

# Progress Toward Rationally Designed Small-Molecule Peptide and Peptidomimetic CXCR4 Antagonists

Jon Våbenø,<sup>\*,a</sup> Bengt Erik Haug<sup>b</sup> & Mette M. Rosenkilde<sup>c</sup>

<sup>a</sup> *Department of Pharmacy, Faculty of Health Sciences, UiT The Arctic University of Norway, Breivika, NO-9037 Tromsø, Norway*

<sup>b</sup> *Department of Chemistry and Centre for Pharmacy, University of Bergen, Allégaten 41, NO-5007 Bergen, Norway*

<sup>c</sup> *Laboratory for Molecular Pharmacology, Department of Neuroscience and Pharmacology, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3, DK-2200 Copenhagen, Denmark*

\*Author for correspondence:

Tel.: +47 77 62 09 09

Fax: +47 77 64 61 51

E-mail: [jon.vabeno@uit.no](mailto:jon.vabeno@uit.no)

## ABSTRACT

Over the last five years, X-ray structures of CXC chemokine receptor 4 (CXCR4) in complex with three different ligands (the small-molecule antagonist IT1t, the polypeptide antagonist CVX15, and the viral chemokine antagonist vMIP-II) have been released. In addition to the inherent scientific value of these specific X-ray structures, they (i) provide a reliable structural foundation for studies of the molecular interactions between CXCR4 and its key peptide ligands (CXCL12 and HIV-1 gp120); and (ii) serve as valuable templates for further development of small-molecule CXCR4 antagonists with therapeutic potential. We here review recent computational studies of the molecular interactions between CXCR4 and its peptide ligands – based on the X-ray structures of CXCR4 – and the current status of small-molecule peptide and peptidomimetic CXCR4 antagonists.

## **DEFINED KEY TERMS** [underlined in main text]

- 1) **Peptidomimetic:** A peptidomimetic is defined by IUPAC as “*a compound containing non-peptidic structural elements that is capable of mimicking or antagonizing the biological action(s) of a natural parent peptide.*” Further, “*a peptidomimetic does no longer have classical peptide characteristics such as enzymatically scissile peptidic bonds*” [1].
- 2) **Isostere:** In the context of this review, an isostere is defined as any functional group or moiety that is included in a peptide sequence as a replacement of an amide bond.
- 3) **Scaffold:** The term scaffold is used for rigid (normally cyclic) structures onto which the functional groups of amino acid side chains can be introduced.
- 4) **Structure-based and ligand-based design:** In structure-based design, the 3D structure of the target is known and guides the design of active compounds. When the 3D structure of the target is unknown, indirect information has to be used in order to design/optimize compounds that bind to the target. This information is normally obtained through SAR studies and pharmacophore modeling, and the overall approach is known as ligand-based design.
- 5) **7TM receptors:** As signalling via G proteins is a common feature for seven-transmembrane domain (7TM) receptors, they are often referred to as G protein-coupled receptors (GPCRs). However, as G protein-independent signalling pathways also exist, e.g. through  $\beta$ -arrestin recruitment, “7TM receptors” is today considered to be a more appropriate name for this receptor superfamily.
- 6) **Polypeptide:** The term polypeptide is only loosely defined by IUPAC-IUB as a peptide with more than 10-20 amino acids [2]. As 10 amino acids correspond to a molecular weight (MW) of approximately 1000, we here define polypeptides as having 10-50 amino acids, which enables differentiation between polypeptides (MW > 1000 Da) and small-molecules (MW < 1000 Da).
- 7) **Small-molecule:** When referring to molecular size, the word “small” will have different meanings in different scientific disciplines. In the field of

medicinal chemistry, the term “small-molecule” typically refers to an organic compound with MW < 1000 Da.

- 8) **Alanine scan:** In order to establish the relative importance of the side chains in a bioactive peptide, a series of analogs where each individual residue is replaced by an Ala residue is synthesized and tested. This is known as an “alanine-scan”; Ala is used because it is non-functionalized and has the same conformational preferences as all non-Gly/Pro residues.
- 9) **Retro-inverso peptides:** In a retro-inverso analog, the N-to-C direction and stereochemistry of the parent peptide are simultaneously changed, which has the potential of resulting in a peptide with overall similar topology with respect to side chain orientation; see reference [3] for a review. However, due to the reversed N-to-C direction, the positioning of the backbone amide carbonyl (H-bond acceptor) and NH (H-bond donor) groups relative to the side chains will obviously be different. Also, the energetically preferred backbone conformations are unlikely to be identical.
- 10) **Peptoid:** A peptoid is defined by IUPAC as “a peptidomimetic that results from the oligomeric assembly of N-substituted glycines” [1]. In a typical peptoid peptidomimetic, the side chain of each residue in the parent peptide is moved from C $\alpha$  to the amide nitrogen, meaning that the N-substituents of the peptoid are the “side chains”.

## 1. INTRODUCTION

The important roles of endogenous peptides in processing (substrates) and signaling (ligands) mean that both proteolytic enzymes and peptide-binding (peptidergic) receptors are attractive targets for peptidomimetic ligands.

Ripka and Rich have classified peptidomimetics into three main types: peptide backbone mimetics (type-I), functional mimetics (type-II), and topographical mimetics (type-III) [4]. In type-I mimetics, an amide bond of the parent peptide is typically replaced with an isostere, e.g. a transition-state isostere as seen for the HIV protease inhibitors. Type-II mimetics are structurally unrelated to the parent peptide, and only mimic (or antagonize) its function, as exemplified by the angiotensin-II receptor antagonists. Type-III mimetics contain the essential functional groups of the parent peptide, with the 3D organization (topography) maintained by a non-peptide template/scaffold instead of the peptide backbone.

While rational design of many peptidomimetic protease inhibitors has benefited from the availability of X-ray structures of the enzyme targets (i.e. structure-based design), design of peptidomimetic ligands for peptidergic 7TM receptors has traditionally relied on ligand-based approaches due to the problems associated with structure determination of membrane-bound targets. However, this situation changed in 2010, when the X-ray structures of the chemokine receptor CXCR4 in complex with the polypeptide antagonist CVX15 and the small-molecule antagonist IT1t were reported (ligand structures are shown in Figure 1) [5]. This represented a milestone not only in the chemokine field, but also in a wider sense, as these were the first experimental structures of a peptidergic 7TM receptor. Very recently, an X-ray structure of CXCR4 in complex with the viral chemokine vMIP-II was also published [6], providing further structural insight into the molecular recognition of large chemokine ligands. Importantly, these structures also provide a reliable structural foundation for studies of the molecular interactions between CXCR4 and other ligands of interest, including small-molecules with therapeutic potential.

The main focus of the present review will be on (i) recent studies of the molecular interactions between CXCR4 and its key peptide ligands (CXCL12 and the HIV-1 surface protein gp120) based on the X-ray structures of CXCR4; and (ii) the current status of small-molecule peptide and peptidomimetic CXCR4 antagonists.

## **2. CXCR4: BIOLOGY AND PHARMACOLOGY**

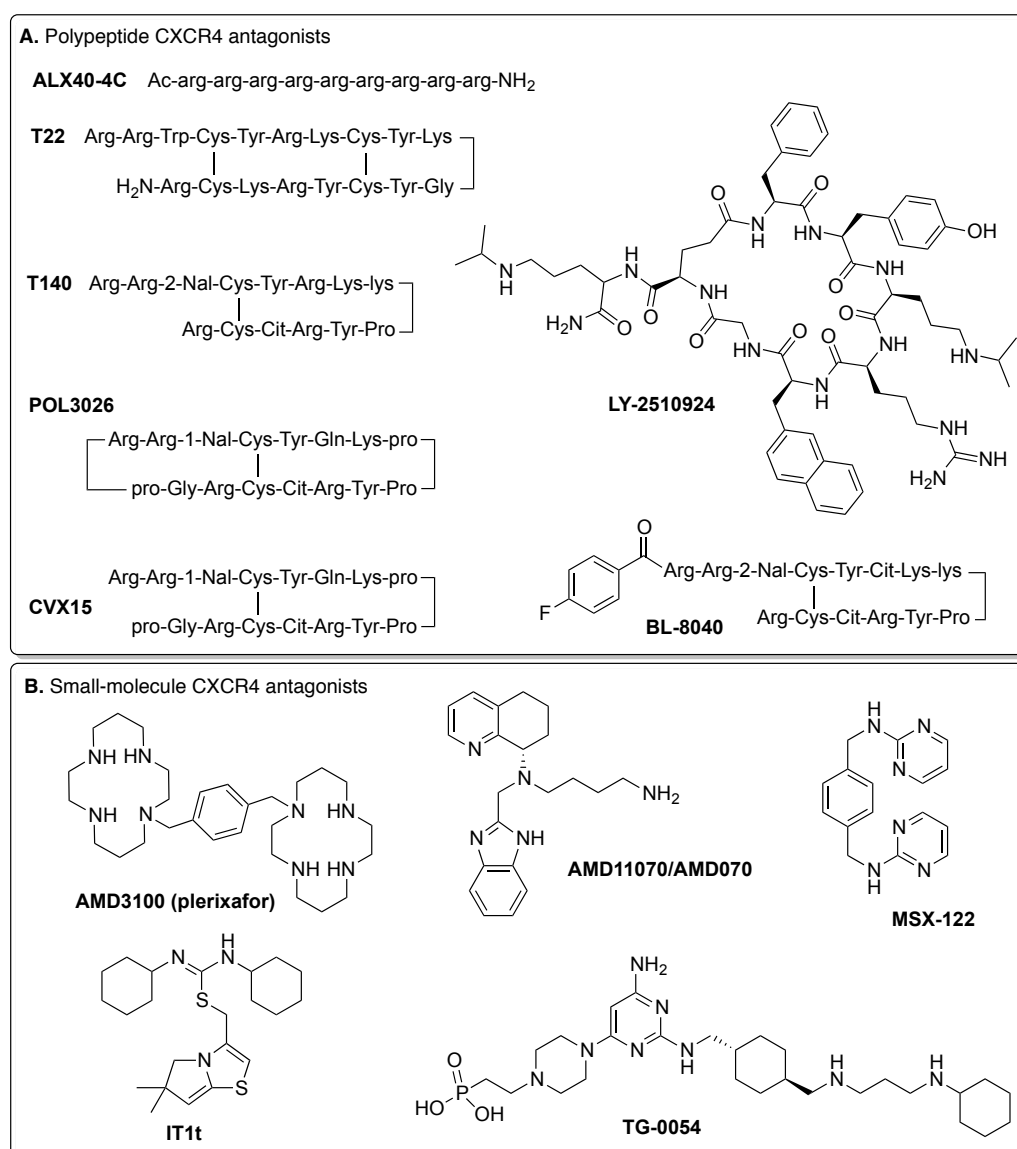
According to the International Union of Basic and Clinical Pharmacology (IUPHAR) around 800 human 7<sup>TM</sup> receptors have been identified [7]. About half of these have sensory functions (olfaction, taste, light and pheromone signaling); the remaining non-sensory receptors (356 in total) are activated by a variety of ligands, including ions, amino acids, monoamines, peptides, lipids, and glycopeptides. Presently, 121 receptors remain orphan, i.e. no endogenous ligand has been identified.

CXCR4 belongs to the subfamily of chemokine receptors, which orchestrate leukocyte migration during homeostasis as well as inflammation [8]. CXCR4 (initially named both LESTR and fusin) was discovered based on its function as a co-receptor for HIV-entry [9-11], and the 68-residue chemokine CXCL12 (initially named SDF-1) was subsequently identified as its endogenous ligand [12]. The monogamous relationship between CXCR4 and CXCL12 is atypical for the otherwise promiscuous chemokine system, which consists of ~25 receptors and >50 ligands [13]. In contrast to most other chemokine receptors, CXCR4 is not only expressed on leukocytes, but also on many other differentiated and non-differentiated cell types outside the hematopoietic compartment, including the brain and the cardiovascular system [14-16]. Consistent with this broad expression pattern, targeted knock-out of either CXCR4 or CXCL12 results in lethality *in utero* [17]. In fact, CXCR4 is the only chemokine receptor essential for life. Furthermore, CXCR4 is expressed on many cancer cells, including breast cancer, ovarian cancers, brain tumors and a variety of hematological cancers, where it has been suggested to influence metastasis as well as tumor growth and angiogenesis [18-23]. The tumor-expression of CXCR4 has been exploited from a

molecular imaging perspective by labeling CXCR4 ligands as PET radiotracers; for recent reviews see references [24-26]. Moreover, the CXCR4:CXCL12 axis controls release of myeloid and lymphoid hematopoietic stem cells from the bone marrow [27], and recent studies suggest that it may also play a role in recruitment of skeletal muscle progenitor cells during myogenesis [28, 29]. Within the immune system, CXCR4 and CXCL12 regulate the migration and maturation of a variety of immune cells (T- and B-lymphocytes, monocytes, macrophages, neutrophils and eosinophils) and are thereby essential for immune surveillance [13]. The broad expression of CXCR4:CXCL12 within (and outside) the immune system implies that novel functional properties of CXCR4 are still to be discovered. One such role is in autoimmunity, and several recent reports suggest that CXCR4 is a biomarker for autoimmunity in e.g. type 1 diabetes [30], autoimmune myopathy [31], and systemic lupus erythematosus [32, 33].

The multiple physiological and pathophysiological roles of CXCR4 have stimulated an intensive search for CXCR4 antagonists. The first CXCR4 antagonist to be tested in the clinic was the N-acetylated nona-D-arginine amide ALX40-4C (Figure 1A). This polycationic peptide was initially designed as an inhibitor of the HIV-1 Tat-TAR interaction, but it was also shown to interfere with HIV entry [34]. Clinical investigations were initiated before the discovery of CCR5 and CXCR4 as the co-receptors of HIV, and it was later shown that ALX40-4C inhibited HIV infection by blocking viral interaction with CXCR4 [35]. ALX40-4C was found to be well tolerated; however, it did not result in a significant reduction in viral load [36]. In 1992 the polyphemusin II-derived 18-mer polypeptide T22 (Figure 1A) was reported to show anti-HIV activity, apparently through inhibition of virus-cell fusion [37]. Following the discovery of the role of CXCR4 in HIV-1 entry four years later [38], T22 was soon shown to be an antagonist for CXCR4 [39]. By the end of the decade, structure-activity relationship (SAR) and downsizing studies of T22 had resulted in the identification of the 14-mer antagonist T140 (Figure 1A) [40], which today is considered as the “prototype” polypeptide CXCR4 antagonist. At this point,

analogs [41] and fragments/dimers [42] of CXCL12 had also been shown to be CXCR4 antagonists.



**Figure 1.** (A) Sequences/structures of selected polypeptide CXCR4 antagonists, and (B) structures of selected small-molecule CXCR4 antagonists (functional mimetics).

The first small-molecule (MW < 1000 Da) non-peptide CXCR4 antagonists were also developed during the 1990s, including the bicyclam AMD3100 (Figure 1B) [43]. These compounds have no structural resemblance with known peptide ligands and can be considered as functional (type-II) mimetics. In 2008, AMD3100 (plerixafor, Mozobil) became the first, and still the only, marketed CXCR4 antagonist. It is currently approved for stem cell mobilization in patients

with non-Hodgkin's lymphoma and multiple myeloma, but clinical trials for other indications are ongoing [44]. Other small-molecule functional mimetics that have been tested clinically include AMD11070/AMD070 (Genzyme), MSX-122 (Metastatix), and TG-0054 (TaiGen) (Figure 1B). The different classes of small-molecule CXCR4 antagonists have been extensively reviewed by Neamati and co-workers [45, 46]. Similarly, several polypeptide CXCR4 antagonists have undergone clinical testing, including BL-8040/BKT140 (BioLineRx, Biokine), POL6326 (Polyphor), and LY2510924 (Lilly) (Figure 1A). The polypeptide (polyphemusin II- and CXCL12-derived) CXCR4 antagonists have recently been reviewed by Oishi and Fujii [47].

It should be noted that subsequent pharmacodynamics studies of some of these compounds showed that T140 had inverse agonistic properties in addition to the antagonistic actions on CXCL12-induced CXCR4 activity, and that AMD3100 and ALX40-4C were weak partial agonists [48].

While progress had been made for both the polypeptide antagonists and the functional mimetics during the 1990s, the foundation for rational design of topographical (type-III) mimetics was not laid until 2003, when Fujii et al. reported a series of cyclic pentapeptides as potent CXCR4 antagonists [49]. These small-molecule peptides (MW = 729 Da) bridged the gap between the large polypeptide antagonists (MW > 2000 Da) and the small-molecule functional mimetics, and are further discussed in section 4.1.



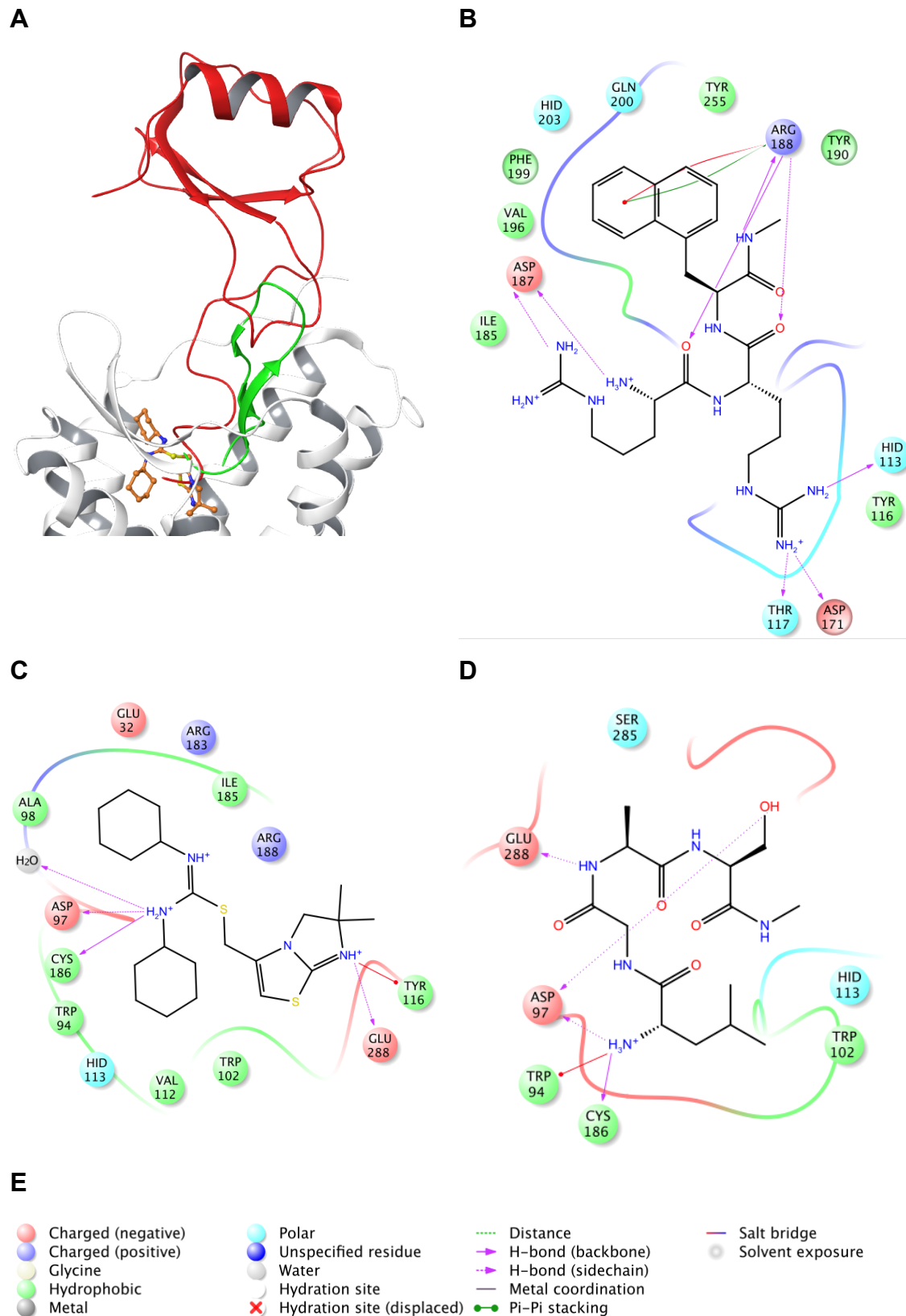
### **3. MOLECULAR INTERACTIONS BETWEEN CXCR4 AND ITS KEY PEPTIDE LIGANDS**

The 7TM receptors are characterized by seven  $\alpha$ -helices (TM1-7) that span the cell membrane. The helices are connected by three extracellular loops (ECL1-3) and three intracellular loops (ICL1-3) and flanked by an extracellular N-terminus and an intracellular C-terminus; in some 7TM receptors an additional helix (H8) is found in the C-terminus.

Before the first X-ray structures of CXCR4 were reported, the only available 7TM structures were light-activated (rhodopsin), aminergic, and nucleoside binding receptors. Prior to the public release of the experimental CXCR4 structures, a community-wide assessment (GPCR Dock 2010) was conducted [50], where the scientific community was challenged to predict the structures of these solved complexes. The results of this assessment showed that the detailed structure of CXCR4 itself (and hence, the binding pocket) was quite difficult to accurately model based on the already known 7TM receptor structures [50]. This was due to several distinctive structural features in the X-ray structures of CXCR4, including shifting, rotation, and extension of transmembrane helices as well as the folding/position of ECL2 [5]. Also, partly as a consequence, the prediction of correct binding mode for the CXCR4 antagonists by docking to homology models proved extremely difficult, especially for the large CVX15 ligand [50]. For this reason, we herein limit the discussion to the X-ray structures of CXCR4 and proposed peptide ligand: CXCR4 complexes that are based on these experimental structures, i.e. homology models published before 2010 are not considered.

#### **3.1. Experimental peptide ligand: CXCR4 complexes**

Two of the experimental CXCR4 complexes published to date contain a peptide/protein ligand: the CVX15: CXCR4 co-crystal structure (PDB: 3OE0) [5] and the vMIP-II: CXCR4 structure (PDB: 4RWS) [6]. The third ligand that has been co-crystallized with CXCR4 is the small-molecule isothiourea-derivative IT1t (PDB: 3ODU) [5].



**Figure 2.** (A) Schematic presentation of the binding modes for the polypeptide antagonist CVX15 (green ribbon; PDB: 30E0) [5], the non-peptide small-molecule antagonist IT1t (ball-and-stick, orange carbon atoms; PDB: 30DU) [5], and the viral chemokine antagonist vMIP-II (red ribbon;

PDB: 4RWS) [6]. For clarity, only the receptor structure of the CVX15:CXCR4 complex (white ribbons; PDB: 30E0) is shown. The figure was created with Maestro [51] by superimposing the C $\alpha$ -atoms of the transmembrane bundle (residues 34-62, 72-99, 105-139, 145-174, 193-224, 239-266, 273-301). (B) Ligand interaction diagram (LID) for the three N-terminal residues (Arg<sup>1</sup>-Arg<sup>2</sup>-1-Nal<sup>3</sup>) of CVX15. (C) LID for IT1t. (D) LID for the four N-terminal residues (Leu<sup>1</sup>-Gly<sup>2</sup>-Ala<sup>3</sup>-Ser<sup>4</sup>) of vMIP-II. (E) Legend for LID. The LIDs were created with Maestro [51] using a cutoff of 3 Å.

CVX15 is a 16-mer opened analog of the head-to-tail cyclized POL-3026 (Figure 1A) [52], which was developed from the 14-mer “prototype” polypeptide CXCR4 antagonist T140.  $\beta$ -Turns are known to be binding motifs for peptidergic 7TM receptors [53], and NMR studies have previously shown that T140 indeed contains a  $\beta$ -turn around positions 8 and 9 (D-Lys<sup>8</sup>-Pro<sup>9</sup>) [54]. However, the crystal structure of the CVX15:CXCR4 complex (Figure 2A) shows that the corresponding D-Pro<sup>8</sup>-Pro<sup>9</sup> turn of CVX15 is not embedded in the ligand binding pocket within the transmembrane bundle, but is oriented towards the extracellular side. As the 24 N-terminal residues of CXCR4 are missing in the co-crystal structure due to lack of interpretable densities, this complex does not reveal all potential interactions between CVX15 and the CXCR4 N-terminus. The receptor interactions are mainly found within the so-called major binding pocket (delimited by TMs 3-6), and are formed by the N- and C-terminal ends of CVX15 (Figure 2B), which contain the four pharmacophoric residues (Arg<sup>2</sup>, Nal<sup>3</sup>, Tyr<sup>5</sup>, and Arg<sup>14</sup>) of the polypeptide antagonists [55]. The CVX15:CXCR4 complex is highly relevant for further development of the T140-derived CXCR4 antagonists, including the cyclopentapeptides (section 4.1).

In contrast to CVX15, which mainly binds within the major binding pocket of CXCR4, the small-molecule non-peptide antagonist IT1t binds within the minor binding pocket, which is delimited by TMs 1, 2, 3, and 7 (Figure 2C). Based on the different binding modes of the polypeptide CVX15 and the non-peptide antagonist IT1t, it is tempting to classify the major and minor binding pocket of CXCR4 as the “peptide” and “non-peptide” pocket, respectively; however, they should not be considered as such. Early mutagenesis studies in CXCR4 predicted the major binding pocket to be the main pocket for the binding of bicyclams and

monocyclams, exemplified by AMD3100 and AMD3465, respectively [56-61]. In both ligand classes, Asp171 (TM4) was suggested to be the anchor point for one cyclam ring. For the bicyclams, the other cyclam ring was suggested to be sandwiched between TM6 and TM7 (Asp262 and Glu288), whereas the non-cyclam end of the monocyclams - in AMD3465, a simple N-pyridinylmethylene moiety - had an expanded interaction pattern that also included residues located in the extracellular segments of TM6 (Ile259) and TM7 (His281) [60]. Also, a recent binding mode study of the non-peptide small-molecule CXCR4 antagonist AMD11070 showed that poses generated by docking to the 3OE0 (peptide ligand) structure were in better agreement with experimental data than poses generated with the 3ODU (non-peptide) structure, and that Asp171 (TM4) in the major binding pocket was involved in strong interactions with this non-peptide ligand [62].

vMIP-II is a viral chemokine that is secreted by human herpesvirus 8, and has been shown to be a potent *antagonist* for several human chemokine receptors, including CXCR4 [63]. Activation of chemokine receptors by their endogenous chemokine agonists is generally considered to be a two step process, where separate regions of the large chemokine ligand are involved in initial binding to and activation of the receptor; the corresponding receptor regions are referred to as chemokine recognition site (CRS) 1 and 2 [64].

The successful crystallization of vMIP-II: CXCR4 was achieved by formation of a covalent complex by disulfide-trapping of the engineered cysteine mutants D187C (CXCR4) and W5C (vMIP-II) [6]. The X-ray structure shows that the vMIP-II core (specifically residues 13-16 and 49-51) makes extensive contacts with the CXCR4 N-terminus, while the vMIP-II N-terminus (residues 1-10) reaches into the binding pocket within the transmembrane bundle. Specifically, the structure identifies residues 23-27 in the CXCR4 N-terminus as CRS1 (the ultimate N-terminal residues 1-22 are not visible in the structure), and CXCR4 residues in TM2 (Trp94, Asp97), TM3 (His113), ECL2 (Asp187), TM6 (Asp262), and TM7 (Glu277, His281, Glu288) as key interaction partners for the N-terminal tip (residues 1-7) of vMIP-II, i.e. CRS2 (Figure 2D). Also, there is an intermediate

region (termed CRS1.5) that involves additional interactions between the N-terminus of CXCR4 (residues 27-31) and vMIP-II (residues 8-12). Interestingly, as the N-terminal tip of vMIP-II mainly occupies the minor binding pocket of CXCR4, the spatial overlap between vMIP-II and CVX15 is quite limited; instead, the ultimate N-terminal residues of vMIP-II overlap with the small-molecule IT1t that binds to the same receptor region (Figure 2A).

Even if the complex between the viral chemokine antagonist vMIP-II and CXCR4 represents an inactive receptor state, the X-ray structure still provides valuable insight into the molecular recognition between CXCR4 and its endogenous chemokine agonist CXCL12 (section 3.2). The structure is also a valuable foundation for design/optimization of peptide/peptidomimetic antagonist based on vMIP-II fragments, e.g. the Trp<sup>5</sup>-His<sup>6</sup>-Arg<sup>7</sup> based peptides reported by Portella et al. (section 4.3.2).

### **3.2. Proposed CXCL12: CXCR4 complexes**

As activation of CXCR4 by its 68-mer protein ligand CXCL12 is a key event in several pathological processes, including cancer metastasis, rational design of e.g. anti-cancer drugs would benefit from a better understanding of the CXCL12: CXCR4 interactions and activation mechanism.

Despite the huge structural differences between the ligands for class A 7TM receptors, it is generally acknowledged that all class A receptor subclasses are activated by the same overall helical movements [65]. At present, >100 crystal structures of ~20 7TM receptors have been reported [66]. Some of these have been crystallized in an active conformation, thereby shedding light on the structural requirements for receptor activity, including the conformation of molecular micro-switches. For a recent review on activation of 7TM receptors, see reference [67].

After the identification of CXCL12 as the endogenous agonist for CXCR4 [12], the first NMR [41] and X-ray [68] structures of CXCL12 soon followed. The solution structure [41] revealed that CXCL12 adopts the common tertiary chemokine-

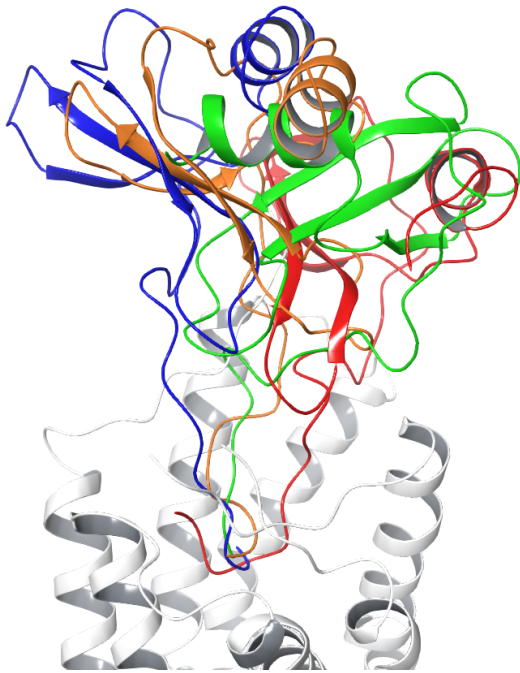
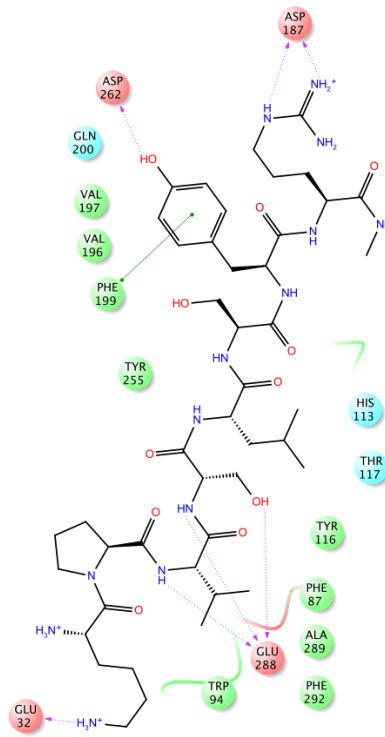
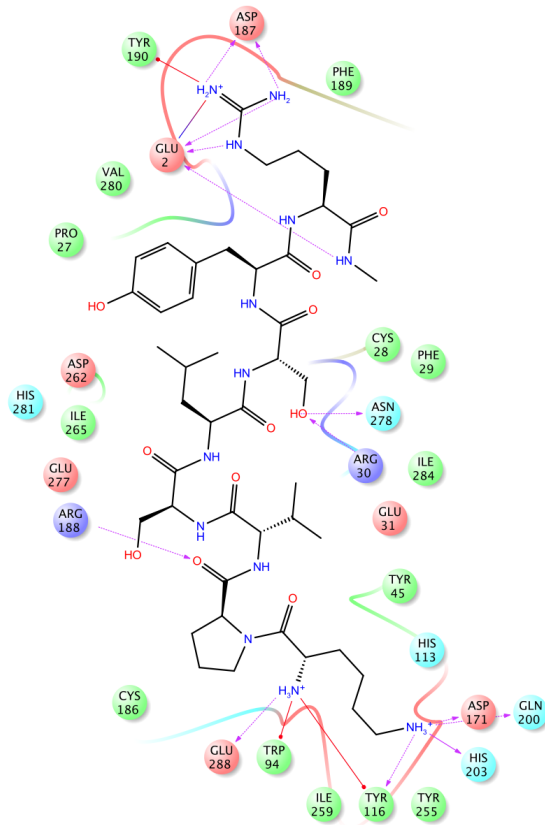
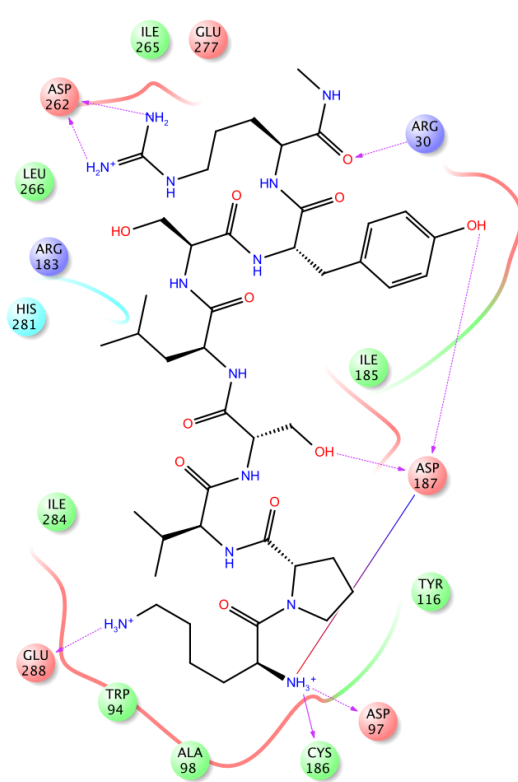
fold, with a flexible N-terminus, an extended loop (N-loop), three antiparallel  $\beta$ -strands ( $\beta_1$ - $\beta_3$ ) connected by two  $\beta$ -turns, and an  $\alpha$ -helix, which is connected to the  $\beta_3$ -strand by another  $\beta$ -turn (later referred to as the 50s loop). [65][70-73] The “two-step” activation model has also been proposed for CXCL12:CXCR4, and early functional studies of CXCL12-analogs indicated that the RFFESH-motif (residues 12-17) in the N-loop had a key role in the initial binding step (CRS1 interaction), while the flexible N-terminal region (residues 1-8), especially Lys<sup>1</sup> and Pro<sup>2</sup>, was responsible for receptor activation (CRS2 interaction) [41]. However, it should be noted that the “two-step” model probably is a simplification, as several minor steps could be anticipated to take place during chemokine binding and subsequent receptor activation [69], as shown for other class A receptors [70-72]. Accordingly, advanced NMR studies have shown that large parts of the CXCL12 core structure, including the N-loop, the 50s loop, and the  $\beta$ -sheet, are involved in the initial binding step [73].

With respect to the stoichiometry of the CXCL12:CXCR4 complex, several different alternatives have been envisioned (1:1, 1:2, 2:1, 2:2); however, recent studies by Kufareva et al. show that the 1:1 complex is the functional unit [6, 74].

In their report of the first X-ray structures of CXCR4 [5], Wu et al. suggested that the co-crystallized ligands CVX15 and IT1t, at least to some extent, occupied the binding site of the activating CXCL12 N-terminus, i.e. CRS2. Five computational models of the CXCL12:CXCR4 interaction have since been reported, four of which have been based on the CVX15/IT1t X-ray structures. Xu et al. combined protein-protein docking, molecular dynamics (MD) simulations, and free energy calculations, starting from one CXCL12 structure and one CXCR4 structure [75]. Tamamis and Floudas published a more extensive study, starting from 22 CXCL12 structures and 17 CXCR4 structures [76]. In both these studies, the missing N-terminal residues of the CXCR4 X-ray structure were constructed by utilizing the NMR structure of CXCL12 complexed with the CXCR4 N-terminus [77]. Costantini et al. docked a crystallographic CXCL12 structure to a CXCR4 structure where the missing N-terminus was modeled based on a rhodopsin structure [78]. The Abagyan/Handel group first generated a model of the

CXCL12:CXCR4 complex by employing an engineered and experimentally validated disulfide bond (CXCL12 S16C to CXCR4 K25C) as restraint [74]. Based on their experimental vMIP-II:CXCR4 structure, the same group recently proposed a further refined CXCL12:CXCR4 model, using the position of the vMIP-II core as a guide for the CRS1 interaction [6]. Both models from the Abagyan/Handel group contain a truncated CXCR4 N-terminus as no attempts were made to construct the missing N-terminal residues.

Figure 3A shows a schematic comparison of four of the CXCL12:CXCR4 complexes (we did not have access to the coordinates of the model proposed by Costantini et al). While experimental data (e.g. from site-directed mutagenesis studies) have been considered in all cases, the figure clearly shows that there are significant differences between the models, both with respect to interactions between the CXCL12 core and the extracellular domains of CXCR4 (CRS1) and the interactions between the CXCL12 N-terminus and the transmembrane bundle (CRS2). Again, this reflects the difficulties in modeling binding of large and complex peptide/protein ligands to their receptors.

**A****B****C****D**

**Figure 3.** (A) Comparison of the CXCL12:CXCR4 complexes suggested by Xu et al. (red) [75], Tamamis and Floudas (green) [76], Kufareva et al. (blue) [74], and Qin et al. (orange) [6]. For



clarity, only the receptor structure of the complex proposed by Kufareva et al. (white ribbons) is shown. The PDB-files were kindly provided by Drs. Tingjun Hou, Christodoulos A. Floudas, and Irina Kufareva (two structures). The figure was created as described in Figure 2A. (B-D) Ligand interaction diagrams for the eight N-terminal CXCL12 residues (KPVLSYR) in the models by Xu et al. (B), Tamamis and Floudas (C), and Qin et al. (D). The LIDs were created as described in Figure 2.

The model proposed by Xu et al. involves electrostatic interactions between Asp262/Glu277 and Lys271 of CXCR4 with Arg<sup>12</sup> and Glu<sup>15</sup> of CXCL12, respectively, thus identifying the top of TM6 and TM7, including ECL3, as CRS1. With respect to CRS2 (Figure 3B), the model suggests that the N-terminus of CXCL12 is curled up in the transmembrane binding pocket, with the tip of the loop pointing up towards the extracellular receptor domains. Arg<sup>8</sup> at the base of the CXCL12 N-terminus is involved in a salt bridge with Asp187 (ECL2), while H-bond interactions between Val<sup>3</sup>/Ser<sup>4</sup> and Glu288 (TM7) anchor CXCL12 to the bottom of the pocket. However, Lys<sup>1</sup> of CXCL12, which is known to be important for activation of CXCR4, reaches out of the pocket and interacts with Glu32 in the CXCR4 N-terminus.

In contrast, Tamamis and Floudas identified the N-terminus and ECL2 (connecting TM4 and TM5) of CXCR4 as CRS1, specifically residues Glu2, Ile6-Tys12, Glu14, Met16, and Tyr190-Asp193. Salt bridges were seen between Arg<sup>12</sup> of CXCL12 and Glu2, Tys7, and Asp193 of CXCR4. They also identified Asp187 (ECL2) as the main interaction partner for Arg<sup>8</sup> of CXCL12 (Figure 3C), but suggest that the N-terminus points straight down into the helical bundle, and that residues Lys<sup>1</sup>-Leu<sup>5</sup> are practically buried in the transmembrane pocket (CRS2). Here, Lys<sup>1</sup> forms salt bridges with both Asp171 (TM4) and Glu288 (TM7) at the bottom of the pocket.

The model recently proposed by Qin et al. [6] appears to be the most mature, as structural information from the experimental vMIP-II: CXCR4 complex has been employed to generate the CXCL12: CXCR4 complex. This model identifies the CXCR4 N-terminus and the top of TM6 and TM7, including ECL3, as CRS1/1.5. The proximal N-terminus (residues 21-25; residues 1-20 are not included in the

model) is in extensive contact with several domains in the CXCL12 core, including the N-loop and the  $\beta_2$ - $\beta_3$  loop. At the interface between the extracellular receptor domains and the transmembrane pocket (CRS1.5) electrostatic interactions are found between Asp262:Arg<sup>8</sup> and Glu277:Arg<sup>12</sup>. With respect to CRS2, Qin et al. suggest that the N-terminus of CXCL12 is oriented towards the minor binding pocket (Figure 3D). Specifically, Lys<sup>1</sup> of CXCL12 forms salt bridges with Asp97 (TM2) via the positively charged N-terminal amino group, and with Glu288 (TM7) via the side chain amino group. The side chains of Ser<sup>4</sup> and Tyr<sup>7</sup> are both involved in H-bonds to Asp187 (ECL2).

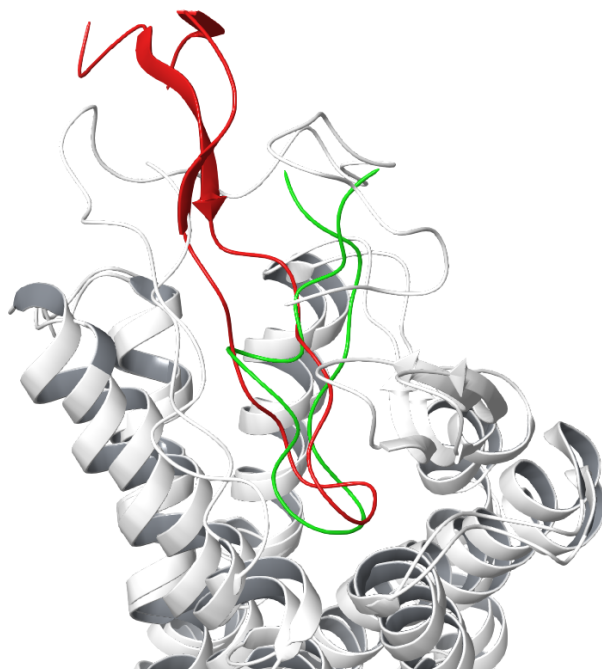
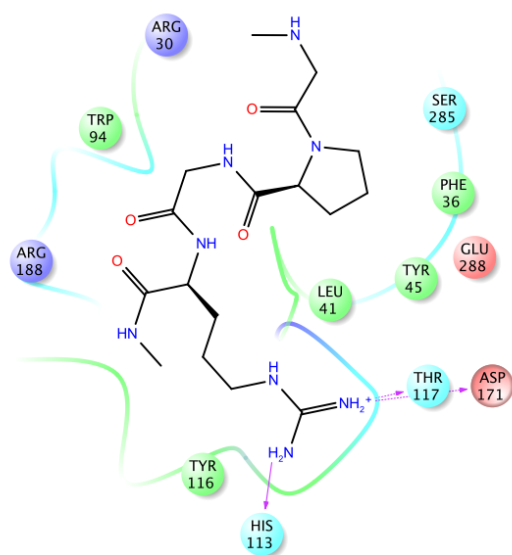
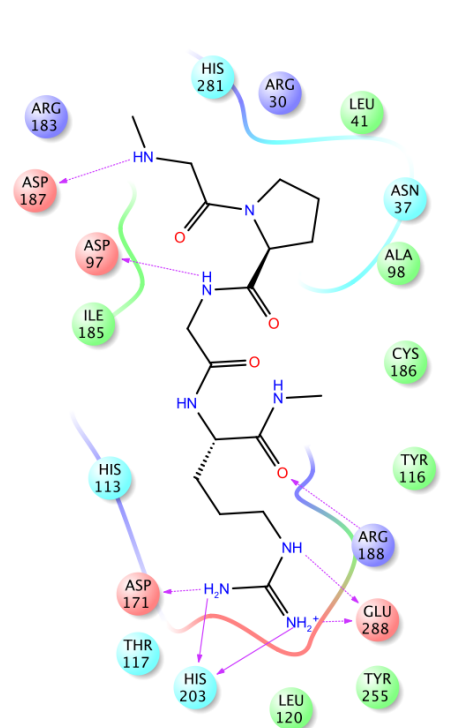
While these models of the CXCL12:CXCR4 complex are intriguing, it should be noted that the experimental CXCR4 structures that have been used to generate them are in inactive (antagonist-bound) states, and thus not ideally suited as templates for modeling of the active (agonist-bound) state of CXCR4. In the absence of further structural knowledge about the activated CXCR4 state, it is difficult to judge how well the proposed complexes describe the “true” CXCL12:CXCR4 interactions. Consequently, rational design of small-molecule peptidomimetic ligands based on short CXCL12-motifs still remains a challenging task.

### 3.3. Proposed HIV-1 V3:CXCR4 complexes

By now, the role of CXCR4 as co-receptor for HIV-1 entry into human cells is well established [79]. Specifically, this involves binding of the third variable loop (V3) of the HIV-1 envelope glycoprotein gp120 to CXCR4. Thus, understanding the detailed molecular interactions between the V3 loop and CXCR4 is important for development of small-molecule anti-HIV drugs.

The V3 loop consists of a base, a stem, and a tip, and when the first structures of V3 were published, it was believed that the base and stem bind to the N-terminus of the HIV-1 co-receptors CXCR4/CCR5 while the highly conserved V3 tip (GPGR  $\beta$ -turn motif) interacts with the extracellular loops [80, 81], i.e. that binding did not involve the transmembrane pocket. However, the first X-ray structures of CXCR4 provided some indications that the V3 loop could penetrate down into this pocket [5]. Two computational models of the V3:CXCR4 interaction have since been proposed [82, 83], both of which involve the transmembrane pocket (Figure 4A).

In the supporting information of their report of the X-ray structure of CCR5, which is the other co-receptor for HIV entry, Tan et al. presented possible structural complexes for the HIV V3 loop bound to CXCR4 and CCR5 [82]. The V3:CXCR4 complex was generated by docking of the terminally constrained 20 residue peptide fragment **TRKR<sup>306</sup>IR<sup>308</sup>IQR<sup>311</sup>GPGR<sup>315</sup>AFVTIGK<sup>322</sup>**, corresponding to residues 303-322 in the V3 loop of the T-tropic (CXCR4-using) HIV-1 strain HXBc2. In the suggested V3:CXCR4 complex (Figures 4A and B), the following salt-bridges were observed: Arg<sup>306</sup>/Lys<sup>322</sup> to Asp193, Arg<sup>308</sup> to Asp262/Glu277, Arg<sup>311</sup> to Asp97/Asp187, and Arg<sup>315</sup> to Asp171. However, it should be noted that the N-terminal residues that are missing in the X-ray structures of the receptor were not taken into consideration.

**A****B****C**

**Figure 4.** (A) Comparison of the HIV-1 V3:CXCR4 complexes suggested by Tan et al. [82] (green) and Tamamis and Floudas [83] (red). The figure was created as described in Figure 2A [53]. The PDB-files were kindly provided by Professors Beili Wu and Christodoulos A. Floudas, respectively. (B) Ligand interaction diagram (LID) for the conserved HIV-1 V3 GPGR<sup>315</sup> motif in

the model by Tan et al. (C) LID for the same GPGR<sup>18</sup> motif in the model by Tamamis and Floudas.

The LIDs were created as described in Figure 2.

Tamamis and Floudas have recently proposed a model of the V3:CXCR4 complex based on a more comprehensive protocol, which included docking, MD simulations and free energy calculations [83]. They used the entire V3 sequence (CTR<sup>3</sup>PNNNTRK<sup>10</sup>RVSL**GPGR**<sup>18</sup>VWYTTGQIVGDIR<sup>31</sup>KAHC) of a dual-tropic (CXCR4- and CCR5-using) HIV strain, and also constructed the missing N-terminal of CXCR4. Analysis of the final complex (Figures 4A and C) showed that most of the V3 loop was buried in the receptor, with the N-terminal of CXCR4 curved around V3. Extensive contacts were found for all TMs, although to different extents, and also for all extracellular domains except ECL1. The strongest intermolecular interaction in the entire complex was the salt bridge between Arg<sup>18</sup> in V3 and Asp171 (TM4) and Glu288 (TM7) (Figure 4C). Other key V3:CXCR4 interactions included Arg<sup>3</sup>:Asp22/Glu268 (N-terminus and ECL3, respectively), Lys<sup>10</sup>:Asp193 (ECL2), and Arg<sup>31</sup>:Glu14/Asp20 (both in N-terminus), i.e. salt bridges between positively charged V3 residues and negatively charged CXCR4 residues.

Due to the differences in length and sequence of the V3 loops used in the two studies, it is difficult to compare the specific residue interactions directly. However, as the overlay of the two complexes (Figure 4A) shows, the conserved GPGR tip is positioned similarly in both cases, with the Arg residue anchored to Asp171 (TM4) (Figures 4B and C). Still, the orientation of the stem and base of the V3 loops differ significantly.

The ligand-receptor complexes for CXCL12:CXCR4 (Figure 3C) and V3:CXCR4 (Figure 4C) published by Tamamis and Floudas show a substantial overlap of the binding sites, where the N-terminus of CXCL12 and the tip of the V3 loop occupy the same area. Specifically, Lys<sup>1</sup> of CXCL12 and Arg<sup>18</sup> of the V3 loop both interact with Asp171 (TM4) and Glu288 (TM7).

As is apparent from the above, the predominance of negatively charged residues in the extracellular regions and the transmembrane binding pocket of CXCR4 means that the molecular interactions are dominated by salt bridges with positively charged residues in the peptide ligands. For peptide ligands in the major binding pocket of CXCR4, Asp171 (TM4) is typically the key interaction site for a ligand arginine residue. Asp171 has also been shown to be an anchor residue for the small-molecule functional mimetics AMD3100 [58], AMD3465 [60] and AMD11070 [62]. Ligands in the minor pocket normally engage in a salt bridge to Asp97 (TM2) and/or the nearby Asp187 (ECL2). Glu288 (TM7), which sits centrally and bridges the major and minor pocket, is implicated in all of the suggested binding modes for peptide CXCR4 ligands (Figures 3 and 4). This residue (GluVII:06) is highly conserved in the chemokine receptor family, and has been proposed to serve as an anchor point for positively charged small-molecule ligands for chemokine receptors [84].

#### 4. SMALL-MOLECULE PEPTIDE AND PEPTIDOMIMETIC CXCR4 ANTAGONISTS

The cyclopentapeptide CXCR4 antagonists are especially interesting lead compounds since cyclic pentapeptides are known to mimic peptide turns [85]. Such reverse-turn motifs appear to be a universal ligand recognition element for peptidergic 7<sup>TM</sup> receptors [53], and the HIV V3: CXCR4 interaction (section 3.3) indeed involves a  $\beta$ -turn motif. Thus, the cyclopentapeptide CXCR4 antagonists provide an excellent starting point for rational design of topographical  $\beta$ -turn mimetics, and now that experimental structures of CXCR4 have become available, further progress in this field is expected.

Marshall has proposed a 6-step hierarchical approach to rational design of peptidomimetics from a parent bioactive peptide [86]:

- (1) Identify critical side chain residues (alanine scan)
- (2) Define active core (size reduction)
- (3) Define local conformational parameters, e.g. probable turns (D-amino acid scan, unusual amino acid scan)
- (4) Generate active constrained analogs (cyclization, introduction of turn mimetics, amide bond modification)
- (5) Generate hypothesis for receptor-bound conformation (conformational analysis, physical studies)
- (6) Arrive at constrained peptidomimetic analogs (design novel compounds that mimic the critical 3D elements)

In this section, we first summarize the extensive studies of the cyclopentapeptide CXCR4 antagonists (roughly corresponding to steps 1-5 above, but not performed in this exact order) and then describe the attempts that have been made to exploit this information for design of peptidomimetic analogs (step 6). Finally, other small-molecule peptide and peptidomimetic CXCR4 antagonists, not inspired by the cyclopentapeptides, are discussed.

## 4.1. Cyclic pentapeptides and derivatives

### 4.1.1. Discovery of the cyclopentapeptide CXCR4 antagonists

The cyclopentapeptide antagonists [49] (Figure 5A) were developed from the macrocyclic 14-mer polypeptide lead compound T140 (Figure 1A) [40], and the downsizing strategy was based on combining the four pharmacophoric residues of T140 (Arg<sup>2</sup>, 2-Nal<sup>3</sup>, Tyr<sup>5</sup>, and Arg<sup>14</sup>) [55] with a Gly spacer to facilitate head-to-tail cyclization. Of the 192 potential cyclopentapeptides (12 sequences x 16 stereoisomers), 60 compounds were synthesized and tested, resulting in the identification of cyclo(-L-Arg<sup>1</sup>-L-Arg<sup>2</sup>-L-2-Nal<sup>3</sup>-Gly<sup>4</sup>-D-Tyr<sup>5</sup>-), later known as FC131 (Figure 5A) as the most potent CXCR4 antagonist with IC<sub>50</sub> = 0.004 μM (inhibition of <sup>125</sup>I-SDF-binding) and EC<sub>50</sub> = 0.038 μM (inhibition of HIV-induced pathogenicity) [49]. The D-Arg<sup>1</sup>-epimer (later known as FC092) was also shown to be a potent antagonist (IC<sub>50</sub> = 0.008 μM and EC<sub>50</sub> = 0.11 μM). Importantly, the linear and capped analogs of the most potent cyclopentapeptides showed significantly reduced activity.

A solution structure for FC131 based on <sup>1</sup>H-NMR studies in DMSO was also reported [49]. While the exact spatial orientation of the relatively flexible side chains could not be determined, the reported backbone conformation is consistent with later NMR studies of the bioactive (receptor-bound) conformation for the cyclopentapeptide antagonists [87-91].

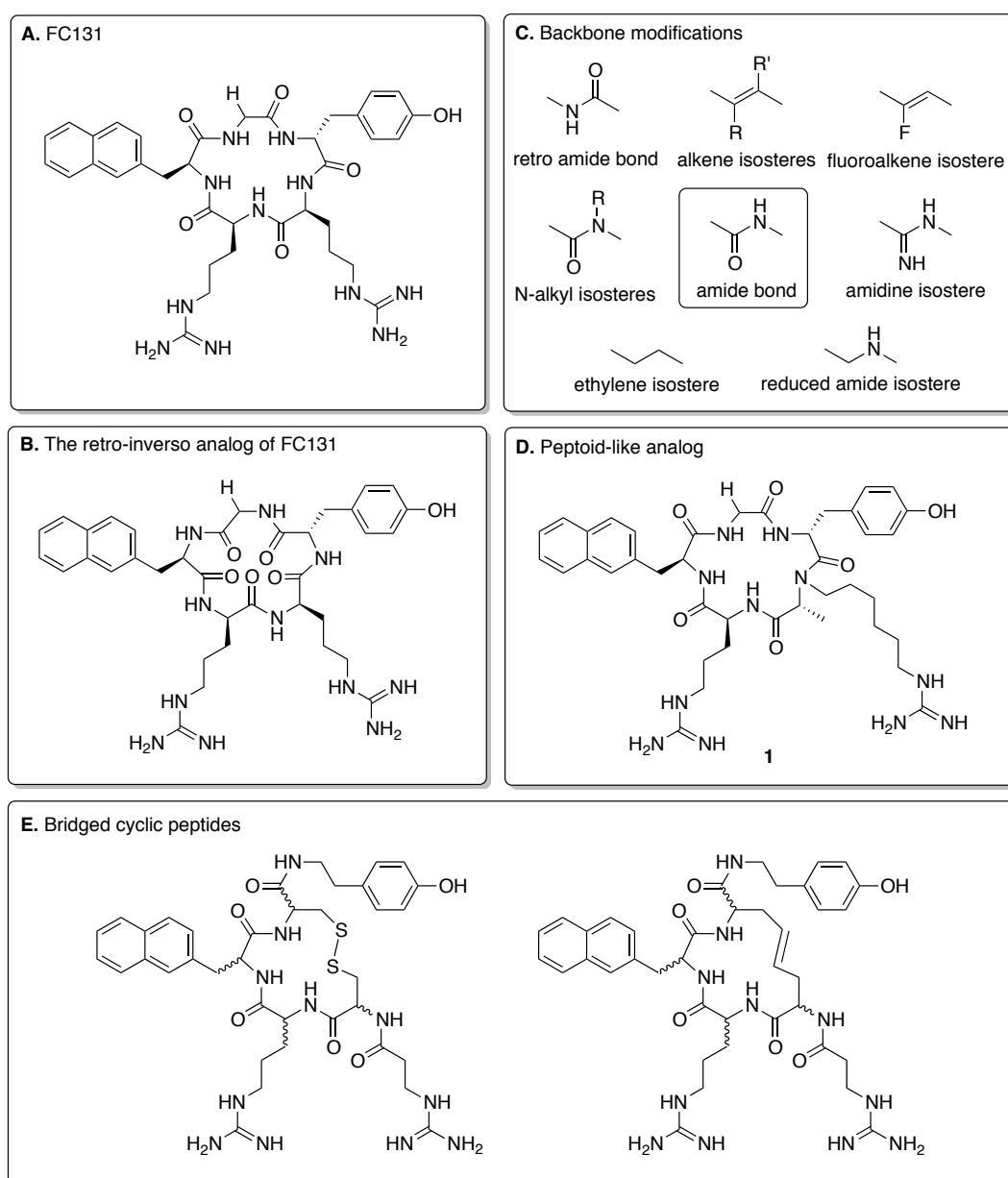
Thus, this key paper [49] demonstrated the simultaneous importance of sequence, stereochemistry, and cyclic constraint for CXCR4 antagonism, and also revealed the presumed bioactive backbone conformation for the lead cyclopentapeptide antagonist FC131.

### 4.1.2. Backbone modifications

**Retro-inverso analogs.** In addition to the “conformation-based” and “sequence-based” libraries in the original paper [49], biological data were soon reported for a third cyclopentapeptide library, consisting of retro-inverso analogs (Figure 5B) [92]. However, the retro-inverso analog with highest potency (EC<sub>50</sub> = 1.7 μM)



was 19-fold less potent than FC131 ( $EC_{50} = 0.088 \mu\text{M}$ ). Also, the three compounds with highest potency ( $EC_{50} < 5 \mu\text{M}$ ) were retro-inverso analogs of FC131 stereoisomers with low potency ( $EC_{50} > 5 \mu\text{M}$ ). The generally low activity of the retro-inverso analogs compared to FC131 indicates an important role of the backbone amide bonds and/or conformation for CXCR4 antagonism.



**Figure 5.** Structures of (A) the lead cyclopentapeptide CXCR4 antagonist FC131, (B) its retro-inverso analog, (C) the investigated amide bond isosteres, (D) the high-affinity peptid-like analog **1** reported by Demmer et al. [91], and (E) the two different classes of bridged cyclic peptides reported by Tamamura et al. [88].

**Amide bond isosteres.** Several studies on backbone modifications have since followed, mainly by replacement of the amide bonds with isosteres, including alkene [87, 90, 93], N-alkyl [89, 91], fluoroalkene [93, 94], amidine [95], reduced amide [87], and ethylene [90] isosteres (Figure 5C).

Of these, the most successful strategy in terms of activity has been introduction of the amidine isostere [95]. The most active analog contained an amidine isostere in the Nal<sup>3</sup>-Gly<sup>4</sup> fragment and had 30-fold higher affinity for CXCR4 than FC131 (IC<sub>50</sub>-values of 4.2 and 126 nM, respectively). The Gly<sup>4</sup>-D-Tyr<sup>5</sup>, Arg<sup>2</sup>-Nal<sup>3</sup>, and Arg<sup>1</sup>-Arg<sup>2</sup> analogs were also highly active (IC<sub>50</sub> < 16 nM), while the D-Tyr<sup>5</sup>-Arg<sup>1</sup> analog showed lower affinity (IC<sub>50</sub> = 679 nM) than FC131. The general success of the amidine strategy was hypothesized to result from favorable ionic interactions between the positively charged amidine group and negatively charged Asp/Glu residues in the binding pocket. However, as the authors noted, the same result was not achieved by using the reduced amide isostere [87], which is also positively charged. This observation suggests an additional beneficial conformational effect of the amidine isostere, which is more rigid than the reduced amide isostere.

In contrast to the introduction of an amidine isostere in the D-Tyr<sup>5</sup>-Arg<sup>1</sup> fragment, modification of this amide bond through N-alkylation has proven fruitful. Ueda et al. reported the N-Me-D-Arg<sup>1</sup> analog FC122, which showed increased affinity compared to FC131 (IC<sub>50</sub>-values of 3 and 4 nM, respectively [89]; Inokuchi et al. have later reported values of 37 and 126 nM, respectively [95]). N-methylation of the other four amide bonds resulted in significantly reduced activity [89], i.e. a completely opposite trend than observed for the corresponding amidine isosteres [95], as discussed above. Collectively, this shows that the carbonyl oxygen in the D-Tyr<sup>5</sup>-Arg<sup>1</sup> amide bond plays a special role, while the amide NH is not needed for activity. The beneficial effect of replacing L-Arg<sup>1</sup> with N-Me-D-Arg<sup>1</sup> was attributed to conformational stabilization of the amide bond, and NMR studies confirmed that FC131 and FC122 display similar backbone conformations [89].

This concept was further developed by Demmer et al. who created a peptoid-like structure by replacing Arg<sup>1</sup> with D-Ala<sup>1</sup> and moving the side chain to the amide nitrogen [91]. Fine-tuning of the N-alkyl chain length and variation of the positively charged functional group (amine or guanidine) resulted in the highly active analog **1** (Figure 5D) with an IC<sub>50</sub>-value of 0.04 nM. Again, the increased activity was explained by conformational effects, as NMR studies showed that the structural modifications effectively “freeze” the backbone in the bioactive conformation.

The remaining backbone modifications listed above have generally resulted in reduced activity. Even if designed to resemble an amide bond, an isostere will have different structural and electronic properties, e.g. geometrical, steric, conformational, and H-bond donor/acceptor properties. It is therefore difficult to determine the exact reason for the reduced activity of these modified analogs.

#### **4.1.3. Side chain modifications**

With respect to the relative importance of the side chains for activity, an alanine scan of FC131 and the D-Arg<sup>1</sup> epimer FC092 [89] showed that Arg<sup>2</sup> and 2-Nal<sup>3</sup> were more important than L-/D-Arg<sup>1</sup> and D-Tyr<sup>5</sup>. Extensive SAR studies of FC131 involving more conservative side chain modifications have been performed, and due to the large number of analogs, we only summarize the main findings.

- Arg<sup>1</sup> is relatively tolerant to a range of structural modifications, indicating that it participates in non-specific receptor interactions [88, 96, 97].
- Arg<sup>2</sup> is very sensitive to the same modifications, indicating that it is involved in highly specific receptor interactions [97, 98].
- 2-Nal<sup>3</sup> is also sensitive to modifications, and the distal aromatic ring has been shown to be especially important [98, 99].
- Gly<sup>4</sup> was originally introduced as a spacer for synthetic feasibility; however, replacement of Gly with other small non-polar residues was later shown to result in reduced activity [89].
- D-Tyr<sup>5</sup> remains the best of the investigated phenyl-substituted analogs: 4-OH > H > 2-F > 3-F > 4-NH<sub>2</sub> > 4-F > 4-OMe > 4-Cl > 4-Br [96, 98-100].

The collective SAR data for the cyclopentapeptide CXCR4 antagonists indicate a “sidedness” in the molecule, where one half (Arg<sup>2</sup>-2-Nal<sup>3</sup>) provides the strongest, most specific interactions with the receptor, and the other half (D-Tyr<sup>5</sup>-Arg<sup>1</sup>) less specific, but still beneficial, interactions.

#### **4.1.4. Global modifications**

Also, introduction of larger structural changes while keeping the four side chain functionalities of the cyclopentapeptide ligands – here referred to as global modifications – has been attempted. Tamamura et al. reported two different classes of bridged cyclic peptides that were designed to mimic FC131 (Figure 5E) [88]. These compounds are essentially tetrapeptides based on an R<sup>1</sup>-Xaa<sup>1</sup>-Arg<sup>2</sup>-2-Nal<sup>3</sup>-Xaa<sup>4</sup>-R<sup>2</sup> motif, where R<sup>1</sup> and R<sup>2</sup> contain the two other pharmacophoric groups (positively charged group and phenol group). Cyclization between the Xaa<sup>1</sup> and Xaa<sup>4</sup> side chains, either via a disulfide bridge or an olefin bridge, provides the cyclic constraint. However, in both cases, the activity was generally low, the most active compound being a disulfide-bridged analog with an IC<sub>50</sub>-value of 0.54 μM, which was approximately 130-fold lower than for FC131. Also, the absence or presence of a tyramine (R<sup>2</sup>) did not significantly affect the activity. Collectively, the findings suggest a suboptimal orientation of the four side chains compared to the parent cyclopentapeptides, probably as a result of the relatively extensive shuffling of functional groups in the R<sup>1</sup>-Xaa<sup>1</sup> and Xaa<sup>4</sup>-R<sup>2</sup> fragments.

#### **4.1.5. Pharmacophore Model**

Based on an extensive exploration of the conformational space for a series of reported cyclopentapeptide CXCR4 antagonists, Våbenø et al. have proposed a minimalistic 3D pharmacophore model for this compound class [101]. The identified features included the spatial arrangement of the pharmacophoric side chains as well as the optimal conformation of the cyclopentapeptide backbone, which was consistent with the experimental solution structure for FC131 [49].

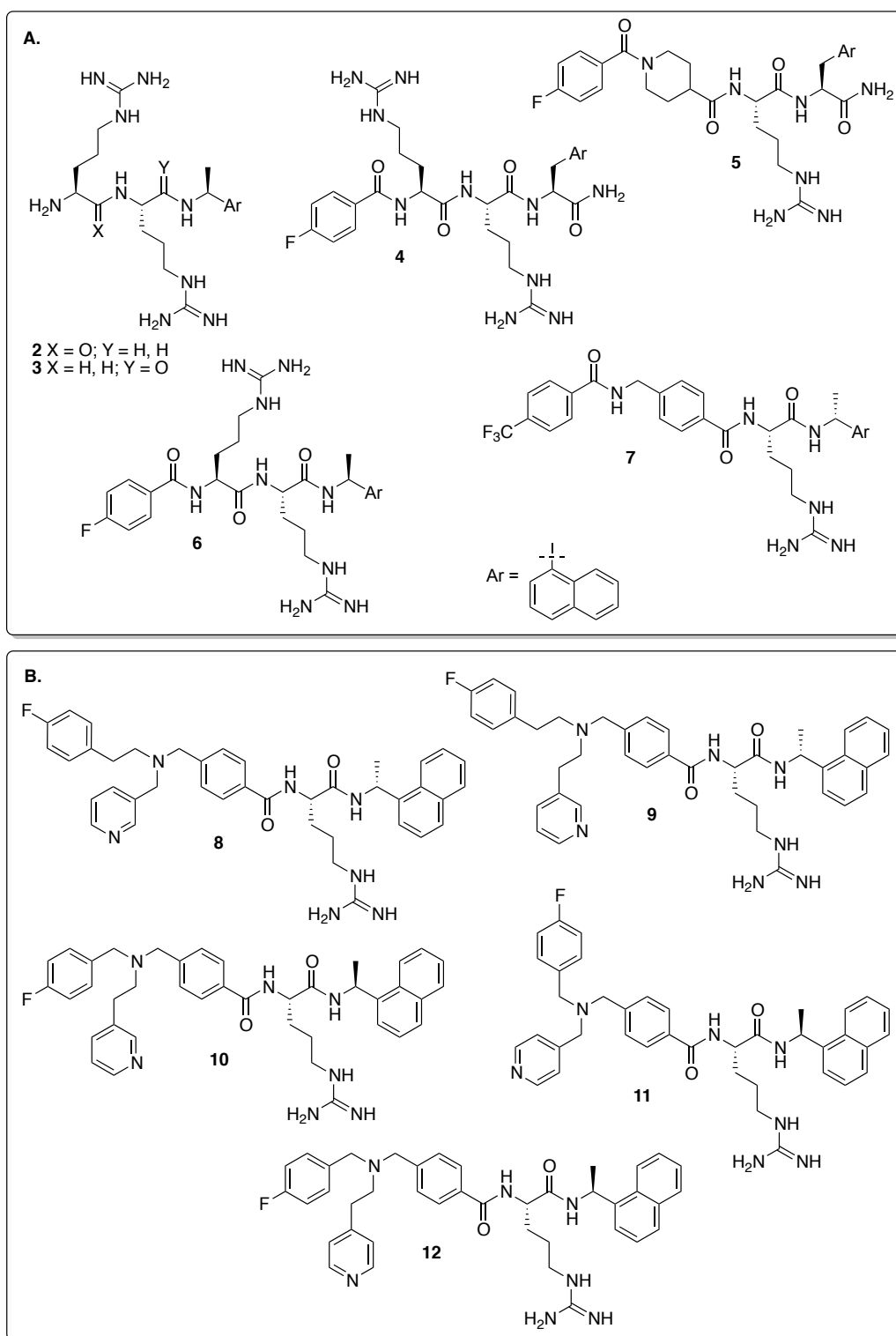
#### **4.1.6. Binding Mode**

Several computational models of the complex between CXCR4 and the cyclopentapeptide antagonist FC131 have been reported [99, 102, 103], and despite coming from three different groups, the suggested binding modes were quite consistent. Importantly, none of these studies were accompanied by *in vitro* experiments that verified the proposed binding mode. However, a recent binding mode study of FC131 based on extensive site-directed mutagenesis [104] supports the previously suggested binding models. The collective picture that has emerged from these models is generally consistent with SAR (see section 4.1.3) and suggests a critical interaction between Arg<sup>2</sup> of FC131 and His113 (TM3) and Asp171 (TM4). Thus, the essential Arg<sup>2</sup> residue of FC131 appears to have the same binding role as Arg<sup>2</sup> of CVX15 (Figure 2B) and Arg in the GPGR motif of the HIV V3 loop (Figures 4B and C).

## **4.2. Small-Molecule Peptidomimetics Derived from FC131**

### **4.2.1. Linear Peptidomimetics**

Tamamura et al. have performed extensive SAR studies on tri- and tetrapeptide mimetics starting from an Arg-Arg dipeptide (Figure 6A) [105]. The C-terminus was modified by addition of a 1-(1-naphthyl)ethylamine group, which presumably mimics the naphthalene groups found in both T140 and FC131, and also in the independently developed KRH-1636 [106] (see section 4.3.1 and Figure 8A below). They found that analogs **2** and **3**, where either of the amide bonds was replaced with a reduced amide isostere, displayed increased anti-HIV activity albeit similar inhibition of [<sup>125</sup>I]-CXCL12 binding to CXCR4 [105]. Further, they found that addition of an N-terminal Tyr residue, inspired by FC131, or particularly addition of a 4-fluorobenzoyl group, which had earlier been described as a beneficial N-terminal modification for T140 [107], increased anti-HIV activity and for some analogs also antagonistic activity against CXCR4. A series of N-terminally modified Arg-Arg-Nal-NH<sub>2</sub> and Arg-Nal-NH<sub>2</sub> peptides (e.g. **4** and **5**) were also evaluated for anti-HIV and CXCR4 antagonistic activity. For a number of the analogs prepared, no significant CXCR4 binding affinity could be observed at 1 μM, however significant anti-HIV activity was found for **6** and **7** (Figure 6A).



**Figure 6.** Structures of linear peptidomimetics reported by (A) Tamamura et al. [105], and (B) Narumi et al. [108]. See Figure 8A for the structure of KRH-1636.

Compounds containing two Arg side chains were generally found to display higher CXCR4 binding affinity compared to analogs with only one Arg residue.

Interestingly, as the authors note, the optimal stereochemistry of the 1-(1-naphthyl)ethylamide was found to differ for this compounds series compared to that of KRH-1636 (see Figure 8A). The differences between anti-HIV potency and affinity for CXCR4 seen throughout the series of compounds are explained by the differences found in the interaction of HIV and CXCL12 with the receptor. None of the compounds showed similarly high anti-HIV potency or CXCR4 inhibitory activity as FC131 or T140.

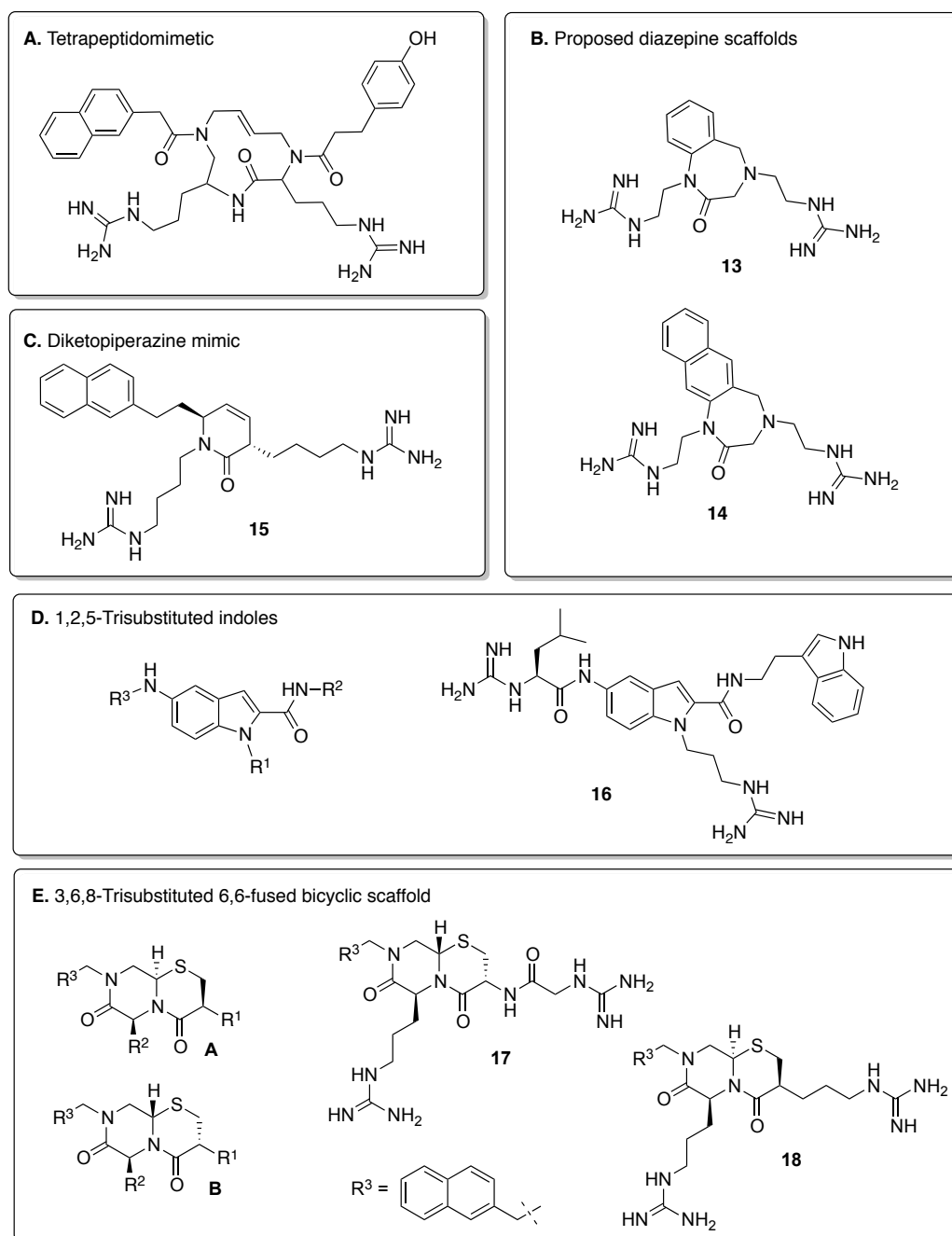
In a later SAR study by Narumi et al., compounds with only one Arg residue, i.e. analogs of KRH-1636 and **7**, were further modified [108]. Here, the stereochemistry of the 1-(1-naphthyl)ethylamide at the C-terminus was varied and different 4-fluorophenyl containing and 2-, 3- or 4-pyridyl containing substituents were introduced at the benzylic N-position (**8-12**, Figure 6B). The inhibitory activity against CXCR4 was also compared to that of KRH-1636, FC131 and T140, and while none of the analogs displayed the same high anti-HIV activity or CXCR4 inhibitory activity, several interesting observations were made. As for earlier analogs (vide supra), it was found that these low molecular weight compounds show two types of recognition modes for CXCR4. Compounds **8**, **9** and **10** were found to have highest anti-HIV activity, whereas compounds **11** and **12** displayed lower anti-HIV activity but similar or higher CXCR4 inhibition.

#### **4.2.2. Scaffold-Based Peptidomimetics**

The successful downsizing of the polypeptide CXCR4 antagonists to the cyclopentapeptide FC131 has spurred further efforts toward development of antagonists based on various scaffolds onto which the peptide pharmacophoric groups can be grafted.

**Tetrapeptidomimetics.** Cluzeau et al. have reported a series of tetrapeptidomimetics based on a 1,2,5,7-tetrasubstituted (*E*)-1,4,7-triazacycloundec-9-en-3-one scaffold (Figure 7A) [109]. These structures were designed to mimic FC131; however, the most active compound had 800-fold

reduced activity compared to FC131 (IC<sub>50</sub>-values of 3.2 and 0.004  $\mu$ M, respectively). No further studies on this compound class have been reported.



**Figure 7.** Structures of reported scaffold-based (type III) peptidomimetics.

**Tripeptidomimetics.** Marshall suggested early on that benzodiazepine and naphthyldiazepine structures (**13** and **14**, Figure 7B) could function as suitable scaffolds for Arg-Arg-Nal tripeptidomimetics [110], however, details regarding biological activity are not available.



Niida and co-workers have showed that the Arg-Arg-Nal motif of FC131 could be grafted onto a 3,6-dihydropyridin-2-on scaffold, which was intended as a highly functionalized diketopiperazine mimic, to give a low molecular weight CXCR4 antagonist **15** (Figure 7C) [111]. No further investigation into this scaffold has been published to date.

Ueda et al. have reported a series of 1,2,5-trisubstituted indole derivatives (Figure 7D), where each of the three substituents was introduced as mimics of the Arg<sup>1</sup>-Arg<sup>2</sup>-Nal<sup>3</sup> or Arg<sup>2</sup>-Nal<sup>3</sup>-D-Tyr<sup>5</sup> side chains of FC131 and derivatives [112]. The indole was chosen as scaffold in a ligand-based design approach, where low energy conformations of FC131 and 5-acetamido-1-methylindole-2-carboxamide were compared. The distances between the N-atoms of the amide substituents and the methyl group of the indole derivative was found to match well with the distances between the three  $\beta$ -carbon atoms of the important side chains. The choice of the indole scaffold offers a modular approach to the synthesis of derivatives, and the most active derivative **16**, which contains two guanidine groups and an indole side chains, displayed an IC<sub>50</sub> of 1.2  $\mu$ M. This is still 150-fold less active than FC131, and was proposed by the authors to serve as a useful lead.

Zachariassen et al. have recently reported on the synthesis of scaffold-based tripeptidomimetic CXCR4 antagonists [113], which were designed to mimic the Arg<sup>1</sup>-Arg<sup>2</sup>-Nal<sup>3</sup> motif of the cyclopentapeptides discussed above. After comparison of the proposed bioactive backbone conformation of cyclopentapeptide CXCR4 antagonists with low-energy conformations of the 3,6,8-trisubstituted bicyclic scaffold **A**, CXCR4 antagonists based on **A** and its diastereoisomer **B** (Figure 7E) were pursued. While bicyclic compounds containing an amide bond in the Arg<sup>1</sup> side chain (**17** and its diastereoisomer based on scaffold **A**, Figure 7E) showed no activity within the concentration range tested, the two analogs with an unmodified Arg<sup>1</sup> side chain (**18** and its diastereoisomer based on scaffold **B**, Figure 7E) showed antagonistic activity in the micromolar range in a cell based functional assay for CXCR4. Thus, the

bicyclic scaffolds provide an interesting new starting point for CXCR4 antagonist design.

### **4.3. Small-Molecule Peptides and Peptidomimetics not Derived from FC131**

#### **4.3.1. KRH-series**

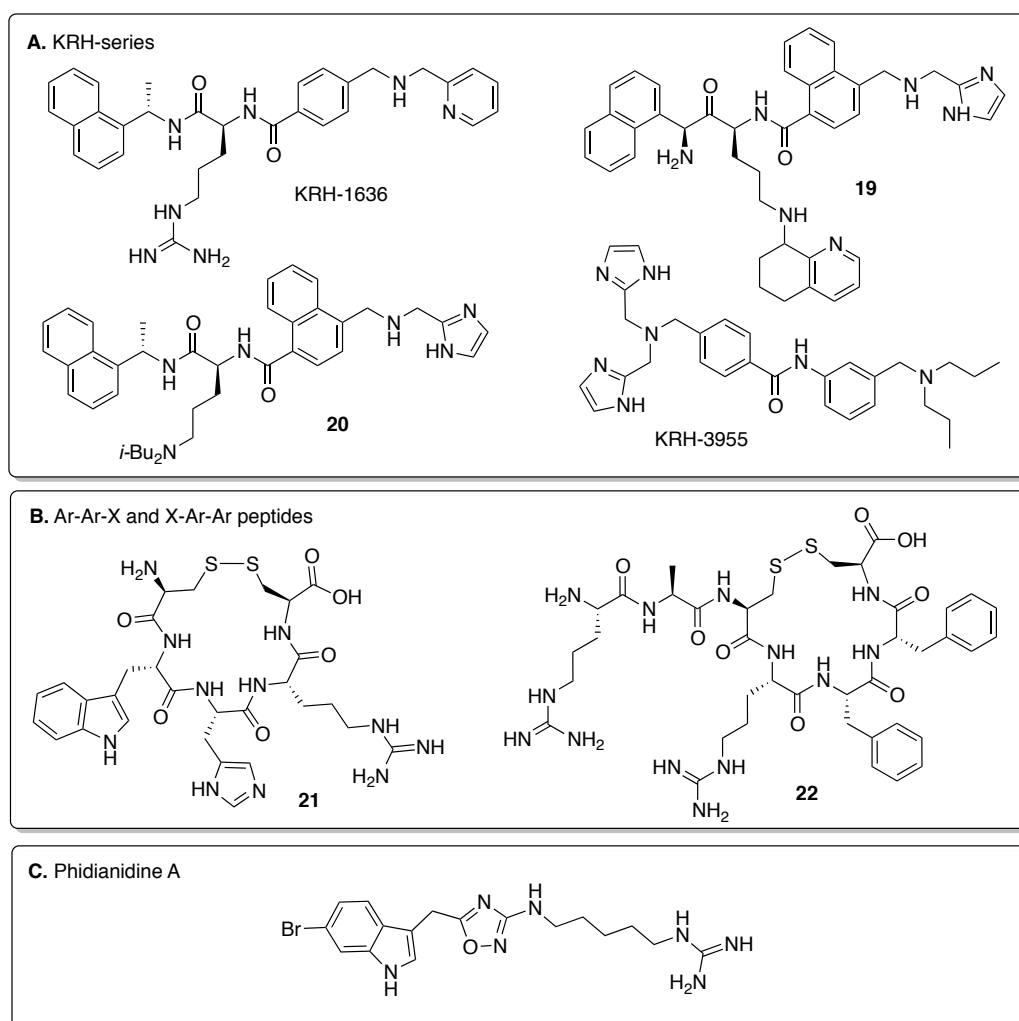
In 2003, Ichiyama and co-workers reported KRH-1636 (Figure 8A) as a non-peptide CXCR4 antagonist with highly potent and selective anti-HIV-1 activity [106]. Structural comparison of low-energy conformations of KRH-1636 with the 3D-pharmacophore model for cyclopentapeptide antagonists later suggested that the (pyridin-2-ylmethyl)amino, guanidine and naphthyl groups of KRH-1636 could mimic the Arg<sup>1</sup>-Arg<sup>2</sup>-2-Nal<sup>3</sup> fragment of FC131 and KRH-1636 is thus considered as a tripeptidomimetic [101]. In the original study, KRH-1636 was found to be duodenally absorbable and was considered as a promising lead. However at a later stage further progression of this compound was abandoned due to low oral bioavailability [114].

After the initial disclosure of KRH-1636, further development of several analogs has been reported, e.g. KRH-2731 and KRH-3148 [115], but the literature is unclear regarding the structure of several of these. Examples of structures that have been assumed as lead structures (**19** and **20**) are shown in Figure 8A [116]. Eventually KRH-3955 (Figure 8A), which does not contain any amino acids, was reported to be an orally bioavailable highly potent inhibitor of X4 HIV-1 replication and CXCL12 mediated chemotaxis [114, 115]. Thus, the KRH-series of compounds have moved from initial peptide-like leads into a more cyclam-like clinical candidate.

#### **4.3.2. Ar-Ar-X and X-Ar-Ar peptides**

Recently, Portella et al. reported an array of short cyclic peptides containing Ar-Ar-Arg or Arg-Ar-Ar motifs as a novel class of CXCR4 antagonists, and several of these (Figure 8B) were found to inhibit growth of cancer cells and reduce the number of metastases in mice [117]. The peptide library was designed based on

the identification of two similar tripeptide motifs (although in reversed order) in the N-terminal tail of vMIP-II (Trp<sup>5</sup>-His<sup>6</sup>-Arg<sup>7</sup>)[64] and in the N-loop of CXCL12 (Arg<sup>12</sup>-Phe<sup>13</sup>-Phe<sup>14</sup>). Cys residues were added at either end of these tripeptide motifs and the structures were cyclized through disulfide formation. Based on modeling studies, the details of which were not revealed, the authors suggest that these peptides bind to the intrahelical site of CXCR4. With respect to the original design concept, the experimental vMIP-II: CXCR4 structure (section 3.1) now provides the details of the interaction between the Trp<sup>5</sup>-His<sup>6</sup>-Arg<sup>7</sup> motif and CXCR4, while the proposed CXCL12: CXCR4 complexes (section 3.2) describe plausible orientations of the Arg<sup>12</sup>-Phe<sup>13</sup>-Phe<sup>14</sup> motif. However, there is no apparent overlap between these two motifs.



**Figure 8.** Linear peptidomimetics and side chain cyclized peptides.

## 5. FUTURE PERSPECTIVE

The literature on CXCR4 is rapidly expanding, defining new roles of CXCR4, or deciphering older roles in more molecular details, in particular within cancer cell migration, i.e. metastasis [118, 119].

Even if X-ray structures of CXCR4 have been available since 2010, the major classes of small-molecule peptide/peptidomimetic CXCR4 antagonists reported to date (section 4) have been developed without knowledge/consideration of the CXCR4 structure, i.e. by ligand-based design.

The experimental CXV15: CXCR4 [5] and vMIP-II: CXCR4 [6] complexes (section 3.1) obviously serve as structural templates for rational design of small-molecule peptidomimetics, and have not yet been properly exploited for this purpose. Similarly, based on the extensive studies of the binding mode for the cyclopentapeptide FC131, the proposed FC131: CXCR4 complex [104] should be a viable foundation for structure-based downsizing and optimization of peptidomimetic ligands. Also, our understanding of the molecular interactions between CXCR4 and the HIV-1 V3 loop has now matured to the point where the recently suggested computational V3: CXCR4 complexes (section 3.3) represent viable starting points for rational design of peptidomimetics based on short V3-motifs.

The continued studies of the interactions between CXCL12 and CXCR4 (section 3.2) will also feed important insight into antagonist design, however; so far the lack of experimental data poses as a challenge for the rational design of small-molecule peptidomimetic ligands. However, based on their model of the CXCL12: CXCR4 complex, Costantini [78] identified Ac-SLSYRC-NH<sub>2</sub> and Ac-YRCPGRF-NH<sub>2</sub> (corresponding to N-terminal residues 4-9 and 7-13 of CXCL12, respectively) as short peptides able to bind to CXCR4 with promising affinity.

The non-liganded X-ray structures of CXCR4 can also be used to identify new hit compounds by virtual screening; the structural nature of such hits will of course

depend on the (virtual) compound library that is screened. For example, Mysinger et al. reported the identification of four novel small-molecule CXCR4 antagonists by virtual screening of 4.2 million molecules from the ZINC database [120]; however, none of these bear any structural resemblance to known peptide ligands and must be considered as functional mimetics. Interestingly, Vitale et al. recently identified the natural product phidianidine A (Figure 8C) as a potential ligand for CXCR4 by employing a partly receptor-based pharmacophore model as well as virtual screening; subsequent *in vitro* studies confirmed phidianidine A to be a CXCR4 antagonist [121]. Structurally, phidianidine A can be considered as a topographical Arg-Trp dipeptidomimetic based on a 3,5-disubstituted 1,2,4-oxadiazole scaffold. Molecular docking studies indicated that the cationic (guanidine) moiety of phidianidine A may interact with Asp97, Asp187, Asp262 or Glu288 while the hydrophobic (bromoindole) moiety involves in aromatic interactions with TM1 and TM2. Thus, phidianidine A shows similar structural features and binding interactions as the peptide CXCR4 antagonists FC131 and T140/CVX15. This study shows that virtual screening of peptide (or peptide-like) libraries also holds potential for identification of novel peptidomimetic CXCR4 antagonists.

## 6. EXECUTIVE SUMMARY

- The established role of the peptidergic 7TM chemokine receptor CXCR4 in e.g. HIV-entry, cancer, and inflammation means that CXCR4 is an attractive target for peptidomimetic drugs.
- The X-ray structures of CXCR4 have provided a reliable structural foundation for computational studies of the molecular interactions between CXCR4 and its endogenous agonist CXCL12 as well as the HIV V3 loop.
- The cyclopentapeptide CXCR4 antagonists represent an excellent starting point for rational development of small-molecule peptidomimetics, and the SAR and binding mode of this compound class have been extensively studied.
- Several attempts to develop smaller and more drug-like scaffold-based peptidomimetic CXCR4 antagonists from the cyclopentapeptides have been reported; however, this transition has proven challenging.
- Future efforts toward peptidomimetic CXCR4 antagonist will benefit from integration of the ligand-based knowledge with the structural information that now is available, i.e. the experimental complexes between CXCR4 and CVX15 and vMIP-II, and the proposed complexes between CXCR4 and CXCL12, HIV V3, and FC131.

## 7. REFERENCES

1. Wermuth G, Ganellin CR, Lindberg P, Mitscher LA. Glossary of terms used in medicinal chemistry (IUPAC recommendations 1998). *Pure Appl. Chem.* 70(5), 1129-1143 (1998).
2. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN). Nomenclature and symbolism for amino acids and peptides. Recommendations 1983. *Eur. J. Biochem.* 138(1), 9-37 (1984).
3. Chorev M. The partial retro-inverso modification: A road traveled together. *Biopolymers* 80(2-3), 67-84 (2005).
4. Ripka AS, Rich DH. Peptidomimetic design. *Curr. Opin. Chem. Biol.* 2(4), 441-452 (1998).
5. Wu B, Chien EYT, Mol CD *et al.* Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* 330(6007), 1066-1071 (2010).  
**\*\* The first experimental structures of a peptidergic 7<sup>TM</sup> receptor, revealing notable local differences compared to already known 7<sup>TM</sup> receptor structures.**
6. Qin L, Kufareva I, Holden LG *et al.* Crystal structure of the chemokine receptor CXCR4 in complex with a viral chemokine. *Science* 347(6226), 1117-1122 (2015).  
**\*\* The first crystal structure of a chemokine receptor (CXCR4) in complex with a chemokine (the viral vMIP-II).**
7. <http://www.guidetopharmacology.org>
8. Luster AD. Chemokines - chemotactic cytokines that mediate inflammation. *New Engl. J. Med.* 338(7), 436-445 (1998).
9. Nagasawa T, Nakajima T, Tachibana K *et al.* Molecular cloning and characterization of a murine pre-B-cell growth-stimulating factor/stromal cell-derived factor 1 receptor, a murine homolog of the human immunodeficiency virus 1 entry coreceptor fusin. *Proc. Natl. Acad. Sci. USA* 93(25), 14726-14729 (1996).
10. Endres MJ, Clapham PR, Marsh M *et al.* CD4-independent infection by HIV-2 is mediated by fusin/CXCR4. *Cell* 87(4), 745-756 (1996).
11. Lapham CK, Ouyang J, Chandrasekhar B, Nguyen NY, Dimitrov DS, Golding H. Evidence for cell-surface association between fusin and the CD4-gp120 complex in human cell lines. *Science* 274(5287), 602-605 (1996).
12. Oberlin E, Amara A, Bachelier F *et al.* The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature* 382(6594), 833-835 (1996).
13. Bachelier F, Ben-Baruch A, Burkhardt AM *et al.* International Union of Pharmacology. LXXXIX. Update on the extended family of chemokine receptors and introducing a new nomenclature for atypical chemokine receptors. *Pharmacol. Rev.* 66(1), 1-79 (2014).
14. Bajetto A, Bonavia R, Barbero S, Florio T, Costa A, Schettini G. Expression of chemokine receptors in the rat brain. *Ann. N.Y. Acad. Sci.* 876(1), 201-209 (1999).
15. Han Y, Wang J, He T, Ransohoff RM. TNF- $\alpha$  down-regulates CXCR4 expression in primary murine astrocytes. *Brain Res.* 888(1), 1-10 (2001).

16. Barbero S, Bajetto A, Bonavia R *et al.* Expression of the chemokine receptor CXCR4 and its ligand stromal cell-derived factor 1 in human brain tumors and their involvement in glial proliferation *in vitro*. *Ann. N.Y. Acad. Sci.* 973(1), 60-69 (2002).
17. Ma Q, Jones D, Borghesani PR *et al.* Impaired B-lymphopoiesis, myelopoiesis, and derailed cerebellar neuron migration in CXCR4- and SDF-1-deficient mice. *Proc. Natl. Acad. Sci. USA* 95(16), 9448-9453 (1998).
18. Scotton CJ, Wilson JL, Milliken D, Stamp G, Balkwill FR. Epithelial cancer cell migration: A role for chemokine receptors? *Cancer Res.* 61(13), 4961-4965 (2001).
19. Kulbe H, Levinson NR, Balkwill F, Wilson JL. The chemokine network in cancer - much more than directing cell movement. *Int. J. Dev. Biol.* 48(5-6), 489-496 (2004).
20. Balkwill F. The significance of cancer cell expression of the chemokine receptor CXCR4. *Semin. Cancer Biol.* 14(3), 171-179 (2004).
21. Balkwill F. Cancer and the chemokine network. *Nat. Rev. Cancer* 4(7), 540-550 (2004).
22. Muller A, Homey B, Soto H *et al.* Involvement of chemokine receptors in breast cancer metastasis. *Nature* 410(6824), 50-56 (2001).
23. Terasaki M, Sugita Y, Arakawa F, Okada Y, Ohshima K, Shigemori M. CXCL12/CXCR4 signaling in malignant brain tumors: A potential pharmacological therapeutic target. *Brain Tumor Pathol.* 28(2), 89-97 (2011).
24. Kuil J, Buckle T, Van Leeuwen FWB. Imaging agents for the chemokine receptor 4 (CXCR4). *Chem. Soc. Rev.* 41(15), 5239-5261 (2012).
25. Knight JC, Wuest FR. Nuclear (PET/SPECT) and optical imaging probes targeting the CXCR4 chemokine receptor. *Med. Chem. Commun.* 3(9), 1039-1053 (2012).
26. George GP, Pisaneschi F, Nguyen QD, Aboagye EO. Positron emission tomographic imaging of CXCR4 in cancer: Challenges and promises. *Mol. Imaging* 13, 1-19 (2014).
27. Larochelle A, Krouse A, Metzger M *et al.* AMD3100 mobilizes hematopoietic stem cells with long-term repopulating capacity in nonhuman primates. *Blood* 107(9), 3772-3778 (2006).
28. Masyuk M, Abduelmula A, Morosan-Puopolo G *et al.* Retrograde migration of pectoral girdle muscle precursors depends on CXCR4/SDF-1 signaling. *Histochem. Cell Biol.* 142(5), 473-488 (2014).
29. Masyuk M, Brand-Saberi B. *Recruitment of skeletal muscle progenitors to secondary sites: A role for CXCR4/SDF-1 signalling in skeletal muscle development.* In: *Vertebrate Myogenesis*, Brand-Saberi B (Ed.). Springer, Berlin Heidelberg, 1-23 (2015).
30. Glawe JD, Mijalis EM, Davis WC *et al.* SDF-1-CXCR4 differentially regulates autoimmune diabetogenic T cell adhesion through ROBO1-SLIT2 interactions in mice. *Diabetologia* 56(10), 2222-2230 (2013).
31. Fujiyama T, Ito T, Ogawa N, Suda T, Tokura Y, Hashizume H. Preferential infiltration of interleukin-4-producing CXCR4+ T cells in the lesional muscle but not skin of patients with dermatomyositis. *Clin. Exp. Immunol.* 177(1), 110-120 (2014).



32. Wang A, Guilpain P, Chong BF *et al.* Dysregulated expression of CXCR4/CXCL12 in subsets of patients with systemic lupus erythematosus. *Arthritis Rheum.* 62(11), 3436-3446 (2010).
33. Launay O, Paul S, Servettaz A *et al.* Control of humoral immunity and autoimmunity by the CXCR4/CXCL12 axis in lupus patients following influenza vaccine. *Vaccine* 31(35), 3492-3501 (2013).
34. O'Brien WA, Sumner-Smith M, Mao SH, Sadeghi S, Zhao JQ, Chen IS. Anti-human immunodeficiency virus type 1 activity of an oligocationic compound mediated via gp120 V3 interactions. *J. Virol.* 70(5), 2825-2831 (1996).
35. Doranz BJ, Grovit-Ferbas K, Sharron MP *et al.* A small-molecule inhibitor directed against the chemokine receptor CXCR4 prevents its use as an HIV-1 coreceptor. *J. Exp. Med.* 186(8), 1395-1400 (1997).
36. Doranz BJ, Filion LG, Diaz-Mitoma F *et al.* Safe use of the CXCR4 inhibitor ALX40-4C in humans. *AIDS Res. Hum. Retroviruses* 17(6), 475-486 (2001).
37. Nakashima H, Masuda M, Murakami T *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.* 36(6), 1249-1255 (1992).
38. Feng Y, Broder CC, Kennedy PE, Berger EA. HIV-1 entry cofactor: Functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* 272(5263), 872-877 (1996).
39. Murakami T, Nakajima T, Koyanagi Y *et al.* A small molecule CXCR4 inhibitor that blocks T cell line-tropic HIV-1 infection. *J. Exp. Med.* 186(8), 1389-1393 (1997).
40. Tamamura H, Xu Y, Hattori T *et al.* A low-molecular-weight inhibitor against the chemokine receptor CXCR4: A strong anti-HIV peptide T140. *Biochem. Biophys. Res. Commun.* 253(3), 877-882 (1998).
41. Crump MP, Gong JH, Loetscher 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.* 16(23), 6996-7007 (1997).
42. Loetscher P, Gong J-H, Dewald B, Baggiolini M, Clark-Lewis I. N-terminal peptides of stromal cell-derived factor-1 with CXC chemokine receptor 4 agonist and antagonist activities. *J. Biol. Chem.* 273(35), 22279-22283 (1998).
43. De Clercq E. The bicyclam AMD3100 story. *Nat. Rev. Drug Discov.* 2(7), 581-587 (2003).
44. <https://clinicaltrials.gov>
45. Grande F, Garofalo A, Neamati N. Small molecules anti-HIV therapeutics targeting CXCR4. *Curr. Pharm. Des.* 14(4), 385-404 (2008).
46. Debnath B, Xu S, Grande F, Garofalo A, Neamati N. Small molecule inhibitors of CXCR4. *Theranostics* 3(1), 47-75 (2013).
47. Oishi S, Fujii N. Peptide and peptidomimetic ligands for CXC chemokine receptor 4 (CXCR4). *Org. Biomol. Chem.* 10, 5720-5731 (2012).
48. Zhang WB, Navenot JM, Haribabu 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.* 277(27), 24515-24521 (2002).

49. Fujii N, Oishi S, Hiramatsu K *et al.* Molecular-size reduction of a potent CXCR4-chemokine antagonist using orthogonal combination of conformation- and sequence-based libraries. *Angew. Chem. Int. Ed.* 42(28), 3251-3253 (2003).
- \*\* Reports the discovery of the cyclopentapeptide CXCR4 antagonists as well as the first structure-activity data and a solution structure for the lead compound FC131.**
50. Kufareva I, Rueda M, Katritch V, Stevens RC, Abagyan R. Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure* 19(8), 1108-1126 (2011).
51. Maestro, version 10.0, Schrödinger, LLC, New York, NY, 2014.
52. Demarco SJ, Henze H, Lederer A *et al.* Discovery of novel, highly potent and selective  $\beta$ -hairpin mimetic CXCR4 inhibitors with excellent anti-HIV activity and pharmacokinetic profiles. *Bioorg. Med. Chem.* 14(24), 8396-8404 (2006).
53. Ruiz-Gomez G, Tyndall JD, Pfeiffer B, Abbenante G, Fairlie DP. Update 1 of: Over one hundred peptide-activated G protein-coupled receptors recognize ligands with turn structure. *Chem. Rev.* 110(4), PR1-41 (2010).
54. Tamamura H, Sugioka M, Odagaki Y *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.* 11(3), 359-362 (2001).
55. Tamamura H, Omagari A, Oishi S *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.* 10(23), 2633-2637 (2000).
56. Gerlach LO, Skerlj RT, Bridger GJ, Schwartz TW. Molecular interactions of cyclam and bicyclam non-peptide antagonists with the CXCR4 chemokine receptor. *J. Biol. Chem.* 276(17), 14153-14160 (2001).
57. Hatse S, Princen K, Vermeire K *et al.* Mutations at the CXCR4 interaction sites for AMD3100 influence anti-CXCR4 antibody binding and HIV-1 entry. *FEBS Lett.* 546(2-3), 300-306 (2003).
58. Rosenkilde MM, Gerlach L-O, Jakobsen JS, Skerlj RT, Bridger GJ, Schwartz TW. Molecular mechanism of AMD3100 antagonism in the CXCR4 receptor: Transfer of binding site to the CXCR3 receptor. *J. Biol. Chem.* 279(4), 3033-3041 (2004).
59. Fricker SP, Anastassov V, Cox J *et al.* Characterization of the molecular pharmacology of AMD3100: A specific antagonist of the G-protein coupled chemokine receptor, CXCR4. *Biochem. Pharmacol.* 72(5), 588-596 (2006).
60. Rosenkilde MM, Gerlach L-O, Hatse S *et al.* Molecular mechanism of action of monocyclam *versus* bicyclam non-peptide antagonists in the CXCR4 chemokine receptor. *J. Biol. Chem.* 282(37), 27354-27365 (2007).
61. Wong RSY, Bodart V, Metz M, Labrecque J, Bridger G, Fricker SP. Comparison of the potential multiple binding modes of bicyclam, monocyclam, and noncyclam small-molecule CXC chemokine receptor 4 inhibitors. *Mol. Pharmacol.* 74(6), 1485-1495 (2008).
62. Cox BD, Prosser AR, Katzman BM *et al.* Anti-HIV small-molecule binding in the peptide subpocket of the CXCR4:CVX15 crystal structure. *ChemBioChem* 15(11), 1614-1620 (2014).

63. Kledal TN, Rosenkilde MM, Coulin F *et al.* A broad-spectrum chemokine antagonist encoded by Kaposi's sarcoma-associated herpesvirus. *Science* 277(5332), 1656-1659 (1997).
64. Scholten DJ, Canals M, Maussang D *et al.* Pharmacological modulation of chemokine receptor function. *Br. J. Pharmacol.* 165(6), 1617-1643 (2012).  
\*\* **Comprehensive review of the molecular pharmacology of the chemokine receptor family.**
65. Schwartz TW, Frimurer TM, Holst B, Rosenkilde MM, Elling CE. Molecular mechanism of 7TM receptor activation - a global toggle switch model. *Annu. Rev. Pharmacol. Toxicol.* 46(1), 481-519 (2006).
66. Venkatakrishnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF, Babu MM. Molecular signatures of G-protein-coupled receptors. *Nature* 494(7436), 185-194 (2013).
67. Manglik A, Kobilka B. The role of protein dynamics in GPCR function: Insights from the  $\beta$ 2AR and rhodopsin. *Curr. Opin. Cell Biol.* 27, 136-143 (2014).
68. Dealwis C, Fernandez EJ, Thompson DA, Simon RJ, Siani MA, Lolis E. Crystal structure of chemically synthesized [N33A] stromal cell-derived factor 1 $\alpha$ , a potent ligand for the HIV-1 "fusin" coreceptor. *Proc. Natl. Acad. Sci. USA* 95(12), 6941-6946 (1998).
69. Thiele S, Rosenkilde MM. Interaction of chemokines with their receptors - from initial chemokine binding to receptor activating steps. *Curr. Med. Chem.* 21(31), 3594-3614 (2014).
70. Dror RO, Pan AC, Arlow DH *et al.* Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl. Acad. Sci. USA* 108(32), 13118-13123 (2011).
71. Dror RO, Arlow DH, Maragakis P *et al.* Activation mechanism of the  $\beta$ 2-adrenergic receptor. *Proc. Natl. Acad. Sci. USA* 108(46), 18684-18689 (2011).
72. Shan J, Khelashvili G, Mondal S, Mehler EL, Weinstein H. Ligand-dependent conformations and dynamics of the serotonin 5-HT<sub>2A</sub> receptor determine its activation and membrane-driven oligomerization properties. *PLoS Comput. Biol.* 8(4), e1002473 (2012).
73. Kofuku Y, Yoshiura C, Ueda T *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.* 284(50), 35240-35250 (2009).
74. Kufareva I, Stephens BS, Holden LG *et al.* Stoichiometry and geometry of the CXC chemokine receptor 4 complex with CXC ligand 12: Molecular modeling and experimental validation. *Proc. Natl. Acad. Sci. USA* 111(50), E5363-E5372 (2014).
75. Xu L, Li Y, Sun H, Li D, Hou T. Structural basis of the interactions between CXCR4 and CXCL12/SDF-1 revealed by theoretical approaches. *Mol. BioSyst.* 9(8), 2107-2117 (2013).
76. Tamamis P, Floudas CA. Elucidating a key component of cancer metastasis: CXCL12 (SDF-1 $\alpha$ ) binding to CXCR4. *J. Chem. Inf. Model.* 54(4), 1174-1188 (2014).

77. Veldkamp CT, Seibert C, Peterson FC *et al.* Structural basis of CXCR4 sulfotyrosine recognition by the chemokine SDF-1/CXCL12. *Sci. Signal.* 1(37), ra4 (2008).
  78. Costantini S, Raucci R, Colonna G *et al.* Peptides targeting chemokine receptor CXCR4: Structural behavior and biological binding studies. *J. Pept. Sci.* 20(4), 270-278 (2014).
  79. Klasse PJ. The molecular basis of HIV entry. *Cell. Microbiol.* 14(8), 1183-1192 (2012).
  80. Huang C-C, Tang M, Zhang M-Y *et al.* Structure of a V3-containing HIV-1 gp120 core. *Science* 310(5750), 1025-1028 (2005).
  81. Huang CC, Lam SN, Acharya P *et al.* Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. *Science* 317(5846), 1930-1934 (2007).
  82. Tan Q, Zhu Y, Li J *et al.* Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science* 341(6152), 1387-1390 (2013).
  83. Tamamis P, Floudas CA. Molecular recognition of CXCR4 by a dual tropic HIV-1 gp120 V3 loop. *Biophys. J.* 105(6), 1502-1514 (2013).
  84. Rosenkilde MM, Schwartz TW. GluVII:06 - a highly conserved and selective anchor point for non-peptide ligands in chemokine receptors. *Curr. Top. Med. Chem.* 6(13), 1319-1333 (2006).
  85. Stradley SJ, Rizo J, Bruch MD, Stroup AN, Gierasch LM. Cyclic pentapeptides as models for reverse turns: Determination of the equilibrium distribution between type I and type II conformations of Pro-Asn and Pro-Ala beta-turns. *Biopolymers* 29(1), 263-287 (1990).
  86. Marshall GR. A hierarchical approach to peptidomimetic design. *Tetrahedron* 49(17), 3547-3558 (1993).
  87. Tamamura H, Hiramatsu K, Ueda S *et al.* Stereoselective synthesis of [L-Arg-L/D-3-(2-naphthyl)alanine]-type (E)-alkene dipeptide isosteres and its application to the synthesis and biological evaluation of pseudopeptide analogues of the CXCR4 antagonist FC131. *J. Med. Chem.* 48(2), 380-391 (2005).
  88. Tamamura H, Araki T, Ueda S *et al.* Identification of novel low molecular weight CXCR4 antagonists by structural tuning of cyclic tetrapeptide scaffolds. *J. Med. Chem.* 48(9), 3280-3289 (2005).
  89. Ueda S, Oishi S, Wang ZX *et al.* Structure-activity relationships of cyclic peptide-based chemokine receptor CXCR4 antagonists: Disclosing the importance of side-chain and backbone functionalities. *J. Med. Chem.* 50(2), 192-198 (2007).
  90. Kobayashi K, Oishi S, Hayashi R *et al.* Structure-activity relationship study of a CXC chemokine receptor type 4 antagonist, FC131, using a series of alkene dipeptide isosteres. *J. Med. Chem.* 55(6), 2746-2757 (2012).
  91. Demmer O, Frank AO, Hagn F *et al.* A conformationally frozen peptoid boosts CXCR4 affinity and anti-HIV activity. *Angew. Chem. Int. Ed.* 51(32), 8110-8113 (2012).
- \* Elegant example of affinity/potency increase for the cyclopentapeptide CXCR4 antagonists, achieved through conformational stabilization.**

92. Tamamura H, Mizumoto M, Hiramatsu K *et al.* Topochemical exploration of potent compounds using retro-enantiomer libraries of cyclic pentapeptides. *Org. Biomol. Chem.* 2(8), 1255-1257 (2004).
  93. Narumi T, Hayashi R, Tomita K *et al.* Synthesis and biological evaluation of selective CXCR4 antagonists containing alkene dipeptide isosteres. *Org. Biomol. Chem.* 8(3), 616-621 (2010).
  94. Narumi T, Tomita K, Inokuchi E *et al.* Diastereoselective synthesis of highly functionalized fluoroalkene dipeptide isosteres and its application to Fmoc-based solid-phase synthesis of a cyclic pentapeptide mimetic. *Tetrahedron* 64(19), 4332-4346 (2008).
  95. Inokuchi E, Oishi S, Kubo T *et al.* Potent CXCR4 antagonists containing amidine type peptide bond isosteres. *ACS Med. Chem. Lett.* 2(6), 477-480 (2011).
  96. Tanaka T, Nomura W, Narumi T *et al.* Structure-activity relationship study on artificial CXCR4 ligands possessing the cyclic pentapeptide scaffold: The exploration of amino acid residues of pentapeptides by substitutions of several aromatic amino acids. *Org. Biomol. Chem.* 7(18), 3805-3809 (2009).
  97. Mungalpara J, Thiele S, Eriksen Ø, Eksteen J, Rosenkilde MM, Våbenø J. Rational design of conformationally constrained cyclopentapeptide antagonists for C-X-C chemokine receptor 4 (CXCR4). *J. Med. Chem.* 55(22), 10287-10291 (2012).
  98. Tamamura H, Esaka A, Ogawa T *et al.* Structure-activity relationship studies on CXCR4 antagonists having cyclic pentapeptide scaffolds. *Org. Biomol. Chem.* 3(24), 4392-4394 (2005).
  99. Mungalpara J, Zachariassen ZG, Thiele S, Rosenkilde MM, Våbenø J. Structure-activity relationship studies of the aromatic positions in cyclopentapeptide CXCR4 antagonists. *Org. Biomol. Chem.* 11, 8202-8208 (2013).
  100. Tanaka T, Tsutsumi H, Nomura W *et al.* Structure-activity relationship study of CXCR4 antagonists bearing the cyclic pentapeptide scaffold: Identification of the new pharmacophore. *Org. Biomol. Chem.* 6(23), 4374-4377 (2008).
  101. Våbenø J, Nikiforovich GV, Marshall GR. A minimalistic 3D pharmacophore model for cyclopentapeptide CXCR4 antagonists. *Biopolymers* 84(5), 459-471 (2006).
  102. Demmer O, Dijkgraaf I, Schumacher U *et al.* Design, synthesis, and functionalization of dimeric peptides targeting chemokine receptor CXCR4. *J. Med. Chem.* 54(21), 7648-7662 (2011).
  103. Yoshikawa Y, Kobayashi K, Oishi S, Fujii N, Furuya T. Molecular modeling study of cyclic pentapeptide CXCR4 antagonists: New insight into CXCR4-FC131 interactions. *Bioorg. Med. Chem. Lett.* 22(6), 2146-2150 (2012).
  104. Thiele S, Mungalpara J, Steen A, Rosenkilde MM, Våbenø J. Determination of the binding mode for the cyclopentapeptide CXCR4 antagonist FC131 using a dual approach of ligand modifications and receptor mutagenesis. *Br. J. Pharmacol.* 171(23), 5313-5329 (2014).
- \* The first binding mode study of FC131 that is based on *in vitro* experiments.**

105. Tamamura H, Tsutsumi H, Masuno H *et al.* Development of a linear type of low molecular weight CXCR4 antagonists based on T140 analogs. *Org. Biomol. Chem.* 4(12), 2354-2357 (2006).
106. Ichiyama K, Yokoyama-Kumakura S, Tanaka Y *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. USA* 100(7), 4185-4190 (2003).
107. Tamamura H, Hiramatsu K, Mizumoto M *et al.* Enhancement of the T140-based pharmacophores leads to the development of more potent and bio-stable CXCR4 antagonists. *Org. Biomol. Chem.* 1(21), 3663-3669 (2003).
108. Narumi T, Tanaka T, Hashimoto C *et al.* Pharmacophore-based small molecule CXCR4 ligands. *Bioorg. Med. Chem. Lett.* 22(12), 4169-4172 (2012).
109. Cluzeau J, Oishi S, Ohno H *et al.* Design and synthesis of all diastereomers of cyclic pseudo-dipeptides as mimics of cyclic CXCR4 pentapeptide antagonists. *Org. Biomol. Chem.* 5(12), 1915-1923 (2007).
110. Marshall G: US20050192272. (2005).
111. Niida A, Tanigaki H, Inokuchi E *et al.* Stereoselective synthesis of 3,6-disubstituted-3,6-dihydropyridin-2-ones as potential diketopiperazine mimetics using organocopper-mediated anti-S<sub>N</sub>2' reactions and their use in the preparation of low-molecule CXCR4 antagonists. *J. Org. Chem.* 71(10), 3942-3951 (2006).
112. Ueda S, Kato M, Inuki S *et al.* Identification of novel non-peptide CXCR4 antagonists by ligand-based design approach. *Bioorg. Med. Chem. Lett.* 18(14), 4124-4129 (2008).
113. Zachariassen ZG, Thiele S, Berg EA *et al.* Design, synthesis, and biological evaluation of scaffold-based tripeptidomimetic antagonists for CXC chemokine receptor 4 (CXCR4). *Bioorg. Med. Chem.* 22(17), 4759-4769 (2014).
114. Murakami T, Kumakura S, Yamazaki 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.* 53(7), 2940-2948 (2009).
115. Iwasaki Y, Akari H, Murakami T *et al.* Efficient inhibition of SDF-1 $\alpha$ -mediated chemotaxis and HIV-1 infection by novel CXCR4 antagonists. *Cancer Sci.* 100(4), 778-781 (2009).
116. Mosley CA, Wilson LJ, Wiseman JM, Skudlarek JW, Liotta DC. Recent patents regarding the discovery of small molecule CXCR4 antagonists. *Expert Opin. Ther. Pat.* 19(1), 23-38 (2009).
117. Portella L, Vitale R, De Luca S *et al.* Preclinical development of a novel class of CXCR4 antagonist impairing solid tumors growth and metastases. *PloS One* 8(9), e74548 (2013).
118. Cojoc M, Peitzsch C, Trautmann F, Polishchuk L, Telegeev GD, Dubrovskaya A. Emerging targets in cancer management: Role of the CXCL12/CXCR4 axis. *OncoTargets Ther.* 6, 1347 (2013).
119. Chatterjee S, Behnam Azad B, Nimmagadda S. The intricate role of CXCR4 in cancer. *Adv. Cancer Res.* 124, 31-82 (2014).

120. Mysinger MM, Weiss DR, Ziarek JJ *et al.* Structure-based ligand discovery for the protein-protein interface of chemokine receptor CXCR4. *Proc. Natl. Acad. Sci. USA* 109(14), 5517-5522 (2012).
121. Vitale RM, Gatti M, Carbone M *et al.* Minimalist hybrid ligand/receptor-based pharmacophore model for CXCR4 applied to a small-library of marine natural products led to the identification of phidianidine A as a new CXCR4 ligand exhibiting antagonist activity. *ACS Chem. Biol.* 8(12), 2762-2770 (2013).

**\* Recent example of identification of a new peptidomimetic ligand for CXCR4 by combining ligand- and structure-based approaches.**