 5 μM . ICL Ga36 also significantly inhibited cell proliferation with an IC_{50} value of 12.5 μM . ICL Ga25 only marginally inhibited cell proliferation of MCF7 cells at high concentrations.

In MDA-MB-231 cells as depicted in [Figure 63c](#), ICL Ga9 inhibited cell proliferation with an IC_{50} value of 5 μM . ICL Ga25 significantly inhibited proliferation of MDA-MB-231 cells; with an IC_{50} value of 10 μM . ICL Ga36 significantly inhibited cell proliferation of MDA-MB-231 cells, with an IC_{50} value of 30 μM .

Finally, in A549 cells as depicted in [Figure 63d](#), ICL Ga9 inhibited cell proliferation with an IC_{50} value of 7.5 μM . ICL Ga25 significantly inhibited cell proliferation of A549 cells, with an IC_{50} value of 10 μM . ICL Ga36 significantly inhibited cell proliferation of A549 cells, with an IC_{50} value of 22.5 μM .

The three compounds illustrate three avenues of proliferation inhibition. ICL Ga9 is generally inhibitory, affecting both normal and cancer cells at levels lower than the cytotoxicity levels seen in Figure 62. ICL25 inhibited the growth of metastatic cancer cell lines A549 and MDA-MB-231, but not HS27 cells or non-metastatic MCF7 cells. Finally, ICL36 inhibited the growth of non-metastatic MCF7 cells, metastatic cancer cell lines to a lesser extent, and did not inhibit proliferation of HS27 cells. These results are intriguing as a CXCR4 inhibitor need not be cytotoxic to function as a metastasis inhibitor.

Analysis of anti-CXCR4 compound mechanism through molecular dynamics simulations

As demonstrated by Holland *et al*, heterotrimeric G-protein interaction with CXCR4 occurs in metastatic breast cancer cell lines but not non-metastatic lines [137]. We have postulated that this behavior is due to differential activation of CXCR4 in the metastatic cell lines producing altered CXCR4 conformation, a hypothesis that is supported by the findings of Dr. Napier that *in silico* the constitutively active N119S conformation of CXCR4 is capable of interacting with the G-protein heterotrimer while the native protein conformation is not.

The initial native CXCR4 and N119S mutant structures were developed by Dr. Kelby Napier via homology modeling utilizing the high resolution rhodopsin crystal structure [288]. When this work was conducted, the x-ray crystal structure for CXCR4 had not yet been released, so this structure was not a viable option for these studies.

Molecular modeling of CXCR4 and ICLGa complexes

We hypothesize that the binding of ICLGa complexes with anti-CXCR4 activity induces a conformational change in the receptor, impeding signaling. Complexes of active ICLGa compounds and CXCR4 were investigated by molecular dynamics. The highest scoring conformations of the ICLGa compounds were utilized as a guide to initial placement of the inhibitor. Surflex-Dock [234, 244] was used to dock 12 ICLGa compounds with anti-CXCR4 activity at 10 μ M or less and without excessive cytotoxicity into the defined intracellular surface of the native structure of CXCR4. All twelve compounds

were successfully docked into the structure regardless of their initial status as CXCR4-targeted or Gai targeted compounds. Fully solvated lipid bilayer molecular dynamics simulations of native CXCR4 bound with active ICLGa compounds were performed using the AMBER suite of programs using the parm99 force field [284].

Some ICLGa compounds do not bind to the Native conformation of CXCR4 in molecular dynamics simulations.

The final averaged structures of native CXCR4 in complex with each of the anti-CXCR4 compounds from the 5 ns molecular dynamics simulation are shown in **Figures 64**. Of the twelve docked compounds, stable CXCR4: compound complexes were generated for compounds ICLGa8, ICLGa12, ICLGa25, ICLGa28, ICLGa29 and ICLGa37. ICLGa5, ICLGa9, ICLGa13, ICLGa17, ICLGa22, and ICLGa36 did not form stable complexes with the native conformation of CXCR4.

Inverse agonist, antagonist and weak partial agonist inhibitors all potentially can bind to different structures of CXCR4. Weak partial agonists should prefer to interact with inactive conformations of the receptor, where their partial agonistic activity is more capable of inducing a partially active conformation. Conversely, inverse agonists interact with constitutively active receptors and receptors with basal activity suggesting that these inhibitors should preferentially interact with an inactive conformation of the receptor, in order to shift the conformation to an inactive state. The fact that these 6 active ICLGa compounds do not bind to the native conformation of the receptor suggests that

Figure 64: Final averaged structure of Native CXCR4 dynamically bound to ICLGa Compounds. Average structure of CXCR4 presented in ribbon. ICLGa compounds are in their dynamically determined binding position in red.

Figure 64: Final averaged structure of Native CXCR4 dynamically bound to ICLGa Compounds.

these active anti-CXCR4 compounds may act as inverse agonists, and may be unable to interact with the inactive conformation of the receptor.

The ICLGa compounds which did form stable complexes with native CXCR4 were energetically stable throughout the production run as can be seen in **Figures 65**. Conformational stability of the bound CXCR4-compound complexes was investigated using RMSD analysis, seen in **Figures 66**. RMSD values remained stable for the majority of the production runs, with the final average structure representative of the trajectory after approximately 1.5 ns. The helices were stable through the production run and the majority of the variation was seen in the loops. The conformation assumed by ICLGa29 was the closest to that of uncomplexed native CXCR4. The RMSD of the ICLGa29 complex was 3.1, compared to RMSD values of 3.2 for ICLGa28, 3.3 for ICLGa8, 3.4 for ICLGa37, 4.0 for ICLGa25 and 4.5 for ICLGa12.

The dynamic binding site of ICLGa8, shown in **Figure 67**, consists of ICL1 superior to the compound, ICL2 to the right, and the C-terminal tail to the left and posterior to the compound. The ethoxybenzene component of ICLGa8 is stabilized by hydrophobic interactions with the aliphatic portions of the Lys149 and Ser71 side chains. The central imidazole component is stabilized by interactions with the Lys 68 side chain, as well as the backbones of Thr342 and Glu343. The anisole group is interposed between folds of the C-terminal tail, interacting with a hydrophobic network composed of the side chain of Val340, the backbone of Ser338, and the aliphatic portion of the side chains of Glu345, Ser 346 and Ser348.

Figure 65: Energy plot of ICLGa compounds bound to native CXCR4.

Energy values across the trajectory for total energy (blue), kinetic energy (red) and potential energy (green) are presented. CXCR4 was found to be energetically stable and maintained equilibrium across the production run. A) ICLGa8, B) ICLGa12, C) ICLGa25 D) ICLGa28 E) ICLGa29 F) ICLGa37

Figure 65: Energy plot of ICLGa compounds bound to native CXCR4. A)

ICLGa8, B) ICLGa12, C) ICLGa25 D) ICLGa28 E) ICLGa29 F) ICLGa37

Figure 66: RMSD values of CXCR4 bound to ICLGa Compounds across 5 ns of the production trajectory. Total RMSD calculated at the C α position presented in green. Helical domains are presented in blue, and loop regions are presented in red. The CXCR4 structure was noted to be stable, and the average structure was noted to represent the conformation present in the trajectory after 0.5 ns of simulation time. A) ICLGa8, B) ICLGa12, C) ICLGa25 D) ICLGa28 E) ICLGa29 F) ICLGa37

Figure 66: RMSD values of CXCR4 bound to ICLGa Compounds across 5 ns of the production trajectory. A) ICLGa8, B) ICLGa12, C) ICLGa25 D)

ICLGa28 E) ICLGa29 F) ICLGa37

B

D

E

F

Figure 67: Dynamic binding site of ICLGa8 to native CXCR4. Critical residues for ICLGa8 binding to CXCR4 are highlighted. ICLGa is depicted in red ball-stick representation, while CXCR4 residues within 3 angstroms of the compound are presented in stick representation and superstructure of CXCR4 is presented as teal ribbon. ICLGa8 interacts with the ICL1, ICL2, and C-terminal regions of the intracellular surface.

Figure 67: Dynamic binding site of ICLGa8 to native CXCR4.

The dynamic binding site of ICLGa12, shown in **Figure 68** consists of ICL1 superior to the compound, ICL2 to the right, and the C-terminal tail to the left and posterior to the compound. The anisole and quinolone components of ICLGa12 are oriented in a planar conformation. The quinolone portion lies between folds of the C-terminal tail and ICL1, in a conformation stabilized by interactions with the aliphatic portion of the Lys68 and thr342 side chains and the backbone of Ser344 above and the backbones of Ser 348, 351 and 352 below. The aromatic region of the anisole component is stabilized by interactions with the aliphatic portion of the Lys149 side chain and Pro147.

The dynamic binding site of ICLGa25, shown in **Figure 69** consists of ICL1 superior to the compound, ICL2 to the right, and the C-terminal tail to the left and posterior to the compound. The nitrobenzene and dimethylaniline components of ICLGa25 are in a stacked conformation. The nitrobenzene component is stabilized by hydrogen bonding to the side chain of Arg334 which is further stabilized by intramolecular hydrogen bonding between Arg334 and Glu345. The dimethylaniline is further stabilized by hydrophobic interaction with Phe349. The central quinoxaline component of ICLGa25 is projected between ICL1 and ICL3 and is stabilized by hydrogen bonding to Lys68, as well as a hydrophobic network composed of the backbone of Ser 71, the side chain of Leu150, and the aliphatic portion of the Lys149 side chain.

The dynamic binding site of ICLGa28, shown in **Figure 70** consists of ICL1 superior to the compound, ICL2 to the right, the C-terminus anterior to the compound, the ICL3 posterior to the compound. The phenol group of ICLGa28 is

Figure 68: Dynamic binding site of ICLGa12 to native CXCR4. Critical residues for ICLGa12 binding to CXCR4 are highlighted. ICLGa12 is depicted in red ball-stick representation, while CXCR4 residues within 3 angstroms of the compound are presented in stick representation and superstructure of CXCR4 is presented as magenta ribbon. ICLGa12 interacts with the ICL1, ICL2, and C-terminal regions of the intracellular surface.

Figure 68: Dynamic binding site of ICLGa12 to native CXCR4

Figure 69: Dynamic binding site of ICLGa25 to native CXCR4. Critical residues for ICLGa25 binding to CXCR4 are highlighted. ICLGa25 is depicted in red ball-stick representation, while CXCR4 residues within 3 angstroms of the compound are presented in stick representation and superstructure of CXCR4 is presented as orange ribbon. ICLGa25 interacts with the ICL1, ICL2, and C-terminal regions of the intracellular surface.

Figure 69: Dynamic binding site of ICLGa25 to native CXCR4.

Figure 70: Dynamic binding site of ICLGa28 to native CXCR4. Critical residues for ICLGa28 binding to CXCR4 are highlighted. ICLGa28 is depicted in red ball-stick representation, while CXCR4 residues within 3 angstroms of the compound are presented in stick representation and superstructure of CXCR4 is presented as white ribbon. ICLGa28 interacts with the ICL1, ICL2, ICL3 and C-terminal regions of the intracellular surface.

Figure 70: Dynamic binding site of ICLGa28 to native CXCR4.

interposed between the C-terminus and ICL3 stabilized by interactions with Lys236 of ICL3, and Lys308 and His337 of the C-terminus. The isoquinolone structure extends towards ICL1, and is stabilized by a hydrophobic network composed of the aliphatic portions of the Lys68 and Ser71 side chains, as well as the backbone of Arg70 and the side chain of Leu150. The dimethyl benzene component is stabilized by a separate hydrophobic network composed of the aliphatic portion of the Lys149 side chain, Pro147 and the backbone of Arg148.

The dynamic binding site of ICLGa29, shown in **Figure 71** consists of ICL1 superior to the compound, ICL2 to the right, and the C-terminal tail to the left and posterior to the compound. The orientation of the compound in the binding site is unique among the ICLGa compounds in that the compound is oriented vertically in the binding site whereas the other compounds were oriented horizontally in the site. The compound is stabilized by hydrogen bonding between linker chain carboxyl groups and the side chain amines of Lys68 and Lys149. The 2,4 dimethyl-3-nitro-pyrazole projects towards ICL1, and is stabilized by a hydrophobic network composed of the aliphatic portion of the Lys68, Lys149 side chains as well as the backbone of Arg70, Ser71 and the side chain of Leu150.

The dynamic binding site of ICLGa37, shown in **Figure 72** consists of ICL1 superior to the compound, and ICL2 and the C-terminus posterior to the compound. The 3,4 dimethyl-benzene group is stabilized by hydrophobic interactions with the aliphatic portions of the side chains of CXCR4 residues Lys 68, Ser71, Leu150 and Ser348. The benzyl portion of the meta-trifluoromethyl benzene component of the compound is stabilized by hydrophobic interactions

Figure 71: Dynamic binding site of ICLGa29 to native CXCR4. Critical residues for ICLGa29 binding to CXCR4 are highlighted. ICLGa29 is depicted in red ball-stick representation, while CXCR4 residues within 3 angstroms of the compound are presented in stick representation and superstructure of CXCR4 is presented as yellow ribbon. ICLGa29 interacts with the ICL1, ICL2 and C-terminal regions of the intracellular surface.

Figure 71: Dynamic binding site of ICLGa29 to native CXCR4.

Figure 72: Dynamic binding site of ICLGa37 to native CXCR4. Critical residues for ICLGa37 binding to CXCR4 are highlighted. ICLGa37 is depicted in red ball-stick representation, while CXCR4 residues within 3 angstroms of the compound are presented in stick representation and superstructure of CXCR4 is presented as blue ribbon. ICLGa37 interacts with the ICL1, ICL2 and C-terminal regions of the intracellular surface.

Figure 72: Dynamic binding site of ICLGa37 to native CXCR4.

with the aliphatic portion of the side chain of Lys149 and Leu150. In the center of the compound, hydrogen bonds form between Lys149 and the thiazole of the compound.

These structures shed light on the potential mechanism of the activity of the bound ICLGa compounds. All of the compounds interacted with the aliphatic components of the side chains and backbones of ICL1 and ICL2, suggesting that the compounds could act as bridging elements between these two regions, stabilizing their position and preventing signal transduction or producing a more closed intracellular surface potentially incapable of binding Gai. Interaction by the compounds with the C-terminal region could potentially stabilize TM7, or interfere with the interface between the C-terminus and Gai. This could be especially true in the case of ICLGa28, where interposition of the compound between the C-terminus and ICL3 could restrict the movement of both regions and directly impede the cleft which would be occupied by bound Gai.

ICLGa compound binding alters the conformation of the extracellular and intracellular surfaces of CXCR4.

Given the interactions of the ICLGa compounds with the intracellular surface of CXCR4, we sought to investigate what effect inhibitor binding could have on overall receptor conformation. To do this, we sought to compare the structure of the ICLGa complexed conformations of CXCR4 to the inactive native conformation and the constitutively active N119S conformations of the protein. The extracellular surface of CXCR4 is highly conserved between ICLGa8 and native CXCR4. Although the structure of the loops is conserved between the

three conformations, there are several differences in positioning. In the N-terminal region, the anti-parallel β -sheet is more compact and the distance between ECL1 and the N-terminal region is decreased, mimicking the N119S conformation, while the entire region is displaced away from the extracellular surface of the TM regions. ECL2 retains the native conformation and positioning, while ECL3 is outwardly displaced from the center of the receptor. At the intracellular surface of CXCR4, the ICLGa conformation retains the native conformation and position of ICL1, ICL2 and the C-terminus. A loss of one turn of helical structure at the cytoplasmic surface of TM5 produces a displacement of ICL3 towards the C-terminal tail in the ICLGa8 complexed structure.

Several significant structural changes characterize the extracellular surface the ICLGa12 complex conformation. In the N-terminal region, the three-strand β sheet of the native conformation is converted into a two-strand sheet by extension at Thr8 rather than flexion. The proximal portion of the N-terminus also takes on an additional turn of α -helical structure. ECL1 and ECL2 adopt the native conformation of the loops, while the apex of the ECL3 loop is displaced towards the base of the N-terminal strand as in the constitutively active N119S conformation. At the intracellular surface of the receptor, ICL1 adopts the native conformation, while the plane of the ICL2 loop is inclined towards the center of the receptor, rather than away from the center as in the native conformation. ICL3 adopts the unkinked structure of the native conformation, and is displaced away from the center of the receptor as in the N119S conformation of CXCR4. Helix 8 structure and positioning is conserved between ICLGa12 and native,

while a section of the C-terminal tail between Ser319 and Ser330 takes on a more helical conformation and is inclined towards the center of the receptor when compared to the native conformation.

The extracellular surface of CXCR4 is highly conserved between ICLGa25 and native CXCR4. The structures of the N-terminal region and the three ECL structures are nearly identical to that of the native conformation. However at the ICL surface, while ICL1 structure is conserved between the ICLGa25 and native conformations of the receptors, significant differences are apparent for the other intracellular structures. The apex of ICL2 is located at the midpoint of the loop as in the native conformation, while the angle of inclination of ICL2 is closer towards the center of the receptor as in the N119S conformation. The distance between the apex of ICL3 and Helix 8 is also significantly decreased when compared to the native structure. Finally, the tail of the C-terminal region is in a more vertical orientation than in the native or N119S conformations.

Several significant structural changes characterize the extracellular surface the ICLGa28 complex conformation. The N-terminal β -sheet is more compactly packed in the ICLGa28 complexed conformation as in the N119S structure. The proximal portion of the N-terminal tail is displaced towards ECL3. ECI1, ECL2 and ECL3 strongly conserve the native conformation. At the intracellular surface of the receptor, ICL1 maintains the native conformation, while the apexes of ICL2 and ICL3 are displaced towards the center of the receptor, and Helix 8 is outwardly displaced from the cytoplasmic terminal of TM1.

The extracellular surface of CXCR4 is highly conserved between ICLGa29 and native CXCR4. The structures and positioning of the N-terminal region, ECL2 and ECL3 maintain the native conformation of the receptor. There is also a slight displacement of ECL1 towards the N-terminal region. At the intracellular surface of the receptor, ICL1 conformation and positioning maintains the native conformation, while the angle of inclination for ICL2 is decreased, displacing the apex of this loop towards the center of the receptor. ICL3 adopts the linear conformation of the native receptor, while the apex of ICL3 is displaced towards the apex of ICL2. Finally, although the structure of the C-terminal tail is largely preserved between the ICLGa29 complexed conformation and the native structure, the distal portion of the tail is brought into close contact with the apexes of ICL2 and ICL3.

The extracellular surface of CXCR4 is highly conserved between ICLGa37 and native CXCR4. The distal N-terminal β -sheet adopts the compact conformation of the N119S structure, and the proximal region adopts the displacement towards ICL1 as in the N119S conformation. ECL2 is maintained in the native conformation, while ECL3 is outwardly displaced from the center of the receptor. At the intracellular surface of the receptor, ICL1 and C-terminal region conformation and positioning largely maintains the native structure. The angle of inclination for ICL2 is decreased, displacing the apex of this loop towards the center of the receptor and the apex of ICL3 is displaced away from the center of the receptor.

The overall implication from these ICLGa complex conformations is that the receptor adopts either the inactive native conformation (ICLGa25, ICLGa29) or an intermediary structure between the inactive and active conformations (ICLGa8, ICLGa12, ICL28, ICLGa37).

ICLGa compound binding alters the conformation of the transmembrane helices of CXCR4.

Given the interactions of the ICLGa compounds with the intracellular surface of CXCR4 and the altered conformations of the intracellular and extracellular surfaces, we investigated the effect of inhibitor binding on the structure of the helices. To do this, we compared the reduced representations of the transmembrane helices for each of the ICLGa complexed conformations of CXCR4 to the inactive native conformation and the constitutively active N119S conformation of the protein. Comparison of multiple receptor conformations to the native CXCR4 structure and the constitutively active N119S variant was facilitated by the helixanal program developed by Dr. Trent.

In investigating the reduced representations of the transmembrane region of the complexes produced by ICLGa8, ICLGa12, ICLGa25, ICLGa29, and ICLGa37, we found that they were similar in conformation to native CXCR4 as depicted in Figure 73. The N-terminal portion of the TM region is highly conserved compared to the native structure, and as a consequence, there are multiple deviations by the complexed structures from the structure of the constitutively active N119S mutant. The bulge of TM5 produced by loss of helical character between residues Val206 and Ile215 present in the native conformation

Figure 73: Transmembrane regions of ICLGa compound-CXCR4 complexes adopt a similar conformation to native CXCR4. Reduced representations of receptor transmembrane helices are presented: native CXCR4 in green, N119S in red, and the ICLGa8 conformation in teal, ICLGa12 conformation in magenta, ICLGa25 conformation in orange, ICLGa28 conformation in yellow, ICLGa29 conformation in white, ICLGa37 conformation in purple. The ICLGa8, ICLGa12, ICLGa25, ICLGa29, and ICLGa37 conformations adopt the transmembrane structure of the native receptor. The ICLGa28 conformation adopts an intermediate conformation between the active and inactive conformations of the receptor.

Figure 73: Transmembrane regions of ICLGa compound-CXCR4 complexes adopt a similar conformation to native CXCR4.

was present in the complexed structures. Conversely, the kink of TM1 at Val54, the more linear conformation of TM2, the outward flexion of TM3, the greater angle of inclination of TM4, the axial rotation of TM6, and the outward shift of TM7 present in the N119S conformation were not present in the complexed structures. This suggests that ICLGa8, ICLGa12, ICLGa25, ICLGa29 and ICLGa37 stabilize the TM structures in the inactive native conformation of the receptor potentially acting as inverse agonists.

The transmembrane region of the complex produced by ICLGa28 seems to be an intermediate form of the receptor between the native and N119S conformations as depicted in **Figure 73**. TM1 is more linear in the complexed conformation and lacks a kink at Ser46 present in the native conformation. TM2 and TM3 structure and positioning is conserved between the complexed conformation and the native form. The complexed TM4 and TM6 structures are more closely aligned with the structure of the N119S active conformation. The TM5 bulge present in the native structure is present in the complexed structure, but the helix seems to be rotated towards TM3. Finally, the cytoplasmic portion of TM7 is rotated outward from the center of the receptor. In total, ICLGa28 drives CXCR4 into an intermediate conformation with elements of the inactive native structure and the active N119S configuration, suggesting that this compound may act as a weak partial agonist.

Further examination of the transmembrane helices involved assessment of the distances between the intracellular and extracellular termini of the interhelical segments. At the extracellular surface, as seen in **Table 16**, variation

Table 16: ICLGa induced conformations adopt native interhelical distances at the extracellular surface of CXCR4. Interhelical distances measured at extracellular termini of transmembrane domains of native CXCR4, constitutively active N119s, and compound docked structures. Distances measured at the extracellular termini of reduced representations of the averaged structures. Illustrated interhelical distances demonstrate 1.5 angstrom or greater deviation from N119S structure, demonstrating that at the extracellular surface, the bound ICLGa structures maintain or induce the native conformation of CXCR4.

Table 16: ICLGa induced conformations adopt native interhelical distances at the extracellular surface of CXCR4.

TMA	TMB	Native	N119S	ICL8	ICL12	ICL25	ICL28	ICL29	ICL37
5	6	11.2	13.1	11.4	11.9	11.3	11.3	11.7	11.6
4	7	14.3	17.5	14.4	14.9	14.3	14.2	14.6	13.8
1	7	15.8	14.6	15.9	15.9	15.5	16.0	16.3	16.6
3	7	18.8	21.3	18.6	18.7	18.8	18.3	19.3	18.5
5	7	21.4	24.7	20.9	21.3	21.4	21.0	21.3	21.3

from N119S interhelical distance between TM7 and helices one, three, four and five was a consistent finding in all of the active CXCR4 inhibitors investigated.

There was a similarly uniform decrease in the TM5-6 interhelical distance when compared to N119S. All of these interhelical distances were compatible with the expected distances for the native receptor. ICLGa37 did produce a unique 1.8 angstrom increase in the interhelical distance between TM1-2, which is an unusual finding in that this is the shortest interhelical distance on the receptor and is highly conserved across conformations. In this case it could represent a shift between the native conformation of the N-terminal and ECL1 conformations and the N119S conformation. There was an additional decrease in the interhelical distance in TM2-7 in the conformation of ICLGa12. This decrease could be the source of the altered conformation of ECL3.

At the intracellular surface, as seen in Table 17, several key similarities are noted between the complexed conformations and the native structure of CXCR4. The TM3-6 interhelical distance is uniformly decreased in all of the active receptor structures. Interhelical distance between transmembrane regions TM1-5, TM2-3, TM2-5, TM2-6, TM3-4, TM3-5, and TM4-6 is decreased compared to the conformation of N119S, and reflects the expected difference between the native conformation of the receptor and the N119S conformation. This relationship is also reflected in the distance between TM1-7 in the conformations produced by compounds ICLGa8, ICLGa12, ICLGa25 and ICLGa37. These compounds all also exhibit an increased interhelical distance between TM4-7 when compared to the native structure.

Table 17: ICLGa induced conformations adopt native interhelical distances at the intracellular surface of CXCR4. Interhelical distances measured at cytoplasmic termini of transmembrane domains of native CXCR4, constitutively active N119s, and compound docked structures. Distances measured at the intracellular termini of reduced representations of the averaged structures. Illustrated interhelical distances demonstrate 1.5 Å or greater deviation from either the Native or N119S structure.

Table 17: ICLGa induced conformations adopt native interhelical distances at the intracellular surface of CXCR4.

TMA	TMB	Native	N119S	ICL8	ICL12	ICL25	ICL28	ICL29	ICL37
2	6	10.7	12.9	11.1	10.9	10.8	10.3	10.7	10.7
3	5	11.5	16.5	11.6	11.9	12.0	11.6	11.8	11.2
3	6	12.0	13.4	11.3	11.1	11.2	11.1	11.4	11.8
1	7	12.2	13.6	12.1	12.3	12.6	13.7	13.4	12.1
2	3	15.6	17.9	15.4	15.3	15.2	15.9	15.8	15.6
3	4	16.0	12.6	16.2	15.9	16.0	15.1	15.5	15.6
4	6	19.5	17.6	20.1	19.2	19.3	17.9	18.7	19.2
4	7	22.3	21.0	22.5	23.1	23.1	24.2	25.2	21.9
2	5	22.5	25.3	23.3	22.6	22.7	22.8	22.9	22.7
1	5	32.2	33.9	31.5	32.1	31.5	32.4	32.4	31.8

In addition to these alterations, compounds ICLGa28 and ICLGa29 also produce additional alterations in the interhelical distances between TM7 and the other helices as seen in **Table 18**. The interhelical distance between TM7 and TM2, TM3, TM5 and TM6 was significantly increased compared to both native and N119S conformations. The TM4-7 was also increased to a greater extent than seen in the other compounds, while the TM1-7 interhelical distance in the conformation produced by these two compounds closely reflects that of the N119S conformation.

This data reflects that the majority of compound complex conformations reflect the native conformation. In addition, each of the compounds produced a decrease in the interhelical distance between TM3-6 suggesting increased stability of the inactive conformation. The complexed structures produced by ICLGa28 and ICLGa29, produced significant variation in the position of TM7 reflecting significant structural variation from the native structure and N119S, suggesting an intermediate conformation of the receptor.

Interhelical hydrogen bonding characterizes differences between ICLGa compound binding.

To continue to investigate the differences in interhelical movements produced by compound binding, receptor hydrogen bonding was analyzed. As the N119 residue was located on TM3, and we previously demonstrated significant variation in the interhelical distances of TM7, our efforts were focused on the hydrogen bonds formed by TM3 and TM7. Hydrogen bonding in each of the compound-CXCR4 complex molecular dynamics trajectories at equilibrium

Table 18: ICLGa28 and ICLGa29 alter the position of TM7 at the intracellular surface of CXCR4. Interhelical distances measured at cytoplasmic termini of transmembrane domains of native CXCR4, constitutively active N119s, and compound docked structures. Distances measured at the intracellular termini of reduced representations of the averaged structures. Illustrated interhelical distances demonstrate 1.5 angstrom or greater increase from the Native and N119S structures, indicating significant alterations of the structure of TM7 by compounds ICLGa28 and ICLGa29.

**Table 18: ICLGa induced conformations adopt native interhelical distances
at the intracellular surface of CXCR4.**

TMA	TMB	Native	N119S	ICL28	ICL29
2	7	11.7	12.3	13.8	14.5
6	7	12.3	11.8	15.8	16.3
3	7	21.5	22.2	24.7	25.3
5	7	24.8	23.7	27.8	28.6

was determined using a PTRAJ program to process and analyze sets of 3-D coordinates and trajectory files. Occupancy is defined as the prevalence of a hydrogen bond across the 5 ns trajectory, as is used to quantify the influence of a hydrogen bond on receptor structure.

The hydrogen bonding by TM7 residues in each of the compound-CXCR4 complexes was analyzed. Bonds between TM7 and TMs 2, 3, 6, and ECL2 were identified for each of the conformations residues as illustrated in [Table 19](#). Hydrogen bonding between Asn298-Tyr76, Asp187-Ser285, and Asp187-Thr287 was consistent between all of the compound-CXCR4 complexes. Compounds ICLGa12, ICLGa25, ICLGa28 and ICLGa29 each formed stable hydrogen bonds between Asp84-Asn298 and Asn119-His294, but the hydrogen bond occupancy of Asp84-Asn298 for ICLGa28 was significantly reduced as was the hydrogen bond occupancy of Asn119-His294 for ICLGa25. In addition, the hydrogen bond occupancy between Thr240 and Tyr302 was 70.2% for the ICLGa25 complex while this value was significantly reduced for ICLGa28 and ICLGa29, and there was no bond at all in the ICLGa12 complex. Hydrogen bonding between Asn288 and Lys282 also significantly varied between compound complexes, with ICLGa12 and ICLGa28 producing almost full occupancy of this hydrogen bond across the trajectory, while ICLGa25 and ICLGa29 produced significant reductions in the occupancy of this hydrogen bond. ICLGa12, ICLGa25, ICLGa28 and ICLGa29 each exhibit hydrogen bonding between His294 and Trp252, although this hydrogen bond is much more prevalent in the ICLGa28 conformation.

Table 19: Hydrogen bonds formed by the side chains of TM7 in anti-CXCR4 compound bond molecular dynamics simulations. Hydrogen bonds mediated by TM7 residues were found to be conserved in each of the derived conformations. % indicates percent occupancy of the hydrogen bond.

Table 19: Hydrogen bonds formed by the side chains of TM7 in anti-CXCR4 compound bond molecular dynamics simulations.

ICL Ga12			ICL Ga25		
Donor	Acceptor	%	Donor	Acceptor	%
84 OD1	298 ND2	89.6	84 OD1	298 ND2	89.76
119 OD1	294 NE2	73.12	119 OD1	294 NE2	48.44
187 OD2	285 OG	99.6	187 OD2	285 OG	99.64
187 OD2	287 OG1	82.52	187 OD2	287 OG1	92.24
288 OE2	282 NZ	53.56	240 OG1	302 OH	70.24
288 OE2	282 NZ	35.2	288 OE2	282 NZ	52.12
288 OE2	282 NZ	8.44	288 OE1	282 NZ	8.16
294 ND1	252 NE1	28.4	294 ND1	252 NE1	22.24
298 OD1	76 OH	86.84	298 OD1	76 OH	82.52
ICL Ga28			ICL Ga29		
Donor	Acceptor	%	Donor	Acceptor	%
84 OD1	298 ND2	60.04	84 OD1	298 ND2	89.6
119 OD1	294 NE2	80.6	119 OD1	294 NE2	68.28
187 OD2	285 OG	99.88	187 OD2	285 OG	99.8
187 OD2	287 OG1	82.2	187 OD2	287 OG1	81.2
240 OG1	302 OH	17.4	240 OG1	302 OH	24.76
288 OE2	282 NZ	41.08	288 OE1	282 NZ	32
288 OE2	282 NZ	32.24	288 OE2	282 NZ	17.72
288 OE2	282 NZ	17.84	294 ND1	252 NE1	29.64
294 ND1	252 NE1	57.56	298 OD1	76 OH	81.76
298 OD1	76 OH	89.16			

The hydrogen bonding by TM1 residues in each of the compound-CXCR4 complexes was also analyzed in [Table 20](#). Only one hydrogen bond involving TM1 was present in all four complexed structures, the bond between TM1 residue Tyr65 and the C-terminus residue Ser319. Although present over 30-40% of the trajectory for compounds ICLGa28 and ICLGa29, this bond was nearly non-existent for residues ICLGa12 and ICLGa25.

Significant differences in TM3 mediated interhelical hydrogen bonding in each of the compound-CXCR4 complexes is illustrated in [Table 21](#). Hydrogen bonding between Asp171-His113 and Asp187-Tyr116 was consistent for all of the residue complexes.

However hydrogen bond occupancy was reduced for the bonds between Tyr116-Tyr255 and Asn119-His294 for compounds ICLGa25 and ICLGa29. Occupancy of the Thr117-Arg188 bond was increased for the ICLGa12 complex and decreased for the ICLGa29 complex. Occupancy of the Tyr121-Ala188 bond was increased by compounds ICLGa28 and ICLGa29. Finally, the occupancy between Asp84-Asn119 was decreased by ICLGa12 and increased by ICLGa28.

Significant differences in TM4 mediated interhelical hydrogen bonding in each of the compound-CXCR4 complexes is illustrated in [Table 22](#). All of the complexes consistently formed hydrogen bonds between Glu153-Tyr157. However, ICLGa28 produced a significantly greater occupancy of the hydrogen bond between Tyr157-Lys75, a bond which was barely present in the other conformations. Also, hydrogen bonds between Asp171-His113 and Asp171-Lys110 were occupied in the conformations produced by ICLGa12,

Table 20: Hydrogen bonds formed by the side chains of TM1 in anti-CXCR4 compound bond molecular dynamics simulations. Hydrogen bonds mediated by TM1 residues were found to be conserved in each of the derived conformations. % indicates percent occupancy of the hydrogen bond.

Table 20: Hydrogen bonds formed by the side chains of TM1 in anti-CXCR4 compound bond molecular dynamics simulations.

	Donor	Acceptor	%
ICL12	319 OG	65 OH	0.04
ICL25	319 OG	65 OH	0.08
ICL28	319 OG	65 OH	41.72
ICL29	319 OG	65 OH	30.44

Table 21: Hydrogen bonds formed by the side chains of TM3 in anti-CXCR4 compound bond molecular dynamics simulations. Hydrogen bonds mediated by TM3 residues were found to be conserved in each of the derived conformations. % indicates percent occupancy of the hydrogen bond.

Table 21: Hydrogen bonds formed by the side chains of TM3 in anti-CXCR4 compound bond molecular dynamics simulations.

ICL Ga12			ICL Ga25		
Donor	Acceptor	%	Donor	Acceptor	%
84 OD2	119 ND2	49.72	84 OD2	119 ND2	34.24
84 OD1	119 ND2	14.64	84 OD1	119 ND2	26.24
84 OD2	119 ND2	3.76	84 OD2	119 ND2	13.88
116 OH	255 OH	77.52	116 OH	255 OH	51.4
117 OG1	188 NE	48.28	117 OG1	188 NH2	20.4
117 OG1	188 NH2	3.16	117 OG1	188 NE	3.52
119 OD1	294 NE2	73.12	119 OD1	294 NE2	48.44
121 OH	188 NH2	2.36	121 OH	188 NH2	2.52
121 OH	188 NH1	1.64	121 OH	188 NH2	1.04
171 OD2	113 NE2	15.52	171 OD2	113 NE2	16.8
187 OD1	116 OH	99.04	187 OD1	116 OH	99.32
ICL Ga28			ICL Ga29		
Donor	Acceptor	%	Donor	Acceptor	%
84 OD2	119 ND2	92.76	84 OD2	119 ND2	57.6
84 OD1	119 ND2	1	84 OD1	119 ND2	11.24
84 OD2	119 ND2	0.44	84 OD2	119 ND2	5.2
116 OH	255 OH	69.72	116 OH	255 OH	59.72
117 OG1	188 NE	20.16	117 OG1	188 NE	0.28
119 OD1	294 NE2	80.6	119 OD1	294 NE2	68.28
121 OH	188 NH2	21.32	121 OH	188 NH1	42.16
121 OH	188 NH2	2.04	121 OH	188 NH2	2.36
171 OD2	113 NE2	21.16	171 OD2	113 NE2	17.24
187 OD1	116 OH	98.76	187 OD1	116 OH	95.32

Table 22: Hydrogen bonds formed by the side chains of TM4 in anti-CXCR4 compound bond molecular dynamics simulations. Hydrogen bonds mediated by TM4 residues were found to be conserved in each of the derived conformations. % indicates percent occupancy of the hydrogen bond.

Table 22: Hydrogen bonds formed by the side chains of TM4 in anti-CXCR4 compound bond molecular dynamics simulations.

ICL Ga12			ICL Ga25		
Donor	Acceptor	%	Donor	Acceptor	%
153 OE1	157 OH	46.72	153 OE1	157 OH	31.24
153 OE2	157 OH	52.8	153 OE2	157 OH	68.16
157 OH	75 NZ	0.04	157 OH	75 NZ	0.16
171 OD2	113 NE2	15.52	157 OH	75 NZ	0.04
171 OD2	110 NZ	12.12	157 OH	75 NZ	0.04
171 OD2	110 NZ	11.28	171 OD2	113 NE2	16.8
171 OD1	110 NZ	11.16	171 OD2	110 NZ	27.92
171 OD1	110 NZ	8.48	171 OD2	110 NZ	21
171 OD1	110 NZ	7.24	171 OD2	110 NZ	19.12
171 OD2	110 NZ	10.84	171 OD1	110 NZ	4
			171 OD1	110 NZ	2.88
			171 OD1	110 NZ	2.56
ICL Ga28			ICL Ga29		
Donor	Acceptor	%	Donor	Acceptor	%
153 OE1	157 OH	96.12	153 OE1	157 OH	41.56
153 OE2	157 OH	0.56	153 OE2	157 OH	56.72
157 OH	75 NZ	9.08	157 OH	75 NZ	0.16
157 OH	75 NZ	9	157 OH	75 NZ	0.04
157 OH	75 NZ	8.6	157 OH	75 NZ	0.04
171 OD2	113 NE2	21.16			
171 OD2	110 NZ	26.08			
171 OD2	110 NZ	19.92			
171 OD2	110 NZ	16.76			
171 OD1	110 NZ	2.12			
171 OD1	110 NZ	1.72			
171 OD1	110 NZ	1.64			

ICLGa25 and ICLGa28 but completely absent in the conformation formed by ICLGa29.

This hydrogen bonding data suggests that movements of TM7 induced by ICLGa28 and ICLGa29 produce an alteration of the hydrogen bonding network, which is distinct from ICLGa12 and ICLGa25. The ICLGa28 and ICLGa29 hydrogen bonding networks may produce a CXCR4 conformation with less internal stability due to decreased hydrogen bonding by TM3, TM4 and TM7.

Discussion

Our goal was to target the CXCR4:G-protein interface to discover novel small molecules which could selectively inhibit receptor function and be used as scaffolds for development of clinically active anti-metastatic compounds. Initial screening of the thirty-seven ICLGa compounds at 50 μ M identified twenty-four compounds which inhibited chemotaxis (Figure 55). This hit rate of 64.9 % at 50 μ M for initial screening of our thirty-seven ICLGA compounds represents a significant increase over contemporary studies [314-316]. Ten of the twenty-four active compounds had chemotaxis IC₅₀ values of 10 μ M or less (Figure 56), for an exceptional hit rate of 27%. Of the ten ICLGa compounds inhibiting CXCR4-mediated chemotaxis at 10 μ M or less, only five exhibited any anti-BLT1 activity, and none of these possessed an anti-BLT1 IC₅₀ of less than 25 μ M, suggesting that the most active compounds are highly specific.

Subsequent investigations of the impact of the ICLGa compounds on signal transduction by CXCR4 revealed that twelve ICLGa compounds produced significant reductions in calcium release (Figure 58), and six of the twelve active

compounds had calcium mobilization IC₅₀ values of 10 μ M or less (Figure 59). Selectivity screening of the ICLGa compounds with anti-calcium release activity revealed that six of the nine ICLGa compounds with anti-BLT1 activity had IC₅₀ values that were lower for BLT1 than for CXCR4, suggesting that the anti-calcium release activity produced by these compounds was not specific for CXCR4. However, three of the most active compounds had significantly greater anti-CXCR4 activity than anti-BLT1 activity, suggesting that anti-calcium mobilization activity in these compounds is selective for CXCR4.

The effects of the twenty-six active anti-CXCR4 compounds were produced in the absence of significant cytotoxicity, as only seven of the compounds produced elevated cytotoxicity at 50 μ M, and only four of these compounds produced any cytotoxicity at less than 10 μ M.

Sorting of the thirteen ICLGa compounds with anti-CXCR4 activity at 10 μ M or less in chemotaxis and calcium mobilization allows the compounds to be divided into four categories (Table 23). Seven compounds exhibit CXCR4-specific and exclusive anti-chemotaxis activity. Three compounds exhibit CXCR4-specific anti-chemotaxis and anti-calcium mobilization activity. Two compounds exhibit non-specific exclusive anti-calcium mobilization activity. The remaining compound exhibits CXCR4-specific and exclusive anti-calcium mobilization activity. The ICLGa compounds could be interfering with the interaction between CXCR4 and G α i, completely preventing signal transmission through the receptor inhibiting both chemotaxis and calcium mobilization as is the case with ICLGa25, ICLGa28 and ICLGa36.

Table 23: Classification of ICLGa Compounds with anti-CXCR4 activity of 10 μ M or less. ICLGa compounds with anti-CXCR4 activity can be classified by their primary mode of inhibition; either chemotaxis or calcium mobilization. Additional subclassification based on molecular dynamics studies suggest that compounds which do not successfully interact with the native structure can be classified as potential inverse agonists, while dynamically bound compounds can be classified as potential antagonists or weak partial agonists.

Table 23: Classification of compounds with anti-CXCR4 activity at 10 μ M or less.

	Chemotaxis IC50 (μM)	Calcium IC50 (μM)	Activity	Target	Mode
ICL Ga4	10	35	Chemotaxis	Gai	
ICL Ga5	5	45	Chemotaxis	Gai	Inverse Agonist
ICL Ga8	3	--	Chemotaxis	CXCR4	Weak Partial Agonists
ICL Ga9	45	1	Calcium	CXCR4	Inverse Agonist
ICL Ga12	12.5	5	Calcium	CXCR4	Weak Partial Agonists
ICL Ga13	12.5	5	Calcium	CXCR4	Inverse Agonist
ICL Ga17	10	--	Chemotaxis	CXCR4	Inverse Agonist
ICL Ga22	10	--	Chemotaxis	Gai	Inverse Agonist
ICL Ga25	10	1	Dual	CXCR4	Antagonists
ICL Ga28	1	10	Dual	CXCR4	Weak Partial Agonists
ICL Ga29	10	--	Chemotaxis	CXCR4	Antagonists
ICL Ga36	1	6	Dual	CXCR4	Inverse Agonist
ICL Ga37	5	--	Chemotaxis	Gai	Weak Partial Agonists

Assessment of inhibition of tumor proliferation by three of the most active ICLGa compounds by MTT screening revealed that the three compounds produce three different avenues of proliferation inhibition. ICLGa9 is generally inhibitory, affecting both normal and cancer cells at levels lower than the cytotoxicity levels seen in **Figure 62**. ICL25 inhibited the growth of metastatic cancer cell lines A549 and MDA-MB-231, but not normal cells or non-metastatic MCF7 cells. Finally, ICL36 inhibited the growth of non-metastatic MCF7 cells, metastatic cancer cell lines to a lesser extent, and did not inhibit proliferation of HS27 cells. These results are intriguing as a CXCR4 inhibitor need not be cytotoxic to function, for instance ICL36 inhibits chemotaxis but does not affect metastatic cell proliferation.

Of the twelve docked compounds, stable CXCR4: compound complexes were generated for compounds ICLGa8, ICLGa12, ICLGa25, ICLGa28, ICLGa29 and ICLGa37. ICLGa5, ICLGa9, ICLGa13, ICLGa17, ICLGa22, and ICLGa36 did not form stable complexes with the native conformation of CXCR4. As these compounds were initially designed to target the constitutively active N119S conformation of CXCR4, it may be that ICLGa5, ICLGa9, ICLGa13, ICLGa17, ICLGa22, and ICLGa36 are only capable of interacting with the receptor in active conformation, suggesting that these compounds could be acting as inverse agonists.

The bound ICLGa compounds were able to demonstrate several significant differences in CXCR4 structure. All of the compounds interacted with the aliphatic components of the side chains and backbones of ICL1 and ICL2,

suggesting that the compounds could act as bridging elements between these two regions, stabilizing their position and preventing signal transduction or producing a more closed intracellular surface potentially incapable of binding G α i. Interaction by the compounds with the C-terminal region could potentially stabilize TM7, or interfere with the interface between the C-terminus and G α i. This could be especially true in the case of ICLGa28, where interposition of the compound between the C-terminus and ICL3 could restrict the movement of both regions and directly impede the cleft which would be occupied by bound G α i.

Comparisons of the intracellular and extracellular surfaces of CXCR4 produced by the ICLGa compounds revealed that the receptor adopts either the native conformation (ICLGa25, ICLGa29) or an intermediary structure between the inactive native conformation and the constitutively active conformation (ICLGa8, ICLGa12, ICL28, ICLGa37), suggesting that the receptor is being stabilized in an inactive state by the ICLGa compounds, either by an inverse agonist or by a weak partial agonist. This stabilization is further suggested by the uniformly decreased interhelical distance TM3-6 in the compound conformations. In prior lipid bilayer molecular dynamics simulations, a decrease in TM3-6 distance has been associated with T140 acting as an inverse agonist, while AMD3100 acting as a weak partial agonist did not produce this effect [200]. This suggests that these compounds are acting as inverse agonists, and producing similar alterations to the inverse agonist T140.

Thus, we have developed a series of compounds with anti-CXCR4 activity by targeting the intracellular CXCR4:G α i interface. This study represents the first

successful attempt to generate potential pharmaceuticals by targeting this interface that has produced phenotypical changes in the GPCR, rather than just conformational changes of the receptor. Molecular dynamics demonstrated that several of these compounds produced changes in CXCR4 consistent with inverse agonist activity. As CXCR4 represents one of the central mechanisms in cancer metastasis, these novel compounds could represent lead templates for the possible treatment of metastatic disease.

CHAPTER IV

VALIDATION OF INHIBITOR COMPOUNDS IDENTIFIED BY VIRTUAL SCREENING

The process of structure based drug design

Although high-throughput screening methodologies have provided the traditional pipeline for lead compounds for pharmaceutical development [322, 323], SBDD has produced some notable successes. Some of the most notable successes have involved the development of the inhibitors targeting HIV protease [324, 325] and the neuraminidase inhibitors for treating influenza [326].

SBDD is an iterative process, illustrated in **Figure 74**. The process begins with identification of a structural basis for a pathological phenomenon. Compounds are screened targeting this process, through high-throughput or virtual screening to identify lead compounds which modulate a phenotype. Active compounds are then tested for toxicity and specificity. Compound and receptor interactions are then modeled. The basic structure of the compound is then investigated through the development of derivatives based on the initial model of compound:receptor interaction. Quantitative structure activity relationship analysis (QSAR) can then be developed to characterize compound activity based on modifications to select structural regions. In addition, mutational analysis can be used to determine particular regions of the target essential to compound

Figure 74: Progression in the SBDD protocol for developing inhibitors of

CXCR4. Surflex-Dock was used to virtually re-dock ECLVS and ICLGa compounds into their respective surfaces of CXCR4 to identify potential interaction residues between the receptors and compounds. Molecular dynamics based on *in silico* mutants were used to determine if structural modifications of the target sites would induce an altered conformation of CXCR4. Site directed mutagenesis would also be used to develop mutants to identify the *in vitro* effects of modification of the active site residues. Finally, the most active anti-CXCR4 compounds will be tested in *in vivo* assays to determine if they can inhibit metastasis.

Figure 74: Progression in the SBDD protocol for developing inhibitors of CXCR4.

function or to further characterize the structure of the target. Additional derivatives can be generated based on the regions described in QSAR analysis to test a new hypothesis of interaction in an effort to uncover the mechanism of action for the lead compounds. Well-characterized and strongly active compounds can then be submitted for animal testing in an effort to translate findings from the laboratory to the clinical setting.

In the case of our work, the identification of metastatic cells expressing elevated concentrations of cell surface CXCR4, combined with migration by these cells in relation to a CXCL12 gradient by Müller [9], highlighted CXCR4 as a potential target for inhibiting metastasis. We utilized virtual screening targeted at separate regions or targets of CXCR4 to identify two series of active inhibitors of the receptor, using CXCR4-mediated chemotaxis as a model for metastasis. Our compounds are highly specific in their anti-CXCR4 activity, and lack non-specific cytotoxicity, as discussed in Chapters II and III. Thus, continuing development of our identified compounds will involve reflecting back on our mechanistic model, as well as looking forward to testing of the compounds *in vivo* in animal models of pathology.

***In silico* methods can be high-throughput methodologies for SBDD.**

Virtual screening of compounds through molecular docking methods such as Surflex[234], Autodock[227], and DOCK[226] is based on the theory that the suggested compound placement produced by the docking program corresponds to the compound conformation adopted in the cellular milieu. Several studies have confirmed that virtual docking studies can reproduce the active site

placement of the compounds seen in the crystal structure [327-329] with particular success in carbohydrate ligands [330], factor Xa inhibitors [331], acetylcholinesterase [332] and cytochrome P450 [333].

Evaluation of successful compound placement is largely based on evaluation of electrostatic and hydrophobic interactions between the compound and the receptor, as well as steric factors such as clash scores which quantify penetration by the placement of the compound in the surface of the target protein. The scores used to evaluate the docked conformations can either be proprietary to the programs as in Surflex[234], DOCK[226] or Autodock[227] or ligand conformations can be output by the docking program and a separate scoring program such as FlexX[240], SCORE[226], and Ligandfit[241] can be used to evaluate the compounds.

Surflex-Dock [234, 244] screening begins with the development of a protomol, a series of pseudo-atoms used to define the target for the docking sequence. Ligand based protomols based on ligands, inhibitors or partner protein substructures allow specific interaction residues to be highlighted by the docking process. The same is true for residue based protomols, which can be generated based on series of defined residues to target specific regions of an active site. Proper placement and focus of a target site for a virtual screening program can make the difference between a screen that identifies active inhibitors and one that returns nothing but false positives [334].

The results of a Surflex-Dock virtual screen are not simply a list of single compounds with an associated score. The program generates multiple initial

fragmentations of the compounds in its flexible docking protocol. These initial fragments are scored and used as a starting position for construction of the full virtual compound. Greater results variety and exploration of the space of the active site can be achieved by randomizing the initial conformation of the compound and varying the site for initial placement of the compound fragments. Randomization of the compound starting conformation causes the docking program to consider different fragmentation schemes, and thus attempts build the flexible ligand in the active site in different fashions between different conformations. Varying the initial starting position of the compound fragments causes different chemical elements of the active site to come into play in building the rest of the compound depending on where the process is begun. As a result of this, hundreds of unique compound conformations can be produced for a single compound, and by expanding the sampling pool the highest scoring compound conformations from a true sampling of the entire space of the active site can be identified.

Mutagenesis assays can be used to characterize receptor structure and function.

SBDD can also be approached from the perspective of the receptor and not just the compound, in order to gain greater understanding of receptor structure. Mutagenesis studies have been extensively used to characterize GPCR functional mechanisms. There are three primary mechanisms by which mutation can alter receptor function: 1, the mutation can directly interfere with ligand binding; 2, the mutation can result in a conformational shift altering the

ligand binding pocket; and 3, the mutation can modulate allosteric movement altering signal transduction.

One of the best examples of mutational studies directly interfering with ligand binding corresponds to studies concerning the mechanism of ligand binding for CCR5. CCR5 is the GPCR co-receptor used for plasma membrane fusion in M-tropic strains of HIV-1 [335]. Scanning alanine site directed mutagenesis has been used to identify acidic residues (Asp2, Asp11 and Glu18) which are essential to chemokine and gp120 binding, and by extension, HIV entry [336]. Further study of this region by this technique added the aromatic residues Tyr3 and Tyr10 to the residues which are essential to chemokine and gp120 binding [337, 338]. Further analysis of the extracellular surface of CCR5 including the extracellular loops identified Lys171 and Lys191 as essential elements in the binding of ligands MIP-1 α and RANTES [339].

In GPCRs, mutations modulating allosteric sites and preventing signal transduction usually correspond to mutations of the TMs. Basu *et al* in the Trent and Bodduluri laboratories utilized a series of mutants targeting the TM region of the Leukotriene B4 receptor BLT1 to characterize the allosteric effects of that mutations to this region had on receptor function. Mutations of the TM1 residue Asn36, TM3 residue Asp64, and the TM7 residues Ser277, Ser278, Ser279 and Asn281 produced a significant reduction in LTB4 mediated chemotaxis and calcium mobilization. Specifically, mutation of these residues prevented interactions between Asp64 and Asn281 as well as Asp64 and Asn36, which were required for transition of the molecule to its active state [151]. Similar

studies involving the thyrotropin receptor [340] and histamine H1 receptor [341] revealed that mutation of TM regions could prevent transition of the receptors to the active state.

Mutational analysis has helped to characterize the structure and function of CXCR4. Chimeras replacing the extracellular domains of CXCR4 with comparable domains of CXCR2 were used in some of the earliest studies of CXCR4 structure to characterize interactions between the extracellular surface of the receptor at the distal N-terminus and ECL3 and CXCL12. [342] These studies also demonstrated that the serine and threonine residues of the carboxyl terminus were essential to CXCL12-induced signaling, through a series of truncation mutants, which reduced the length of the C-terminus. Mutational analysis has also been used to characterize the antagonist binding site of anti-CXCR4 compounds T140 and AMD3100 [200]. Alanine scanning mutations identified Asp171, Arg188, Tyr190 and Asp262 as essential to binding of T140 to the extracellular surface of the CXCR4 at ECL2. Additional alanine scanning mutations identified Phe189, Tyr190, Asp262 and Glu288 as essential to the binding of AMD3100 to the base of ECL3.

Biophysical studies can be used to characterize GPCR structure, but analysis of transmembrane proteins remains difficult.

Biophysical analysis is another method which can be used to define and develop characteristics of protein structure, although these methods are not without their difficulties when applied to transmembrane proteins. Crystallography remains the gold standard of biophysical and structural characterization of

proteins. However, growth of well-ordered three dimensional crystals remains the major constraint of generating high-resolution structural data. Novel crystallization techniques have allowed the crystal structures of the β 2-adrenergic receptor [108], α 2A-adrenergic receptor [111] and recently, CXCR4 [257] to be characterized. These techniques have essentially been designed to minimize the mobility and flexibility of the protein including removal of the N-terminal and C-terminal subunits as well as substitution of ICL3 with T4L [109, 110] which exchanges a mobile and flexible unit for a solid and immobile element which is a desirable characteristic in the development of well-ordered three dimensional crystals. In spite of the good resolution of these novel crystallization studies, one has to question the impact of removing essential elements of protein structure, and what impact this would have on the derived crystals of the remaining structure.

Spin labeling is another method which can be used to monitoring the structure and dynamics of soluble and membrane bound proteins. In this method, sulfhydryl groups are modified by a nitroxide reagent to produce a nitroxide probe on substructures of a larger group. This nitroxide probe possesses unique paramagnetic properties which can be used in conjunction with electron paramagnetic resonance spectroscopy to define the solvent accessibility, mobility, polarity index of the adjacent environment, and the distance between the probe and the paramagnetic center of the protein, all vital information for the biophysical characterization of biomolecules [343]. Some of the earliest characterizations of GPCR structure were derived from experimentation based

on incorporation of rhodopsin into lipid bilayers with motion of TM segments assessed by the spin-labeled phosphatidylcholine which made up the bilayers [344]. Subsequent development of site-directed mutagenesis allowed development of a site-directed spin labeling technique for assessing biophysical properties of proteins [345]. In this method, site directed mutagenesis is used to replace protein residues with cysteine, which can be subsequently used to react with the nitroxide probe of the spin-labeling reaction allowing characterization of specific protein subunits. The developer of this method, Dr. Wayne L Hubbell, has utilized the method primarily to detect structural changes between the active and inactive conformations of rhodopsin. Site directed spin labeling has been used to characterize the mobility of the C-terminus region, describing the structure as a disordered and dynamic unit tethered at one end of the protein, with no additional stability provided by palmitoylation of the protein at residues 322 or 323 [346]. Spin-labeling of N-terminal and C-terminal elements allowed the active state of the receptor to be characterized by movement of TM7 away from TM1 when compared to the inactive state of the receptor [347]. This method is not without its problems in assessing biophysical characteristics of transmembrane proteins. As the paramagnetic signal is highly sensitive to its surrounding environment, the protein should be placed in purified micelles rather than the cell membrane to directly assess the influence of the protein on the paramagnetic signal rather than the surrounding elements of the cellular milieu. It is difficult to purify sufficient volumes of TM protein to perform these assays, as well as place the proteins in micelles while maintaining some semblance of

structural elements of an active receptor. In addition, as GPCR structure is highly dependent on several conserved cysteine linkages, addition of other cysteine elements could produce abnormal structural conformations by incorrect linkages with substituted cysteine residues. The bulky nitroxide labels provide another potential source of error in the form of disrupted native structure, and conformational flexibility of the labels themselves leading to poor resolution [348].

***In vivo* analysis of inhibition of metastasis in animal models**

Given the myriad factors and conditions involved in successful metastasis of neoplastic cells, accurate *in vivo* models of metastasis are exceptionally difficult to develop. In xenograft models, neoplastic cells are directly injected into proposed sites of metastasis, such as long bones like the femur in an orthotopic model [349] or into the mammary fat pad [350], and tumor growth in the novel microenvironment is assessed. Non-small cell lung cancer (NSCLC) cells can also be used in the assessment of metastasis. One such method injected A549 cells (a NSCLC cell line) directly into SCID mouse lung tissue and assessed distant organ and contralateral lung metastasis development. Philips *et al* used this method to identify increased cell surface expression of CXCR4 in A549 cells with increased metastasis to distant organs and the contralateral lung, a process which could be inhibited by knockdown of CXCL12 expression [32]. Another common model of metastasis used in conjunction with pharmaceutical development is the tail vein injection method. In this method, severe combined immune deficient mice are injected in their large diameter tail veins with cells from tumor cell lines such as MDA-MB-231 [9, 351]. This method was utilized in

the Müller paper to identify favorable microenvironments and homing locations for metastatic cells expressing cell surface CXCR4. This method has also been utilized to demonstrate the anti-metastatic properties of T140 derivative TN14003 [351]. Other models of metastasis focus on the development of bone metastasis. In this model, tumor cell lines such as MDA-231-BSC12 are directly injected into the left ventricle of athymic mice and cells preferentially metastasize to sites in the bone marrow [350]. This model has been successfully used to demonstrate the anti-metastatic properties of anti-CXCR4 peptide CTCE-9908 [350].

Tail vein injection remains one of the most popular methods for assessing *in vivo* metastasis, primarily because of its ease of use and draw backs of the other methods. Direct injection of cells into the lung parenchyma runs the risk of producing pneumonia in the injected lung, and carries an increased risk of bleeding and pneumothorax. Direct injection of cells into the left ventricle carries its own difficulties, with an experienced and steady hand required to successfully place the needle in the required location, and an increased risk of mortality in the injected animals due to arrhythmia or cardiac tamponade due to disruption of the ventricular chamber. Tail vein injections are not without their own drawbacks, as they suffer from a significant first pass effect, as the cells must pass through the portal and pulmonary circulation, increasing the risk of entrapment of tumor cells in the attendant capillary beds of these systems. One of the main benefits of the bone metastasis model is that it avoids this first pass effect by directly injecting the cells into the systemic circulation, and thus provides at least one full pass of the cells through the systemic circulation.

In this chapter, we will seek to continue the SBDD process for our anti-CXCR4 compounds targeting the extracellular and intracellular surfaces of the receptor. We will use virtual screening in a novel manner, taking advantage of the additional structural information gained from the virtual screening process, to identify the sites of our targets on the extracellular and intracellular surfaces of the receptor where our compounds most frequently interacted in virtual screens. In this manner, we will identify the residues where interaction with the anti-CXCR4 compounds produced an altered phenotype of the receptor. We will then utilize molecular dynamics to demonstrate that alteration of these ECL2, ECL3, ICL3 and C-terminal residues modulate the TM regions of the receptors, and affect the conformation of CXCR4. We will then use mutagenesis analysis to determine the effects of these mutations on CXCR4 mediated signaling and migration. We will also advance further in the SBDD process and test the anti-metastatic activity of the compounds in *in vivo* metastatic models.

Methods

Virtual re-docking

Surflex-Dock [234, 244] was used to generate high-accuracy docked conformations of the ICLGa compounds, as well as the ECLVS and EVS14 compounds. Compound structures were generated in Sybyl (Tripos) and subjected to 2500 iterations of steepest descent minimization. Compound conformation was randomized using the Surflex-Dock “random” command to generate ten random conformations of each compound. The ligand-based protocol targeting the intracellular surface of the constitutively active N119S CXCR4 described in the virtual screening protocol in Chapter III was reused for this procedure for ICLGa compounds. A novel protocol targeting the residues listed in **Table 24** was designed to dock the compounds targeting the extracellular surface of CXCR4. Surflex-Dock [234, 244] was used to dock the ten random conformations of each compound using 5 randomized starting positions. The top 200 scored poses produced by the Surflex-Dock [234, 244] protocol were collected for each compound. The top 5 for each compound were then subjected to distance analysis on either the extracellular or intracellular surface and used to identify CXCR4 residues within 3 Å of the compounds.

Transfection and generation of 300.19 stable cell lines expressing mutations of hCXCR4-GFP

Site-specific mutations of CXCR4 and their effects on inhibitor function and cell phenotype were investigated using 300.19 cell lines overexpressing stably transfected hCXCR4-GFP mutants. Residues identified in the virtual re-

Table 24: Extracellular Residues of CXCR4 used in protomol development in Surflex. Residues on CXCR4 within 15 Å of a centroid placed on the extracellular surface of CXCR4 were used to develop a residue based protomol targeting the extracellular surface of the receptor.

**Table 24: Extracellular Residues of CXCR4 used in protomol development
in Surflex.**

Residue	Amino Acid	Residue	Amino Acid	Residue	Amino Acid
1	Met	171	Asp	194	Leu
2	Glu	172	Phe	195	Trp
3	Gly	173	Ile	196	Val
4	Ile	174	Phe	197	Val
5	Ser	175	Ala	198	Val
21	Tyr	176	Asn	199	Phe
22	Asp	177	Val	200	Gln
23	Ser	178	Ser	203	His
24	Met	179	Glu	263	Ser
25	Lys	180	Ala	266	Leu
26	Glu	181	Asp	267	Leu
27	Pro	182	Asp	270	Ile
28	Cys	183	Arg	271	Lys
29	Phe	184	Tyr	276	Phe
31	Glu	185	Ile	277	Glu
102	Trp	186	Cys	280	Val
104	Phe	187	Asp	281	His
105	Gly	188	Arg	282	Lys
106	Asn	189	Phe	283	Trp
107	Phe	190	Tyr	284	Ile
109	Cys	191	Pro	285	Ser
110	Lys	192	Asn	286	Ile
113	His	193	Asp	288	Glu
116	Tyr				

screening protocol for site directed mutagenesis are listed in Table 25. Additional mutants were developed based on distance calculations between the intracellular surface of CXCR4 and the comparable interface surface of Gai, which are also listed in Table 25. PCR-based site-directed mutagenesis was utilized to generate sequence confirmed plasmids using primers listed in Tables 26 and 27. PCR-based site-directed mutagenesis was carried out using the followed parameters for each sample: 5 µl 10X Thermo Pol Buffer, 1 µl 100 mM MgSO₄, 2.5 mM dNTPs, 1 µl Deep Vent Polymerase, 2 µl 50 ng/µl Template DNA, 1 µl each of Sense and Anti-sense primer and 31 µl Millipore H₂O. PCR was run using the following specifications: an initial denaturing step of 4 min at 95 °C, followed by 20 repetitions of 45 s 95 °C, 1 min 54 °C, 6 min 72 °C, and a final annealing step of 15 min at 72 °C. 300.19 cells were maintained as suspension cultures in RPMI supplemented with 10% FBS, 100 U/ml penicillin and 100 µg/ml streptomycin, and 55 µM of 2-mercaptoethanol. Transfections were performed by electroporation. The stable cell lines are maintained in the presence of 1 mg/ml of G418. Cells are cultured in complete medium for 24 hrs after transfection and used for microscopy. The cell surface expression of hCXCR4 is also determined by incubating receptor or parental cells with 12G5 Ab followed by PE labeled secondary Ab as well as detection of protein GFP tag and expression was analyzed by FACS (BD Biosciences).

Chemotaxis analysis

Chemotaxis analysis for suspension 300.19 cells expressing hCXCR4-GFP mutants were performed as previously described in chapter 2.

Table 25: Residues identified for SDM. Residues identified in the virtual redocking protocol which established close contacts with the ECLVS compounds are listed in the first column. Close contact residues interacting with the ICLVS compounds are listed in the second column. Residues which were identified as within 2 Å of the Gai G-protein subunit in the N119S-G-protein hetero trimer homology model are listed in the third column.

Table 25: Residues identified for SDM.

ECL	ICL	Gai
Val177	Lys239	Lys234
Ser178	Thr240	Arg322
Glu179	Ala307	Lys327
Phe189	Lys308	Lys333
Asn192	Phe309	Ser341
Leu194	Ser338	Ser351
Phe276		Ser352
His281		

Table 26: CXCR4 ECL SDM plasmids. Plasmids utilized in PCR based site directed mutagenesis SDM of CXCR4 to mutate residues of the extracellular region of CXCR4 identified by virtual redocking. Listed with mutant and sense and anti-sense plasmids for development of SDM for each residue.

Table 26: CXCR4 ECLSDM plasmids.

Mutant	Direction	Plasmid
V177A	Sense	5'-CTTTGCCAACGCCAGTGAGGCAGA-3'
	Anti-sense	5'-TCT GCC TCA CTG GCG TTG GCA AAG-3'
N192A	Sense	5'-TTC TAC CCC GCT GAC TTG TGG GTG G-3'
	Anti-sense	5'-CCA CCC ACA AGT CAG CGG GGT AGA A-3'
F276A	Sense	5'-GGT GTG AGG CTG AGA ACA CTG TGC A-3'
	Anti-sense	5'-TGC ACA GTG TTC TCA GCC TCA CAC C-3'
H281A	Sense	5'-GAA CAC TGT GGC CAA GTG GAT TTC CA-3'
	Anti-sense	5'-TGG AAA TCC ACT TGG CCA CAG TGT TC-3'
S178A-E179A	Sense	5'-TTG CCA ACG TCG CTG CGG CAG ATG-3'
	Anti-sense	5'-CAT CTG CCG CAG CGA CGT TGG CAA-3'
F189A-L194A	Sense	5'-GAC CGC GCC TAC CCC AAT GAC GCG TGG GTG-3'
	Anti-sense	5'-CAC CCA CGC GTC ATT GGG GTA GGC GCG GTC-3'

Table 27: CXCR4 ICL SDM plasmids. Plasmids utilized in PCR based site directed mutagenesis SDM of CXCR4 to mutate residues of the intracellular region of CXCR4 identified by virtual redocking. Listed with mutant and sense and anti-sense plasmids for development of SDM for each residue.

Table 27: CXCR4 ICL SDM plasmids.

Mutant	Direction	Plasmid
K234A	Sense	5'-GGC CAC CAG GCG CGC AAG GCC CT-3'
	Anti-sense	5'-CCG GTG GTC CGC GCG TTC CGG GA-3'
K239A	Sense	5'-AAG GCC CTC GCG ACC ACA ATC-3'
	Anti-sense	5'-TTC CGG GAG CGC TGG TGT TAG-3'
K239A-T240A	Sense	5'-GCG TTC CGG GAG CGC CGG TGT TAG TAG GGG TA-3'
	Anti-sense	5'-CGC AAG GCC CTC GCG GCC ACA ATC ATC CCC AT-3'
T240A	Sense	5'-AAG GCC CTC AAG GCC ACA ATC ATC C-3'
	Anti-sense	5'-TTC CGG GAG TTC CGG TGT TAG TAG G-3'
A307E	Sense	5'-GCT TTC CTT GGA GAG AAA TTT AAA-3'
	Anti-sense	5'-CGA AAG GAA CCT CTC TTT AAA TTT-3'
K308A	Sense	5'-TTC CTT GGA GCC GCA TTT AAA ACC-3'
	Anti-sense	5'-AAG GAA CCT CGG CGT AAA TTT TGG-3'
F309A	Sense	5'-TTG GAG CTA AGG CTA AAA CCT CTC C-3'
	Anti-sense	5'-AAC CTC GGC TTC GAT ATT GGA GAT G-3'
A307E-K308A	Sense	5'-TTT CCT TGG AGA GGC ATT TAA AAC CTC TGC CG-3'
	Anti-sense	5'-AAA GGA ACC TCT CCG TAA ATT TTG GAG ACG GG-3'
K308A-F309A	Sense	5'-TTT CCT TGG AGC CGC AGC TAA AAC CTC TGC CC-3'
	Anti-sense	5'-AAA GCA ACC TCG GCG TCG ATT TTG GAG ACG GG-3'
A307E-K308A-F309A	Sense	5'-TTT CCT TGG AGA GGC AGC TAA AAC CTC TGC CC-3'
	Anti-sense	5'-AAA GGA ACC TCT CCG TCG ATT TTG GAG ACG GG-3'
R322A	Sense	5'-GTG AGC GCA GGG TCC AGG CT-3'
	Anti-sense	5'-CAC TCG CGT CCC AGG TCG GA-3'
K327A	Sense	5'-GGG TCC AGC CTC GCG ATC CTC TCC AAA-3'
	Anti-sense	5'-CCC AGG TCG GAG CGC TAG GAG AGG TTT-3'
K333A	Sense	5'-CCC AGG TCG GAG CGC TAG GAG AGG TTT-3'
	Anti-sense	5'-TAG GAG AGC --- CCT TTC GCT G-- -3'
S338A	Sense	5'-GGT GGA CAT GCA TCT GTT TCC ACT-3'
	Anti-sense	5'-CCA CCT GTA CGT AGA CAA AGG TGA-3'
S341A	Sense	5'-TCA TCT GTT GCC ACT GAG TCT G-3'
	Anti-sense	5'-AGT AGA CAA CGG TGA CTC AGA C-3'
S351A	Sense	5'-AGT TTT CAC GCC AGC CTA CG-3'
	Anti-sense	5'-TCA AAA CTG CGG TCG GAT GT-3'
S352A	Sense	5'-TTT CAC TCC GCC CTA CAT CAT-3'
	Anti-sense	5'-AAA GTG AGG CGG GAT GTA GTA-3'
S351A-S352A	Sense	5'-CAA GTT TTC ACG CCG CCC TAC ATC ATC AT-3'
	Anti-sense	5'-GTT CAA AAG TGC GGC GGG ATG TAG TAG TA-3'

Cell signaling analysis

Cell signaling by CXCR4 was investigated by assessment of intracellular calcium mobilization as seen in chapter 2.

Molecular dynamics

Fully solvated lipid bilayer molecular dynamics simulations of mutated CXCR4 structures were performed using the AMBER suite of programs using the parm99 force field [284]. As previously described in chapter 2. Calculations were performed on servers from the Cardinal Research Cluster.

Development of animal model for CXCR4 dependent cancer metastasis in lungs.

Lewis lung carcinoma, B16F10 melanoma and TC-1 cervical cancer carcinoma, mouse cell lines known to metastasize to lungs in mice models, were screened for CXCR4 cell surface expression using flow cytometry based surface receptor staining using a flurochrome labeled antibody to mouse CXCR4 receptor. TC-1 cervical cancer carcinoma cell lines were selected for tail vein injection studies.

Toxicity of lead compounds was assessed by intraperitoneal injection of compounds at an initial dose of 5 mg kg⁻¹ body weight, followed by a second dose two days later of either 20 mg kg⁻¹ body weight or 40 mg kg⁻¹ body weight. Toxicity was assessed by loss of body weight and development of signs of physical distress including lethargy and signs of illness over a period of 3 weeks.

Antimetastatic compound activity was assessed using a mouse pulmonary metastatic cancer model. In this model, 21 days of daily intra-peritoneal injection

of AMD3100, ECLVS14, ICLGa28 or ICLGa36 (5 mg kg^{-1} body weight) were followed by injection of 5×10^5 TC-1 cells intravenously into tail veins of C57Bl/6 mice. On day 21 post tumor challenge, mice were sacrificed and lungs were perfused and stored in formalin for 16 hours prior to transfer to 70% ethanol. Tumors were counted by magnifying glass, microscopy and size was measured by digital caliper. The total number of tumors counted in all the five lung lobes of an individual mouse was presented as tumor incidence and addition of the mean tumor sizes were done to calculate cumulative tumor size which was presented as tumor burden for each mouse.

Results

Virtual Re-docking analysis identifies potential critical residues on the extracellular and intracellular surfaces of CXCR4.

Virtual screening can produce more information than just the placement of a single compound in a binding site. The results of a virtual screen are frequently presented by the program as a list of the top scored conformations of the individual compounds, but the number of compounds returned in a given screen is a variable that can be set by the user. As can be seen in Figure 75, there is significant diversity in the poses produced by docking a single compound into a receptor. Our hypothesis is that by docking multiple conformations of active compounds into the active site of the receptors, we will be able to identify specific sites which are essential to receptor activity. In essence we will be virtually screening the residues of the active sites to identify potential targets for site directed mutagenesis analysis under the assumption that because the natural function of these residues is interrupted by the anti-CXCR4 compounds, their interruption by site directed mutagenesis could produce similar results.

To this end, we employed a high-intensity docking protocol using Surflex-Dock [234, 244] to identify CXCR4 amino acid residues within two Å of the top five scored conformations of each compound. The high-intensity docking protocol incorporated ten randomized starting conformations for each compound and five unique starting positions for the initial fragment placement in the flexible docking protocol of Surflex-Dock. The extracellular-targeted compounds most frequently interacted with nine residues illustrated in Figure 76. Compounds primarily

Figure 75: Virtual screening can be used to assess the interaction of multiple conformations of a single compound within an active site. By inspecting the position of multiple compounds in an active site, additional information about the receptor can be obtained, identifying specific interaction residues which can affect phenotypical changes through many different compounds. A) The top 10 scored conformations of ECLVS14 produced by Surflex-Dock, placed in the extracellular surface of the receptor. B) The top 10 scored conformations of ICLGa36 produced by Surflex-Dock, placed in the extracellular surface of the receptor.

Figure 75: Virtual screening can be used to assess the interaction of multiple conformations of a single compound within an active site.

Figure 76: Virtual rescreening of extracellular loop targeted compounds identifies specific contact residues. Scores are presented as percentage of compound poses within two Å of the residue for the entire set of compounds (in blue), those with anti-chemotaxis activity (in red) and those with anti-calcium mobilization activity (in green). Compounds primarily interacted with nine specific residues of ECL2 and ECL3.

Figure 76: Virtual rescreening of extracellular loop targeted compounds identifies specific contact residues.

contacted the second and third extracellular loops along with the N-terminal methionine. The most frequently interacting residues were Val177, Ser178, Glu179 on the N-terminal side of the second extracellular loop and Phe189, Pro191, Asn192 and Leu 194 on the C-terminal side of the loop. Phe276 and His281 on the third extracellular loop also frequently interacted with the compounds. It is important to note that there was relatively little overlap between the compounds with anti-chemotaxis activity and those with anti-calcium signaling activity.

The intracellular-targeted compounds primarily interacted with six residues illustrated in **Figure 77**. The six most common contact points were located on 3 regions of the intracellular surface. Lys239 and Thr240 were located on the intracellular surface of TM6. Ala307, Lys308 and Phe309 are located on the intracellular surface of TM7. Ser338 was located on the C-terminal tail of CXCR4. As these compounds represent a structurally diverse initial virtual screen, the contact frequencies seen in the virtual re-docking were substantially decreased when compared to screening results of the structurally similar extracellular-targeting compounds. There is a wider diversity in contact residues between compounds with anti-chemotaxis activity and anti-calcium signaling, with anti-calcium signaling compounds coming into more frequent contact with Lys239, Ala307 and Phe309. In contrast, contact with residues Lys308 and Ser338 seems to favor anti-chemotaxis activity. It is also important to note that the docked position of these compounds is in direct conflict with the docked position of a key projection of G-protein Gai into the intracellular surface of CXCR4.

Figure 77: Virtual rescreening of intracellular loop targeted compounds identifies specific contact residues. Scores are presented as percentage of compound poses within two Å of the residue for the entire set of compounds (in Blue), those with anti-chemotaxis activity (in Red) and those with anti-calcium mobilization activity (in Green). Compounds primarily interacted with six residues on the intracellular surface of TM6 and TM7 and the C-terminal tail of the protein.

Figure 77: Virtual rescreening of intracellular loop targeted compounds identifies specific contact residues.

We also sought to investigate the interaction between CXCR4 and Gai. Using the model developed for this interaction, we sought to identify close contact residues of CXCR4 important to the interaction of the two proteins. Seven CXCR4 residues were identified that were within 2 Å of Gai: ICL3 residue Lys 234, and C-terminus residues Arg322, Lys327, Lys333, Ser 341, Ser 351, and Ser 352.

Analysis of CXCR4 mutants through molecular dynamics simulations

The virtual redocking protocol identified several amino acid residues which were consistent interaction points for our anti-CXCR4 compounds. Modulation of CXCR4 through interaction with these residues produced significant alterations in the phenotype of the receptor. Thus, in line with our SBDD method, we sought to further investigate and characterize the structure of our receptor by *in silico* mutagenesis of the native structure of CXCR4, to investigate whether mutations of these residues would similarly alter its conformation. As mentioned in Chapters II and III, the native CXCR4 and N119S structures were developed by Dr. Napier via homology modeling based upon the high resolution crystal structure of rhodopsin [288].

Molecular Dynamics simulations of intracellular CXCR4 mutants

We hypothesize that disruption of residues which are purported to interact with compounds targeting the intracellular and extracellular surfaces of the receptor as well as the residues identified as being within binding distance of the Gai G-protein subunit will produce significant alteration of the structure of CXCR4. To investigate this hypothesis, mutations of the native CXCR4 structure

were investigated by molecular dynamics as described in Chapters II and III. However, rather than the ten 500 ps production runs conducted in the compound docking dynamics studies in the previous Chapters, twenty production runs of 500 ps were conducted for a 10 ns total production run for each mutant to assess structural conformation.

Molecular dynamics simulations of CXCR4 intracellular and extracellular surface residue mutations form stable conformations of the receptor.

The final averaged structures of the mutated CXCR4 conformations from the 10 ns molecular dynamics simulations are shown in **Figure 78** along with the location of the mutated residues. The mutated CXCR4 conformations were energetically stable throughout the production run (data not shown), indicating that the simulation was well-equilibrated prior to production analysis.

Conformational stability of the mutated CXCR4 structures were investigated using root mean square deviation (RMSD) analysis, (data not shown). RMSD values as measured at the C α position remained stable for the majority of the production runs, with the final average structure representative of the trajectory after approximately 1.5 ns. The helical domains were stable through the production run, and the majority of structural variation was in the loop regions.

Conformation of the mutant receptors will be presented in reference to the inactive native structure and the constitutively active N119S conformations. As discussed in Chapters II and III, significant differences exist between these structures. Briefly, the N-terminal β -sheet is more compact for N119S than native. The proximal portion the N-terminus is displaced towards ECL1 for

Figure 78: Final averaged structures of molecular dynamics of mutated

CXCR4. Average structure of CXCR4 presented in cyan ribbon. Mutant residues are highlighted in ball-stick rendering. A) V177A B) S178A C) E179A D) F189A E) N192A F) L194A G) K234A H) K239A I) T240A J) F276A K) H281A L) A307E M) K308A N) R322A O) K327A P) K333A Q) S338A R) S341A S) S351A T) S352A U) V177A-F189A-L194A V) V177A-F189A-L194A-F276A W) V177A-F276A X) S178A-E179A Y) S178A-E179A-H281A Z) S178A-E179A-N192A AA) S178A-E179A-N192A-H281A AB) F189A-L194A AC) F189A-L194A-F276A AD) N192A-H281A AE) K239A-T240A AF) A307E-K308A AG) A307E-F309A AH) A307E-K308A-F309A AI) K308A-F309A AJ) S351A-S352A

Figure 78: Final averaged structure of molecular dynamics of mutated CXCR4. A) V177A B) S178A C) E179A D) F189A E) N192A F) L194A G) K234A H) K239A I) T240A J) F276A K) H281A L) A307E M) K308A N) R322A O) K327A P) K333A Q) S338A R) S341A S) S351A T) S352A U) V177A-F189A-L194A V) V177A-F189A-L194A-F276A W) V177A-F276A X) S178A-E179A Y) S178A-E179A-H281A Z) S178A-E179A-N192A AA) S178A-E179A-N192A-H281A AB) F189A-L194A AC) F189A-L194A-F276A AD) N192A-H281A AE) K239A-T240A AF) A307E-K308A AG) A307E-F309A AH) A307E-K308A-F309A AI) K308A-F309A AJ) S351A-S352A

N119S. ECL1 is conserved as a short bridging loop between TM2 and TM3 and ECL2 is conserved as a two-stranded anti-parallel β -sheet structure. ECL3 is a disordered loop between TM6 and TM7. The apex of this loop is in close proximity to the distal β sheet structure of the N-terminus in the native conformation, but is displaced towards the proximal portion of the N-terminus in the N119S conformation. As a result, constitutively-active N119S is more open at the extracellular surface, allowing greater access to the acidic residues of the extracellular termini of the TM regions. ICL1 forms a short loop joining TM1 and TM2, and the apex is displaced towards the cytoplasmic terminus of TM3 for N119S. ICL2 is a more extended loop that joins TM3 and TM4, with an apex that occurs towards the midpoint of the loop in the native conformation and more proximally in the N119S conformation. The apex of the loop is oriented towards ICL1 in the native conformation, but is displaced away from the center of the receptor in the N119S conformation. Kinks in the structure of ICL3 in the N119S conformation produce increased angles at residues Lys 230 and Ala 237, causing the apex of ICL3 to be displaced upward and away from the center of the receptor. The C-terminal region produces an eighth helix aligned perpendicular to the rest of the helices and parallel with the lipid surface. In the N119S conformation torsion at Ser313 causes this eighth helix to lose α -helix conformation, and the distal C-terminal is brought into much closer contact with the cytoplasmic termini of TM1 and TM4.

Mutation of the distal end of ICL3 alters CXCR4 structure.

Significant variations characterized the four mutants targeting ICL3: K234A, K239A, T240A and K239A-T240A. Although the majority of the structure mirrored the native conformation of the receptor, the mutants exhibited differences in the structure and positioning of the N-terminus, ICL1, ICL2, ICL3 and ECL3. The N-terminus of the K234A mutant adopted the more compact structure and orientation. This region in the K239 mutant was more distended than the native structure converting the three strand anti-parallel terminal portion of the structure to adopt a two strand anti-parallel conformation. The T240 mutant N-terminus was displaced towards ECL3 as in the N119S conformation, but the β -sheet adopted the native conformation. The mutants primarily maintained the conserved ICL1 loop, although the K239A-T240A mutant shifted the apex of this loop towards the center of the receptor. K234A and K239A-T240A mutants maintained the native conformation of ICL2, but the apex of the loop was shifted towards ICL1 by K239A and T240A. The T240A mutant shifted the apex of ICL3 towards the C-terminus and the center of the receptor, while the K234A, K239A, and K239A-T240 mutants shifted the apex of this loop away from the C-terminal region and towards ICL2. Finally, the K239A-T240A mutant shifted the apex of ECL3 towards the N-terminal region, while T240A caused an opposite effect, shifting the Apex of the loop away from the N-terminus of the receptor.

These surface changes of the receptor suggest that the mutations are causing CXCR4 to adopt different conformations. There were relatively few modifications of receptor structure caused by the K234A mutation, but those changes resulted in a more compact N-terminus producing a more exposed

CXCL12 binding site and a more open Gai binding site at the intracellular surface. The K239A mutation seemed to primarily affect the N-terminal region of the receptor, resulting in a more open intracellular surface, but primarily maintaining the native conformation. The T240A mutation seems to drive CXCR4 more towards a more inactive form, shifting the N-terminus and ECL3 to produce a distortion of the extracellular surface, but producing greater shifting of the intracellular loops toward ICL1 and the center of the receptor than in the native receptor, producing obstructions to G-protein binding. K239A-T240A mutation produces greater interaction between the N-terminus and ECL3, producing greater exposure of the CXCL12 ligand binding surface, but greater interaction between ICL2 and ICL3, suggesting that while the extracellular surface is more receptive to ligand binding, the intracellular surface would be less receptive to signal transduction.

Conformational differences were noted with respect to the intermembranous segments of the mutant CXCR4 structures as well. Although the conformations primarily conformed to the native structure of the receptor, several key differences were apparent. As depicted in **Figure 79a**, the K234A mutant produced no gross deviations from native CXCR4 TM structure. The same was true for the K239A mutant shown in **Figure 79b**, which aside from a more linear TM4 conformation produced little variation from the native structure. The T240A and K239A-T240A mutants produced more significant variations in **Figure 79c-79d**. The T240A mutant produced a shift at the apex of TM4 though additional angularity of the helix at Thr168. An additional kink in TM5 was

Figure 79: Transmembrane regions of the ICL3 mutants adopt a similar conformation to native CXCR4. Reduced representations of receptor transmembrane helices are presented: native CXCR4 in green, N119S in red, the K234A conformation in teal, K239A conformation in white, T240A conformation in magenta, and the K239A-T240A conformation in blue. A) K234A, B) K239A, C) T240A, D) K239A-T240A

Figure 79: Transmembrane regions of the ICL3 mutants adopt a similar conformation to native CXCR4. A) K234A B) K239A C) T240A D) K239A-T240A

produced at Pro211 while in the K239A-T240A mutant this helix was tilted towards TM4. The basal termini of TM6 and TM7 in both mutants were shifted towards the intracellular surface of the membrane.

The TM regions of the receptors were further characterized by measurement of interhelical distances as depicted in [Table 28](#). At the extracellular surface, K234A mutation primarily altered TM1 and TM5 positioning, modulating distances based on these TM regions and reproducing the N119S interhelical distance for TM1-2, TM1-3, and TM1-4. K239A mutation seemed to have widespread effects, modulating interhelical distances corresponding to TM1, TM4 and TM6, reproducing the N119S interhelical distance for TM1-2, TM1-4, TM1-5, TM2-4, and TM5-6. T240A and K239A-T240A did not produce as much change at the extracellular surface of the receptor, primarily modulating TM6 position.

Interhelical distances at the intracellular surface of the receptor are illustrated in [Table 29](#). K234A mutation had relatively few effects on intracellular interhelical distances, decreasing the TM4-7 distance compared to native and causing TM1-6 to approach the N119S interhelical distance. K239A produced similar light effects at the intracellular surface, causing a mild increase in TM6-7 distance, and a decrease in the TM5-6 distance. A more substantial difference was noted in the T240A mutant, where mutation produced increased interhelical distances for TM1-7, TM2-7, and TM6-7, and decreases in the TM5-6 distance. Very substantial modifications were noted with the K239A-T240A mutant, as this mutant produced significant alteration in the positioning of TM7, leading to an

Table 28: Interhelical distances at the extracellular surface for the ICL3 mutants. Interhelical distances measured at extracellular termini of transmembrane domains of native CXCR4, constitutively active N119s, and ICL3 mutant structures measured in averaged conformations of dynamically determined structures. At the extracellular surface, the K234A and K239A mutants alter the position of TM1.

Table 28: Interhelical distances at the extracellular surface for the ICL3 mutants.

	Native	N119S	K234A	K239A	T240A	K239A-T240A
TM1-2	6.5	7.6	7.7	7.4	7.1	6.9
TM 1-3	20.9	22.2	22.4	21.5	21.4	21.7
TM 1-4	21.9	23.3	23.1	23.6	22.7	22.3
TM 1-5	33.6	34.8	33.6	34.5	33.5	34.6
TM 1-6	26.0	25.4	26.7	26.3	26.4	27.6
TM 1-7	15.8	14.6	17.1	17.2	15.3	16.2
TM 2-6	28.7	28.0	28.6	29.3	29.3	29.9
TM 4-6	15.9	15.4	16.7	17.1	17.0	16.3
TM 5-6	11.2	13.1	11.4	12.8	12.3	11.3
TM 6-7	12.1	13.1	11.4	11.1	12.3	12.8

Table 29: Interhelical distances at the intracellular surface for the ICL3 mutants. Interhelical distances measured at intracellular termini of transmembrane domains of native CXCR4, constitutively active N119s, and ICL3 mutant structures of averaged structures dynamically determined structures. Mutations of residues 239 and 240 altered the position of TM6 and TM7.

Table 29: Interhelical distances at the extracellular surface for the ICL3 mutants.

	Native	N119S	K234A	K239	T240A	K239A-T240A
TM 1-5	32.2	33.9	32.7	30.8	31.9	33.6
TM 1-6	20.5	21.3	21.2	20.3	20.7	21.3
TM 1-7	12.2	13.6	11.5	11.9	13.1	13.1
TM 2-5	22.5	25.3	23.0	22.4	23.1	23.8
TM 2-7	11.7	12.3	10.4	11.9	12.7	14.3
TM 3-4	16.0	12.6	15.7	16.2	15.8	14.8
TM 3-6	12.0	13.4	11.6	11.4	11.2	10.7
TM 3-7	21.5	22.2	20.6	21.8	24.3	26.0
TM 4-7	22.3	21.0	20.4	22.3	23.7	24.6
TM 5-6	13.3	12.7	13.3	12.7	12.8	13.1
TM 5-7	24.8	23.7	24.8	24.6	27.3	29.6
TM 6-7	12.3	11.8	12.6	13.4	15.1	17.5

average increase of 3.4 Å for the interhelical distances involving this helix when compared to the native conformation. This effect seems to correlate across the mutants, as the interhelical distance between TM6-7 for the K239A mutant exhibited a mild increase, which was greater in the T240A mutant and greater still in the double mutant. This suggests that the base of TM6 where K239 and T240 are located is essential to the maintenance of TM7 position and that these modifications do substantially affect CXCR4 structure.

Mutations proximal to Helix eight alter TM7 position.

Mutations of the proximal portion of the C-terminal region corresponding to the beginning of helix eight primarily adopt the native conformation of the receptor, including maintenance of helical structure in helix 8 in all mutants. There were a few notable exceptions to conservation of native conformation. At the N-terminus the A307E mutant adopted a more extended conformation of the β sheet, while the A307E-F309A mutant adopts a more compact conformation. Several of the mutants also altered the conformation of ECL1, with the A307E mutant producing a shift of the apex of the loop towards the N-terminus and the center of the receptor. The K308A mutant adopted a similar conformation, although the apex of the loop was directed more towards the base of the N-terminus rather than the center of the receptor. Finally, the A307E-K308A mutant shifted ECL1 opposite the other two mutants and outward from the center of the receptor. ECL3 was altered in the A307E, A307E-F309A, K308A, and K308A-F309A mutants to adopt a conformation more in line with the N119S mutant but shifted away from the n-terminus and outward from the center of the receptor.

Thus, while the majority of the structures of the extracellular terminus are maintained in their native conformation in many of the mutants, the A307E-F309A, K308A, and K308A-F309A produce a mild distortion of the CXCL12 binding site by altering the position of ECL3. The A307E mutant also seems to adopt a more inactive conformation, with distortions of the β -sheet structure of the N-terminus in addition to the distortion of ECL3 conformation.

On the intracellular surface of the receptor, mutants K308A, A307E-F309A, A307E-K308A and K308A-F309A all shifted the apex of ICL2 opposite of the N119S conformation and towards ICL1 rather than ICL3 and the C-terminus. The A307E-K308A, K308A-F309A and A307E-K308A-F309A mutants shifted the apex of ICL3 away from the C-terminus and towards ICL2, while the K308A mutant shifted the apex of the ICL3 loop towards the C-terminus and the center of the receptor. The structure of the intracellular surface of the mutant conformations once again primarily maintained the native conformation of the receptor with a few key differences. The A307E-K308A and K308A-F309A mutants shifted both ICL2 and ICL3 away from the C-terminal region, suggesting that the G-protein binding site would be more open. K308A shifted ICL2 away from the C-terminus but ICL3 towards the C-terminus, distorting the G-protein binding site.

Structural modifications were seen on the TM regions of CXCR4 in several of the Helix 8 mutants illustrated in Figure 80. In TM1 of the K308A-F309A conformation, the base of the helix is displaced by an additional kink at Asn56. The TM4 of the A307E-K308A mutant adopted an intermediate conformation

Figure 80: Transmembrane regions of the proximal C-terminal mutants adopt a similar conformation to native CXCR4. Reduced representations of receptor transmembrane helices are presented: native CXCR4 in green, N119S in red, the A307E conformation in white, K308A conformation in blue, A307E-K308A conformation in orange, A307E-F309A conformation in magenta, K308A-F309A conformation in yellow, and the A307E-K308A-F309A conformation in tan. The A307E conformation adopts alterations of TM5 and TM6. The K308A conformation adopts alterations of TM1. The A307E-K308A conformation adopts alterations of TM4 and TM5. The A307E-F309A conformation adopts alterations of TM5 and TM6. The K308A-F309A conformation adopts alterations of TM4 and TM6. The A307E-K308A-F309A conformation adopts the transmembrane structure of the native receptor. A) A307E B) K308A C) A307E-K308A D) A307E-F309A E) K308A-F309A F) A307E-K308A-F309A

Figure 80: Transmembrane regions of the proximal C-terminal mutants adopt a similar conformation to native CXCR4. A) A307E B) K308A C) A307E-K308A D) A307E-F309A E) K308A-F309A F) A307E-K308A-F309A

between that of the native and N119S conformations of the receptor, while an additional kink displaced the base of the helix in the K308A-F309A mutant. The characteristic bulge of TM5 was rotated in the A307E, A307E-F309A and A307E-K308A mutants, suggesting rotation of the helix as a whole. TM6 in the A307E and A307E-F309A mutants was in a more linear conformation than the native TM6, while the apex of the helix was shifted towards TM5 in the K308A-F309A mutant. Given their proximity to the base of TM7, it was interesting that none of these mutants produced any gross deviation of this helix.

The interhelical distances of the proximal C-terminal mutants measured at the extracellular surface are depicted in [Table 30](#). The A307E mutation produced the most significant variation from the native conformation, primarily altering increasing interhelical distances or shifting them to coincide with the expected value for the N119S conformation by altering positioning of TM1 and TM2. Less substantial alterations were seen for the K308A mutation which only increased the interhelical distances between TM4-6 and decreased the TM4-5 distance to coincide with the N119S conformation. The A307E-K308A mutant primarily modulated the TM6 position, increasing the interhelical distance for TM4-6, TM1-6, and TM2-6. The A307E-F309A mutation shifted the interhelical distances of TM1-2 and TM5-6 to coincide with the N119S conformation. The K308A-F309A mutation acted primarily by modulating the position of TM1, resulting in increased interhelical distances between TM1-2, TM1-7 and TM1-6. A307E-K308A-F309A produced only minor variations of the extracellular surfaces of the TM regions, increasing TM4-6 and decreasing TM5-7 and TM6-7.

Table 30: Interhelical distances at the extracellular surface for the helix 8 mutants. Interhelical distances measured at extracellular termini of transmembrane domains of native CXCR4, constitutively active N119s, and helix 8 mutant structures measured in averaged conformations of dynamically determined structures. At the extracellular surface, A307E altered TM1 and TM2, A307E-K308A altered TM6, K308A-F309A altered TM1 and A307E-K308A-F309A altered TM6.

Table 30: Interhelical distances at the extracellular surface for the ICL3 mutants.

TM	Native	N119S	A307E	K308A	A307E-K308A	A307E-F309A	K308A-F309A	A307E-K308A-F309A
1-2	6.5	7.6	7.6	6.7	6.8	7.6	7.3	6.7
5-6	11.2	13.1	12.2	11.2	11.3	13.0	11.4	12.1
6-7	12.1	13.1	12.0	12.2	12.8	11.7	12.7	11.2
1-7	15.8	14.6	17.0	15.7	17.8	15.5	16.7	16.2
4-6	15.9	15.4	15.4	16.6	16.9	15.3	15.9	17.1
4-5	16.8	15.6	15.7	15.7	17.3	16.8	16.4	16.3
2-4	20.6	21.6	22.3	21.0	22.1	20.9	21.3	20.9
1-3	20.9	22.2	22.7	21.3	21.0	20.3	20.8	21.8
5-7	21.4	24.7	21.8	20.7	21.8	22.9	21.8	20.3
1-6	26.0	25.4	27.8	26.6	27.7	24.4	27.1	25.8
2-6	28.7	28.0	30.5	29.6	30.3	28.2	30.1	28.6

The interhelical distances of the proximal C-terminal mutants measured at the intracellular surface are depicted in **Table 31**. The A307E mutation had relatively few effects on interhelical distances, increasing TM6-7 and decreasing TM1-4 to coincide with the N119S conformation. The K308A mutant primarily altered TM7 position, increasing TM3-7, TM4-7, TM5-7, and TM6-7 interhelical distances. The A307E-K308A mutation mildly altered TM1 position, resulting in decreases of the TM1-4 and TM1-5 interhelical distances. The A307E-F309A mutant primarily altered TM7 position, increasing TM3-7, TM5-7 and TM6-7 interhelical distances, and decreasing the TM1-7 distance. K308a-F309A mutation primarily altered TM6 position, increasing interhelical distance of TM1-6, TM2-6, TM4-6 and TM6-7. The A307E-K308A-F309A mutant produced significant variation of TM6 and TM7 position, increasing the distances of TM1-7, TM2-7, TM3-7, and TM4-7 and decreasing the distance of TM2-6 and TM4-6.

Helix 8 mutations primarily maintained the inactive conformation of the receptor, although several of the mutations possessed alterations of TM structures. At the extracellular surface, A307E altered TM1 and TM2, A307E-K308A altered TM6, K308A-F309A altered TM1 and A307E-K308A-F309A altered TM6. At the intracellular surface, K308A, A307E-F309A altered TM7, A307E-K308A altered TM1, K308A-F309A altered TM6 and A307E-K308A-F309A altered TM6 and TM7. This suggests that the conformations adopted by the mutant structures could be more inactivated than the native structure due to altered interhelical distances, especially in mutants A307E-F309A and A307E-K308A-F309A where disruption of TM7 could alter signal transduction.

Table 31: Interhelical distances at the intracellular surface for the helix 8 mutants. Interhelical distances measured at intracellular termini of transmembrane domains of native CXCR4, constitutively active N119s, and helix 8 mutant structures of averaged structures dynamically determined structures. At the intracellular surface, K308A, A307E-F309A altered TM7, A307E-K308A altered TM1, K308A-F309A altered TM6 and A307E-K308A-F309A altered TM6 and TM7.

Table 31: Interhelical distances at the intracellular surface for the helix 8 mutants.

TM	Native	N119S	A307E	K308A	A307E-K308A	A307E-F309A	K308A-F309A	A307E-K308A-F309A
2-6	10.7	12.9	10.5	10.9	10.7	11.3	13.4	8.9
2-7	11.7	12.3	12.0	15.4	11.6	11.8	13.0	12.9
3-6	12.0	13.4	11.4	11.3	11.2	11.3	11.4	13.2
1-7	12.2	13.6	12.3	12.5	11.6	11.2	12.7	13.2
6-7	12.3	11.8	13.3	17.6	12.8	15.1	14.0	12.7
1-4	18.2	17.7	17.3	17.4	17.5	17.4	18.2	16.8
4-6	19.5	17.6	18.3	18.6	18.9	19.0	21.3	17.9
1-6	20.5	21.3	20.7	19.9	20.5	21.1	21.7	18.5
3-7	21.5	22.2	21.5	26.7	20.4	23.2	22.6	24.4
4-7	22.3	21.0	22.0	25.4	21.5	22.1	23.5	23.8
5-7	24.8	23.7	24.4	28.6	24.4	26.9	25.2	24.7
1-5	32.2	33.9	31.9	30.7	30.9	32.6	31.6	30.3

Mutations of the distal C-terminus alter the structure of CXCR4 at TM7.

Mutations of the distal tail of the C-terminus primarily preserved the structure of the native receptor. However, several of the mutants produced variations of the native conformation at the extracellular and intracellular surfaces of the receptor. The N-terminus of the receptor was predominantly preserved amongst the distal C-terminal mutants, except for the R322A mutant who altered the N-terminal β -sheet to a more compact structure, much like the N119S conformation. ECL1 was also altered by several of the mutants, with the apex of the loop being repositioned towards the apex of TM1 in the S338A mutant. The S341A and S351A mutants both shifted the apex of the loop towards the N-terminus, while the S351A-S352A mutant shifted the apex of this loop away from the N-terminus. In addition, K327A, S351A, S352A, and S351A-S352A mutants all caused ECL3 to adopt a conformation reminiscent of the N119S conformation, but shifted away from the N-terminus. Thus, K327A and S352A feature disruption of the CXCL12 binding pocket by distortion of ECL3, while S351A and S351A-S352A also disrupt this structure while also modulating ECL1 structure, producing a more open, but potentially more unstable extracellular surface.

At the intracellular surface of the receptor, K333A shifts the apex of ICL1 downwards while S351A-S352A displaces the apex of the loop towards the C-terminus. ICL2 adopts a variety of conformations in the different mutants. R322A and S341A shift the apex of ICL2 into a conformation opposite that of the N119S structure, and away from ICL3. K327A, K333A and S352A all feature a more compact conformation of the loop, which is displaced downwards from the native

position. There is a similar variety of positions for ICL3, with K327A, R322A, and K333A shifting the apex of the loop towards the C-terminus and the center of the receptor, while S338A, S351A, and S352A each shift the apex of the loop away from the C-terminus and towards ICL2. Finally, mutants R322A, S341A, S351A, and S351A-S352A produce an increased angle in Helix 8, shifting the distal end of the helix away from the cell membrane, and displacing the loop outwards from the native conformation.

Modifications of CXCR4 structure were also seen in the TM regions of the distal C-terminal mutations depicted in [Figure 81](#). In the TM1 of the R322A mutant, the apex of the helix was displaced by a kink at Asn56, while the S352A mutant TM1 was displaced towards TM2. TM2 is also modulated by several of the mutants. In the conformation of the S338A mutant, the apex of the helix is shifted outwards away from the center of the receptor. In the S352A mutant conformation, the helix is more bowed in the direction of TM4. In the S351A and S351A-S352A mutant conformations, the base of the helix aligns with the native conformation, but the apex of the helix is aligned with the N119S conformation. In the conformation of TM4 in the R322A mutant, a kink displaces the base of the helix, while an increased angle of TM4 at Leu165 results in an outward displacement of the helix in the S338A and S341A mutants. The position of TM7 is shifted towards TM1 in the S338A mutant.

The interhelical distances of the native, N119S and distal C-terminal mutants at the extracellular surface are listed in [Table 32](#). The R322A mutant had almost no position on the interhelical distances, essentially maintaining the

Figure 81: Transmembrane regions of the C-terminal mutants adopt a similar conformation to native CXCR4. Reduced representations of receptor transmembrane helices are presented: native CXCR4 in green, N119S in red, the R322A conformation in teal, K327A conformation in magenta, K333A conformation in white, S338A conformation in yellow, S341A conformation in orange, S351A conformation in blue, S352A conformation in tan, and the S351A-S352A conformation in violet. The R322A, K327A, K333A, S338A, S341A, S351A, S352A, and S351A-S352A mutant conformations adopt the transmembrane structure of the native receptor. A) R322A B) K327A C) K333A D) S338A E) S341A F) S351A G) S352A H) S351A-S352A

Figure 81: Transmembrane regions of the C-terminal mutants adopt a similar conformation to native CXCR4. A) R322A B) K327A C) K333A D) S338A E) S341A F) S351A G) S352A H) S351A-S352A

Table 32: Interhelical distances at the extracellular surface for the distal C terminus mutants. Interhelical distances measured at extracellular termini of transmembrane domains of native CXCR4, constitutively active N119s, and **distal C terminus** mutant structures measured in averaged conformations of dynamically determined structures. At the extracellular surface, K237A altered TM1 and TM2, K333A alter TM1, S338A alters TM1 and TM2, S341A alters TM3, S351A alters TM1 and TM2, S351A-S352A alters TM1 and TM2.

Table 32: Interhelical distances at the extracellular surface for the distal C terminus mutants.

TM	Native	N119S	R322A	K327A	K333A	S338A	S341A	S351A	S352A	S351A-S352A
1-2	6.5	7.6	6.7	8.0	6.9	7.3	7.1	10.8	7.0	8.6
1-7	15.8	14.6	16.6	16.5	17.0	19.4	16.4	16.6	16.1	18.2
2-3	18.3	18.6	18.4	18.8	18.0	20.0	19.1	18.1	18.4	18.1
2-7	19.1	19.3	19.8	20.5	19.7	22.6	19.9	21.8	19.7	22.2
2-4	20.6	21.6	21.1	20.8	21.1	23.6	21.4	20.9	20.9	21.0
1-3	20.9	22.2	21.0	22.3	21.5	22.6	21.8	22.5	21.5	22.1
1-4	21.9	23.3	22.3	22.4	23.0	24.8	22.7	22.8	22.4	22.8
1-6	26.0	25.4	26.7	26.7	26.4	28.6	26.5	27.6	27.2	27.1
2-6	28.7	28.0	29.2	29.8	28.7	31.3	29.3	31.6	29.8	30.0
1-5	33.6	34.8	34.3	34.8	33.2	36.0	34.4	35.0	33.8	34.2

native conformation of the receptor. More substantial alterations were seen in the K237A mutant, where TM1 and TM2 were displaced to produce increased distances in TM1-2, TM1-3, TM1-5, TM2-6, and TM2-7. K333A mutation altered TM1 position, increasing the TM1-4 and TM1-7 distance. S338A significantly altered the extracellular termini of TM1 and TM2, producing an average increased interhelical distance of 2.5 Å for each of the helices involving TM1 and TM2. S341A mutation altered the position of TM3, increasing the distance of TM1-3 and TM2-3. S351A mutation altered the positions of TM1 and TM2, increasing the distance of TM1-2, TM1-3, TM1-5, TM1-6, TM2-6 and TM2-7. S352A mutation had only minor effects on the extracellular surface of the receptor, increasing the distance only for TM1-6. S351A-S352A mutation altered the positions of TM1 and TM2 in a manner similar to the S351A mutant, increasing the distance of TM1-2, TM1-3, TM1-6, TM1-7, TM2-6 and TM2-7.

The interhelical distances of the native, N119S and distal C-terminal mutants at the intracellular surface are listed in [Table 33](#). Two of the mutants produced relatively little alteration of interhelical distance at the intracellular surface. The K327A mutant primarily altered the position of TM6, producing increased interhelical distances for TM3-6, TM5-6 and TM3-7, while the K333 mutant increased the interhelical distance of TM1-6, while decreasing the interhelical distance of TM4-7. The rest of the distal C-terminal mutants altered the position of TM7, producing significant interhelical distance variation between TM7 and each of the other helices.

Table 33: Interhelical distances at the intracellular surface for the distal C terminus mutants. Interhelical distances measured at extracellular termini of transmembrane domains of native CXCR4, constitutively active N119s, and distal C terminus mutant structures measured in averaged conformations of dynamically determined structures. At the intracellular surface, R322A, S338A, S341A, S351A, S352A and S351A-S352A alter TM7.

Table 33: Interhelical distances at the intracellular surface for the distal C terminus mutants

TM	Native	N119S	R322A	K327A	K333A	S338A	S341A	S351A	S352A	S351A-S352A
1-4	18.2	17.7	17.0	17.5	17.9	18.3	17.5	17.9	18.2	19.2
1-6	20.5	21.3	21.2	19.5	21.3	21.9	21.1	21.2	21.3	21.0
2-7	11.7	12.3	12.4	12.1	10.9	14.1	12.5	13.7	13.2	13.1
3-6	12.0	13.4	11.1	13.2	11.4	11.2	11.4	11.5	11.4	11.4
3-7	21.5	22.2	22.5	22.4	19.7	25.6	22.4	25.1	25.2	22.1
4-7	22.3	21.0	22.0	22.3	21.0	24.4	22.7	23.9	24.0	23.9
5-6	13.3	12.7	12.6	14.3	12.9	13.3	13.1	12.6	12.6	12.8
5-7	24.8	23.7	25.5	25.5	23.5	29.7	26.5	28.2	27.6	25.5
6-7	12.3	11.8	14.4	12.2	12.2	17.2	14.6	16.6	16.2	13.9

While the distal C-terminal mutants maintain the native conformation of the receptor, the increased distance between TM7 and the other helices suggests that the receptor may have difficulty successfully transducing signals from the extracellular surface to the intracellular surface due to displacement of the TM7.

Mutations of polar extracellular mutants alter the intracellular position of TM7 and the extracellular position of TM1 and TM2

Mutations of the polar and charged groups of the extracellular surface of CXCR4 produced significant alteration of several facets of the extracellular surface. Mutation of E179A and S178A produced a more compact N-terminal conformation matching the N119S conformation the receptor. In contrast, S178A-E179A mutation produced a less compact conformation than the native receptor. The N192A-H281A mutation produced a compact N-terminal region matching the N119S conformation, as well as displacing the N-terminus of the receptor towards the extracellular surface of the TM region as in the N119S conformation. The E179A, H281A, N192A-H281A, S178A and S178A-E179A-N192A-H281A mutations shifted the apex of ECL3 away from the N-terminus and outward from the center of the receptor.

Several mutations of the polar and charged extracellular residues were demonstrated to produce a shift of the first intracellular loop towards the C-terminus of the receptor, including mutations involving S178A and E179A, and the N192A-H281A mutations. E179A, H281A, N192A-H281A, S178A-E179A, S178A-E179A-N192A and S178A-E179A-N192A-H281A mutations shifted the

apex of the ICL2 away from the C-terminus and towards TM1, in a conformation that seems to be opposite of the N119S conformation. Two different conformations of ICL3 were also seen in the mutations of the polar and charged residues. The E179A, H281A and S178A-E179A-N192A-H281A mutations each shifted ICL3 towards the C-terminus and the center of the receptor. In contrast, the N192A, N192A-H281A, S178A, S178A-E179A, S178A-E179A-N192A and S178A-E179A-H281A mutations caused ICL3 to be shifted away from the C-terminus and downward away from the cytoplasmic membrane. Finally, mutants H281A, N192A-H281A and S178A produce an increased angle in Helix 8, shifting the distal end of the helix away from the cell membrane, and displacing the loop outwards from the native conformation.

Alterations of the conformation of the TM regions are also seen in the mutants of the charged and polar extracellular residues, as seen in Figure 82. Mutation of E179A alters TM1 such that there is increased angularity at residues Tyr45 and Gly55 producing angularity of the base and apex of the helix. The S178A mutation produces a more extended conformation of TM1. The H281A and N192A mutations cause the apex of TM2 to be displaced towards TM2, while the N192A-H281A and E179A mutants possess decreased angularity at Asp84, producing a displacement of the apical portion of the helix. Decreased angularity at Leu120 displaces the base of TM3 in the E179A mutant, while the Apex of the mutant is displaced towards TM4 in the S178A-E179A-N192A mutant. The N192A, H281A, S178A-E179A-N192A and S178A-E179A-H281A mutations each produce an outward tilt of TM4. Multiple conformations of tTM5 are adopted by

Figure 82: Transmembrane regions of the polar/charged ECL mutants adopt a similar conformation to native CXCR4. Reduced representations of receptor transmembrane helices are presented: native CXCR4 in green, N119S in red, S178A conformation in yellow, E179A conformation in teal, N192A conformation in white, H281A conformation in blue, S178A-E179A conformation in tan, S178A-E179A-N192A conformation in teal, S178A-E179A-H281A conformation in yellow, S178A-E179A-N192A-H281A conformation in teal, and the N192A-H281A conformation in white. The S178A, E179A, N192A, H281A, S178A-E179A, S178A-E179A-N192A, S178A-E179A-H281A, S178A-E179A-N192A-H281A, and N192A-H281A conformations adopt the transmembrane structure of the native receptor. A) S178A B) E179A C) N192A D) H281A E) S178A-E179A F) S178A-E179A-N192A G) S178A-E179A-H281A H) S178A-E179A-N192A-H281A I) N192A-H281A

Figure 82: Transmembrane regions of the polar/charged ECL mutants adopt a similar conformation to native CXCR4. A) S178A B) E179A C) N192A D) H281A E) S178A-E179A F) S178A-E179A-N192A G) S178A-E179A-H281A H) S178A-E179A-N192A-H281A I) N192A-H281A

the polar extracellular mutants. An additional kink at Pro211 displaces the apex of TM5 outwards in the S178A-E179A-H281A mutant, while the entire helix is displaced towards TM4 in the S178A-E179A mutant. The H281A and S178A-E179A-N192A-H281A mutations each cause the TM5 bulge to reproduce the more extended structure of the N119S conformation. In contrast, the N192A and S178A mutants cause the TM5 bulge to adopt a tighter conformation than the native structure, producing outward displacement of the apical and basal termini of the helix. Modulation of TM6 structure is also affected by the mutants. S178A, N192A, H281A, S178A-E179A-N192A and S178A-E179A-H281A mutations are characterized by an additional kink at the base of the helix, shifting the terminus towards TM5. The S178A-E179A and N192A-H281A mutations each produced an additional angularity at Tyr255, shifting the apex of the helix towards TM5 and producing a more complex apical loop. S178A, E179A, S178A-E179A, N192A, S178A-E179A-N192A and S178A-E179A-H281A mutants each featured an additional angularity of the base of TM7 producing an outward displacement of the terminus.

Several of the mutations of polar extracellular residues produced relatively little impact on the intrahelical distances at the extracellular termini of the TM regions as seen in [Table 34](#). S178A produced increases of the TM1-2 and TM1-3 distances, and a decrease of the TM3-5 distance, causing these distances to coincide with the expected interhelical distance of N119S. S178A-E179A-H281A and S178A-E179A-N192A-H281A had similar sparse influence on the extracellular termini of the helices, in the case of S178A-E179A-H281A

Table 34: Interhelical distances at the extracellular surface for the polar extracellular mutants. Interhelical distances measured at extracellular termini of transmembrane domains of native CXCR4, constitutively active N119s, and polar extracellular mutant structures measured in averaged conformations of dynamically determined structures. At the extracellular surface, K237A altered TM1 and TM2, K333A alter TM1, S338A alters TM1 and TM2, S341A alters TM3, S351A alters TM1 and TM2, S351A-S352A alters TM1 and TM2.

Table 34: Interhelical distances at the extracellular surface for the polar extracellular mutants.

TM	Native	N119S	S178A	E179A	N192A	H281A	S178A- E179A	N192A- H281A	S178A E179A N192A
1-2	6.5	7.6	7.8	8.8	7.9	6.9	8.2	11.9	7.5
1-3	20.9	22.2	22.2	22.3	21.5	21.6	20.7	22.0	22.0
1-4	21.9	23.3	22.6	22.4	22.3	22.7	22.4	23.5	23.3
1-5	33.6	34.8	34.1	34.8	34.3	35.6	34.2	34.1	32.5
1-6	26.0	25.4	27.0	27.7	25.3	27.6	26.8	27.7	25.6
1-7	15.8	14.6	15.6	17.0	16.2	15.2	17.9	17.7	15.9
2-3	18.3	18.6	18.6	19.8	20.9	18.7	17.9	18.1	18.8
2-4	20.6	21.6	20.8	22.1	23.9	21.5	21.3	21.2	21.4
2-5	34.5	34.9	34.2	36.2	37.6	36.4	35.5	34.9	33.0
2-6	28.7	28.0	29.4	31.7	30.8	30.3	30.2	32.3	28.5
2-7	19.1	19.3	19.2	22.3	22.2	19.0	21.7	23.1	19.5

increasing TM2-7, but decreasing TM3-5, while S178A-E179A-N192A-H281A decreased TM2-5 and increased TM4-6 to coincide with the N119S distances.

In contrast several other mutations significantly altered these interhelical distances. E179A mutation significantly altered the position of TM1 and TM2, producing an average deviation of the interhelical distances involving these to helices by 1.6 and 2.2 Å respectively. Similar effects were seen in N192A, where alteration of TM2 position increased the interhelical distances involving this helix by an average of 2.6 Å. H281 altered the TM5 and TM6, increasing TM1-5 and TM2-5 by an average of 2 Å, and TM1-6, TM2-6, and TM6-7 by an average of 1.5 Å. S178A-E179A mutation caused TM1-2, TM1-7, TM2-6 and TM2-7 to increase by an average of 2 Å. The N192A-H281A mutation significantly altered TM1 at the extracellular surface, increasing interhelical distances involving this helix by an average of 2.3 Å. S178A-E179A-N192A seemed to have a different effect from the rest of the mutations, in that it modulated TM5 position, producing and average decrease in interhelical distance of this helix by 1.8 Å.

Alterations of the interhelical distances at the intracellular surface are listed in **Table 35**. S178A, N192A, S178A-S179A, S178A-E179A-N192A, S178A-E179A-H281 and S178A-E179A-N192A-H281A alter the position of the base of TM7, altering the average interhelical distance involving this helix by 2.9 Å, 4.5 Å, 2.9 Å, -1.4 Å, 3.1 Å and 2.6 Å, respectively.

This suggests that mutations of the extracellular polar residues alter the conformation of CXCR4, primarily by displacing TM7 at the intracellular surface, and displacement of TM1 and TM2 at the extracellular surface.

Table 35: Interhelical distances at the intracellular surface for the polar extracellular mutants. Interhelical distances measured at intracellular termini of transmembrane domains of native CXCR4, constitutively active N119s, and polar extracellular mutant structures measured in averaged conformations of dynamically determined structures. At the intracellular surface, N192A altered TM6 and TM7, H281A altered TM3 and TM6, N192A-H281A altered TM1, S178A, S178A-E179A, S178A-E179A-H281A, S178A-E179A-N192A, and S178A-E179A-N192A-H281 altered TM7.

Table 35: Interhelical distances at the intracellular surface for the polar extracellular mutants.

TM	Native	N119S	S178A	E179A	N192	H281A	S178A E179A	N192A H281A	S178A E179A H281A	S178A E179A N192A	S178A E179A N192A H281A
1-4	18.2	17.7	17.8	18.4	17.2	16.7	18.8	17.1	18.0	18.0	17.6
1-6	20.5	21.3	21.7	20.7	22.4	20.7	20.5	21.7	21.8	21.0	20.0
2-7	11.7	12.3	11.3	11.5	12.6	11.2	13.5	11.8	10.5	12.9	13.1
3-5	11.5	16.5	11.7	11.5	11.9	14.3	12.4	11.5	11.5	12.2	11.7
3-7	21.5	22.2	23.3	21.2	25.2	20.5	24.3	21.3	19.9	25.5	25.0
4-6	19.5	17.6	18.7	18.5	19.0	20.1	19.1	19.5	19.1	17.7	18.4
4-7	22.3	21.0	21.8	21.9	22.7	21.1	24.4	21.5	20.8	23.3	24.4
5-7	24.8	23.7	28.4	25.8	29.5	25.0	28.2	25.0	24.1	28.1	27.1
6-7	12.3	11.8	15.6	12.5	17.3	13.1	15.6	13.3	12.3	16.3	16.0

Mutations of non-polar extracellular mutants alter the extracellular positions of TMs.

Mutations of the polar and charged groups of the extracellular surface of CXCR4 produced significant alteration of several facets of the extracellular and intracellular surface. V177A-F276A, F189A-L194A, V177A-F189A-L194A, V177A-F189A-L194A-F276A mutations produced a more compact distal N-terminus resembling the N119S conformation. The L194A, V177A and F189A-L194A mutations each shifted the apex of ECL1 towards the N-terminus. Substantial modifications of ECL3 were noted in conjunction with the V177A-F276A, F189A-L194A, V177A-F189A-L194A, V177A-F189A-L194A-F276A mutations, shifting the apex of the loop away from the N-terminus and outwards from the center of the receptor.

At the intracellular surface of the receptor, mutations altered the positioning of several loop regions. V177A, F189A, L194A, F189A-L194A, V177A-F189A-L194A and V177A-F276A mutations produced a shift of ICL2 tangential to the TM helices, but in the opposite direction from the N119S mutation. F276A or F189A-L194A-F276A mutation produces an intermediate conformation for this loop between the native and N119S positions, while directing the apex of the loop towards ICL3. The F189A, L194A, F189A-L194A and V177A-F189A-L194A mutations each shifted the apex of ICL3 towards the C-terminus and the center of the receptor. In contrast, V177A, F276A, F189A-L194A-F276A and V177A-F189A-L194A-F276A mutations shifted the apex of ICL3 away from the C-terminus and downward away from the cytoplasmic

membrane. V177A mutation was associated with a more extended and distended C-terminal region. F189A, L194A, V177A-F276A and V177A-F189A-L194A mutations were also associated with an increased angle in Helix 8, shifting the distal end of the helix away from the cell membrane, and displacing the loop outwards from the native conformation.

These mutations also significantly altered the conformations of the TM regions, as depicted in **Figure 83**. The F189A and V177A-F276A mutations produced greater displacement of TM1 towards TM2. In the F276A and V177A-F189A-L194A mutations, decreased angularity of Asp84 at the base of the helix displaces the apical termini. Several of the mutations, including V177A, L194A, F276A, F189A-L194A, F189A-L194A-F276A, V177A-F276A and V177A-F189A-L194A-F276A displaced the apex and base of TM4 by producing an increased outward tilt of the helix. In the L194A and F276A mutations, the bulge of TM5 was in a tighter conformation than the native, producing an outward displacement of the basal terminus. In contrast, the F189A-L194A-F276A mutation produced a more extended bulge in the mode of N119S. The F276A and F189A-L194A mutations exhibited increased angle at Tyr255, shifting the apex of the helix towards TM5. L194A and V177A-F276A mutations exhibited an additional angularity of Phe249, shifting the base of the helix towards TM5. Finally, V177A, F189A, L194A, F276A, V177A-F189A-L194A and F189A, L194A, V177A-F276A mutations exhibited an additional angularity at Pro299, shifting the base of the helix outwards from the center of the receptor.

Figure 83: Transmembrane regions of the non-polar ECL mutants adopt a similar conformation to native CXCR4. Reduced representations of receptor transmembrane helices are presented: native CXCR4 in green, N119S in red, the V177A conformation in teal, F189A conformation in yellow, L194A conformation in white, F276A conformation in magenta, V177A-F276A conformation in orange, V177A-F189A-L194A conformation in teal, V177A-F189A-L194A-F276A conformation in tan, F189A-L194A conformation in violet, and the F189A-L194A-F276A conformation in aqua. The V177A, F189A, L194A, F276A, V177A-F276A, V177A-F189A-L194A, V177A-F189A-L194A-F276A, F189A-L194A, and F189A-L194A-F276A conformations adopt the transmembrane structure of the native receptor. A) V177A B) F189A C) L194A D) F276A E) V177A-F276A F) V177A-F189A-L194A G) V177A-F189A-L194A-F276A H) F189A-L194A I) F189A-L194A-F276A

Figure 83: Transmembrane regions of the non-polar ECL mutants adopt a similar conformation to native CXCR4. A) V177A B) F189A C) L194A D) F276A E) V177A-F276A F) V177A-F189A-L194A G) V177A-F189A-L194A-F276A H) F189A-L194A I) F189A-L194A-F276A

Several of the mutations of non-polar extracellular residues significantly altered intrahelical distances at the extracellular termini of the TM regions as seen in [Table 36](#). Apical TM1 position was altered by the V177A, L194A, V177A-F276A and V177A-F189A-L194A-F276A mutations producing an average increase in interhelical distance of 1.5, 1.9, 1.8, and 2.3 Å, respectively. The apex of TM2 was significantly altered by the F189A-L194A mutation increasing the intrahelical distances by an average of 1.8 Å. The F189A and F189A-L194A-F276A mutations each altered the apical position of TM5, decreasing interhelical distances by 1.3 and 1.2 Å, while L194A mutation increased the interhelical distances involving TM5 by an average of 2 Å.

Interhelical distances produced by the non-polar mutants at the intracellular surface are depicted in [Table 37](#). Mutation of V177A produces an average decrease of interhelical distances involving TM3 of 1.5 Å. The mutation of F276A modulates the position of TM5, increasing interhelical distances involving this helix by an average of 2.1 Å. The F189A mutation displaces TM7 to increase its interhelical distances by an average of 3.4 Å. In contrast, L194A and V177A-F189A-L194A-F276A mutations each decrease the TM7 associated interhelical distances by 2.1 and 1.2 Å respectively. The remaining mutants maintain the native interhelical distances.

Although the non-polar mutants of the extracellular surface of CXCR4 mutants primarily maintain the native conformation of the receptor, several key differences exist. Although the conformations of the extracellular loops and transmembrane helices are primarily in the native conformation, several of the

Table 36: Interhelical distances at the extracellular surface for the non-polar extracellular mutants. Interhelical distances measured at extracellular termini of transmembrane domains of native CXCR4, constitutively active N119s, and non-polar extracellular mutant structures measured in averaged conformations of dynamically determined structures. A) V177A, L194A, F189A-L194A, V177A-F276A, and V177A-F189A-L194A-F276A alter TM1 conformation. B) F189A, L194A, F276A, F189A-L194A, V177A-F276A, F189A-L194A-F276A, V177A-F189A-L194A, and V177A-F189A-L194A-F276A alter TM5 conformation.

Table 36A: Interhelical distances at the extracellular surface for the polar extracellular mutants.

TM	Native	N119S	V177A	L194A	F189A-L194A	V177A-F276A	V177A-F189A-L194A-F276A
1-2	13.4	11.9	13.4	13.3	12.3	13.7	13.5
1-3	24.6	25.7	23.0	24.1	24.2	24.6	24.2
1-4	18.2	17.7	17.8	18.0	16.9	18.2	16.7
1-5	32.2	33.9	31.0	31.9	31.4	32.2	31.3
1-6	20.5	21.3	20.2	21.0	20.6	20.9	20.7
1-7	12.2	13.6	12.0	12.0	12.2	11.8	12.7

Table 36B: Interhelical distances at the extracellular surface for the polar extracellular mutants.

TM	Native	N119S	F189A	L194A	F276A	F189A L194A	V177A F276A	F189A L194A F276A	V177A- F189A- L194A	V177A- F189A- L194A- F276A
1-5	33.6	34.8	33.1	36.1	33.6	33.5	35.8	32.6	33.9	33.4
3-5	20.5	19.9	19.1	20.7	19.1	18.5	19.5	19.4	19.2	19.0
4-5	16.8	15.6	15.6	17.4	15.9	16.0	16.7	16.4	17.4	16.1
5-7	21.4	24.7	20.7	23.2	22.3	20.1	21.3	20.0	22.2	20.9

Table 37: Interhelical distances at the intracellular surface for the non-polar extracellular mutants. Interhelical distances measured at extracellular termini of transmembrane domains of native CXCR4, constitutively active N119s, and non-polar extracellular mutant structures measured in averaged conformations of dynamically determined structures. A) V177A, V177A-F189A-L194A-F276A, F189A, F189A-L194A, F189A-L194A-F276A, and F276A each alter the position of TM1. B) V177A, F189A-L194A, F189A-L194A-F276A, F276A, F189A, L194A, and V177A-F189A-L194A-F276A each alter the position of TM7.

**Table 37A: Altered TM1 interhelical distances at the extracellular surface
for the polar extracellular mutants.**

TM	Native	N119S	V177A	V177A F189A L194A F276A	F189A	F189A- L194A	F189A L194A F276A	L194A	F276A
1-4	18.2	17.7	17.8	16.7	17.2	16.9	18.1	18.0	17.9
1-5	32.2	33.9	31.0	31.3	30.7	31.4	31.8	31.9	34.6
1-6	20.5	21.3	20.2	20.7	21.0	20.6	21.6	21.0	22.1

**Table 37B: Altered TM7 interhelical distances at the extracellular surface
for the polar extracellular mutants.**

TM	Native	N119S	V177A	F189A L194A	F189A L194A F276A	F276A	F189A	L194A	V177A F189A L194A F276A
2-7	11.7	12.3	11.8	12.6	11.5	11.5	14.9	11.0	10.6
3-7	21.5	22.2	20.1	22.7	20.4	21.8	25.3	19.1	20.4
4-7	22.3	21.0	22.0	22.4	21.8	21.5	25.0	21.5	20.9
5-7	24.8	23.7	24.0	25.3	24.5	27.6	27.3	22.9	23.3
6-7	12.3	11.8	12.3	13.9	12.7	14.0	16.8	12.0	11.6

loops are modified the conformations of the extracellular and intracellular loops. The cases of the F189A, L194A, V177A-F276A and V177A-F189A-L194A mutations are particularly interesting as these mutations result in a significant increase in the angularity of Helix 8, shifting the base of the helix away from the intracellular surface of the receptor. These changes are reflected in alterations of the interhelical distances of the TM1 and TM5 of the extracellular surface of the receptor, and Tm1 and Tm7 on the intracellular surface of the receptor.

Influence of ECL Mutants on CXCL12-mediated chemotaxis.

In light of the modifications of native CXCR4 structure produced by the mutations of ECL residues, we attempted to generate stable cell lines expressing mutated strains of CXCR4 by site directed mutagenesis. Unfortunately, we were only able to generate stably transfected cell lines for the N192A, F276A, H281A, V177A-F276A, V177A-F189A-L194A, S178A-E179A-H281A, and S178A-E179A-N192A-H281A mutants.

We did find that several of these mutants displayed significant reductions in chemotaxis as seen in **Table 38**. Dose response analysis of the mutant CXCR4 expressing cell lines to treatment with CXCL12, demonstrated that the maximal chemotaxis of the V177A-F276A, S178A-E179A-H281A, S178A-E179A-N192A-H281A, N192A and H281A each produced significant reductions in chemotaxis when compared to wild type CXCR4. Maximal chemotaxis response of the N192A mutant was produced by treatment with 10 nM treatment with CXCL12, producing a maximal response 37% of the wild type maximal response. Maximal chemotaxis response of the V177A-F276A, S178A-E179A-H281A and

Table 38: Mutation of extracellular residues of CXCR4 alters chemotaxis.

Significant reductions of CXCL12-induced chemotaxis were produced by V177A-F276A, S178A-E179A-H281A, S178A-E179A-N192A-H281A, and N192A mutations. V177A-F189A-L194A, F276A, and H281A mutations maintained WT amplitude of chemotaxis. .

Table 38: Mutation of extracellular residues of CXCR4 alters chemotaxis.

Strain	Maximal Concentration	Chemotaxis index	St. Dev.
WT	10	1	0.27
V177A-F189A-L194A	20	1.01	0.19
V177A-F276A	20	0.20	0.19
S178A-E179A-H281A	20	0.44	0.014
S178A-E179A-N192A-H281A	20	0.10	0.13
N192A	10	0.37	0.17
F276A	20	1.47	1.32
H281A	10	1.29	0.82

S178A-E179A-N192A-H281A mutants was achieved with 20 nM of CXCL12, producing a response which was 17%, 44% and 12% of WT. Each of these mutations had produced alterations of TM7 interhelical distances at the intracellular surface of the receptor suggesting impaired signal transduction capabilities in the molecular dynamics studies.

These poor chemotaxis responses were in contrast to the results seen for the V177A-F189A-L194A, F276A, and H281A mutations. In these cases, maximal chemotaxis response at 20 nM produced migrations of 100%, 125%, and 150% of wild type. In the case of H281A, this mutant shared many dynamics characteristics with the inactive S178A-E179A-N192A-H281A in that both mutants primarily maintained the native conformation of the receptor and shifted the apex of ICL3 towards the C-terminus and the center of the receptor. The only noted differences were that H281A shifted the apex of the N-terminus outwards from the center of the receptor and away from the N-terminus and the angle taken at the origin of Helix 8 shifted the base of that helix away from the cytoplasmic membrane. The V177A-F189A-L194A and F276A mutants did not share any common features in the dynamics simulations either.

Effects of ICL mutants on CXCL12-mediated chemotaxis.

We also attempted to generate site directed mutations of ICL residues highlighted in the virtual redocking program. However, difficulties in plasmid generation, and cell line stability meant that the mutated cell lines available for initial CXCL12 dose response testing were limited to K234A, K239A, K239A-T240A, A307E-K308A, A307E-K308A-F309A, S341A, S351A and S351A-S352A.

Several of these mutants were found to have a significantly reduced chemotaxis response to CXCL12, as seen in Table 39. The K234A, K239A-T240A, A307E-K308A, A307E-K308A-F309A, and S351A mutants each expressed significant reduction in chemotaxis response when compared to WT. The A307E-K308A and S351A-S352A mutants produced no significant chemotaxis at any of the screened concentrations of CXCL12.

The C-terminal serine residues are essential to CXCR4 function through truncation mutants [342] and site directed mutagenesis [352], so it is not surprising that loss of these residues would impair chemotaxis. Mutations of helix eight have not been demonstrated to adversely affect CXCR4 function, and in fact disruption of this helix is one of the characteristics of the constitutively active mutants. But perhaps it is because in the constitutively active mutants the center of the helix is disrupted, while the mutations of the apex of the helix near the base of TM7 continue to maintain helical conformation.

The K234A, K239A-T240A and A307E-K308A-F309A mutants each migrate in response to CXCL12 although at a much reduced rate compared to the WT strain, with maximal chemotaxis at 20 nM CXCL12 producing chemotaxis which was only 40 % and 5% of WT for the K234A and A307E-K308A-F309A mutants and maximal response at 10 nM for the K239A-T240A of only 20% of WT. This suggests that ICL2 residues K234, K239 and T240 play a significant role in transducing the signal from the extracellular surface to the intracellular.

Effects of CXCR4 mutants on CXCL12-mediated calcium mobilization.

Table 39: Mutation of intracellular residues of CXCR4 alters chemotaxis.

Significant reductions of CXCL12-induced chemotaxis were produced by K234A, K239A-T240A, A307E-K308A, A307E-K308A-F309A, and S351A mutations.

K239A mutation maintained the amplitude of the CXCL12-induced chemotaxis response, while S341A and S351A-S352A mutation significantly increased CXCL12 induced Chemotaxis.

Table 39: Mutation of extracellular residues of CXCR4 alters chemotaxis.

Strain	Maximal Concentration	Chemotaxis index	St. Dev.
WT	10	1	0.27
K234A	20	0.40	0.30
K239A	20	1.1	0.78
K239A-T240A	10	0.20	0.063
A307E-K308A	20	0.0053	0.0033
A307E-K308A-F309A	10	0.054	0.060
S341A	10	4.30	0.93
S351A	1	0.095	0.14
S351A-S352A	10	1.53	0.20

Several hCXCR4-GFP mutants were generated targeting residues identified in the virtual re-docking analysis to see what impact loss of these contact residues would have on the activity of active CXCR4 inhibitors. We investigated the impact of these mutations on cell signaling via intracellular calcium mobilization. We were able to generate cell lines and perform intracellular calcium mobilization testing for the following extracellular hCXCR4-GFP mutants: N192A, F276A, H281A, V177A-F276A, V177A-F189A-L194A, S179A-E179A-H281A, and S179A-E179A-N192A-H281A. We were also able to generate cell lines and perform intracellular calcium mobilization testing for the following intracellular mutants: K234A, K239A, K239A-T240A, A307E-K308A, A307E-K308A-F309A, S341A, S351A and S351A-S352A. The impact of the extracellular mutants on CXCL12 mediated signaling was assessed in terms of signal amplitude and maximal CXCL12 concentration. Table 40 depicts the maximal CXCL12 concentrations and invoked calcium responses for the extracellular mutants. Three of these mutants, N192A, F276A and V177A-F189A-L194A, maintained a maximal CXCL12 concentration of 10 nM. The H281A, V177A-F276A, S179A-E179A-H281A and S179A-E179A-N192A-H281A mutations each increased maximal CXCL12 concentration to 20 nM. Calcium response was then assessed, and normalized to the maximal response produced by WT CXCR4. Maximal calcium response was reduced by the N192A, F276A, V177A-F276A, S179A-E179A-H281A and S179A-E179A-N192A-H281A. The H281A and V177A-F189A-L194A mutants both maintained the calcium mobilization of WT CXCR4.

Table 40: Mutation of extracellular residues of CXCR4 alters CXCL12 – induced calcium mobilization. Calcium response is presented as maximum amplitude of calcium signal normalized to the maximal response of WT CXCR4. Calcium response was reduced by N192A, F276A, V177A-F276A, S179A-E179A-H281A and S179A-E179A-N192A-H281A mutations. H281A and V177A-F189A-L194A mutants both maintained the calcium mobilization of WT CXCR4.

Table 40: Mutation of extracellular residues of CXCR4 alters CXCL12 – induced calcium mobilization.

Mutant	Maximal concentration	Calcium Mobilization	St. Dev.
WT	10	0.92	0.22
N192A	10	0.37	0.17
F276A	10	0.78	0.46
H281A	20	1.24	1.15
V177A-F276A	10	0.26	0.18
V177A-F189A-L194A	20	1.01	0.19
S179A-E179A-H281A	20	0.44	0.01
S179A-E179A-N192A-H281A	20	0.10	0.13

The impact of the intracellular mutants on CXCL12 mediated signaling was assessed in terms of signal amplitude and maximal CXCL12 concentration. **Table 41** depicts the maximal CXCL12 concentrations and invoked calcium responses for the extracellular mutants. Three of these mutants, K239A-T240A, A307E-K308A-F309A and S351A-S352A maintained a maximal CXCL12 concentration of 10 nM. The H K234A, K239A, A307E-K308A, S341A and S351A mutations each increased maximal CXCL12 concentration to 20 nM. Calcium response was then assessed, and normalized to the maximal response produced by WT CXCR4. Maximal calcium response was reduced by the K234A, K239A, K239A-T240A, and A307E-K308A-F309A. The S341A and S351A-S352A mutants both enhanced calcium mobilization when compared to WT CXCR4. A307E-K308A and S351A mutants did not effectively transmit calcium signaling. The impact of hCXCR4-GFP extracellular mutants on the activity of active CXCR4 inhibitors of intracellular calcium mobilization was then assessed. Compounds were tested using the maximal CXCL12 concentration for each mutant, and results were normalized for each mutant to a DMSO alone control. Results were prepared for available mutants; however, due to difficulty in successful generation of PCR based site directed mutant plasmid generation and difficulties in generation and maintenance of single cell derived cell lines stably expressing the plasmid protein, only select mutants were available for study.

The V177A-F276A mutation produced an interesting phenotypical difference when compared to wild type, as depicted in **Figure 84**. Mutation of these two residues resulted in active inhibition of intracellular calcium

Table 41: Mutation of intracellular residues of CXCR4 alters calcium mobilization. Significant reductions of CXCL12-induced chemotaxis were produced by K234A, K239A-T240A, A307E-K308A, A307E-K308A-F309A, and S351A mutations. K239A mutation maintained the amplitude of the CXCL12-induced chemotaxis response, while S341A and S351A-S352A mutation significantly increased CXCL12 induced Chemotaxis.

Table 41: Mutation of intracellular residues of CXCR4 alters calcium mobilization.

Mutant	Maximal concentration	Calcium Mobilization	St. Dev.
WT	10	0.92	0.22
K234A	20	0.40	0.30
K239A	20	0.65	0.19
K239A-T240A	10	0.18	0.08
A307E-K308A	20	0.01	0.00
A307E-K308A-F309A	10	0.05	0.06
S341A	20	2.07	0.41
S351A	20	0.01	0.00
S351A-S352A	10	1.53	0.20

Figure 84: ECLVS14 and EVS14-2-AB produce greater calcium inhibition in V177A-F276A mutant CXCR4. V177A-F276A mutation produces greater inhibition of calcium mobilization by ECLVS14 and EVS14-2-AB. EVS-2-AB demonstrated a significant reduction in intracellular calcium mobilization, with a nearly two-fold decrease in signaling at 50 μ M. Unpaired samples denoted by * have no WT counterpart because lack of screening activity did not warrant dose response testing.

Figure 84: ECLVS14 and EVS14-2-AB produce greater calcium inhibition in V177A-F276A mutant CXCR4.

mobilization by ECLVS14, a compound with no such activity in wild type cells. In V177A-F276A CXCR4 cells, the IC₅₀ of this activity is roughly 40 μ M, compared to no significant inhibition in Wild type cells, suggesting that interaction with the valine and phenylalanine residues in the wild type protein influences the manner in which ECLVS14 alters downstream signaling. Similarly, while not achieving significant inhibition of intracellular calcium mobilization, the V177A-F276A mutation produced a greater than two fold decrease in signal amplitude for compound EVS14-2-AB at 50 μ M. In contrast, ECLVS59 and ECLVS82 both inhibited intracellular calcium mobilization in V177A-F276A CXCR4 cells, while neither EVS14-21 nor EVS14-22 displayed a substantial increase in inhibition.

The V177A-F189A-L194A mutation produced several noteworthy phenotypical alterations, as depicted in [Figure 85](#). Unlike the V177A-F276A mutation, this mutation did not produce significant inhibition of intracellular calcium mobilization by ECLVS14, although signaling was reduced almost two-fold by V177A-F189A-L194A mutation. A similar increase in compound mediated inhibition was seen for EVS-2-AB, with an almost two fold increase in signaling inhibition, but no significant inhibition of intracellular calcium mobilization. Conversely, this mutation caused compound EVS-21 to produce a significant increase in intracellular calcium mobilization, with a greater than 2.5 fold increase in signaling compared to DMSO alone control. This mutation also produced a reduction in activity of compound ECLVS59, which while remaining a significant inhibitor of calcium mobilization, saw a decrease in IC₅₀ from 5 μ M for Wild type CXCR4 to 10 μ M for the V177A-F189A-L194A mutant.

Figure 85: EVS14-21 produces greater calcium mobilization in V177A-F189A-L194A mutant CXCR4. EVS14-21 treatment in cells expressing the V177A-F189A-L194A mutant produce an almost two and a half fold increase in CXCL12-induced signaling. Inhibitory activity of ECLVS14 and EVS-2-AB is increased almost two-fold compared to WT. Unpaired samples denoted by * have no WT counterpart because lack of high concentration activity did not warrant dose response testing.

Figure 85: EVS14-21 produces greater calcium mobilization in V177A-F189A-L194A mutant CXCR4.

N192A mutation produced relatively few phenotypic changes in compound function, as depicted in **Figure 86**. The only two differences involved improved but still not significant inhibition of intracellular calcium mobilization by ECLVS14 and EVS-2-AB. Mutation of this residue produced a roughly 1.5 fold decrease in signaling for ECLVS14 and a roughly 2.5 fold decrease in signaling for EVS-2-AB.

F276A mutation produced relatively few phenotypic changes in compound function, as depicted in **Figure 87**. Two differences involved improved but still not significant inhibition of intracellular calcium mobilization by ECLVS14 and EVS-2-AB. Mutation of this residue produced a roughly 1.5 fold decrease in signaling for ECLVS14 and a roughly 2 fold decrease in signaling for EVS-2-AB.

H281A mutation produced relatively few phenotypic changes in compound function, as depicted in **Figure 88**. Two differences involved improved but still not significant inhibition of intracellular calcium mobilization by ECLVS14 and EVS-2-AB. Mutation of this residue produced a roughly 2 fold decrease in signaling for ECLVS14 and a roughly 2.5 fold decrease in signaling for EVS-2-AB.

S178A-E179A-N192A-H281A mutation produced several interesting alterations in compound phenotypes, as depicted in **Figure 89**. While compound ECLVS59 activity remained relatively consistent between mutant and wild type, inhibition by ECLVS14, EVS-21, EVS-22, and EVS-2-AB was significantly enhanced, such that all four compounds produced significant inhibition of intracellular calcium mobilization, when compared to DMSO alone control. ECLVS14 exhibited an IC₅₀ of ~35, and EVS-21, EVS-22, and EVS-2-AB

Figure 86: ECLVS14 and EVS14-2-AB produce greater inhibition of calcium mobilization in N192A mutant CXCR4. ECLVS14 and EVS-2-AB achieve an over one and a half fold reduction in CXCL12 induced calcium signaling and a nearly three-fold reduction in signaling for EVS-2-AB. Unpaired samples denoted by * have no WT counterpart because lack of high concentration activity did not warrant dose response testing.

Figure 86: ECLVS14 and EVS14-2-AB produce greater inhibition of calcium mobilization in N192A mutant CXCR4.

Figure 87: ECLVS14, EVS14-21, EVS14-22, and EVS14-2-AB produce greater inhibition of calcium mobilization in S178A-E179A-N192A-H281A mutant CXCR4. ECLVS14, EVS-21, EVS-22, and EVS-2-AB achieve significant inhibition of CXCL12-mediated intracellular calcium mobilization in cells expressing S178A-E179A-N192A-H281A mutant CXCR4. ECLVS14 was improved to an IC₅₀ of 35μM while EVS-21, EVS-22 and EVS-2-AB each displayed IC₅₀ values of less than 10 μM. Unpaired samples denoted by * have no WT counterpart because lack of high concentration activity did not warrant dose response testing.

Figure 87: ECLVS14, EVS14-21, EVS14-22, and EVS14-2-AB produce greater inhibition of calcium mobilization in S178A-E179A-N192A-H281A mutant CXCR4.

Figure 88: ECLVS14 and EVS14-2-AB produce greater inhibition of calcium mobilization in F276A mutant CXCR4. ECLVS14 and EVS-2-AB achieve significant inhibition of CXCL12-mediated intracellular calcium mobilization in cells expressing F276A mutant CXCR4. ECLVS14 and EVS-2-AB achieve a nearly two-fold reduction in signaling for ECLVS14, and a two-fold reduction in signaling for EVS-2-AB. Unpaired samples denoted by * have no WT counterpart because lack of high concentration activity did not warrant dose response testing.

Figure 88: ECLVS14 and EVS14-2-AB produce greater inhibition of calcium mobilization in F276A mutant CXCR4.

Figure 89: EVS14-2-AB produce greater inhibition of calcium mobilization in H281A mutant CXCR4. EVS-2-AB achieves an over two and a half fold reduction in CXCL12 signaling. Unpaired samples denoted by * have no WT counterpart because lack of high concentration activity did not warrant dose response testing

Figure 89: EVS14-2-AB produce greater inhibition of calcium mobilization in H281A mutant CXCR4.

produced IC₅₀ values of 10 μ M or less. Loss of the charged contact residues in the S178A-E179A-N192A-H281A mutation allows ECLVS14 and its derived compounds to inhibit intracellular calcium mobilization to a substantially greater degree than seen in wild type CXCR4, suggesting that interaction with these residues influences the magnitude of the inhibition by these compounds, as well as the specificity of their activity.

Difficulty in successful generation of site-directed mutant plasmids and stably transfected cell lines meant that only one of the six desired intracellular mutant cell lines, K239A, was generated for compound testing as depicted in **Figure 90**. There was no significant difference in ICLGa9, ICLGa13 or ICLGa36 mediated inhibition of intracellular calcium mobilization in K239A or WT cells. However, K239A mutation produced a 2.4 fold decrease in ICLGa12 activity at 50 μ M, and a 2.7 fold decrease in ICLGa12 activity at 10 μ M. From this limited data, it suggests that Lyseine 239 on the second intracellular loop is important to ICLGa12-mediated inhibition of intracellular calcium mobilization.

Due to difficulty in producing stable CXCR4 mutants, there is a limited volume of mutations from which to compare the data between the experimental protocols exploring the effect of mutation of CXCR4 on chemotaxis and calcium mobilization. However, some important conclusions can be drawn, as can be seen in **Table 42**. The S341A and S351A-S352A mutants each produced increased amplitude of response to CXCL12 induced chemotaxis and calcium mobilization. Like many of the inactive mutations, these mutants featured a displacement of TM7, and helix eight adopted an angle, displacing the distal

Figure 90: ICLGa12 produces greater inhibition of calcium mobilization in K239A mutant CXCR4. Mutation of this residue significantly increased CXCL12-induced signaling for ICLVS12 at 10 μ M and 50 μ M, increasing the IC₅₀ of this compound from 5 μ M in WT CXCR4 to 35 μ M in the K239A strain.

Figure 90: ICLGa12 produces greater inhibition of calcium mobilization in K239A mutant CXCR4.

Table 42: Comparison of Mutant Molecular Dynamics and Experimental

data. Comparing the molecular dynamics and experimental data for the CXCR4 mutants developed by site directed mutagenesis, we find that while the mutants primarily maintained the native conformation of CXCR4, each of the mutants produced alteration of the dynamically determined structure of CXCR4. Of the mutants developed by SDM, only the V177A-F189A-L194A and H281A mutants reproduced WT chemotaxis and calcium mobilization activity. Two of the mutants were hypersensitive to CXCL12, two displayed no activity in Chemotaxis or Calcium mobilization, and the remaining mutants displayed reduced CXCR4 activity.

Table 42: Comparison of Mutant Molecular Dynamics and Experimental data.

	Molecular Dynamics					Experimental	
Mutant	Primary Conformation	ECL	ICL	EC TM	IC TM	Chemotaxis	Calcium Flux
V177A-F189A-L194A	Native	N-term, ECL3	ICL3, C-term	TM5	none	WT	WT
V177A-F276A	Native	N-term, ECL1, ECL3	ICL2, Helix 8	TM1	TM1	decreased	decreased
S178A-E179A-H281A	Native	N-term, ECL1	ICL2, ICL3	none	TM7	decreased	decreased
S178A-E179A-N192A-H281A	Native	none	ICL1, ICL2, ICL3	none	TM7	decreased	decreased
N192A	Native	ECL1	ICL3	TM2	TM6, TM7	decreased	decreased
F276A	Native	none	ICL1, ICL2, ICL3	TM5	none	WT	decreased
H281A	Native	ECL3	ICL2, ICL3, Helix 8	TM3, TM4	TM5, TM6	WT	WT
K234A	Native	N-term	ICL3	TM1, TM5	none	decreased	decreased
K239A	Native	N-term	ICL2, ICL3	none	TM1, TM4, TM6	WT	decreased
K239A-T240A	Native	ECL3	ICL1, ICL2, ICL3	TM6	TM6, TM7	decreased	decreased
A307E-K308A	Native	ECL1	ICL2, ICL3	TM6	TM1	Inactive	Inactive
A307E-K308A-F309A	Native	None	ICL3	TM7	TM6, TM7	Inactive	Inactive
S341A	Native	ECL1,	Helix 8	TM3	TM7	Increased	Increased
S351A	Native	ECL1	ICL2, ICL3, Helix 8	TM1, TM2	TM7	Inactive	Inactive
S351A-S352A	Native	ELC3	ICL1, Helix 8	TM1, TM2	TM7	Increased	Increased

portion of the helix away from the cytoplasmic membrane. However, in contrast to many of the other inactive mutants, these two preserved the structure and conformation of their intracellular loops with minimal displacement from the native conformation. Decreased chemotaxis and calcium mobilization activity. In many of the mutants is associated with alteration of the ICL conformations from the native.

***In vivo* testing of anti-CXCR4 compounds in a metastatic model reveals significant anti-metastatic activity.**

Experiments conducted by Dr. Rajesh Sharma sought to establish the activity of anti-CXCR4 compounds in a CXCR4 dependant lung tumor metastasis model. Initial development of the model centered on identifying CXCR4 cell surface expression by a series of metastatic cell lines including Lewis lung carcinoma cells, B16F10 melanoma cells, and TC-1 cervical carcinoma cells. Of these, TC-1 cells expressed the most cell-surface CXCR4, as illustrated in **Figure 91a**. This result is in line with prior research, demonstrating cidofovir anti-metastatic activity in TC-1 cells via inhibition of CXCR4 signaling [353]. Lewis lung carcinoma cells did not express elevated cell surface concentrations of CXCR4, also illustrated in **Figure 91b**.

Three compounds, ECLVS14, ICLGa28, and ICLGa36 were initially assessed for in vivo cytotoxicity. These three represented the best of the anti-CXCR4 compounds targeting the extracellular and intracellular surface of the receptor. All three produced almost no cytotoxicity and selectively inhibited CXCR4 mediated chemotaxis with chemotaxis IC₅₀ values of 5 μ M for

Figure 91: TC-1 cells express increased cell surface CXCR4. TC-1 cells (A) express increased cell surface CXCR4 when compared to Lewis Lung carcinoma cells (B). Expression of CXCR4 expression of 50 for TC-1 cells compared to 10 for Lewis Lung carcinoma cells suggests that TC-1 cells would be more responsive the CXCR4 targeted inhibitors than Lewis Lung carcinoma cells.

Figure 91: TC-1 cells express increased cell surface CXCR4.

ECLVS14, 1 μ M for ICLGa28 and 1 μ M for ICLGa36. ICLGa28 and ICLGa36 also significantly inhibited CXCL12-induced intracellular calcium mobilization with IC50 values of 10 μ M and 6 μ M respectively. *In vivo* cytotoxicity was initially assessed at 5 mg/kg body weight, 20 mg/kg body weight, and 40 mg/kg body weight. None of these concentrations produced gross toxicity, as would be exhibited by weight loss, physiological distress, overt illness, or lethargy over a period of three weeks.

Further investigation using the Lewis lung carcinoma model revealed that in a cell line with reduced surface expression of CXCR4, anti-CXCR4 compounds do not inhibit metastasis (Figure 92). Investigation of the anti-CXCR4 compounds using the TC-1 metastasis model compared ECLVS14, ICLGa28 and ICLGa36 mediated inhibition of tumor incidence and lung tumor burden to known inhibitor, AMD3100. As seen in Figure 93, all three compounds significantly reduced tumor incidence and tumor burden. ECLVS14 produced a roughly 60% improvement in tumor incidence when compared to AMD3100, while ICLGa28 and ICLGa36 produced roughly equivalent reductions in incidence. Average tumor burden in ECLVS14 treated mice was reduced by more than 60%, while reduction in tumor burden was reduced by nearly 50% for AMD3100, ICLGa28 and ICLGa36. This data confirms that these compounds are effective inhibitors of metastasis.

Summary

In the SBDD process, identification and characterization of active compounds, illustrated in Chapters II and III, is followed by rationalization of target protein structure and testing of lead compounds in *in vivo* models to

Figure 92: Lewis lung cancer cell metastasis is not altered by anti-CXCR4 compounds. Usage of Lewis lung cancer cells in an *in vivo* metastasis model reveals that treatment with anti-CXCR4 compound AMD3100 does not alter tumor incidence or tumor burden.

Figure 92: Lewis lung cancer cell metastasis is not altered by anti-CXCR4 compounds.

Figure 93: Treatment with lead compounds ECLVS14, ICLGa28 and ICLGa36 each reduce tumor incidence and tumor burden of TC-1 cells in a syngenic mouse model of lung cancer metastasis. ECLVS14 reduced tumor burden and incidence to a 60 % greater extent than AMD3100. ICLGa28 and ICLGa36 each inhibited tumor incidence and burden to the same extent as AMD3100, with each compound reducing tumor incidence and tumor burden by 50 %.

Figure 93: Treatment with lead compounds ECLVS14, ICLGa28 and ICLGa36 each reduce tumor incidence and tumor burden of TC-1 cells in a syngenic mouse model of lung cancer metastasis.

potentially translate the research from the laboratory to the clinic. In this chapter, we have used a virtual redocking protocol based on Surflex-Dock to identify residues of CXCR4 where interaction with active anti-CXCR4 compounds could elicit a modification of phenotype. We sought to mutate these residues, *in silico* through molecular dynamics simulations and *in vitro* through site directed mutagenesis, to investigate whether or not mutations of these residues could induce a change in structural conformation or an altered phenotype of CXCR4. Virtual redocking identified eight contact residues for the ECLVS compounds, and six contact residues for the ICLGa compounds. Additional intracellular mutants were generated based on contact residues between CXCR4 and Gai identified in the dynamically docked structure of the two proteins. Stable conformations of thirty-six total mutant structures of CXCR4 were generated to investigate whether mutation of the identified residues in and of themselves could induce an altered conformation of CXCR4.

Overall, the dynamically determined structures of the CXCR4 mutants primarily adopted the conformation of the inactive native receptor, with no gross deviations of transmembrane structure and maintenance of the helical conformation of Helix eight. However, several key differences were noted in some of the mutant structures. The mutations of ICL3 residues K239 and K240 altered the conformations of ICL2 and ICL3 as well as the base of TM6 and the position of the intrahelical terminus of TM7, suggesting that these residues are essential to the maintenance of TM7 position and that mutation substantially affects CXCR4 structure.

Mutations of the initial portion of helix eight were not found to disturb helical character of this region; however, several of the mutations produced alterations of the conformation of ICL3, as well as the position of TM7. Both of the mutants developed by SDM targeting this region, A307E-K308A and A307E-K308A-F309A were inactive and failed to produce a significant response to CXCL12. This suggests that the apex of helix 8 at the base of the C-terminus is essential to CXCR4 function.

Mutations of the distal C-terminus were found to predominantly maintain the native structure of the receptor. However, several of the mutations altered TM7 position, as well as the angle of helix 8 in relation to the cytoplasmic membrane. Interestingly, two of the mutants developed by SDM targeting this region displayed hypersensitivity to CXCL12 in the form of increased response in chemotaxis and calcium mobilization. One of these hypersensitive mutants, S341A, produced a unique series of conformational changes in CXCR4 with modulation of ECL1, helix 8, the extracellular terminus of TM3, and the intracellular terminus of TM7. Further exploration of these mutations would be needed to determine the extent of their hypersensitivity.

Mutations of the extracellular surface of CXCR4 produced a variety of alterations in CXCR4 conformation in molecular dynamics screening. In SDM analysis, these mutants decreased CXCR4 response to CXCL12 in the form of reduced chemotaxis and calcium mobilization. In total, this mutagenesis analysis suggests that the altered CXCR4 conformations produced by mutation of target site residues, that the target sites are essential to CXCR4 function.

Perhaps the most important finding of this chapter though, came in the in vivo testing of three of the most active anti-CXCR4 compounds identified in Chapters II and III in a mouse model of pulmonary metastasis. In this assay, we determined that ICLGa28 and ICLGa36 decreased both tumor incidence and tumor burden to the same extent as AMD3100, without any of the associated toxicity seen with the bicyclam compounds [160]. ECLVS14 produced even greater inhibition of both pulmonary tumor development and tumor proliferation, resulting in incidence and burden values of only 40 % of the values elicited by AMD3100. By comparison, testing of anti-CXCR4 palindromic peptide CTCE-9908 in a similar model of pulmonary metastasis produced a 50% reduction in tumor incidence and tumor burden [176]. An 4F-Benzoyl derivative of T140, 4F-Benzoyl-TN140003, was also found to decrease tumor burden by 37% in a similar study [30]. Thus, these studies demonstrate that the ICLGa compounds ICLGa28 and ICLGa36 each potentially inhibit CXCR4 mediated metastasis to the same extent as existing CXCR4 inhibitors. This is the first evidence that targeting the intracellular surface of a GPCR in order to inhibit G-protein binding can be as effective inhibiting protein function as targeting the extracellular surface. In addition, the 60% increase in inhibition of tumor burden and tumor incidence marks ECLVS14 as one of the most active inhibitors of CXCR4-mediated metastasis developed. In a clinical environment where there are currently no anti-metastatic compounds in clinical trials, development of a non-toxic selective inhibitor of metastasis is a significant development.

CHAPTER V: SUMMARY

Tumor metastasis represents a significant portion of the morbidity and mortality attributable to neoplasia. Bone metastasis is present in 30-50% of patients with highly metastatic visceral cancers such as breast, prostate, and non-small cell lung cancer. The diffuse and systemic nature of metastatic disease renders many conventional treatment modalities ineffective, presenting a particularly perplexing predicament for treatment. Unfortunately there are no primary anti-metastatic treatments approved by the FDA or in clinical trials.

Müller *et al* proposed a theory of metastasis centered on chemoattraction. In this theory, elevated CXCR4 surface expression by neoplastic cells coupled with expression of the ligand of CXCR4, CXCL12, in lung, lymph node, brain, bone marrow and liver tissue, leads to directed migration of neoplastic cells to these tissues. In these tissues, the neoplastic cells find favorable microenvironments from which to spawn new metastatic foci. Elevated CXCR4 expression has been detected in over twenty three varieties of neoplastic cells and is frequently associated with a more aggressive and metastatic phenotype leading to greater patient morbidity and mortality.

CXCR4 is a G-Protein coupled receptor involved in signal transduction and cell migration. It is essential in brain and heart development in the embryo, and plays a critical role in the immune system, governing stem cell migration and sequestration in the bone marrow and the development of lymphocytes and

leukocytes. CXCR4 is also central to the process of HIV-1 infection, as one of the primary coreceptors for HIV-1 entry into T-lymphocytes.

One of the goals of SBDD is to identify structural targets which can be used as a basis for inhibiting a pathological process. There are five available target sites to inhibit a GPCR's signaling. The ligand can be sequestered or inactivated or the ligand interaction surface of the receptor can be altered to prevent interaction. The transmembrane regions of the receptor responsible for signal transduction from the extracellular to the intracellular surface of the cell membrane can be disrupted, or held in an inactive position. Finally, the intracellular surface of the receptor can be prevented from interacting with the coupled G-protein, either by altering or obstructing the intracellular surface of the receptor or the complimentary interacting surface of the G-proteins. Inhibitory efforts targeting CXCR4 have targeted just the extracellular surface of the receptor. The remaining four targets remain largely unexplored. The hypothesis of our research is that inhibition of CXCR4 can result in inhibition of metastasis, and that this inhibition can be effected by targeting multiple sites of CXCR4, both on the intracellular and extracellular surface of the receptor. The goals of this research were to utilize SBDD to rationally target specific sites of CXCR4 to develop inhibitors of metastasis. In this process, we would gain a greater understanding of chemical elements which could induce phenotypical changes produced by inhibitors through the use of QSAR analysis. We would also gain insight into the structural changes elicited in the receptor by interactions between active inhibitors and their targets through the use of molecular dynamics

simulations. We would additionally gain insight into the receptor targets themselves through the use of virtual screening in conjunction with molecular dynamics and site directed mutagenesis.

Our analysis of compounds targeting the extracellular surface of the receptor identified three compounds which inhibited CXCR4 at 10 μ M or less. ECLVS14 stands out amongst these compounds as a selective inhibitor of CXCR4 with minimal cytotoxicity. Subsequent QSAR analysis of ECLVS14 revealed that although conversion of the hydrazine-alkyl linker group increased anti-CXCR4 activity, it also increased cytotoxicity. Lengthening of the hydrazine-alkyl linker moiety was found to reduce anti-CXCR4 activity, as did removal of the 1-BPMAM moiety. Modification of the terminal benzene group by addition of electronegative substituents also decreased the anti-CXCR4 activity of the parent compound. ECLVS14 is unique in its inability to modify calcium signaling, a characteristic that is not shared by any other CXCR4 inhibitors [289-291], including T140 [195] and AMD3100 [275].

Our five ns molecular dynamics simulations have allowed us to gain a finer mechanistic understanding of the interaction between CXCR4 and ECLVS14. We have found that ECLVS14 drives CXCR4 into a unique conformation which seems to be more open at the extracellular surface than seen in the native conformation, but less open than seen in the N119S conformation.

Targeting the interface between the intracellular surface of CXCR4 and the G protein subunit Gai with virtual screening allowed us to identify twenty-four

compounds which significantly inhibited CXCR4-mediated chemotaxis. This initial hit rate of 64.9 % at 50 μ M represents a significant increase over contemporary virtual screening and high-throughput studies. Ten of these compounds inhibited CXCR4 mediated chemotaxis at 10 μ M or less, and only 5 of these compounds displayed any activity in fellow class A GPCR BLT1 demonstrating that the compounds were highly selective for the intracellular interface of CXCR4. Three of the ICLGa compounds were also found to inhibit tumor cell proliferation, an intriguing result, as a CXCR4 inhibitor need not be cytotoxic to function as a metastatic agent, for instance ICLGa36 inhibits chemotaxis but does not affect metastatic cell proliferation.

Molecular dynamics analysis revealed that several of the compounds interacted with components of the side chains and backbones of ICL1 and ICL2, suggesting that the compounds could act as bridging elements between these two regions, stabilizing their position and preventing signal transduction or producing a more closed intracellular surface potentially incapable of binding Gai. Interaction by the compounds with the C-terminal region could potentially stabilize TM7, or interfere with the interface between the C-terminus and Gai. This could be especially true in the case of ICLGa28, where interposition of the compound between the C-terminus and ICL3 could restrict the movement of both regions and directly impede the cleft which would be occupied by bound Gai.

This study represents the first attempt of structure based drug design to successfully target the intracellular surface of a GPCR. Further, we demonstrate that selective inhibition of GPCR function need not only target the extracellular

CXCL12 interface, as has been the objective of all other anti-CXCR4 compound development, but rather that CXCR4 can be effectively inhibited by interfering with the entire sequence of events of its signal transduction apparatus.

Further examination of CXCR4 structure through *in silico* mutation of target site residues identified by our virtual rescreening protocol determined that while the receptors predominantly maintained in the inactive native conformation several significant differences between native and mutant structures were detected. These differences were corroborated by our admittedly limited *in vitro* analysis of mutations of the target site residues produced by site directed mutagenesis. All but two of these mutants displayed reductions of chemotaxis and calcium mobilization by 50% or greater when compared to the wild type receptor. The remaining two mutants were hypersensitive to CXCL12 administration, producing elevated chemotaxis and calcium mobilization responses. This suggests that the target sites selected are essential for CXCR4 function and that compound interactions modulating these sites, could potentially alter the CXCR4 phenotype.

Perhaps the most important findings of this dissertation came in the *in vivo* testing of three of the most active anti-CXCR4 compounds, ECLVS14, ICLGa28 and ICLGa36 in a mouse model of pulmonary metastasis. In finding that ICLGa28 and ICLGa36 both decreased both tumor incidence and tumor burden by 50 %, we demonstrated for the first time that the anti-CXCR4 compounds targeting the intracellular surface could inhibit metastasis to the same extent as extracellular inhibitors of CXCR4 AMD3100, CTCE-9908 and T140 derivative,

4F-Benzoyl-TN140003. In addition, the almost 80% decrease of tumor incidence and tumor burden produced by ECLVS14 represented highly significant inhibition of tumor burden and incidence, a 60% increase over the inhibition produced by the most advanced CXCR4 inhibitor, AMD3100. All three of the screened compounds produced no detectable toxic side effects. In a clinical environment where there are currently limited anti-metastatic compounds in clinical trials, development of three non-toxic selective inhibitors of metastasis is a significant development.

The findings presented in this dissertation provide insight into the process of SBDD targeting a GPCR. We have successfully demonstrated that metastasis can be inhibited by targeting multiple regions of CXCR4. We have used drug discovery to develop sixteen lead compounds which have great potential as scaffolds for inhibiting CXCR4. Building upon these results, further drug discovery focusing on optimization of the lead compounds could allow characterization of essential pharmacophores for lead compound activity. The impact of the compounds on different mechanisms of CXCR4 signaling could further characterize the mechanism of action for each of the lead compounds. Development of additional mutants based on the interaction residues identified in the molecular dynamics study could also allow us to confirm the dynamically determined interaction site for the compounds. Further virtual screening could attempt to develop CXCR4 inhibitors targeting the remaining two potential regions of CXCR4: the intrahelical regions of CXCR4 as well as the CXCL12 interface with the extracellular surface of CXCR4.

REFERENCES

1. Houghton, A.M., et al., *Macrophage elastase (matrix metalloproteinase-12) suppresses growth of lung metastases*. *Cancer Res*, 2006. **66**(12): p. 6149-55.
2. de Boer, M., et al., *Breast cancer prognosis and occult lymph node metastases, isolated tumor cells, and micrometastases*. *J Natl Cancer Inst*, 2010. **102**(6): p. 410-25.
3. Rove, K.O. and E.D. Crawford, *Metastatic cancer in solid tumors and clinical outcome: skeletal-related events*. *Oncology (Williston Park)*, 2009. **23**(14 Suppl 5): p. 21-7.
4. Nicolson, G.L., *Paracrine and autocrine growth mechanisms in tumor metastasis to specific sites with particular emphasis on brain and lung metastasis*. *Cancer Metastasis Rev*, 1993. **12**(3-4): p. 325-43.
5. Yokota, J., *Tumor progression and metastasis*. *Carcinogenesis*, 2000. **21**(3): p. 497-503.
6. Zhang, Y., B. Ma, and Q. Fan, *Mechanisms of breast cancer bone metastasis*. *Cancer Lett*, 2010. **292**(1): p. 1-7.
7. Paget, S., *The distribution of secondary growths in cancer of the breast*. 1889. *Cancer Metastasis Rev*, 1989. **8**(2): p. 98-101.
8. Chen, Z., et al., *Birth of parthenote mice directly from parthenogenetic embryonic stem cells*. *Stem Cells*, 2009. **27**(9): p. 2136-45.
9. Muller, A., et al., *Involvement of chemokine receptors in breast cancer metastasis*. *Nature*, 2001. **410**(6824): p. 50-6.
10. McEarchern, J.A., et al., *Invasion and metastasis of a mammary tumor involves TGF-beta signaling*. *Int J Cancer*, 2001. **91**(1): p. 76-82.
11. Hyder, S.M., C. Chiappetta, and G.M. Stancel, *Pharmacological and endogenous progestins induce vascular endothelial growth factor expression in human breast cancer cells*. *Int J Cancer*, 2001. **92**(4): p. 469-73.
12. Hilakivi-Clarke, L., *Estrogens, BRCA1, and breast cancer*. *Cancer Res*, 2000. **60**(18): p. 4993-5001.
13. Kim, H. and W.J. Muller, *The role of the epidermal growth factor receptor family in mammary tumorigenesis and metastasis*. *Exp Cell Res*, 1999. **253**(1): p. 78-87.
14. Youngs, S.J., et al., *Chemokines induce migrational responses in human breast carcinoma cell lines*. *Int J Cancer*, 1997. **71**(2): p. 257-66.
15. Verbeek, B.S., et al., *Overexpression of EGFR and c-erbB2 causes enhanced cell migration in human breast cancer cells and NIH3T3 fibroblasts*. *FEBS Lett*, 1998. **425**(1): p. 145-50.
16. Liotta, L.A., *An attractive force in metastasis*. *Nature*, 2001. **410**(6824): p. 24-5.
17. Chu, C.Y., et al., *Stromal cell-derived factor-1alpha (SDF-1alpha/CXCL12)-enhanced angiogenesis of human basal cell carcinoma cells involves ERK1/2-NF-kappaB/interleukin-6 pathway*. *Carcinogenesis*, 2009. **30**(2): p. 205-13.
18. Yasuoka, H., et al., *CXCR4 expression in papillary thyroid carcinoma: induction by nitric oxide and correlation with lymph node metastasis*. *BMC Cancer*, 2008. **8**: p. 274.

19. Oliveira-Neto, H.H., et al., *Involvement of CXCL12 and CXCR4 in lymph node metastases and development of oral squamous cell carcinomas*. Tumour Biol, 2008. **29**(4): p. 262-71.
20. Zhang, L., et al., *Tissue microenvironment modulates CXCR4 expression and tumor metastasis in neuroblastoma*. Neoplasia, 2007. **9**(1): p. 36-46.
21. Kajiyama, H., et al., *Involvement of SDF-1alpha/CXCR4 axis in the enhanced peritoneal metastasis of epithelial ovarian carcinoma*. Int J Cancer, 2008. **122**(1): p. 91-9.
22. Zhang, T., et al., *Preferential involvement of CX chemokine receptor 4 and CX chemokine ligand 12 in T-cell migration toward melanoma cells*. Cancer Biol Ther, 2006. **5**(10): p. 1304-12.
23. Reckamp, K.L., R.M. Strieter, and R.A. Figlin, *Chemokines as therapeutic targets in renal cell carcinoma*. Expert Rev Anticancer Ther, 2008. **8**(6): p. 887-93.
24. Lima-e-Silva, R., et al., *The SDF-1/CXCR4 ligand/receptor pair is an important contributor to several types of ocular neovascularization*. FASEB J., 2007. **21**(12): p. 3219-30.
25. Zlotnik, A., *Chemokines in neoplastic progression*. Semin Cancer Biol., 2004. **14**(3): p. 181-5.
26. Zeelenberg, I.S., L. Ruuls-Van Stalle, and E. Roos, *The chemokine receptor CXCR4 is required for outgrowth of colon carcinoma micrometastases*. Cancer Res, 2003. **63**(13): p. 3833-9.
27. Hermann, P.C., et al., *Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer*. Cell Stem Cell, 2007. **1**(3): p. 313-23.
28. Chinni, S.R., et al., *CXCL12/CXCR4 signaling activates Akt-1 and MMP-9 expression in prostate cancer cells: the role of bone microenvironment-associated CXCL12*. Prostate, 2006. **66**(1): p. 32-48.
29. Lapteva, N., et al., *CXCR4 knockdown by small interfering RNA abrogates breast tumor growth in vivo*. Cancer Gene Ther, 2005. **12**(1): p. 84-9.
30. Tamamura, H., et al., *T140 analogs as CXCR4 antagonists identified as anti-metastatic agents in the treatment of breast cancer*. FEBS Lett, 2003. **550**(1-3): p. 79-83.
31. Rahman, K.M., et al., *Therapeutic intervention of experimental breast cancer bone metastasis by indole-3-carbinol in SCID-human mouse model*. Mol Cancer Ther, 2006. **5**(11): p. 2747-56.
32. Phillips, R.J., et al., *The stromal derived factor-1/CXCL12-CXC chemokine receptor 4 biological axis in non-small cell lung cancer metastases*. Am J Respir Crit Care Med, 2003. **167**(12): p. 1676-86.
33. Karnoub, A.E., et al., *Mesenchymal stem cells within tumour stroma promote breast cancer metastasis*. Nature, 2007. **449**(7162): p. 557-63.
34. Leiker, M., et al., *Assessment of a nuclear affinity labeling method for tracking implanted mesenchymal stem cells*. Cell Transplant, 2008. **17**(8): p. 911-22.
35. Zhao, H. and D.M. Peehl, *Tumor-promoting phenotype of CD90(hi) prostate cancer-associated fibroblasts*. Prostate, 2009. **69**(9): p. 991-1000.
36. Muehlberg, F.L., et al., *Tissue-resident stem cells promote breast cancer growth and metastasis*. Carcinogenesis, 2009. **30**(4): p. 589-97.
37. Orimo, A. and R.A. Weinberg, *Stromal fibroblasts in cancer: a novel tumor-promoting cell type*. Cell Cycle, 2006. **5**(15): p. 1597-601.
38. Burger, J.A., M. Burger, and T.J. Kipps, *Chronic lymphocytic leukemia B cells express functional CXCR4 chemokine receptors that mediate spontaneous migration beneath bone marrow stromal cells*. Blood, 1999. **94**(11): p. 3658-67.

39. Sipkins, D.A., et al., *In vivo imaging of specialized bone marrow endothelial microdomains for tumour engraftment*. Nature, 2005. **435**(7044): p. 969-73.
40. Orimo, A., et al., *Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion*. Cell, 2005. **121**(3): p. 335-48.
41. Burger, J.A., et al., *Blood-derived nurse-like cells protect chronic lymphocytic leukemia B cells from spontaneous apoptosis through stromal cell-derived factor-1*. Blood, 2000. **96**(8): p. 2655-63.
42. Loetscher, M., et al., *Cloning of a human seven-transmembrane domain receptor, LESTR, that is highly expressed in leukocytes*. J Biol Chem, 1994. **269**(1): p. 232-7.
43. Murdoch, C., *CXCR4: chemokine receptor extraordinaire*. Immunol Rev, 2000. **177**: p. 175-84.
44. Oberlin, E., et al., *The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1*. Nature, 1996. **382**(6594): p. 833-5.
45. Bleul, C.C., et al., *The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry*. Nature, 1996. **382**(6594): p. 829-33.
46. Picard, L., et al., *Role of the amino-terminal extracellular domain of CXCR-4 in human immunodeficiency virus type 1 entry*. Virology, 1997. **231**(1): p. 105-11.
47. Mondor, I., et al., *Interactions among HIV gp120, CD4, and CXCR4: dependence on CD4 expression level, gp120 viral origin, conservation of the gp120 COOH- and NH2-termini and V1/V2 and V3 loops, and sensitivity to neutralizing antibodies*. Virology, 1998. **248**(2): p. 394-405.
48. Ugolini, S., et al., *HIV-1 gp120 induces an association between CD4 and the chemokine receptor CXCR4*. J Immunol, 1997. **159**(6): p. 3000-8.
49. Valenzuela, A., et al., *Neutralizing antibodies against the V3 loop of human immunodeficiency virus type 1 gp120 block the CD4-dependent and -independent binding of virus to cells*. J Virol, 1997. **71**(11): p. 8289-98.
50. Reeves, J.D., et al., *The second extracellular loop of CXCR4 is involved in CD4-independent entry of human immunodeficiency virus type 2*. J Gen Virol, 1998. **79** (Pt 7): p. 1793-9.
51. De Clercq, E., *The design of drugs for HIV and HCV*. Nat Rev Drug Discov, 2007. **6**(12): p. 1001-18.
52. Nagasawa, T., H. Kikutani, and T. Kishimoto, *Molecular cloning and structure of a pre-B-cell growth-stimulating factor*. Proc Natl Acad Sci U S A, 1994. **91**(6): p. 2305-9.
53. Crosby, H.A., et al., *Adhesion of human haematopoietic (CD34+) stem cells to human liver compartments is integrin and CD44 dependent and modulated by CXCR3 and CXCR4*. J Hepatol, 2009. **51**(4): p. 734-49.
54. Wei, L., et al., *Stimulation of chondrocyte hypertrophy by chemokine stromal cell-derived factor 1 in the chondro-osseous junction during endochondral bone formation*. Dev Biol, 2010. **341**(1): p. 236-45.
55. Bagri, A., et al., *The chemokine SDF1 regulates migration of dentate granule cells*. Development, 2002. **129**(18): p. 4249-60.
56. Lu, M., E. Grove, and R. Miller, *Abnormal development of the hippocampal dentate gyrus in mice lacking the CXCR4 chemokine receptor*. Proc Natl Acad Sci U S A., 2002. **99**(10): p. 7090-5.
57. McGrath, K., et al., *Embryonic expression and function of the chemokine SDF-1 and its receptor, CXCR4*. Developmental Biology, 1999. **213**(2): p. 442-56.

58. Boldajipour, B., et al., *Control of chemokine-guided cell migration by ligand sequestration*. Cell, 2008. **132**(3): p. 463-73.
59. Ma, Q., et al., *Impaired B-lymphopoiesis, myelopoiesis, and derailed cerebellar neuron migration in CXCR4- and SDF-1-deficient mice*. Proc Natl Acad Sci U S A, 1998. **95**(16): p. 9448-53.
60. Belmadani, A., et al., *The chemokine stromal cell-derived factor-1 regulates the migration of sensory neuron progenitors*. J Neurosci, 2005. **25**(16): p. 3995-4003.
61. Bagri, A., et al., *The chemokine SDF1 regulates migration of dentate granule cells*. Development, 2002. **129**(18): p. 4249-60.
62. Li, G., et al., *Regional distribution of cortical interneurons and development of inhibitory tone are regulated by Cxcl12/Cxcr4 signaling*. J Neurosci, 2008. **28**(5): p. 1085-98.
63. Rehim, R., et al., *A novel role of CXCR4 and SDF-1 during migration of cloacal muscle precursors*. Dev Dyn, 2010. **239**(6): p. 1622-31.
64. Cohen, J., *Investigators detail HIV's fatal handshake*. Science, 1996. **274**(5287): p. 502.
65. Ara, T., et al., *The role of CXCL12 in the organ-specific process of artery formation*. Blood, 2005. **105**(8): p. 3155-61.
66. Heidemann, J., et al., *Mucosal angiogenesis regulation by CXCR4 and its ligand CXCL12 expressed by human intestinal microvascular endothelial cells*. Am J Physiol Gastrointest Liver Physiol, 2004. **286**(6): p. G1059-68.
67. Smith, J.M., et al., *CXCL12 activation of CXCR4 regulates mucosal host defense through stimulation of epithelial cell migration and promotion of intestinal barrier integrity*. Am J Physiol Gastrointest Liver Physiol, 2005. **288**(2): p. G316-26.
68. Gulino, A., *WHIM syndrome: a genetic disorder of leukocyte trafficking*. Curr Opin Allergy Clin Immunol., 2003. **3**(6): p. 443-50.
69. McDermott, D.H., et al., *AMD3100 is a Potent Antagonist at CXCR4(R334X) , a Hyperfunctional Mutant Chemokine Receptor and Cause of WHIM Syndrome*. J Cell Mol Med, 2010.
70. Sugiyama, T., et al., *Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches*. Immunity, 2006. **25**(6): p. 977-88.
71. Nagib, P.R., et al., *Thymic microenvironmental alterations in experimentally induced diabetes*. Immunobiology, 2010.
72. Tarlinton, D., et al., *Plasma cell differentiation and survival*. Curr Opin Immunol, 2008. **20**(2): p. 162-9.
73. Mentzer, W.C., Jr., et al., *An unusual form of chronic neutropenia in a father and daughter with hypogammaglobulinaemia*. Br J Haematol, 1977. **36**(3): p. 313-22.
74. Weisel, K.C., et al., *Modulation of CXC chemokine receptor expression and function in human neutrophils during aging in vitro suggests a role in their clearance from circulation*. Mediators Inflamm, 2009. **2009**: p. 790174.
75. Ebert, L.M. and S.R. McColl, *Coregulation of CXC chemokine receptor and CD4 expression on T lymphocytes during allogeneic activation*. J Immunol, 2001. **166**(8): p. 4870-8.
76. Ngai, J., et al., *Interplay between the heterotrimeric G-protein subunits Gα₁₂ and Gα₁₃ sets the threshold for chemotaxis and TCR activation*. BMC Immunol, 2009. **10**: p. 27.
77. Petit, I., et al., *G-CSF induces stem cell mobilization by decreasing bone marrow SDF-1 and up-regulating CXCR4*. Nat Immunol, 2002. **3**(7): p. 687-94.
78. Semerad, C.L., et al., *G-CSF potently inhibits osteoblast activity and CXCL12 mRNA expression in the bone marrow*. Blood, 2005. **106**(9): p. 3020-7.

79. Delano, M.J., et al., *Neutrophil Mobilization from the Bone Marrow during Polymicrobial Sepsis Is Dependent on CXCL12 Signaling*. J Immunol, 2011. **187**(2): p. 911-8.
80. DiPersio, J.F., et al., *Plerixafor*. Nat Rev Drug Discov, 2009. **8**(2): p. 105-6.
81. Dar, A., et al., *Rapid mobilization of hematopoietic progenitors by AMD3100 and catecholamines is mediated by CXCR4-dependent SDF-1 release from bone marrow stromal cells*. Leukemia, 2011.
82. Broxmeyer, H.E., et al., *Rapid mobilization of murine and human hematopoietic stem and progenitor cells with AMD3100, a CXCR4 antagonist*. J Exp Med, 2005. **201**(8): p. 1307-18.
83. Vassart, G. and S. Costagliola, *G protein-coupled receptors: mutations and endocrine diseases*. Nat Rev Endocrinol, 2011.
84. Kolakowski, L.F., Jr., *GCRDb: a G-protein-coupled receptor database*. Receptors Channels, 1994. **2**(1): p. 1-7.
85. Davies, M.N., et al., *Proteomic applications of automated GPCR classification*. Proteomics, 2007. **7**(16): p. 2800-14.
86. Doranz, B., et al., *Identification of CXCR4 domains that support coreceptor and chemokine receptor functions*. J Virol., 1999. **4**(73): p. 2752-61.
87. Joost, P. and A. Methner, *Phylogenetic analysis of 277 human G-protein-coupled receptors as a tool for the prediction of orphan receptor ligands*. Genome Biol, 2002. **3**(11): p. RESEARCH0063.
88. Ali, S. and G. Lazennec, *Chemokines: novel targets for breast cancer metastasis*. Cancer Metastasis Rev, 2007. **26**(3-4): p. 401-20.
89. Vindrieux, D., P. Escobar, and G. Lazennec, *Emerging roles of chemokines in prostate cancer*. Endocr Relat Cancer, 2009. **16**(3): p. 663-73.
90. Premack, B.A. and T.J. Schall, *Chemokine receptors: gateways to inflammation and infection*. Nat Med, 1996. **2**(11): p. 1174-8.
91. Wang, D. and A. Richmond, *Nuclear factor-kappa b activation by the cxc chemokine melanoma growth- stimulatory activity/growth-regulated protein involves the mekk1/p38 mitogen-activated protein kinase pathway*. J Biol Chem, 2001. **276**(5): p. 3650-9.
92. Peiper, S.C., et al., *The Duffy antigen/receptor for chemokines (DARC) is expressed in endothelial cells of Duffy negative individuals who lack the erythrocyte receptor*. J Exp Med, 1995. **181**(4): p. 1311-7.
93. Rosenbaum, D.M., S.G. Rasmussen, and B.K. Kobilka, *The structure and function of G-protein-coupled receptors*. Nature, 2009. **459**(7245): p. 356-63.
94. Hofmann, K.P., et al., *A G protein-coupled receptor at work: the rhodopsin model*. Trends Biochem Sci, 2009. **34**(11): p. 540-52.
95. Scheerer, P., et al., *Crystal structure of opsin in its G-protein-interacting conformation*. Nature, 2008. **455**(7212): p. 497-502.
96. Ahuja, S. and S.O. Smith, *Multiple switches in G protein-coupled receptor activation*. Trends Pharmacol Sci, 2009. **30**(9): p. 494-502.
97. Fanelli, F. and A. Felling, *Dimerization and ligand binding affect the structure network of A(2A) adenosine receptor*. Biochim Biophys Acta, 2010.
98. Lambert, N.A., *GPCR dimers fall apart*. Sci Signal, 2010. **3**(115): p. pe12.
99. Kim, H., et al., *Identification of specific transmembrane residues and ligand-induced interface changes involved in homo-dimer formation of a yeast G protein-coupled receptor*. Biochemistry, 2009. **48**(46): p. 10976-87.
100. Tuteja, N., *Signaling through G protein coupled receptors*. Plant Signal Behav, 2009. **4**(10): p. 942-7.

101. Ravindran, A., P.R. Joseph, and K. Rajarathnam, *Structural basis for differential binding of the interleukin-8 monomer and dimer to the CXCR1 N-domain: role of coupled interactions and dynamics*. Biochemistry, 2009. **48**(37): p. 8795-805.
102. Lezoualc'h, F., R. Jockers, and I. Berque-Bestel, *Multivalent-based drug design applied to serotonin 5-HT(4) receptor oligomers*. Curr Pharm Des, 2009. **15**(6): p. 719-29.
103. Harding, P.J., et al., *Constitutive dimerization of the G-protein coupled receptor, neurotensin receptor 1, reconstituted into phospholipid bilayers*. Biophys J, 2009. **96**(3): p. 964-73.
104. Palczewski, K., et al., *Crystal structure of rhodopsin: A G protein-coupled receptor*. Science, 2000. **289**(5480): p. 739-45.
105. Okada, T., K. Takeda, and T. Kouyama, *Highly selective separation of rhodopsin from bovine rod outer segment membranes using combination of divalent cation and alkyl(thio)glucoside*. Photochem Photobiol, 1998. **67**(5): p. 495-9.
106. Okada, T., et al., *X-Ray diffraction analysis of three-dimensional crystals of bovine rhodopsin obtained from mixed micelles*. J Struct Biol, 2000. **130**(1): p. 73-80.
107. Kobilka, B. and G.F. Schertler, *New G-protein-coupled receptor crystal structures: insights and limitations*. Trends Pharmacol Sci, 2008. **29**(2): p. 79-83.
108. Rasmussen, S.G., et al., *Crystal structure of the human beta2 adrenergic G-protein-coupled receptor*. Nature, 2007. **450**(7168): p. 383-7.
109. Cherezov, V., et al., *High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor*. Science, 2007. **318**(5854): p. 1258-65.
110. Rosenbaum, D.M., et al., *GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function*. Science, 2007. **318**(5854): p. 1266-73.
111. Jaakola, V.P., et al., *The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist*. Science, 2008. **322**(5905): p. 1211-7.
112. Tebben, A.J. and D.M. Schnur, *Beyond rhodopsin: G protein-coupled receptor structure and modeling incorporating the beta2-adrenergic and adenosine A(2A) crystal structures*. Methods Mol Biol, 2011. **672**: p. 359-86.
113. Lodowski, D.T., T.E. Angel, and K. Palczewski, *Comparative analysis of GPCR crystal structures*. Photochem Photobiol, 2009. **85**(2): p. 425-30.
114. Deupi, X. and J. Standfuss, *Structural insights into agonist-induced activation of G-protein-coupled receptors*. Curr Opin Struct Biol, 2011.
115. Piirainen, H., et al., *Structural features of adenosine receptors: from crystal to function*. Biochim Biophys Acta, 2011. **1808**(5): p. 1233-44.
116. Dealwis, C., et al., *Crystal structure of chemically synthesized [N33A] stromal cell-derived factor 1alpha, a potent ligand for the HIV-1 "fusin" coreceptor*. Proc Natl Acad Sci U S A, 1998. **95**(12): p. 6941-6.
117. Veldkamp, C.T., et al., *The monomer-dimer equilibrium of stromal cell-derived factor-1 (CXCL 12) is altered by pH, phosphate, sulfate, and heparin*. Protein Sci, 2005. **14**(4): p. 1071-81.
118. Baryshnikova, O.K. and B.D. Sykes, *Backbone dynamics of SDF-1alpha determined by NMR: interpretation in the presence of monomer-dimer equilibrium*. Protein Sci, 2006. **15**(11): p. 2568-78.
119. Davis, D.A., et al., *Identification of carboxypeptidase N as an enzyme responsible for C-terminal cleavage of stromal cell-derived factor-1alpha in the circulation*. Blood, 2005. **105**(12): p. 4561-8.

120. Tan, Y., et al., *Cloning and characterizing mutated human stromal cell-derived factor-1 (SDF-1): C-terminal alpha-helix of SDF-1alpha plays a critical role in CXCR4 activation and signaling, but not in CXCR4 binding affinity*. Exp Hematol, 2006. **34**(11): p. 1553-62.
121. Crump, M.P., et al., *Solution structure and basis for functional activity of stromal cell-derived factor-1; dissociation of CXCR4 activation from binding and inhibition of HIV-1*. EMBO J, 1997. **16**(23): p. 6996-7007.
122. Gupta, S.K., et al., *Pharmacological evidence for complex and multiple site interaction of CXCR4 with SDF-1alpha: implications for development of selective CXCR4 antagonists*. Immunol Lett, 2001. **78**(1): p. 29-34.
123. Gozansky, E.K., et al., *Mapping the binding of the N-terminal extracellular tail of the CXCR4 receptor to stromal cell-derived factor-1alpha*. J Mol Biol, 2005. **345**(4): p. 651-8.
124. Brelot, A., et al., *Identification of residues of CXCR4 critical for human immunodeficiency virus coreceptor and chemokine receptor activities*. J Biol Chem, 2000. **275**(31): p. 23736-44.
125. Ohnishi, Y., et al., *Crystal structure of recombinant native SDF-1alpha with additional mutagenesis studies: an attempt at a more comprehensive interpretation of accumulated structure-activity relationship data*. J Interferon Cytokine Res, 2000. **20**(8): p. 691-700.
126. Kofuku, Y., et al., *Structural basis of the interaction between chemokine stromal cell-derived factor-1/CXCL12 and its G-protein-coupled receptor CXCR4*. J Biol Chem, 2009. **284**(50): p. 35240-50.
127. Zhang, W.B., et al., *A point mutation that confers constitutive activity to CXCR4 reveals that T140 is an inverse agonist and that AMD3100 and ALX40-4C are weak partial agonists*. J Biol Chem, 2002. **277**(27): p. 24515-21.
128. Thelen, M. and S. Thelen, *CXCR7, CXCR4 and CXCL12: an eccentric trio?* J Neuroimmunol, 2008. **198**(1-2): p. 9-13.
129. Lopez-Illasaca, M., *Signaling from G-protein-coupled receptors to mitogen-activated protein (MAP)-kinase cascades*. Biochem Pharmacol, 1998. **56**(3): p. 269-77.
130. Zhan, W., et al., *Discovery of small molecule CXCR4 antagonists*. J Med Chem, 2007. **50**(23): p. 5655-64.
131. Luo, Y., et al., *SDF1alpha/CXCR4 signaling, via ERKs and the transcription factor Egr1, induces expression of a 67-kDa form of glutamic acid decarboxylase in embryonic hippocampal neurons*. J Biol Chem, 2008. **283**(36): p. 24789-800.
132. Hajishengallis, G., et al., *Pathogen induction of CXCR4/TLR2 cross-talk impairs host defense function*. Proc Natl Acad Sci U S A, 2008. **105**(36): p. 13532-7.
133. Ganju, R.K., et al., *The alpha-chemokine, stromal cell-derived factor-1alpha, binds to the transmembrane G-protein-coupled CXCR-4 receptor and activates multiple signal transduction pathways*. J Biol Chem, 1998. **273**(36): p. 23169-75.
134. Curnock, A.P., M.K. Logan, and S.G. Ward, *Chemokine signalling: pivoting around multiple phosphoinositide 3-kinases*. Immunology, 2002. **105**(2): p. 125-36.
135. Floridi, F., et al., *Signalling pathways involved in the chemotactic activity of CXCL12 in cultured rat cerebellar neurons and CHP100 neuroepithelioma cells*. J Neuroimmunol, 2003. **135**(1-2): p. 38-46.
136. Petty, J.M., et al., *Crosstalk between CXCR4/stromal derived factor-1 and VLA-4/VCAM-1 pathways regulates neutrophil retention in the bone marrow*. J Immunol, 2009. **182**(1): p. 604-12.

137. Holland, J.D., et al., *Differential functional activation of chemokine receptor CXCR4 is mediated by G proteins in breast cancer cells*. Cancer Res, 2006. **66**(8): p. 4117-24.
138. Ateeq, B., et al., *Pharmacological inhibition of DNA methylation induces proinvasive and prometastatic genes in vitro and in vivo*. Neoplasia, 2008. **10**(3): p. 266-78.
139. Shim, H., et al., *Lower expression of CXCR4 in lymph node metastases than in primary breast cancers: potential regulation by ligand-dependent degradation and HIF-1alpha*. Biochem Biophys Res Commun, 2006. **346**(1): p. 252-8.
140. De La Luz Sierra, M., et al., *Transcription factor Gfi-1 induced by G-CSF is a negative regulator of CXCR4 in myeloid cells*. Blood, 2007. **110**(7): p. 2276-85.
141. Roland, J., et al., *Role of the intracellular domains of CXCR4 in SDF-1-mediated signaling*. Blood, 2003. **101**(2): p. 399-406.
142. Cheng, Z.J., et al., *beta-arrestin differentially regulates the chemokine receptor CXCR4-mediated signaling and receptor internalization, and this implicates multiple interaction sites between beta-arrestin and CXCR4*. J Biol Chem, 2000. **275**(4): p. 2479-85.
143. Sun, Y., et al., *Beta-arrestin2 is critically involved in CXCR4-mediated chemotaxis, and this is mediated by its enhancement of p38 MAPK activation*. J Biol Chem, 2002. **277**(51): p. 49212-9.
144. Bhandari, D., et al., *Arrestin-2 interacts with the ubiquitin-protein isopeptide ligase atrophin-interacting protein 4 and mediates endosomal sorting of the chemokine receptor CXCR4*. J Biol Chem, 2007. **282**(51): p. 36971-9.
145. Malik, R. and A. Marchese, *Arrestin-2 interacts with the endosomal sorting complex required for transport machinery to modulate endosomal sorting of CXCR4*. Mol Biol Cell, 2010. **21**(14): p. 2529-41.
146. McCormick, P.J., et al., *Impaired recruitment of Grk6 and beta-Arrestin 2 causes delayed internalization and desensitization of a WHIM syndrome-associated CXCR4 mutant receptor*. PLoS One, 2009. **4**(12): p. e8102.
147. Balabanian, K., et al., *WHIM syndromes with different genetic anomalies are accounted for by impaired CXCR4 desensitization to CXCL12*. Blood, 2005. **105**(6): p. 2449-57.
148. Gulino, A.V., et al., *Altered leukocyte response to CXCL12 in patients with warts hypogammaglobulinemia, infections, myelokathexis (WHIM) syndrome*. Blood, 2004. **104**(2): p. 444-52.
149. Jala, V.R., W.H. Shao, and B. Haribabu, *Phosphorylation-independent beta-arrestin translocation and internalization of leukotriene B4 receptors*. J Biol Chem, 2005. **280**(6): p. 4880-7.
150. Jala, V.R. and B. Haribabu, *Real-time analysis of G protein-coupled receptor signaling in live cells*. Methods Mol Biol, 2006. **332**: p. 159-65.
151. Basu, S., et al., *Critical role for polar residues in coupling leukotriene B4 binding to signal transduction in BLT1*. J Biol Chem, 2007. **282**(13): p. 10005-17.
152. Brown, S.L., et al., *Activation and regulation of platelet-activating factor receptor: role of G(i) and G(q) in receptor-mediated chemotactic, cytotoxic, and cross-regulatory signals*. J Immunol, 2006. **177**(5): p. 3242-9.
153. Lee, H.J., et al., *Chemokine receptor CXCR4 expression, function, and clinical implications in gastric cancer*. Int J Oncol, 2009. **34**(2): p. 473-80.
154. Juarez, J., et al., *Effects of inhibitors of the chemokine receptor CXCR4 on acute lymphoblastic leukemia cells in vitro*. Leukemia, 2003. **17**(7): p. 1294-300.
155. Liesveld, J.L., et al., *Effects of AMD3100 on transmigration and survival of acute myelogenous leukemia cells*. Leuk Res, 2007. **31**(11): p. 1553-63.

156. Scotton, C.J., et al., *Multiple actions of the chemokine CXCL12 on epithelial tumor cells in human ovarian cancer*. *Cancer Res*, 2002. **62**(20): p. 5930-8.
157. Rubin, J.B., et al., *A small-molecule antagonist of CXCR4 inhibits intracranial growth of primary brain tumors*. *Proc Natl Acad Sci U S A*, 2003. **100**(23): p. 13513-8.
158. De Falco, V., et al., *Biological role and potential therapeutic targeting of the chemokine receptor CXCR4 in undifferentiated thyroid cancer*. *Cancer Res*, 2007. **67**(24): p. 11821-9.
159. Li, J.K., et al., *Inhibition of CXCR4 activity with AMD3100 decreases invasion of human colorectal cancer cells in vitro*. *World J Gastroenterol*, 2008. **14**(15): p. 2308-13.
160. Hendrix, C.W., et al., *Safety, pharmacokinetics, and antiviral activity of AMD3100, a selective CXCR4 receptor inhibitor, in HIV-1 infection*. *J Acquir Immune Defic Syndr*, 2004. **37**(2): p. 1253-62.
161. Siwers, B., et al., *Initial clinical trial based on biochemical methodology of zimelidine (a serotonin uptake inhibitor) in depressed patients*. *Clin Pharmacol Ther*, 1977. **21**(2): p. 194-200.
162. Fagius, J., et al., *Guillain-Barre syndrome following zimeldine treatment*. *J Neurol Neurosurg Psychiatry*, 1985. **48**(1): p. 65-9.
163. Bjerkenstedt, L., et al., *Clinical and biochemical effects of citalopram, a selective 5-HT reuptake inhibitor--a dose-response study in depressed patients*. *Psychopharmacology (Berl)*, 1985. **87**(3): p. 253-9.
164. Bjerkenstedt, L., et al., *Relationship between clinical effects, serum drug concentration and serotonin uptake inhibition in depressed patients treated with citalopram. A double-blind comparison of three dose levels*. *Eur J Clin Pharmacol*, 1985. **28**(5): p. 553-7.
165. Beving, H., et al., *The effects of citalopram (Lu 10-171) on the serotonin (5-HT) uptake kinetics in platelets from endogenously depressed patients*. *J Neural Transm*, 1985. **61**(1-2): p. 95-104.
166. Goldberg, H.L. and R. Finnerty, *An open-label, variable-dose study of WY-45,030 (venlafexine) in depressed outpatients*. *Psychopharmacol Bull*, 1988. **24**(1): p. 198-9.
167. Schweizer, E., et al., *An open-label, dose-finding study of WY-45,030, a novel bicyclic antidepressant*. *Psychopharmacol Bull*, 1988. **24**(1): p. 195-7.
168. Muth, E.A., et al., *Antidepressant biochemical profile of the novel bicyclic compound Wy-45,030, an ethyl cyclohexanol derivative*. *Biochem Pharmacol*, 1986. **35**(24): p. 4493-7.
169. Stone, N.D., et al., *Multiple-dose escalation study of the safety, pharmacokinetics, and biologic activity of oral AMD070, a selective CXCR4 receptor inhibitor, in human subjects*. *Antimicrob Agents Chemother*, 2007. **51**(7): p. 2351-8.
170. Gudmundsson, K.S., et al., *Amine substituted N-(1H-benzimidazol-2-ylmethyl)-5,6,7,8-tetrahydro-8-quinolinamines as CXCR4 antagonists with potent activity against HIV-1*. *Bioorg Med Chem Lett*, 2009. **19**(17): p. 5048-52.
171. Gudmundsson, K.S., et al., *Imidazopyridine-5,6,7,8-tetrahydro-8-quinolinamine derivatives with potent activity against HIV-1*. *Bioorg Med Chem Lett*, 2009. **19**(22): p. 6399-403.
172. Hatse, S., et al., *Mutations at the CXCR4 interaction sites for AMD3100 influence anti-CXCR4 antibody binding and HIV-1 entry*. *FEBS Lett*, 2003. **546**(2-3): p. 300-6.

173. Iwasaki, Y., et al., *Efficient inhibition of SDF-1alpha-mediated chemotaxis and HIV-1 infection by novel CXCR4 antagonists*. Cancer Sci, 2009. **100**(4): p. 778-81.
174. Kawatkar, S.P., et al., *Computational analysis of the structural mechanism of inhibition of chemokine receptor CXCR4 by small molecule antagonists*. Exp Biol Med (Maywood), 2011.
175. Tamamura, H., et al., *Development of specific CXCR4 inhibitors possessing high selectivity indexes as well as complete stability in serum based on an anti-HIV peptide T140*. Bioorg Med Chem Lett, 2001. **11**(14): p. 1897-902.
176. Kim, S.Y., et al., *Inhibition of the CXCR4/CXCL12 chemokine pathway reduces the development of murine pulmonary metastases*. Clin Exp Metastasis, 2008. **25**(3): p. 201-11.
177. Richert, M.M., et al., *Inhibition of CXCR4 by CTCE-9908 inhibits breast cancer metastasis to lung and bone*. Oncol Rep, 2009. **21**(3): p. 761-7.
178. Hassan, S., et al., *CXCR4 peptide antagonist inhibits primary breast tumor growth, metastasis and enhances the efficacy of anti-VEGF treatment or docetaxel in a transgenic mouse model*. Int J Cancer, 2011. **129**(1): p. 225-32.
179. Kwong, J., et al., *An antagonist of the chemokine receptor CXCR4 induces mitotic catastrophe in ovarian cancer cells*. Mol Cancer Ther, 2009. **8**(7): p. 1893-905.
180. Porvasnik, S., et al., *Effects of CXCR4 antagonist CTCE-9908 on prostate tumor growth*. Prostate, 2009. **69**(13): p. 1460-9.
181. Jorgensen, W.L., *The many roles of computation in drug discovery*. Science, 2004. **303**(5665): p. 1813-8.
182. Miyata, T., et al., *Antimicrobial peptides, isolated from horseshoe crab hemocytes, tachyplesin II, and polyphemusins I and II: chemical structures and biological activity*. J Biochem, 1989. **106**(4): p. 663-8.
183. Nakashima, H., et al., *Anti-human immunodeficiency virus activity of a novel synthetic peptide, T22 ([Tyr-5,12, Lys-7]polyphemusin II): a possible inhibitor of virus-cell fusion*. Antimicrob Agents Chemother, 1992. **36**(6): p. 1249-55.
184. Weeks, B.S., et al., *Lymphocytes and promonocytes attach to the synthetic [Tyr5,12, Lys7]- polyphemusin II peptide*. Biochem Biophys Res Commun, 1994. **202**(1): p. 470-5.
185. Weeks, B.S., et al., *The synthetic [Tyr5,12,Lys7]-polyphemusin II peptide (T22) binds to the CD4 cell surface molecule*. Biochem Biophys Res Commun, 1995. **215**(2): p. 626-31.
186. Tamamura, H., et al., *An anti-HIV peptide, T22, forms a highly active complex with Zn(II)*. Biochem Biophys Res Commun, 1996. **229**(2): p. 648-52.
187. Tamamura, H., et al., *Analysis of the interaction of an anti-HIV peptide, T22 ([Tyr5, 12, Lys7]-polyphemusin II), with gp120 and CD4 by surface plasmon resonance*. Biochim Biophys Acta, 1996. **1298**(1): p. 37-44.
188. Tamamura, H., et al., *Interaction of an anti-HIV peptide, T22, with gp120 and CD4*. Biochem Biophys Res Commun, 1996. **219**(2): p. 555-9.
189. Murakami, T., et al., *A small molecule CXCR4 inhibitor that blocks T cell line-tropic HIV-1 infection*. J Exp Med, 1997. **186**(8): p. 1389-93.
190. Tamamura, H., et al., *Structure-activity relationships of an anti-HIV peptide, T22*. Biochem Biophys Res Commun, 1994. **205**(3): p. 1729-35.
191. Otaka, A., et al., *Molecular parameters for the anti-human immunodeficiency virus activity of T22 ([Tyr5,12, Lys7]-polyphemusin II)*. Biol Pharm Bull, 1994. **17**(12): p. 1669-72.

192. Tamamura, H., et al., *Pharmacophore identification of a chemokine receptor (CXCR4) antagonist, T22 ([Tyr(5,12),Lys7]-polyphemusin II), which specifically blocks T cell-line-tropic HIV-1 infection*. Bioorg Med Chem, 1998. **6**(7): p. 1033-41.
193. Tamamura, H., et al., *Downsizing of an HIV-cell fusion inhibitor, T22 ([Tyr5,12, Lys7]-polyphemusin II), with the maintenance of anti-HIV activity and solution structure*. Bioorg Med Chem, 1998. **6**(4): p. 473-9.
194. Tamamura, H., et al., *Effective lowly cytotoxic analogs of an HIV-cell fusion inhibitor, T22 ([Tyr5,12, Lys7]-polyphemusin II)*. Bioorg Med Chem, 1998. **6**(2): p. 231-8.
195. Tamamura, H., et al., *A low-molecular-weight inhibitor against the chemokine receptor CXCR4: a strong anti-HIV peptide T140*. Biochem Biophys Res Commun, 1998. **253**(3): p. 877-82.
196. Tamamura, H., et al., *Pharmacophore identification of a specific CXCR4 inhibitor, T140, leads to development of effective anti-HIV agents with very high selectivity indexes*. Bioorg Med Chem Lett, 2000. **10**(23): p. 2633-7.
197. Tamamura, H., et al., *Conformational study of a highly specific CXCR4 inhibitor, T140, disclosing the close proximity of its intrinsic pharmacophores associated with strong anti-HIV activity*. Bioorg Med Chem Lett, 2001. **11**(3): p. 359-62.
198. Tamamura, H., et al., *Synthesis and evaluation of pseudopeptide analogues of a specific CXCR4 inhibitor, T140: the insertion of an (E)-alkene dipeptide isostere into the betall'-turn moiety*. Bioorg Med Chem Lett, 2002. **12**(6): p. 923-8.
199. Tamamura, H., et al., *Certification of the critical importance of L-3-(2-naphthyl)alanine at position 3 of a specific CXCR4 inhibitor, T140, leads to an exploratory performance of its downsizing study*. Bioorg Med Chem, 2002. **10**(5): p. 1417-26.
200. Trent, J.O., et al., *Lipid bilayer simulations of CXCR4 with inverse agonists and weak partial agonists*. J Biol Chem, 2003. **278**(47): p. 47136-44.
201. Tamamura, H., et al., *Synthesis of potent CXCR4 inhibitors possessing low cytotoxicity and improved biostability based on T140 derivatives*. Org Biomol Chem, 2003. **1**(21): p. 3656-62.
202. Tamamura, H., et al., *Enhancement of the T140-based pharmacophores leads to the development of more potent and bio-stable CXCR4 antagonists*. Org Biomol Chem, 2003. **1**(21): p. 3663-9.
203. Tamamura, H., et al., *Identification of a CXCR4 antagonist, a T140 analog, as an anti-rheumatoid arthritis agent*. FEBS Lett, 2004. **569**(1-3): p. 99-104.
204. Tamamura, H., et al., *Development of a linear type of low molecular weight CXCR4 antagonists based on T140 analogs*. Org Biomol Chem, 2006. **4**(12): p. 2354-7.
205. Tamamura, H., et al., *Identification of novel low molecular weight CXCR4 antagonists by structural tuning of cyclic tetrapeptide scaffolds*. J Med Chem, 2005. **48**(9): p. 3280-9.
206. Tamamura, H., et al., *Structure-activity relationship studies on CXCR4 antagonists having cyclic pentapeptide scaffolds*. Org Biomol Chem, 2005. **3**(24): p. 4392-4.
207. Boulais, P.E., et al., *Photolabeling identifies transmembrane domain 4 of CXCR4 as a T140 binding site*. Biochem Pharmacol, 2009. **78**(11): p. 1382-90.
208. Irwin, J.J. and B.K. Shoichet, *ZINC--a free database of commercially available compounds for virtual screening*. J Chem Inf Model, 2005. **45**(1): p. 177-82.

209. Gruneberg, S., B. Wendt, and G. Klebe, *Subnanomolar Inhibitors from Computer Screening: A Model Study Using Human Carbonic Anhydrase II*. Angew Chem Int Ed Engl, 2001. **40**(2): p. 389-393.
210. Doman, T.N., et al., *Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B*. J Med Chem, 2002. **45**(11): p. 2213-21.
211. Baxter, C.A., et al., *New approach to molecular docking and its application to virtual screening of chemical databases*. J Chem Inf Comput Sci, 2000. **40**(2): p. 254-62.
212. Tondi, D., et al., *Structure-based discovery and in-parallel optimization of novel competitive inhibitors of thymidylate synthase*. Chem Biol, 1999. **6**(5): p. 319-31.
213. Schapira, M., et al., *Rational discovery of novel nuclear hormone receptor antagonists*. Proc Natl Acad Sci U S A, 2000. **97**(3): p. 1008-13.
214. Perola, E., et al., *Successful virtual screening of a chemical database for farnesyltransferase inhibitor leads*. J Med Chem, 2000. **43**(3): p. 401-8.
215. Hopkins, S.C., R.D. Vale, and I.D. Kuntz, *Inhibitors of kinesin activity from structure-based computer screening*. Biochemistry, 2000. **39**(10): p. 2805-14.
216. Aronov, A.M., et al., *Rational design of selective submicromolar inhibitors of Tritrichomonas foetus hypoxanthine-guanine-xanthine phosphoribosyltransferase*. Biochemistry, 2000. **39**(16): p. 4684-91.
217. Boehm, H.J., et al., *Novel inhibitors of DNA gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimization. A promising alternative to random screening*. J Med Chem, 2000. **43**(14): p. 2664-74.
218. Filikov, A.V., et al., *Identification of ligands for RNA targets via structure-based virtual screening: HIV-1 TAR*. J Comput Aided Mol Des, 2000. **14**(6): p. 593-610.
219. Kraemer, O., et al., *Virtual screening for inhibitors of human aldose reductase*. Proteins, 2004. **55**(4): p. 814-23.
220. Gao, Y., et al., *Rational design and characterization of a Rac GTPase-specific small molecule inhibitor*. Proc Natl Acad Sci U S A, 2004. **101**(20): p. 7618-23.
221. Lyne, P.D., et al., *Identification of compounds with nanomolar binding affinity for checkpoint kinase-1 using knowledge-based virtual screening*. J Med Chem, 2004. **47**(8): p. 1962-8.
222. Westerfors, M., et al., *Structure-based discovery of a new affinity ligand to pancreatic alpha-amylase*. J Mol Recognit, 2003. **16**(6): p. 396-405.
223. Peng, H., et al., *Identification of novel inhibitors of BCR-ABL tyrosine kinase via virtual screening*. Bioorg Med Chem Lett, 2003. **13**(21): p. 3693-9.
224. Vangrevelinghe, E., et al., *Discovery of a potent and selective protein kinase CK2 inhibitor by high-throughput docking*. J Med Chem, 2003. **46**(13): p. 2656-62.
225. Chen, G.S., et al., *Novel lead generation through hypothetical pharmacophore three-dimensional database searching: discovery of isoflavonoids as nonsteroidal inhibitors of rat 5 alpha-reductase*. J Med Chem, 2001. **44**(23): p. 3759-63.
226. Ewing, T.J., et al., *DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases*. J Comput Aided Mol Des, 2001. **15**(5): p. 411-28.
227. Morris, G.M., et al., *Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function*. J. Comp. Chem., 1998. **19**(14): p. 1639-1662.
228. Jones, G., et al., *Development and validation of a genetic algorithm for flexible docking*. J Mol Biol, 1997. **267**(3): p. 727-48.
229. Zsoldos, Z., et al., *eHiTS: a new fast, exhaustive flexible ligand docking system*. J Mol Graph Model, 2007. **26**(1): p. 198-212.

230. Halgren, T.A., et al., *Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening*. J Med Chem, 2004. **47**(7): p. 1750-9.
231. Friesner, R.A., et al., *Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy*. J Med Chem, 2004. **47**(7): p. 1739-49.
232. Wu, G., et al., *Detailed analysis of grid-based molecular docking: A case study of CDOCKER-A CHARMM-based MD docking algorithm*. Journal of Computational Chemistry, 2003. **24**(13): p. 1549-62.
233. Liu, M. and S. Wang, *MCDOCK: a Monte Carlo simulation approach to the molecular docking problem*. J Comput Aided Mol Des, 1999. **13**(5): p. 435-51.
234. Jain, A.N., *Surflex: fully automatic flexible molecular docking using a molecular similarity-based search engine*. J Med Chem, 2003. **46**(4): p. 499-511.
235. Gehlhaar, D.K., et al., *Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming*. Chem Biol, 1995. **2**(5): p. 317-24.
236. Gehlhaar, D.K., et al., *De novo design of enzyme inhibitors by Monte Carlo ligand generation*. J Med Chem, 1995. **38**(3): p. 466-72.
237. Pelletier, D.J., et al., *Evaluation of a published in silico model and construction of a novel Bayesian model for predicting phospholipidosis inducing potential*. J Chem Inf Model, 2007. **47**(3): p. 1196-205.
238. Bohm, H.J., *The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure*. J Comput Aided Mol Des, 1994. **8**(3): p. 243-56.
239. Muegge, I. and Y.C. Martin, *A general and fast scoring function for protein-ligand interactions: a simplified potential approach*. J Med Chem, 1999. **42**(5): p. 791-804.
240. Rarey, M., et al., *A fast flexible docking method using an incremental construction algorithm*. J Mol Biol, 1996. **261**(3): p. 470-89.
241. Venkatachalam, C.M., et al., *LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites*. J Mol Graph Model, 2003. **21**(4): p. 289-307.
242. Kontoyianni, M., L.M. McClellan, and G.S. Sokol, *Evaluation of docking performance: comparative data on docking algorithms*. J Med Chem, 2004. **47**(3): p. 558-65.
243. Lyne, P.D., *Structure-based virtual screening: an overview*. Drug Discov Today, 2002. **7**(20): p. 1047-55.
244. Jain, A.N., *Surflex-Dock 2.1: robust performance from ligand energetic modeling, ring flexibility, and knowledge-based search*. J Comput Aided Mol Des, 2007. **21**(5): p. 281-306.
245. Thomsen, R. and M.H. Christensen, *MolDock: a new technique for high-accuracy molecular docking*. J Med Chem, 2006. **49**(11): p. 3315-21.
246. Uemura, T., et al., *Effect of YM-254890, a specific Galphaq/11 inhibitor, on experimental peripheral arterial disease in rats*. Eur J Pharmacol, 2006. **536**(1-2): p. 154-61.
247. Uemura, T., et al., *Biological properties of a specific Galpha q/11 inhibitor, YM-254890, on platelet functions and thrombus formation under high-shear stress*. Br J Pharmacol, 2006. **148**(1): p. 61-9.
248. Kawasaki, T., et al., *Pharmacological properties of YM-254890, a specific G(alpha)q/11 inhibitor, on thrombosis and neointima formation in mice*. Thromb Haemost, 2005. **94**(1): p. 184-92.

249. Taniguchi, M., et al., *YM-254890 analogues, novel cyclic depsipeptides with Galpha(q/11) inhibitory activity from Chromobacterium sp. QS3666*. Bioorg Med Chem, 2004. **12**(12): p. 3125-33.
250. Kawasaki, T., et al., *Antithrombotic and thrombolytic efficacy of YM-254890, a G q/11 inhibitor, in a rat model of arterial thrombosis*. Thromb Haemost, 2003. **90**(3): p. 406-13.
251. Taniguchi, M., et al., *YM-254890, a novel platelet aggregation inhibitor produced by Chromobacterium sp. QS3666*. J Antibiot (Tokyo), 2003. **56**(4): p. 358-63.
252. Mathews, J.L., A.V. Smrcka, and J.M. Bidlack, *A novel Gbetagamma-subunit inhibitor selectively modulates mu-opioid-dependent antinociception and attenuates acute morphine-induced antinociceptive tolerance and dependence*. J Neurosci, 2008. **28**(47): p. 12183-9.
253. Bonacci, T.M., et al., *Differential targeting of Gbetagamma-subunit signaling with small molecules*. Science, 2006. **312**(5772): p. 443-6.
254. Smrcka, A.V., D.M. Lehmann, and A.L. Dessal, *G protein betagamma subunits as targets for small molecule therapeutic development*. Comb Chem High Throughput Screen, 2008. **11**(5): p. 382-95.
255. Tang, X., et al., *A critical role of Gbetagamma in tumorigenesis and metastasis of breast cancer*. J Biol Chem, 2011. **286**(15): p. 13244-54.
256. Dessal, A.L., et al., *Rational design of a selective covalent modifier of G protein betagamma subunits*. Mol Pharmacol, 2011. **79**(1): p. 24-33.
257. Wu, B., et al., *Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists*. Science, 2010. **330**(6007): p. 1066-71.
258. Thoma, G., et al., *Orally bioavailable isothioureas block function of the chemokine receptor CXCR4 in vitro and in vivo*. J Med Chem, 2008. **51**(24): p. 7915-20.
259. Vandercappellen, J., J. Van Damme, and S. Struyf, *The role of CXC chemokines and their receptors in cancer*. Cancer Lett, 2008. **267**(2): p. 226-44.
260. Arya, M., et al., *Clinical importance and therapeutic implications of the pivotal CXCL12-CXCR4 (chemokine ligand-receptor) interaction in cancer cell migration*. Tumour Biol, 2007. **28**(3): p. 123-31.
261. Ruffini, P.A., et al., *Manipulating the chemokine-chemokine receptor network to treat cancer*. Cancer, 2007. **109**(12): p. 2392-404.
262. Dewan, M.Z., et al., *Stromal cell-derived factor-1 and CXCR4 receptor interaction in tumor growth and metastasis of breast cancer*. Biomed Pharmacother, 2006. **60**(6): p. 273-6.
263. Balkwill, F., *The significance of cancer cell expression of the chemokine receptor CXCR4*. Seminars in Cancer Biology, 2004. **14**(3): p. 171-179.
264. Franco, R., et al., *CXCL12-binding receptors expression in non-small cell lung cancer relates to tumoral microvascular density and CXCR4 positive circulating tumoral cells in lung draining venous blood*. Eur J Cardiothorac Surg, 2011.
265. Otsuka, S., et al., *CXCR4 Overexpression Is Associated with Poor Outcome in Females Diagnosed with Stage IV Non-small Cell Lung Cancer*. J Thorac Oncol, 2011.
266. Wang, J., et al., *Silencing of CXCR4 Blocks Progression of Ovarian Cancer and Depresses Canonical Wnt Signaling Pathway*. Int J Gynecol Cancer, 2011.
267. Hamaguchi, T., et al., *TNF inhibitor suppresses bone metastasis in a breast cancer cell line*. Biochem Biophys Res Commun, 2011. **407**(3): p. 525-30.
268. Wang, Q., et al., *Regulation of VEGF, MMP-9, and metastasis by CXCR4 in a prostate cancer cell line*. Cell Biol Int, 2011.

269. Zhao, B.C., et al., *CXCR4/SDF-1 axis is involved in lymph node metastasis of gastric carcinoma*. World J Gastroenterol, 2011. **17**(19): p. 2389-96.
270. Wang, H., et al., *Expression and potential role of chemokine receptor CXCR4 in human bladder carcinoma cell lines with different metastatic ability*. Mol Med Report, 2011. **4**(3): p. 525-8.
271. Li, T., et al., *The expression of CXCR4, CXCL12 and CXCR7 in malignant pleural mesothelioma*. J Pathol, 2011. **223**(4): p. 519-30.
272. De Clercq, E., et al., *Highly potent and selective inhibition of human immunodeficiency virus by the bicyclam derivative JM3100*. Antimicrob Agents Chemother, 1994. **38**(4): p. 668-74.
273. Schols, D., et al., *Bicyclams, a class of potent anti-HIV agents, are targeted at the HIV coreceptor fusin/CXCR-4*. Antiviral Res, 1997. **35**(3): p. 147-56.
274. Schols, D., et al., *Inhibition of T-tropic HIV strains by selective antagonization of the chemokine receptor CXCR4*. J Exp Med, 1997. **186**(8): p. 1383-8.
275. Donzella, G.A., et al., *AMD3100, a small molecule inhibitor of HIV-1 entry via the CXCR4 co-receptor*. Nat Med, 1998. **4**(1): p. 72-7.
276. Hendrix, C.W., et al., *Pharmacokinetics and safety of AMD-3100, a novel antagonist of the CXCR-4 chemokine receptor, in human volunteers*. Antimicrob Agents Chemother, 2000. **44**(6): p. 1667-73.
277. Ichiyama, K., et al., *A duodenally absorbable CXC chemokine receptor 4 antagonist, KRH-1636, exhibits a potent and selective anti-HIV-1 activity*. Proc Natl Acad Sci U S A, 2003. **100**(7): p. 4185-90.
278. Murakami, T., et al., *The novel CXCR4 antagonist KRH-3955 is an orally bioavailable and extremely potent inhibitor of human immunodeficiency virus type 1 infection: comparative studies with AMD3100*. Antimicrob Agents Chemother, 2009. **53**(7): p. 2940-8.
279. Tarasova, N.I., W.G. Rice, and C.J. Michejda, *Inhibition of G-protein-coupled receptor function by disruption of transmembrane domain interactions*. J Biol Chem, 1999. **274**(49): p. 34911-5.
280. Bar-Nun, S., Y. Shneyour, and J.S. Beckmann, *G-418, an elongation inhibitor of 80 S ribosomes*. Biochim Biophys Acta, 1983. **741**(1): p. 123-7.
281. Pello, O.M., et al., *Ligand stabilization of CXCR4/delta-opioid receptor heterodimers reveals a mechanism for immune response regulation*. Eur J Immunol, 2008. **38**(2): p. 537-49.
282. Vrba, J., et al., *Cytotoxic activity of sanguinarine and dihydrosanguinarine in human promyelocytic leukemia HL-60 cells*. Toxicol In Vitro, 2009. **23**(4): p. 580-8.
283. Asada, M., et al., *Depletion of serotonin and selective inhibition of 2B receptor suppressed tumor angiogenesis by inhibiting endothelial nitric oxide synthase and extracellular signal-regulated kinase 1/2 phosphorylation*. Neoplasia, 2009. **11**(4): p. 408-17.
284. Cornell, W.D., et al., *A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules*. Journal of the American Chemical Society, 1995. **117**(19): p. 5179-5197.
285. Napier, K.B., et al., *CCR5 interactions with the variable 3 loop of gp120*. J Mol Model, 2007. **13**(1): p. 29-41.
286. Berchiche, Y.A., et al., *Direct assessment of CXCR4 mutant conformations reveals complex link between receptor structure and G(alpha)(i) activation*. J Biol Chem, 2007. **282**(8): p. 5111-5.

287. Evans, B.J., et al., *Expression of CXCR4, a G-protein-coupled receptor for CXCL12 in yeast identification of new-generation inverse agonists*. Methods Enzymol, 2009. **460**: p. 399-412.
288. Stenkamp, R.E., D.C. Teller, and K. Palczewski, *Crystal structure of rhodopsin: a G-protein-coupled receptor*. Chembiochem, 2002. **3**(10): p. 963-7.
289. Skerlj, R., et al., *Synthesis and SAR of novel CXCR4 antagonists that are potent inhibitors of T tropic (X4) HIV-1 replication*. Bioorg Med Chem Lett, 2011. **21**(1): p. 262-6.
290. Bodart, V., et al., *Pharmacology of AMD3465: a small molecule antagonist of the chemokine receptor CXCR4*. Biochem Pharmacol, 2009. **78**(8): p. 993-1000.
291. Princen, K., et al., *Inhibition of human immunodeficiency virus replication by a dual CCR5/CXCR4 antagonist*. J Virol, 2004. **78**(23): p. 12996-3006.
292. Righi, E., et al., *CXCL12/CXCR4 blockade induces multimodal anti-tumor effects and prolongs survival of immunocompetent mice with ovarian cancer*. Cancer Res, 2011.
293. Bermudez, D.M., et al., *Inhibition of stromal cell-derived factor-1alpha further impairs diabetic wound healing*. J Vasc Surg, 2011. **53**(3): p. 774-84.
294. Hachet-Haas, M., et al., *Small neutralizing molecules to inhibit actions of the chemokine CXCL12*. J Biol Chem, 2008. **283**(34): p. 23189-99.
295. Gerlach, L.O., et al., *Molecular interactions of cyclam and bicyclam non-peptide antagonists with the CXCR4 chemokine receptor*. J Biol Chem, 2001. **276**(17): p. 14153-60.
296. Rosenkilde, M.M., et al., *Molecular mechanism of AMD3100 antagonism in the CXCR4 receptor: transfer of binding site to the CXCR3 receptor*. J Biol Chem, 2004. **279**(4): p. 3033-41.
297. Rosenkilde, M.M., et al., *Molecular mechanism of action of monocyclam versus bicyclam non-peptide antagonists in the CXCR4 chemokine receptor*. J Biol Chem, 2007. **282**(37): p. 27354-65.
298. Gilchrist, A., et al., *Antagonists of the receptor-G protein interface block Gi-coupled signal transduction*. J Biol Chem, 1998. **273**(24): p. 14912-9.
299. Takasaki, J., et al., *A novel Galphaq/11-selective inhibitor*. J Biol Chem, 2004. **279**(46): p. 47438-45.
300. Nishimura, A., et al., *Structural basis for the specific inhibition of heterotrimeric Gq protein by a small molecule*. Proc Natl Acad Sci U S A, 2010. **107**(31): p. 13666-71.
301. Prevost, G.P., et al., *Anticancer activity of BIM-46174, a new inhibitor of the heterotrimeric Galpha/Gbetagamma protein complex*. Cancer Res, 2006. **66**(18): p. 9227-34.
302. Taylor, C.M., et al., *Using ligand-based virtual screening to allosterically stabilize the activated state of a GPCR*. Chem Biol Drug Des, 2010. **75**(3): p. 325-32.
303. Taylor, C.M., et al., *Modulating G-protein coupled receptor/G-protein signal transduction by small molecules suggested by virtual screening*. J Med Chem, 2008. **51**(17): p. 5297-303.
304. Trotter, A., *Schools Open Doors to 'Grid Computing'*, in Education Week. 2004. p. 8-9.
305. Morgan, D.M., *Tetrazolium (MTT) assay for cellular viability and activity*. Methods Mol Biol, 1998. **79**: p. 179-83.
306. Bates, P.J., et al., *Antiproliferative activity of G-rich oligonucleotides correlates with protein binding*. J. Biol. Chem., 1999. **274**(37): p. 26369-77.
307. Dapic, V., et al., *Antiproliferative activity of G-quartet-forming oligonucleotides with backbone and sugar modifications*. Biochemistry, 2002. **41**(11): p. 3676-85.

308. Dapic, D., et al., *Biophysical and biological properties of quadruplex oligodeoxyribonucleotides*. Nucleic Acids Res, 2003. **31**: p. 2097-2107.
309. Xu, X., et al., *Inhibition of DNA replication and induction of S phase cell cycle arrest by G-rich oligonucleotides*. J Biol Chem, 2001. **276**(46): p. 43221-30.
310. Bates, P.J., et al., *Antiproliferative activity of G-rich oligonucleotides correlates with protein binding*. J Biol Chem, 1999. **274**(37): p. 26369-77.
311. Mitchell, J.L., et al., *Antizyme induction by polyamine analogues as a factor of cell growth inhibition*. Biochem J, 2002. **366**(Pt 2): p. 663-71.
312. Reddy, V.K., et al., *Conformationally restricted analogues of 1N,12N-bisethylspermine: synthesis and growth inhibitory effects on human tumor cell lines*. J Med Chem, 1998. **41**(24): p. 4723-32.
313. Bhattacharya, S., et al., *Ligand-stabilized conformational states of human beta(2) adrenergic receptor: insight into G-protein-coupled receptor activation*. Biophys J, 2008. **94**(6): p. 2027-42.
314. Polgar, T., et al., *Comparative virtual and experimental high-throughput screening for glycogen synthase kinase-3beta inhibitors*. J Med Chem, 2005. **48**(25): p. 7946-59.
315. Zhou, Y., et al., *Discovery of small molecule inhibitors of integrin alphavbeta3 through structure-based virtual screening*. Bioorg Med Chem Lett, 2006. **16**(22): p. 5878-82.
316. Moro, W.B., et al., *Virtual screening to identify lead inhibitors for bacterial NAD synthetase (NADs)*. Bioorg Med Chem Lett, 2009. **19**(7): p. 2001-5.
317. Klein, R.S., et al., *SDF-1 alpha induces chemotaxis and enhances Sonic hedgehog-induced proliferation of cerebellar granule cells*. Development, 2001. **128**(11): p. 1971-81.
318. Smrcka, A.V., et al., *Regulation of polyphosphoinositide-specific phospholipase C activity by purified Gq*. Science, 1991. **251**(4995): p. 804-7.
319. Exton, J.H., *Regulation of phosphoinositide phospholipases by hormones, neurotransmitters, and other agonists linked to G proteins*. Annu Rev Pharmacol Toxicol, 1996. **36**: p. 481-509.
320. Rhee, S.G., *Regulation of phosphoinositide-specific phospholipase C*. Annu Rev Biochem, 2001. **70**: p. 281-312.
321. Schwindinger, W.F. and J.D. Robishaw, *Heterotrimeric G-protein betagamma-dimers in growth and differentiation*. Oncogene, 2001. **20**(13): p. 1653-60.
322. Golebiowski, A., S.R. Klopfenstein, and D.E. Portlock, *Lead compounds discovered from libraries: part 2*. Curr Opin Chem Biol, 2003. **7**(3): p. 308-25.
323. Golebiowski, A., S.R. Klopfenstein, and D.E. Portlock, *Lead compounds discovered from libraries*. Curr Opin Chem Biol, 2001. **5**(3): p. 273-84.
324. Dorsey, B.D., et al., *L-735,524: the design of a potent and orally bioavailable HIV protease inhibitor*. J Med Chem, 1994. **37**(21): p. 3443-51.
325. Kaldor, S.W., et al., *Viracept (nelfinavir mesylate, AG1343): a potent, orally bioavailable inhibitor of HIV-1 protease*. J Med Chem, 1997. **40**(24): p. 3979-85.
326. Hayden, F.G., et al., *Efficacy and safety of the neuraminidase inhibitor zanamivir in the treatment of influenzavirus infections*. GG167 Influenza Study Group. N Engl J Med, 1997. **337**(13): p. 874-80.
327. Kroemer, R.T., et al., *Assessment of docking poses: interactions-based accuracy classification (IBAC) versus crystal structure deviations*. J Chem Inf Comput Sci, 2004. **44**(3): p. 871-81.
328. Clark, R.D., et al., *Consensus scoring for ligand/protein interactions*. J Mol Graph Model, 2002. **20**(4): p. 281-95.

329. Wallach, I., et al., *Normalizing Molecular Docking Rankings using Virtually Generated Decoys*. J Chem Inf Model, 2011.
330. Agostino, M., et al., *Molecular docking of carbohydrate ligands to antibodies: structural validation against crystal structures*. J Chem Inf Model, 2009. **49**(12): p. 2749-60.
331. Xing, L., et al., *Evaluation and application of multiple scoring functions for a virtual screening experiment*. J Comput Aided Mol Des, 2004. **18**(5): p. 333-44.
332. Zaheer ul, H., et al., *Benchmarking docking and scoring protocol for the identification of potential acetylcholinesterase inhibitors*. J Mol Graph Model, 2010. **28**(8): p. 870-82.
333. Keseru, G.M., *A virtual high throughput screen for high affinity cytochrome P450cam substrates. Implications for in silico prediction of drug metabolism*. J Comput Aided Mol Des, 2001. **15**(7): p. 649-57.
334. Dailey, M.M., et al., *Structure-based drug design: from nucleic acid to membrane protein targets*. Exp Mol Pathol, 2009. **86**(3): p. 141-50.
335. Doranz, B.J., et al., *A dual-tropic primary HIV-1 isolate that uses fusin and the beta-chemokine receptors CKR-5, CKR-3, and CKR-2b as fusion cofactors*. Cell, 1996. **85**(7): p. 1149-58.
336. Dragic, T., et al., *Amino-terminal substitutions in the CCR5 coreceptor impair gp120 binding and human immunodeficiency virus type 1 entry*. J Virol, 1998. **72**(1): p. 279-85.
337. Blanpain, C., et al., *Multiple charged and aromatic residues in CCR5 amino-terminal domain are involved in high affinity binding of both chemokines and HIV-1 Env protein*. J Biol Chem, 1999. **274**(49): p. 34719-27.
338. Farzan, M., et al., *Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry*. Cell, 1999. **96**(5): p. 667-76.
339. Navenot, J.M., et al., *Molecular anatomy of CCR5 engagement by physiologic and viral chemokines and HIV-1 envelope glycoproteins: differences in primary structural requirements for RANTES, MIP-1 alpha, and vMIP-II Binding*. J Mol Biol, 2001. **313**(5): p. 1181-93.
340. Urizar, E., et al., *An activation switch in the rhodopsin family of G protein-coupled receptors: the thyrotropin receptor*. J Biol Chem, 2005. **280**(17): p. 17135-41.
341. Jongejan, A., et al., *Linking agonist binding to histamine H1 receptor activation*. Nat Chem Biol, 2005. **1**(2): p. 98-103.
342. Doranz, B.J., et al., *Identification of CXCR4 domains that support coreceptor and chemokine receptor functions*. J Virol, 1999. **73**(4): p. 2752-61.
343. Hubbell, W.L., D.S. Cafiso, and C. Altenbach, *Identifying conformational changes with site-directed spin labeling*. Nat Struct Biol, 2000. **7**(9): p. 735-9.
344. Hong, K. and W.L. Hubbell, *Preparation and properties of phospholipid bilayers containing rhodopsin*. Proc Natl Acad Sci U S A, 1972. **69**(9): p. 2617-21.
345. Farahbakhsh, Z.T., et al., *Mapping light-dependent structural changes in the cytoplasmic loop connecting helices C and D in rhodopsin: a site-directed spin labeling study*. Biochemistry, 1995. **34**(27): p. 8812-9.
346. Langen, R., et al., *Structural features of the C-terminal domain of bovine rhodopsin: a site-directed spin-labeling study*. Biochemistry, 1999. **38**(25): p. 7918-24.
347. Altenbach, C., et al., *Structure and function in rhodopsin: mapping light-dependent changes in distance between residue 65 in helix TM1 and residues in the sequence 306-319 at the cytoplasmic end of helix TM7 and in helix H8*. Biochemistry, 2001. **40**(51): p. 15483-92.

348. Torres, J., T.J. Stevens, and M. Samso, *Membrane proteins: the 'Wild West' of structural biology*. Trends Biochem Sci, 2003. **28**(3): p. 137-44.
349. Bodenshtein, T.M., et al., *Pre-osteoblastic MC3T3-E1 cells promote breast cancer growth in bone in a murine xenograft model*. Chin J Cancer, 2011. **30**(3): p. 189-96.
350. Huang, E.H., et al., *A CXCR4 antagonist CTCE-9908 inhibits primary tumor growth and metastasis of breast cancer*. J Surg Res, 2009. **155**(2): p. 231-6.
351. Liang, Z., et al., *Inhibition of breast cancer metastasis by selective synthetic polypeptide against CXCR4*. Cancer Res, 2004. **64**(12): p. 4302-8.
352. Orsini, M.J., et al., *Trafficking of the HIV coreceptor CXCR4. Role of arrestins and identification of residues in the c-terminal tail that mediate receptor internalization*. J Biol Chem, 1999. **274**(43): p. 31076-86.
353. Amine, A., et al., *Novel anti-metastatic action of cidofovir mediated by inhibition of E6/E7, CXCR4 and Rho/ROCK signaling in HPV tumor cells*. PLoS One, 2009. **4**(3): p. e5018.

CV

Name: Jason Bradley Meier
Address: 2800 Klondike Ln. Louisville, Ky 40218
DOB: Louisville, Ky – November 3, 1980

EDUCATION: B.S. Chemistry
B.S. Biology with subspecialization in Biochemistry.
University of Louisville
1999 to 2004

Honors Program
Thesis: Virtual screening as drug development
targeting iPFK2 and MIF
University of Louisville
1999 to 2004

M.D. Ph. D student and Ph. D. Candidate
Biochemistry and Molecular Biology
University of Louisville
2004 to Present

Employment: Research technician: Dr. Victor Fingar 05/97 to 08/97
Research technician: Dr. Thomas Geoghegan 5/98 to 8/98
Research technician: Dr. Nancy Martin 1999-2001
Research assistant:
James Graham Brown Cancer Center
2006 to Present

Publications:

1. Dailey MM, Hait C, Holt PA, Maguire JM, Meier JB, Miller MC, Petraccone L, Trent JO. Structure-based drug design: from nucleic acid to membrane protein targets. Exp Mol Pathol. 2009 Jun;86(3):141-50.
2. Winner M, Meier J, Zierow S, Rendon BE, Crichlow GV, Riggs R, Bucala R, Leng L, Smith N, Lolis E, Trent JO, Mitchell RA. A novel, macrophage migration inhibitory factor suicide substrate inhibits motility and growth of lung cancer cells. Cancer Res. 2008 Sep 15;68(18):7253-7.
3. Clem B, Telang S, Clem A, Yalcin A, Meier J, Simmons A, Rasku MA, Arumugam S, Dean WL, Eaton J, Lane A, Trent JO, Chesney J. Small-molecule inhibition of 6-phosphofructo-2-kinase activity suppresses glycolytic flux and tumor growth. Mol Cancer Ther. 2008 Jan;7(1):110-20.

Patents:

PCT Application No. PCT/US07/69672: MACROPHAGE MIGRATION
INHIBITORY FACTOR ANTAGONISTS AND METHODS OF USING
SAME

PCT application # PCT/US08/007594:FAMILY OF PFKFB3 INHIBITORS
WITH ANTI-NEOPLASTIC ACTIVITIES

PCT Application No. PCT/US11/31654 : Anti-metastatic compounds
targeting the GPCR CXCR4