UiT
THE ARCTIC
UNIVERSITY
OF NORWAY

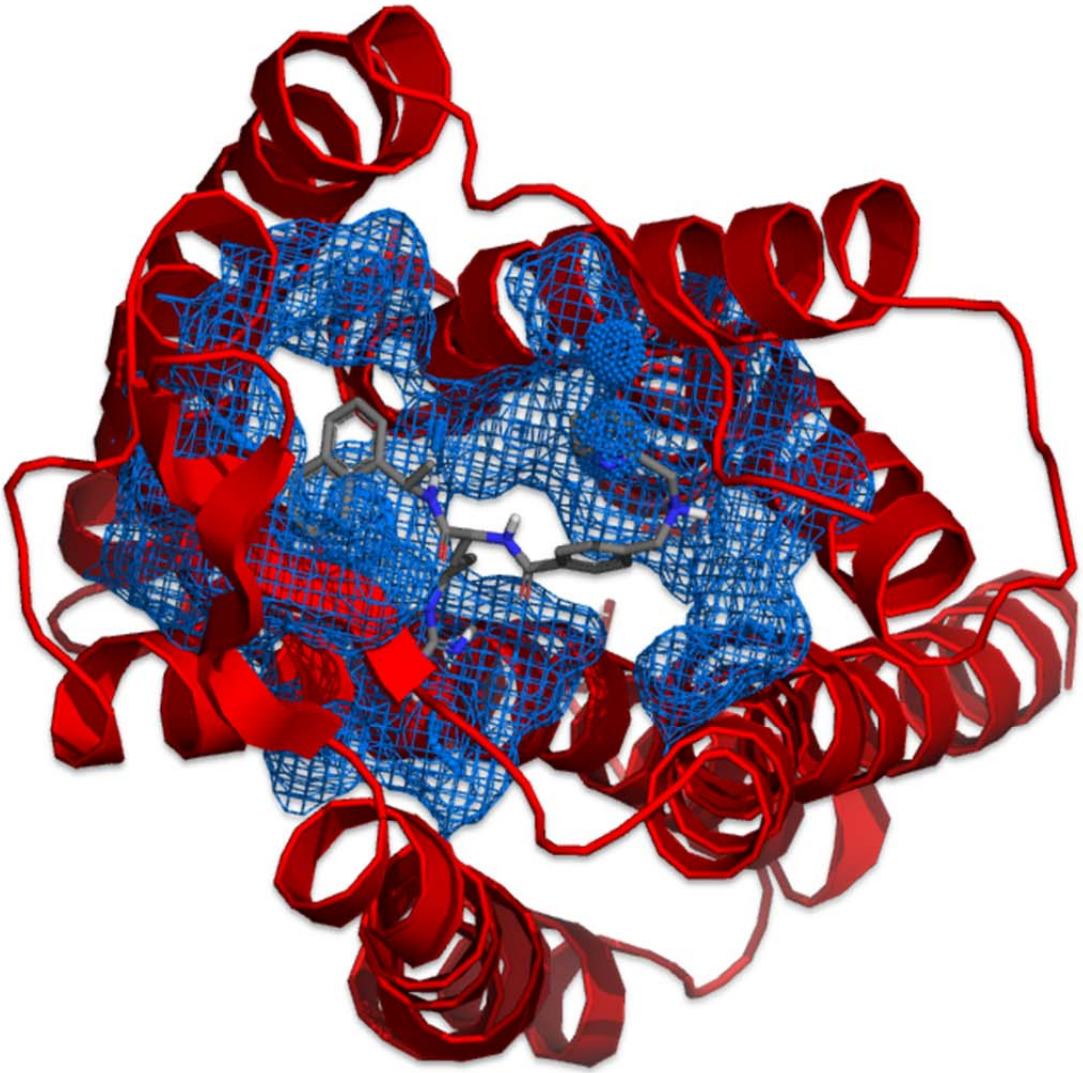
FACULTY OF HEALTH SCIENCES
DEPARTMENT OF PHARMACY

# Toward Tripeptidomimetic CXCR4 Antagonists:

## Design, Synthesis, Biological Evaluation, and Binding Mode Studies

**Zack George Zachariassen**

*A dissertation for the degree of Philosophiae Doctor*
*April 2014*



UNIVERSITY OF TROMSØ

## TABLE OF CONTENTS

**LIST OF PAPERS**

For papers, see the Appendix*.

I. Mungalpara, J.,‡ Zachariassen, Z.G.,‡ Thiele, S., Rosenkilde, M.M. and Våbenø, J. Structure-activity relationship studies of the aromatic positions in cyclopentapeptide CXCR4 antagonists. *Org. Biomol. Chem.* **2013**; 11, 8202-8208. *Published.*

II. Zachariassen, Z.G., Thiele, S., Berg, E.A., Rasmussen, P., Fossen T., Rosenkilde, M.M., Våbenø, J., Haug, B.E. Design, synthesis, and biological evaluation of scaffold-based tripeptidomimetic antagonists for CXC chemokine receptor 4 (CXCR4). *Submitted.*

III. Zachariassen, Z.G., Thiele, S., Haug, B.E., Rosenkilde, M.M., and Våbenø, J. Probing the molecular interactions between C-X-C chemokine receptor 4 (CXCR4) and the tripeptidomimetic antagonist KRH-1636. *Manuscript*.

* Reproduced with permission.

‡ These authors contributed equally.

**ABBREVIATIONS**

| | |
|---|---|
| 1-Nal | L-3-(1-Naphthyl)-alanine |
| 2-Nal | L-3-(2-Naphthyl)-alanine |
| 2Cl-Trt | 2-chlorotrityl |
| 3-Apa | Aminopropanoic acid |
| 4-Aba | Aminobutanoic acid |
| 5-Apa | Aminopentanoic acid |
| 6-Aha | Aminohexanoic acid |
| Aic | 2-Aminoindan-2-carboxylic acid |
| AIDS | Acquired Immune Deficiency Syndrome |
| β-Ala | L-2,4-Diaminobutyric acid |
| Boc | Di-*tert*-butyl dicarbonate |
| Bph | L-4-Benzoyl-phenylalanine |
| Bsa | L-β-Styryl-alanine |
| CD4 | Cluster of differentiation 4 |
| Cha | L-3-Cyclohexyl-alanine |
| Cit | L-Citrulline |
| CXCR4 | CXC chemokine receptor 4 |
| CXCL12 | CXC chemokine ligand 12 |
| DIPEA | *N,N*-Diisopropylethylamine |
| DMF | *N,N*-Dimethylformamide |
| DCM | Dichloromethane |
| ECL | Extracellular loop |
| Fmoc | 9-Fluorenylmethyloxycarbonyl |
| HRMS | High Resolution Mass Spectrometry |

| | |
|---|---|
| HATU | 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate |
| HBTU | O-(Benzotriazol-1-yl)-N,N,N′,N′-tetramethyluronium hexafluorophosphate |
| Hch | 2-Amino-4-cyclohexylbutanoic acid |
| HFIP | 1,1,1,3,3,3-Hexafluoro-2-propanol |
| Hph | L-2-Amino-4-phenylbutanoic acid |
| HTS | High-Throughput Screening |
| NMR | Nuclear magnetic resonance |
| Ppr | 3-phenylproline |
| Phg | α-Phenylglycine |
| Pic | L-Pipecolic acid |
| Pph | L-4-Phenyl-phenylalanine |
| Pbf | 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl |
| PBS | Phosphate buffered saline |
| ROESY | Rotating frame nuclear Overhauser effect spectroscopy |
| RP-HPLC | Reverse-phase High Performance Liquid Chromatography |
| SDF1α | Stromal cell-derived factor-1 α |
| SPA | Scintillation proximity assay |
| SPPS | Solid phase peptide synthesis |
| TFA | Trifluoroacetic acid |
| Tic | 1,2,3,4-Tetrahydroisoquinoline-3-carboxylic acid |
| TIS | Triisopropylsilane |
| TM | Transmembrane |
| TMH | Transmembrane helix |
| WT | Wild type |
| Xaa | Any amino acid |

**SUMMARY**

CXCR4 is a G protein-coupled chemokine receptor that transduces signals of its endogenous ligand CXCL12 (SDF-1α). The involvement of human CXCR4 in several pathological conditions including HIV/AIDS and cancer has stimulated the search for small-molecule CXCR4 antagonists. Cyclopentapeptides based on the $Arg^1$-$Arg^2$-2-$Nal^3$-$Gly^4$-D-$Tyr^5$ sequence are potent CXCR4 antagonists, and an excellent starting point for development of peptidomimetics, *i.e.* compounds that contain non-peptidic structural elements and are capable of mimicking the biological action of a natural parent peptide.

In the present project, the pharmacophore for the lead cyclopentapeptide CXCR4 antagonist FC131 was first refined through structure-activity relationship (SAR) studies of its two aromatic positions. While the D-tyrosine side chain in position 5 was found dispensable, the 2-naphthylalanine side chain in position 3 was shown to be important for the antagonistic activity of the cyclopentapeptide analogues.

Encouraged by this SAR data, which suggest that the activity of cyclopentapeptide CXCR4 antagonists mainly resides in the tripeptide D-/L-Arg-L-Arg-2-Nal fragment, a novel class of scaffold-based tripeptidomimetics were next designed and synthesized. These prototype tripeptidomimetics were found to represent new peptidomimetic hits, and subsequent studies aiming to optimize the prototype compounds have been pursued.

Finally, the binding mode for the known tripeptidic CXCR4 antagonist KRH-1636 was investigated through a ternary strategy combining SAR-, site-directed mutagenesis (SDM) studies, and molecular docking to the X-ray structure of CXCR4. Comparison of the derived binding model for KRH-1636 with the reported binding mode for the cyclopentapeptide antagonist FC131 showed that the two compounds bind to the receptor in different ways; thus, KRH-1636 is not a mimetic of FC131.

Collectively, the findings from the present project provide a foundation for future design of optimized small-molecule peptidomimetic CXCR4 antagonists.

## 1. INTRODUCTION

A number of important physiological and biochemical functions of life are influenced by peptides. Endogenous peptides are involved as neurotransmitters, neuromodulators and as hormones in receptor-mediated signal transduction affecting the nervous and immune system, but also the functions of the intestinal and cardiovascular systems.[1-4] The apparent plethora in the modes of action of bioactive peptides has led to an increased interest in their potential as drugs for the treatment of several pathological conditions.[5-8] A central target in drug discovery are G-protein coupled receptors (GPCRs) and their signaling pathways.[9]

In spite of their wide application today, peptide based drugs have shortcomings, often reflected in poor pharmacokinetic properties. Therefore, peptide mimetic ligands represent an alternative path in drug discovery by providing potential drug candidates with improved properties.[10]

Peptide mimetics can be developed from a bioactive peptide precursor is a systematic manner, and this thesis describes the ligand-based design, synthesis, and SAR for small-molecule mimics derived from a lead cyclopentapeptide antagonist for the G protein-coupled chemokine receptor CXCR4. Moreover, this project extends to provide insights into ligand-receptor interactions by investigating the binding mode of a known potent tripeptidomimetic CXCR4 antagonist.

An introduction to various concepts and approaches applied in drug design is provided in Chapters 2 and 3. Existing background knowledge for the present project is given in Chapters 4-7, and the conducted work is described in Chapters 8-13.

## 2. DRUG DEVELOPMENT APPROACHES

### 2.1 Drug discovery and development

A *drug* can be defined as *"any substance presented for treating, curing or preventing disease in human beings or in animals, for making a medical diagnosis or for restoring, correcting, or modifying physiological functions."*[11] Each drug may also be classified by the chemical type, structure or origin into one or more categories. Protein-based drugs (biologic agents),[12] peptides, and small organic molecule[13, 14] drugs represent some prominent examples.

Drug development involves the discovery or design of chemical compounds that interact with a biological target to produce a beneficial effect. The pharmaceutical industry has embraced more automated drug discovery-approaches such as high-throughput organic synthesis and high-throughput screening (HTS) of large numbers of compounds to a great extent.[15] However, the approach of screening more and more compounds at increasingly faster rates has not turned as fruitful as the industry hoped.[16] Drug development is slow and expensive; in the rare instances that a drug makes it through phase I clinical trials, it was estimated to cost the manufacturer close to a billion US dollars.[9] The current drug discovery paradigm can be synopsized in early phases comprising hit-identification, hit-to-lead optimization and later stages of lead optimization to drug candidates as depicted in

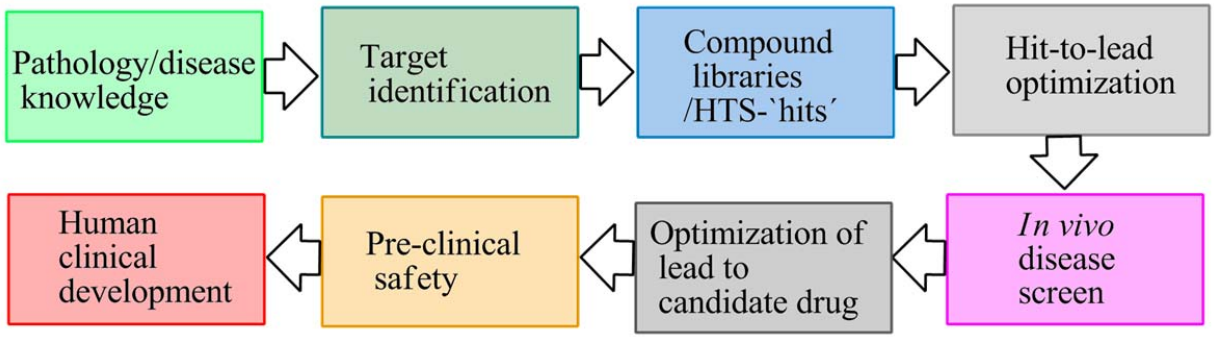



Figure 1.

**Figure 1.** A representative process in drug development.

The `hit´ identification stage refers to molecules that, even with weak activity, represent a useful source to initiate a medicinal chemistry program, while `leads´ represent the compounds which possess a desired but non-optimized biological activity. Subsequently,

drug candidates are optimized leads which fulfill all stereoelectronic, physicochemical, pharmacokinetic and toxicologic requirements for clinical usefulness.[17]

**Rule-of-five:** poor absorption or permeation are more likely when cLogP (the calculated 1-octanol–water partition coefficient) is >5; molecular mass is >500 g/mol; the number of H-bond donors (OH plus NH count) is >5; and the number of H-bond acceptors (O plus N atoms) is >10.

#### 2.1.1 A change of course in drug development

In order to improve the quality of the compounds entering the initial screening, rather than focusing directly on obtaining good drug `candidates,´ the current focus in drug discovery is on doing things earlier by obtaining better quality `leads.´[15, 18] Important steps towards obtaining better lead compounds include the use of Lipinski´s rule-of-five, and the embodiment of alternative approaches summarized in "Rational drug design".

A prominent analysis by Lipinski et al.[19] showed that historically 90% of orally absorbed drugs are far more likely to fall into a category determined by a limited range of physicochemical properties (rule-of-five; see box). Various modifications and alternative definitions have been proposed since Lipinski's rule-of-five,[20-22] however all agreeing that drug-likeness is determined by a set of molecular properties and descriptors.

Rational drug design represents an alternative strategy to the empirically based high-throughput synthesis and -screening. In this approach, bioactive compounds are specifically designed or chosen to interact with the drug target, often assisted by the use of computational modeling techniques.

The concepts and work presented hereafter in the present thesis, give a broad description on several aspects within rational drug design.

### 2.2 General concepts and considerations in rational drug design

The molecular recognition of ligand with the target is an essential event for inducing a biological response. [23] Most of the current approaches in rational drug design may fall into two main categories: ligand-based and structure-based design. The common goal of both

approaches is to suggest novel compounds with better activity profile than the parent compound.

#### 2.2.1 Ligand-based design

In absence of target structure information, the development of drug candidates often begins by optimizing existing bioactive ligands or screening methods to identify a suitable (parent) ligand. In ligand-based design, one proceeds from a parent compound (bioactive `hit´ ligand) with no information about the receptor to eventually determine the pharmacophore, *i.e.* the ensemble of features (steric and electronic) necessary for a drug to possess in order to ensure optimal interactions with the target.[24]

Experimental SAR studies constitute an important part in this process, and if a conformationally constrained molecule is biologically active, it may serve to identify the spatial orientation of the pharmacophoric groups. Aside from pharmacophore modeling (and thereafter screening for potential `hit´ candidates), computationally assisted ligand-based drug design often includes quantitative structure-activity relationship (QSAR) studies in order to provide key insights into potentially favorable ligand-receptor interactions. This further enables the construction of suitable and predictive models for lead discovery and optimization.[25]

Challenges include the difficulty to determine a 3D-pharmacophore due to the inherent conformational flexibility of the ligands (as they exist under physiological conditions as a mixture of interconverting conformations). This issue can often be dealt with by determining the conformational space available to a given ligand, *i.e.* by consideration of the possible conformers available for the ligand; thereby, computational methods such as molecular dynamics and quantum mechanics are often applied to model the ligands.[26-28]

Another way to account for this problem is to identify the most stable conformer based on the assumption that the conformation of minimum energy is that which is receptor-bound (*e.g.* in molecular docking simulations). One has to keep in mind however, that the binding conformation of the ligand with the receptor may not necessarily be the ligand´s minimum energy conformation.[29] Accordingly, placing the ligand in a conformation that is more suitable for interaction with the receptor (based on experimental SAR- or SDM data) one

can obtain better results.[30] Evidently, target structure determination and structure-based design are valuable means to assist in this process.

#### 2.2.2 Structure-based design

In structure-based design, information on the target (e.g., NMR-, X-ray crystallography data) is required, and in addition to pharmacophore definition, the binding mode of the ligand in the receptor binding site has to be determined.

Molecular docking is therefore a widely used method in structure-based design[31] in order to identify and optimize drug candidates, *i.e.* by examining and modeling molecular interactions between ligands and target macromolecules. Scoring functions are applied during docking to evaluate the interactions (binding free energy) and rank the resulting conformations. In this way, a filtering criterion is provided to allow focus on the most promising candidates for ligand optimization.

However, resolved crystal structures (especially for membrane-bound targets) are often not available. Alternatively, information about ligand-receptor interactions can be obtained through SDM. Computational approaches to produce a representative model of the receptor (homology modeling or *de novo* automated design, see section 3.3) are also used. Both approaches have disadvantages compared to the use of input from resolved structures (X-ray, NMR); mutational mapping (SDM) of ligand-receptor interactions may lead to biased information due to protein structure-rearrangements upon mutagenesis, while homology models alone are suboptimal and less accurate representations.

Evidently, challenges within ligand-based or structure-based design are often resolved by combination of strategies to include several methodologies, *e.g*. experimental: (SAR- and SDM studies), biophysical methods: (NMR- X-ray crystallography), calorimetric methods, and computational methods.

## 3. DESIGN OF PEPTIDE DRUGS

### 3.1 Peptides as drugs

Proteins and peptides are oligomers or polymers formed by chains of amino acids (aa) linked to each other through amide bonds (peptide bonds) between the carboxy group of one amino acid and the amino group of the following amino acid (Figure 2).

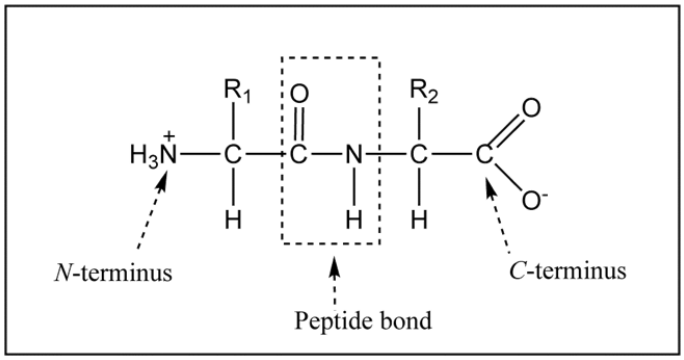



**Figure 2.** Illustration of a peptide bond between two amino acid units; R: side-chain.

An arbitrary distinction places peptides as molecules containing fewer than 50 amino acids (5-50 a.a: ~500-5000 Da) to discriminate them from small organic molecule (<500 Da) drugs and protein based drugs (biologics) of >5000 Da.[32]

At present, biologics constitute a large field of the pharmaceutical industry as a successful class of therapeutics both in treating diseases but also from an economic perspective (*i.e.* Humira Pen and Enbrel in rheumatoid arthritis (RA) treatment with estimated worldwide sales of 9.48 and 8.37 billion dollars respectively in 2012).[33] An important reason for the increasing market share of biologics is their higher target specificity due to their larger size. However, there are also many disadvantages with biologics as drugs such as the lack of membrane permeability, poor oral bioavailability, and lower metabolic stability compared to smaller molecule drugs. In general, biologics disobey every one of "the-rule-of-five" parameters, [19, 34] and as expected they are not suitable for oral administration and normally require injection delivery.

Compared to protein based drugs, peptides have the potential to penetrate deeper into tissues due to their smaller size, and are generally considered to be less immunogenic.[35]

Moreover, several advantages of peptides over small organic molecule drugs include increased selectivity and specificity on binding the desired target,[36] and since their degradation products are amino acids they generally exhibit a reduced risk of toxicity.[37] Thus, peptides represent a class of molecules that have the specificity and potency of the larger size biologics, but are smaller in size, more accessible and cheaper to produce using chemical methods.[38] Therapeutic peptides have been traditionally derived from: i) bioactive natural peptides produced by plants, animals or humans and ii) isolated from genetic or recombinant libraries.[39, 40] However, limited availability of tissue sources, methods of extraction, and increased risks of contamination are reasons as to why the isolation of peptides from natural sources is often problematic. Although a number of peptide based therapeutics have reached the market, their development as drugs has been limited; low systemic stability, high clearance, poor membrane permeability, negligible activity upon oral administration, and high cost of production are some of the challenges to be named.[41, 42]

*Progress in peptide synthesis.* Production of synthetic peptides has become possible for the pharmaceutical industry with automation of solid-phase peptide synthesis (SPPS), initially developed by Merrifield.[43] In comparison to peptides derived from natural sources, chemical synthesis offers access to a much wider structural diversity by use of unnatural amino acids as well as different structural modifications such as amide bond replacement to obtain modified peptides with improved pharmacokinetic profile. Various chemical strategies have been developed in an attempt to overcome the limitations of peptides and increase their *in vivo* plasma residence time. Cyclization of target peptide, modification of peptide bonds (pseudopeptides),[44, 45] and design of peptide mimetics as substitutes for peptides in their interaction with the receptor[46] are some of the approaches to obtain better drug candidates.

To optimize the properties of a lead bioactive peptide and ultimately to rationally design small-molecule peptidomimetics, a systematic approach can be applied[23, 29, 47] (see section 3.3).

### 3.2 A systematic approach to drug design from a lead peptide

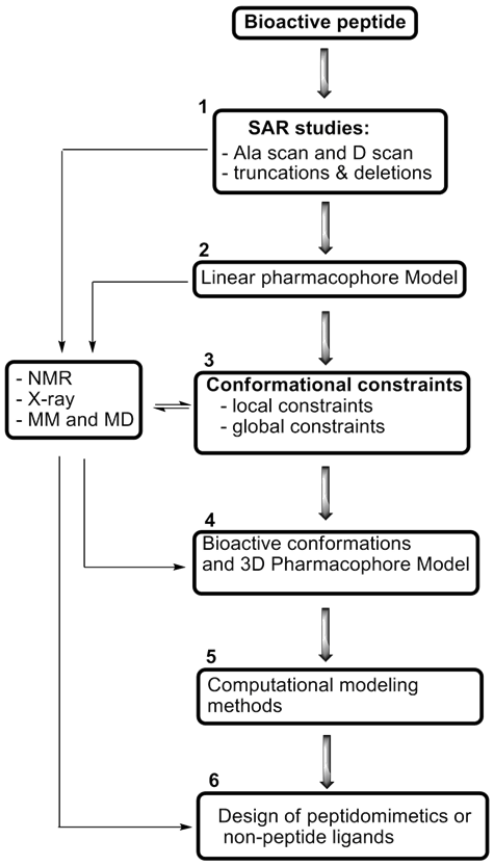



**Figure 3.** Flow-chart of main steps in a systematic approach for drug design from a biologically active peptide. Modified and reproduced with permission.[23]

Once the structure of a bioactive natural peptide ligand is known, an initial step involves the identification of the key aa-side chain residues necessary for receptor recognition, by means of single amino acid-modifications in the ligand. This process usually includes an Ala-scan,[48] (step 1, Figure 3) where each amino acid is systematically substituted by L-alanine followed by biological activity measures (SAR studies) to examine the relative importance of each side-chain group. To obtain initial information on structural arrangement or identify potential turn inducing positions (*i.e.* highly desired secondary protein structure elements for peptide-receptor interactions) in the sequence, a D-scan (step 1, Figure 3) can be performed. Accordingly, the original amino acids (in L-configuration) are replaced by their D-enantiomers.[23]

#### 3.2.1 Conformational constraints

Additionally, other local conformational constraints (step 3, Figure 3) can be applied to the ligand in order to constrain the backbone conformation (φ, ψ, and ω torsional angles, see Figure 4A) to more energetically preferred conformations. For instance, *N*-methylation restricts the amide bond and allows formation of a *cis* bond, while isosteric amide bond replacements and α-substituted amino acids can induce favorable secondary structures (α-helix, β-sheet, reverse turns, etc.) according to their own unique stereostructural properties, often including the nature of the $\chi^1$ group as well.[49-51]

Ultimately, the introduction of a global constraint (Figure 3) by means of peptide cyclization is a crucial step in the sense that an appropriate template for all the elements that make up the pharmacophore is provided. Moreover, cyclization improves the

pharmacokinetic properties of the ligand by reducing H-bonding (rule-of-five), enhances membrane permeability, and increases stability against proteolytic degradation.

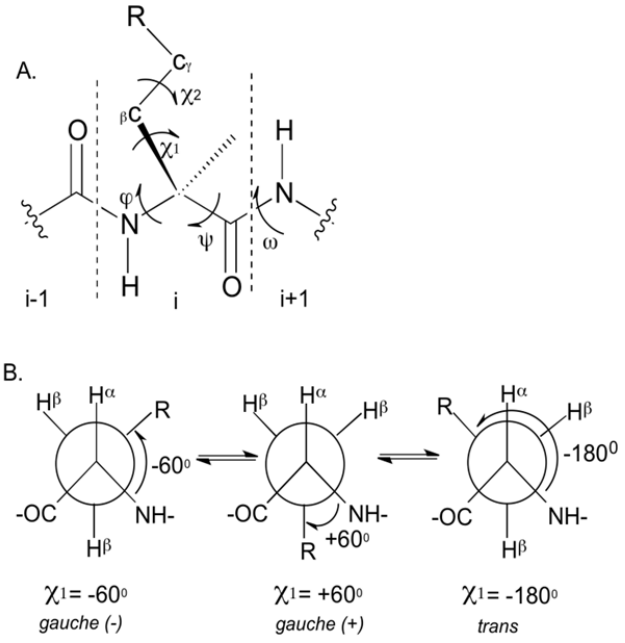



**Figure 4. A.** Backbone torsional angles (φ, ψ, ω) and side-chain dihedral angles χ-angles. **B.** Newman projections of three staggered rotamer conformations in an L-amino acid. Modified, and reproduced with permission.[23]

### 3.2.2 3D-Pharmacophore determination

At this point, alternative strategies to determine the bioactive conformation and 3D-pharmacophore model (step 4, Figure 3) may include synthesis and SAR-studies of analogues of the cyclic peptide. Accordingly, sequential alteration of the aa-sequence order in combination with D-scans contributes to identify the desired turn conformation in the cycle[52] as well as the optimal stereochemistry of the side-chains. The side-chain groups of amino acid residues in a peptide generally have free rotation about the side-chain torsional (or dihedral) angles (for topography in χ-space see Figure 4A, *e.g.* $\chi^1$ for Cα-Cβ bond, $\chi^2$ for Cβ-Cγ bond etc.). It has become increasingly apparent that also the χ-angles (in conjunction with the backbone angles (φ, ψ)), are critical for ligand-receptor interactions.[47]

As depicted in Figure 4B, the three low-energy conformations for $\chi^1$ are referred to as gauche (-), gauche (+), and trans.[53, 54] The challenge seems to lie in determining which of the three low-energy conformations the side-chains adopt upon binding to the receptor, *i.e.* the one that is implicit as part of the pharmacophore. An example of chemical modifications that can be done to define the χ-topography, is the restriction of the rotation around Cα-Cβ bond and Cβ-Cγ bond by incorporating the side-chain of interest into various ring structures. Subsequent SAR studies in conjunction with NMR, and computational methods (Molecular Mechanics and -Dynamics) can then determine the topography or 3D arrangement of critical side-chain groups.

Evidently, at this point in the process a more precise 3D conformation of the pharmacophore is obtained, and in many cases highly potent and efficacious drug candidates are produced in the process. However, this is also a good starting point for development of peptide mimetics (or peptidomimetics) from the derived potent cyclic peptide and/or the defined 3D pharmacophore (step 6, Figure 3).

### 3.3 Design of peptidomimetics

The overall goal with peptidomimetic drug design is to obtain ligands with improved pharmacokinetic profile. Higher stability to biodegradation, good bioavailability, and potential for oral delivery renders the peptidomimetic class of compounds more in agreement with Lipinski`s rule-of-five, than their peptide precursors. If a non-peptide ligand is desired, the proper choice of scaffold (*i.e.* the core structure of the molecule replacing the peptide backbone) that can place the key side-chain residue in 3D-space is the challenge. The determination of a scaffold is particularly important in *de novo* design[17] of novel ligands.[55, 56]

When a 3D pharmacophore or a receptor-binding site is known then it is possible to develop novel ligands with different scaffolds. Information extracted from ligand-receptor interactions constitutes the primary criteria or constraints, *i.e.* physicochemical properties and potential interaction points that contribute to binding affinity, and if this information can be collected from the 3D-receptor structure, the design strategy is receptor-based.

#### 3.3.1 *De novo* structure-based design

Receptor based design using X-ray input of the receptor structure usually starts with the determination of the binding site. Several *de novo* design softwares[56] exist with diversified searching algorithms and scoring criteria; potentially all leading to the determination of the interaction site and the definition of the explicit requirements or primary criteria (*e.g.* H-bonding potential) for increased affinity of ligand binding. In turn, this strategy will narrow down the vast number of possible ligand structures. Briefly, design software programs can grow ligands in the defined receptor binding site using building blocks (functional groups) and linkers from available databases.[57] Despite the apparent difficulty to predict whether a compound can actually be synthesized, a careful selection of building blocks and linkage

rules provides reasonable synthetic feasibility.[58] Docking simulations, calculated scores and visual evaluation provide an additional selection step.

#### 3.3.2 *De novo* ligand-based design

In the absence of a 3D target structure, an alternative strategy is to use the known parent ligand or its 3D pharmacophore model as an input for design of novel compounds. An advantage with this approach is that the topology or the 3D conformation of the known ligand can provide more accurate information in the starting point. The pharmacophore model can be used to obtain a pseudoreceptor model, *i.e.* to computationally generate an artificial protein receptor as a replacement for the 3D structure.[59, 60] The models attempt to capture the shape of the binding site and its interaction points for successful ligand binding. Accordingly, from this point on, the same structure-based strategy described in the previous section can be applied as the derived receptor or pseudoreceptor guides the design of ligand structures that are complementary to the defined primary constraints.

#### 3.3.3 Ligand-based scaffold-hopping

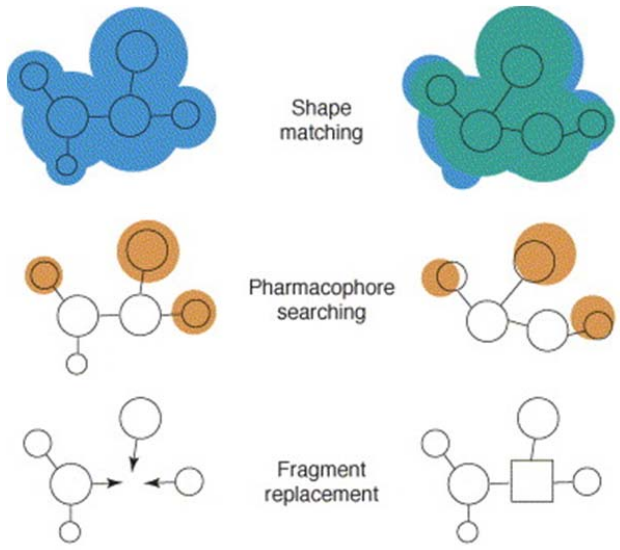



**Figure 5.** Illustration of three approaches to scaffold-hopping. Reprinted with permission.[62]

Alternatively, the 3D pharmacophore can be used in pharmacophore-based virtual screening methods such as scaffold-hopping.[61] Scaffold-hopping refers to the identification of isofunctional but structurally different chemotypes to a query lead ligand by using pharmacophoric features as an input,[62] *i.e.* structurally novel compounds can be pursued by altering the central core structure or template of a known active compound. Of the most attractive scaffold-hops is the transition from peptidic ligands of the receptor to small non-peptide mimetics of the peptidic ligand precursors. There are several computational approaches to scaffold-hopping, and some examples are illustrated in Figure 5. Shape matching and pharmacophore searching are

based on the use of the structure of an active ligand and pharmacophoric features as input, respectively, and appropriate computational programs can search databases for matching structural elements. In fragment replacement, the spatial (distance, angles) relationships between 2 or 3 single bonds (vectors) is used to search a database of chemical structures for suitable alternative fragments fitting onto these vectors.

### 3.4 The key to success: Combining disciplines

Major steps have been done in developing fully automated *de novo* design softwares, and their use and future development is on the rise. However, the use of experimental data either to guide selection of candidates for further structural tuning (SAR studies) or to determine starting conformations (NMR-, X-ray, etc.) of input structures in computational programs are an inherent part to a successful strategy towards drug design. Moreover, experimental determination of receptor binding sites by SDM studies often provides a more accurate input on side-chain coordinates. This offers a higher potential for obtaining hit-candidates than merely relying on fully automated receptor site generation. Ultimately, it remains a medicinal chemists task and judgement to pick the most promising approach, or to combine experimental with computational methods for the most effective outcome.

Despite recent successes of crystallography in resolving target receptor structures, only a small procent of the Protein Data Bank (PDB)[63] entries (<0.1%) are related to GPCRs which does not reflect the fact that GPCRs are the most successful drug targets in terms of therapeutic benefit and potential sales.[64] GPCRs are highly insoluble and very dynamic; their structure is constantly changing during interaction with ligands and proteins rendering them difficult to isolate and crystallize.[9] In the absence of 3D target information, ligand-based design is still the usual way to develop new drug candidates with GPCRs as targets. A short description of recent crystallographic achievements concerning GPCRs is given in section 4.2.

## 4. G PROTEIN-COUPLED RECEPTORS

### 4.1 General overview

GPCRs are the largest family of membrane proteins and their involvement in signal transmission is fundamental for most physiological conditions, ranging from vision, smell and taste to neurological, cardiovascular, endocrine and reproductive functions. Accordingly, the GPCR superfamily is a main target for therapeutic intervention and represents the target directly or indirectly of 50-60% of all current therapeutic agents.[65-67]

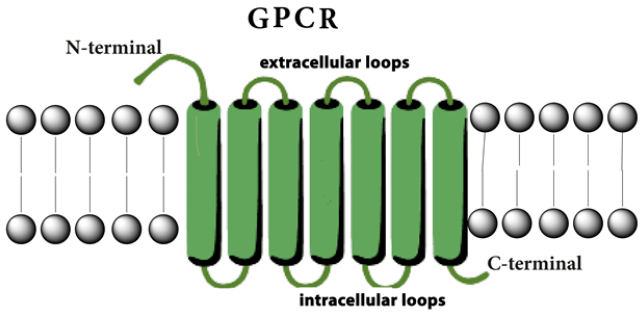



**Figure 6.** GPCR basic structure.

GPCRs share common structural elements of seven hydrophobic transmembrane helices (TMHs) with an extracellular *N*-terminal segment, three extracellular loops (ECLs), three-four intracellular loops (ICLs), and a C-terminal segment (Figure 6). Based on sequence similarity within their 7 TMHs, GPCRs can be clustered into five major families: class A (the rhodopsin), class B (secretin), class C (Glutamate), class D (Fungal pheromone), class E (cAMP) and the Frizzled/smoothened family receptors.[68-70] The rhodopsin family is by far the largest and most diverse family with four main groups (α, β, γ, and δ) and 13 subbranches.[68] Members within a group are characterized by conserved sequence motifs.

### 4.2 X-ray structures

Before 2007, structural insights into the GPCRs were limited to crystal structures of bovine rhodopsin (class A)[71] and to structures of extracellular domains of the secretin (class B)[69] and Glutamate (class C) receptors.[70] Newer developments in the area of X-ray crystallography accelerated the rate of resolved high-resolution structures, and by 2014 a number of structures of different class A GPCRs has been determined. Most of these receptors are aminergic and are subclassified as group α-class A receptors; they bind acetylcholine and monoamine neurotransmitters and examples include β-adrenergic receptors ($\beta_1$AR and $\beta_2$AR)[72, 73] and the dopamine $D_3$ receptor ($D_3$R).[74] Also structures of

peptide-binding receptors (group γ-class A GPCRs) have been solved, including the chemokine receptor CXCR4[75] and δ-opioid receptor (δ-OR).[76] Furthermore, structures of neurotensin receptor 1 (group β-class A)[77] were recently released.

### 4.3 Structural rearrangements in GPCR-conformation upon ligand-induced activation

A number of studies have suggested that most GPCRs exist in a dynamic equilibrium between inactive and active states, and can be further converted to a signaling state in the presence of heterotrimeric G-proteins.[78, 79] Comparisons of active and inactive-state structures indicated common activation-related features based on conformational changes in the intracellular sides of the receptors. Alternatively, different GPCR conformations are related to different signaling activity states,[78, 80] and a number of class A GPCR crystal structures were determined either in the inactive[81, 82] or active state conformation.[83, 84] The majority of the endogenous and synthetic ligands of class A GPCRs are found to bind within the transmembrane helix (TMH)-domain close to ECL-2.[85]

Furthermore, a suggested mechanism at play upon ligand–induced activation, known as the "global toggle switching", claims that an outward "swinging" motion of TMH 6 in accord with TMH 7 takes place upon receptor activation.[86, 87] Subsequently, ligand-induced activation involves spatial TMH-rearrangement, particularly for TMHs 5-7.[81] Moreover, contacts between ECL-2 and extracellular parts of the helices are suggested to take place during ligand induced activation.[88]

### 4.4 Signal transduction

Upon ligand binding to GPCRs, the exposed receptor intracellular sites interact with G-protein heterotrimer (α, β, and γ subunits) which play a crucial role in signal transduction towards second messenger cascades (Figure 7). Notably, the activation of some GPCRs also results in message transmission through arrestins and kinases, *i.e.* through non-G-protein pathways. The main Gα types are $G\alpha_s$, $G\alpha_i$ and $G\alpha_q$ based on the induced effect on secondary messengers (s-stimulation, i-inhibition, q-stimulates phospholipase C pathway). Structural shifts between the G-protein subunits are followed by exchange of GDP for GTP in the Gα

and separation of Gα from Gβγ subunits. Potential contacts with the subunits of other effectors lead to different effects. More detailed understandings however, related to selectivity of G-protein coupling of GPCRs, are not yet available.[89]

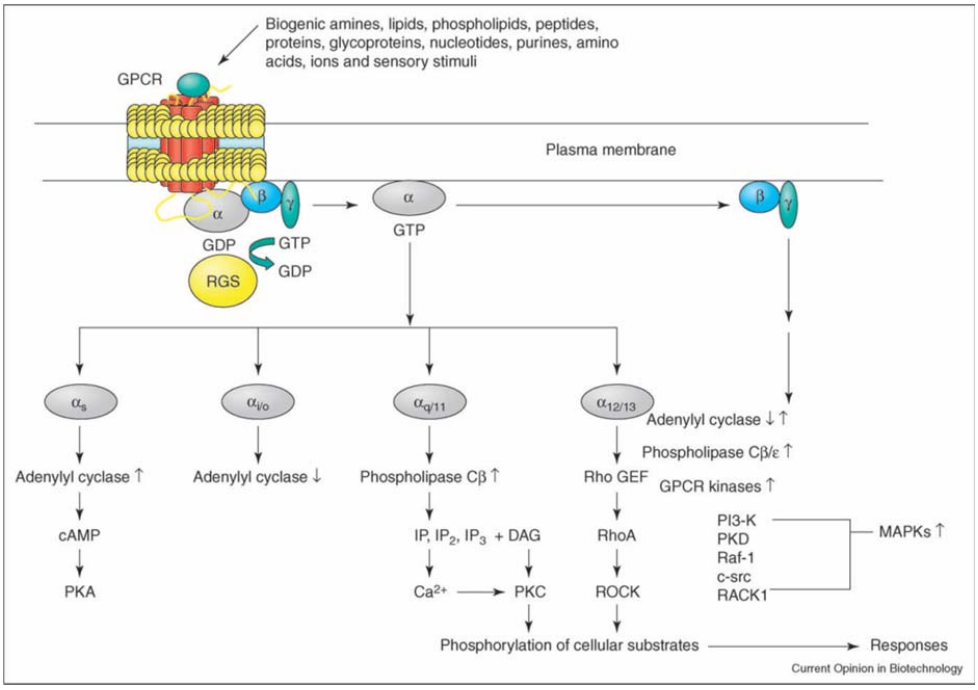



**Figure 7.** GPCR activation and signaling pathways through the heterotrimeric G-protein. Reprinted with permission.[90]

## 5. CHEMOKINES AND CHEMOKINE RECEPTORS

### 5.1 Chemokines

#### 5.1.1 General overview

Chemokines (or chemotactic cytokines) are a family of small secreted proteins 8-14 kDa, and through their receptor binding and mediated effects, they control immune responses with an emphasis on leukocyte trafficking, and maturation.[91] However, they are also known to be involved in growth regulation, hematopoiesis, embryonic development and angiogenesis.[92]

#### 5.1.2 Classification of chemokine ligands

Traditionally, chemokines have been sub-classified as CXC, CC, XC, and CX3C based on the spacing and sequential relationship of the disulfide bridges holding the peptide together (Figure 8); less commonly these groups are referred to as α, β, γ, and δ, respectively.

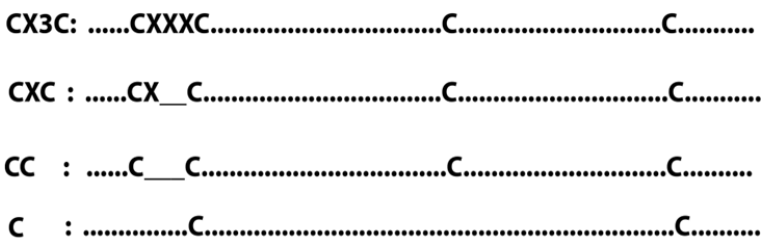



**Figure 8**. Structural classification of the chemokine family by signature cysteines. Underlines indicate gaps in the alignment; X, an amino acid other than cysteine; and dots, other amino acids. Spacing between cysteines is similar in all four groups. The N- and C-termini can vary in length.

The CXC subfamily (α-subfamily) contains two cysteine residue pairs forming two Cys-bridges separated by one nonconserved amino acid. The CC- (β-subfamily) contains two adjacent Cys residues, and the XC- (γ-subfamily) contains only one disulfide bridge. Moreover, the CX3C- (δ-subfamily) contains three residues between the two Cys residues and has only one member known to date. All chemokines bear an L-suffix to denote that they are ligands (e.g., CCL1).[93]

### 5.1.3 Structural features of chemokines

The discovery of neutrophile-targeted chemokine IL-8 (of the C-X-C subfamily) and its structure determination by X-ray and NMR-studies[94, 95] gave the first important insights into the 3D secondary structure elements of chemokines. The crystal structures[94, 95] indicated a heterodimer (Figure 9) stabilized by formation of six-stranded antiparallel β-sheet (three from each monomer) and by hydrophobic interactions with the overlying helices. Structure-activity relationship studies[96] on truncated analogues of IL-8 indicated furthermore, the Glu[4]-Leu[5]-Arg[6] (ELR) motif as essential for the binding and activity of not only IL-8 but also for CXCR1 and CXCR2 chemokine ligands.[97]

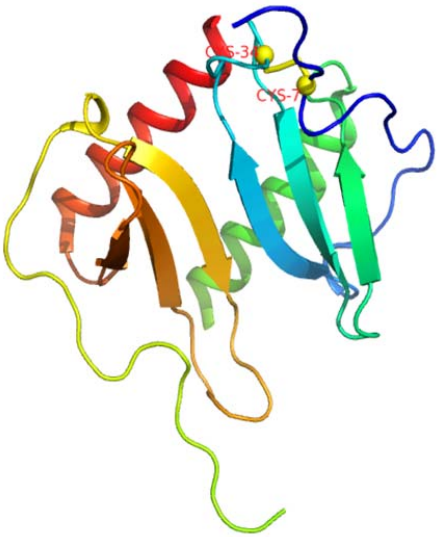



**Figure 9.** 3D structure of IL-8. Generated in PyMOL.[217]

Additional studies[98] indicated that residues 4-22 (*N*-terminus) are essential for receptor binding, and residues 30-35 (turn) contributes through a disulfide (7-34) bridge (Figure 9) to ensure correct conformation of the *N*-terminal region. Moreover, some research groups[99, 100] suggested the monomer might be the biologically active form instead of the dimer. However, the possibility that the dimer might be essential during the binding was not excluded. In general, the essential monomeric structural fold of chemokines is well conserved consisting of a three-stranded antiparallel β-sheet with an α-helix at the C-terminus. The N-terminus is generally disordered but found important for activation.[98, 101] Additionally, an extended loop region leading to a $3_{10}$-helix turn right before the β-sheet was also reported as important for the antiproliferative response of the CCL3 chemokine ligand.[102]

Representative examples of the CXC-, CC-, XC-, and CX3C-subfamilies are illustrated in Figure 10.

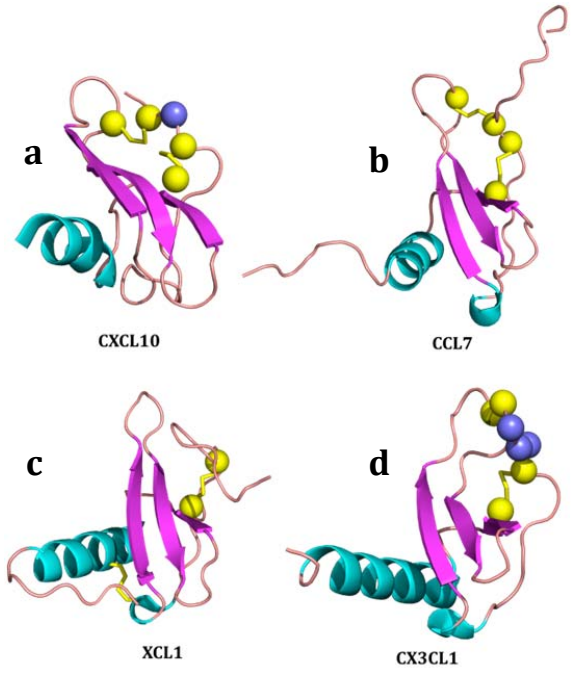



**Figure 10.** Representative examples from the four chemokine subfamilies; **a.** CXCL10 **b.** CCL7 **c.** XCL1 **d.** CX3CL1; Disulfide bonds are shown as yellow sticks. The figure was generated in PyMOL.

## 5.2 Chemokine receptors

### 5.2.1 General overview

Chemokine receptors belong to the rhodopsin family (class A) of GPCRs and are classified according to the class of chemokines that they bind. Accordingly, they bear an R-suffix to indicate receptor (*e.g.* CXCR4).[93] A great deal of promiscuity is evident in the interactions between chemokines and their receptors;[103] some chemokines bind and activate more than one chemokine receptor, and some chemokine receptors can be activated by more than one ligand,[104] whereas others are highly specific.

Apart from CXCR7, which is particularly biased towards β-arrestin mediated signaling,[105] all chemokine receptors transduce signals through heterotrimeric G-proteins.[106] However

three decoy (non G-protein signaling) chemokine receptor (D6, DARC, and CCX CKR) were found to be involved in scavenging inflammatory chemokines from the extracellular microenvironment.[107] With the discovery of chemokine receptors, the interest in chemokines as therapeutic targets increased. It is estimated that about half of the drugs currently in the market are either agonists or antagonists of GPCRs suggesting that at least some members of the chemokine receptor family are "druggable" targets.[104] As an example, the discovery of the role of CCR5 and CXCR4 as co-receptors in HIV infection, stimulated the search for antagonists of those receptors.[108]

However, there are many obstacles in drug discovery of chemokines; one of these issues concerns the lack of selectivity as many small-molecule antagonists cross-react with other GPCRs. Other issues include the lack of relevant animal models and the nature of the screening employed.[104, 109] More detailed understanding of ligand receptor interactions is therefore needed and consequently very few drugs have made the market by now.

#### 5.2.2 Interaction of chemokines with their receptors: the "two-site" model

A "two-site" binding model for the interaction of chemokines with their receptors has been proposed which provides a distinction between the binding and signaling phases. According to the model,[110, 111] an initial interaction takes place between the compact core of the chemokine and the *N*-terminus of the receptor (site I). This is followed by the interaction of the flexible *N*-terminus of the chemokine with site II of the receptor (the latter is formed as a pocket by extracellular loops and membrane-spanning domains). Accordingly, interaction with site II leads to receptor activation.

The two-site model was initially suggested as a general explanation for the interactions of chemokines with their receptors.[112] However, extensive studies on the binding of CXCL12 including mutational analysis[101, 113] on its receptor (CXCR4), showed a good agreement with the two-site model theory. In brief, it was suggested that the CXCL12-RFFESH loop is optimal for the initial binding or docking with the *N*-terminus of CXCR4 receptor (site I); thus, allowing access to the more buried receptor site. Subsequently, the *N*-terminal residues of CXCL12 (KPVSLSYR-CPC) bind to a groove among the helices (site II) (Figure 11).

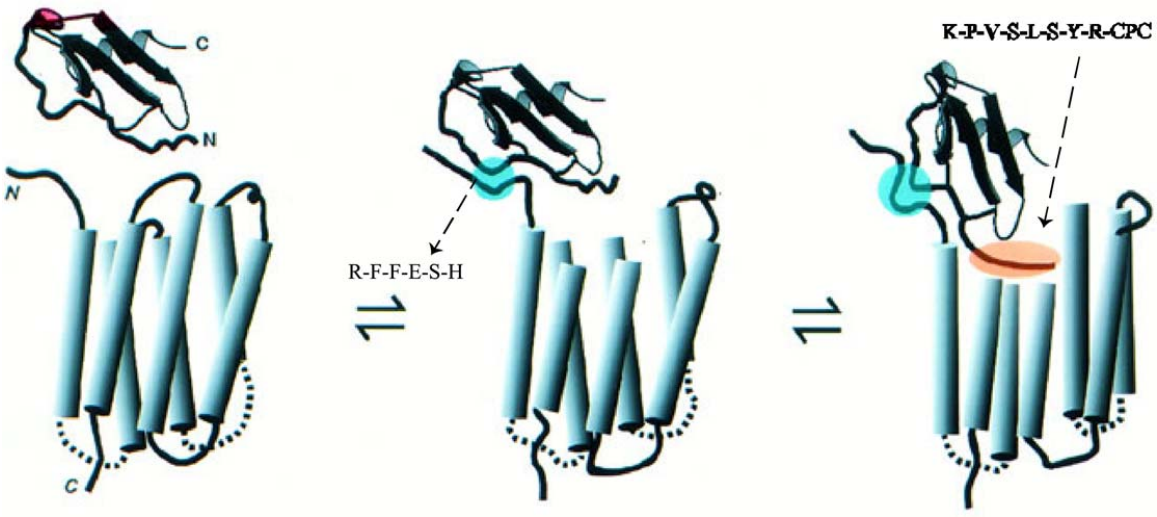



**Figure 11.** Interaction model of CXCL12 with CXCR4 demonstrating the two-site model of binding. Reprinted and modified with permission.[101]

Evidently, a change in the conformation of the receptor´s TM-helices allows intracellular G-protein binding and signaling.[114]

#### 5.2.3 High-resolution structures: support for the two-site model

An NMR structure of CXCL12 in complex with an *N*-terminal peptide part of CXCR4 (Protein Data Bank (PDB)-ID: 2k05)[115] is considered to represent a part of the site-I complex; thus, providing an insight into the ligand-receptor interactions. Further support for the two-site model comes from the recently released X-ray structures of CXCR4 by Wu *et al.*,[75] where four crystal CXCR4 structures were reported bound to a small antagonist IT1t, and one structure bound to a cyclic peptide CVX15 (Figure 12). The CVX15 peptide may trace to some extent the path of *N*-terminal peptide sequence of CXCL12 (KPVSLSYR), and the binding site of IT1t may point to the major anchor region for this domain.

The small ligand IT1t and the CVX15 peptide are both orthosteric competitors of the CXCL12 *N*-terminal-signaling trigger; hence, their binding site in CXCR4 relates to the proposed site II. Moreover, the IT1t ligand showed a unique binding mode, and it is the first

to portray a ligand binding within `the minor ligand pocket´,[75] (see Figure 13) while the CVX15 occupied the complete binding cavity and extended out towards the extracellular side of the protein.

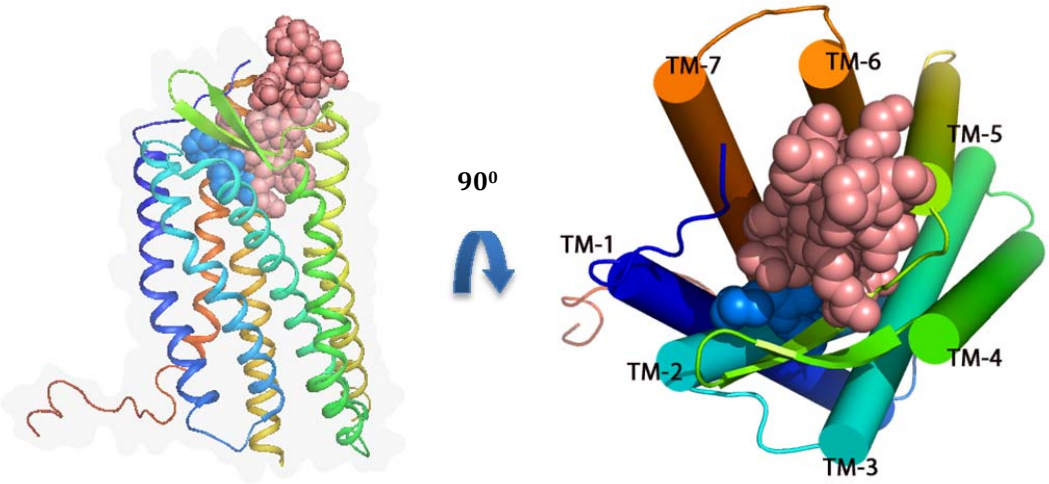



**Figure 12.** Crystal structure of CVX15 (in light brown color) and IT1t (in blue) shown as spheres, in complex with the CXCR4 receptor. Generated in PyMOL.

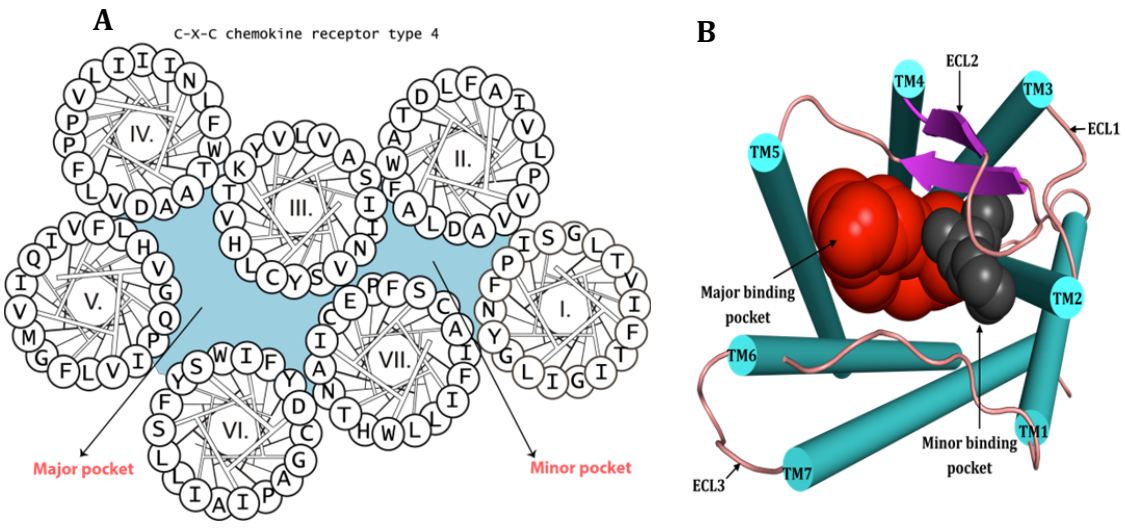



**Figure 13.** The `minor ligand binding pocket´ comprised of TM helices 1-3 and 7, and `major ligand binding pocket´ comprised of TM helices 3-6 and 7;[116] **A.** A helical wheel diagram of CXCR4. **B.** A 3D representation of the minor and major binding pockets illustrated for the class A GPCR receptor CXCR4. Generated in PyMOL.

## 6. CXCR4 AS A THERAPEUTIC TARGET

### 6.1 General overview

The chemokine C-X-C receptor 4 (CXCR4) is comprised of 352 amino acid residues, and displays 33% homology to other CXC and CC members of the chemokine receptor family.[117] CXCR4 has only one known natural ligand, the 68-mer chemokine peptide CXCL12 (SDF-1a) that is rich in basic amino acids (Arg, Lys, and His). CXCR4 itself is however strongly negatively charged in comparison to other chemokine receptors,[118] and has an overall electrostatic surface charge of -9. The CXCR4-CXC12 axis is found to be involved in physiological processes, such as the homing of immune cells (T-cells) to sites of inflammation,[119] growth-regulatory functions,[120] angiogenic[121] and embryonic development.[122]

### 6.2 Pathophysiological role of CXCR4

Besides its expression in normal tissues, CXCR4 has been related to a number of diseases; the receptor was initially reported as a co-receptor for $CD4^+$ T-cell infection of human immunodeficiency virus (HIV) type I,[123] and subsequently in pathogenesis of rheumatoid arthritis[124] as well as multiple types of cancer.[125] Evidently, the involvement of the CXCL12-CXCR4 system in a wide range of physiological and pathological conditions, and its lack of promiscuity is of increasing interest in drug discovery. It is not within the scope or objectives of the present thesis to comprehensively cover the involvement of the CXCL12-CXCR4 system in all the related pathologies. However, a brief introduction in the mechanisms involving CXCL12-CXCR4 in HIV-entry is presented in the following section.

#### 6.2.1 Role of CXCR4 in HIV infection

HIV-entry is a multistep process involving a host surface receptor CD4 and co-receptor, either CCR5 or CXCR4[126] and a viral envelope glycoprotein (Figure 14). Expression of these receptors determines viral tropism, which is related to the capacity of the virus to use CCR5 and /or CXCR4 as coreceptors (Figure 14). CXCR4 is expressed on T-cells and allows entry of T-tropic HIV-1 strains, while M-tropic HIV-1 strains preferentially use CCR5, which is expressed in monocytes-macrophages. Moreover, other viral strains exhibit dual-tropism by using both coreceptors. CXCR4-using T-tropic as well as dual tropic viruses generally

emerge in later stages of infection and are associated with the disease progression to AIDS.[127, 128] Viral entry as depicted for T-tropic strains in Figure 15, involves the binding of the trimeric gp120 viral envelope protein to the CD4 receptor, which in turn induces a conformational change to allow binding of the V3 loop (gp120) with CXCR4. CXCL12 blocks T-tropic HIV-1 from entering cells, and a potential drug target is therefore implicated. Several studies have demonstrated the ability of both CXCL12 and isoforms to block HIV-1 entry via CXCR4.[101, 129] However, the use of chemokines as antiretroviral agents is limited by their short half-life and potential undesirable inflammatory effects.[130] Hence, the rationale behind the development of anti-HIV CXCR4-antagonists as drug candidates, lies in their "non-signal-inducing" block of HIV-entry, limiting therefore undesirable inflammatory responses. Their function does not rely on receptor down-regulation, but on receptor occupancy.

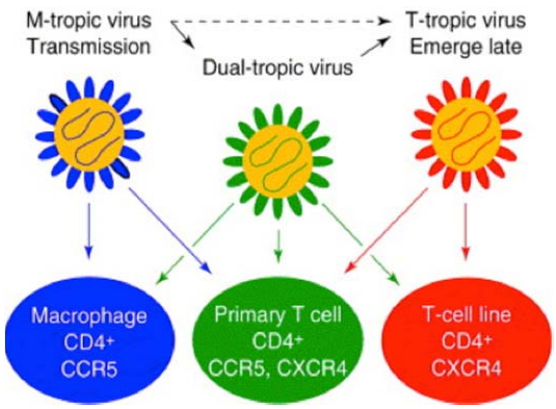



**Figure 14.** HIV-1 tropism; M-tropic viruses use CCR5, T-tropic viruses use CXCR4, and dual-tropic viruses use both co-receptors. Reprinted with permission.[217]

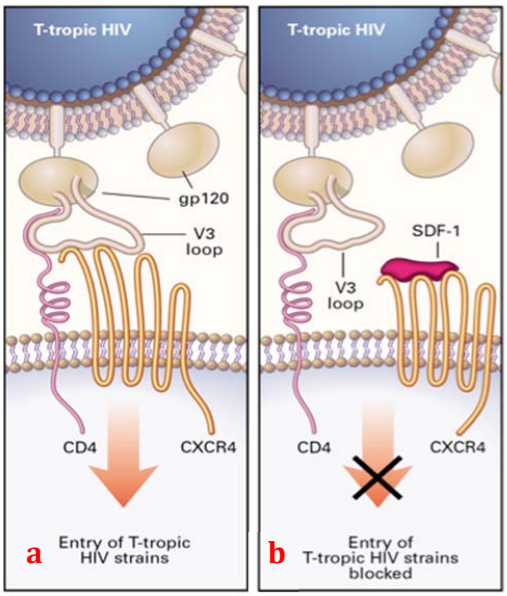



**Figure 15**. **a.** T-tropic HIV-1 entry **b.** Block of entry by CXCL12 (SDF-1). Reprinted with Permission.[218]

## 7. CXCR4 ANTAGONISTS

A number of CXCR4 ligands have been described over the years.[131-133] Most of the drug discovery targeting CXCR4 has focused on the development of antagonists, and initially the focus was turned toward their potential as anti-HIV drugs. However, as the field of drug research quickly expanded, other disease states were shown to involve CXCR4 as well. A number of different chemical classes of CXCR4 antagonists exist, and it is not within the scope of this thesis to cover every class in detail although, two prominent categories can be distinguished: peptide-based-, and small-molecule CXCR4 antagonists.

### 7.1 Peptide-based CXCR4 antagonists

Among the first compounds to be reported as CXCR4 antagonists were peptide derivatives such as the peptide analogue T22, an 18-mer synthetic analogue derived from polyphemusin II (a self-defense peptide isolated from horseshoe crab), and subsequently its shortened 14-mer peptide, T140[134] (Figure 16). T22 and T140 (first and second generation polyphemusin II-derivatives respectively) possess strong anti-HIV activity by blocking X4-HIV-1 entry to the cell and inhibiting $Ca^{+2}$ mobilization normally induced by CXCL12-signaling;[135-137] T140 forms an antiparallel β-sheet structure supported by a disulfide bridge and connected by a βII' turn[138] (Figure 16). Although T140 was found unstable toward biodegradation,[134, 139] the modified T140-analogues 4F-benzoyl-TN14003 and 4F-benzoyl-TE14011 (Figure 16), displayed enhanced biostability and anti-HIV activity ($EC_{50}$ values 0.6 and 1.6 nM respectively),[134, 140, 141] suggesting that the *N*-terminal 4-fluorobenzoyl moiety could be a part of the pharmacophore associated with anti-HIV activity. SAR studies on T140[142] indicated furthermore the four amino acids ($Arg^2$, $Nal^3$, $Tyr^5$, and $Arg^{14}$) as essential for significant activity, and as potential pharmacophoric residues. Following studies which included NMR analysis and MD-calculations, indicated the four essential residues (see Figure 16) of T140[142] to be in close proximity.

Subsequently, in a pharmacophore-based approach of screening cyclic pentapeptidic libraries,[143] the potent CXCR4 antagonist FC131 (Figure 16; a third generation polyphemusin II-derivative) was discovered, shown to be equipotent to T140 ($IC_{50}$ values 0.004 μM,[143]). Unlike T140 however, FC131 is globally constrained and more stable towards biodegradation. In the years following the discovery of FC131, extensive SAR

studies were carried out where several approaches to drug optimization were employed (see section 3.2), including: Ala-scans, D-scans, N-methylations, constraints through use of unnatural amino acids and isosteric replacements.[144-152] Accordingly, derived SAR-data have shed light into the structural requirements of cyclopeptides for CXCR4 antagonistic activity, and in some cases analogues with improved potency were detected. Importantly, these studies laid the foundation for future development of more drug-like mimetics of the cyclic pentapeptide prototype FC131.

In retrospect, the whole body of work from polyphemusin II to development of FC131 analogues represents a perfect example of the systematic approach to design from a lead peptide as described in section 3.2.

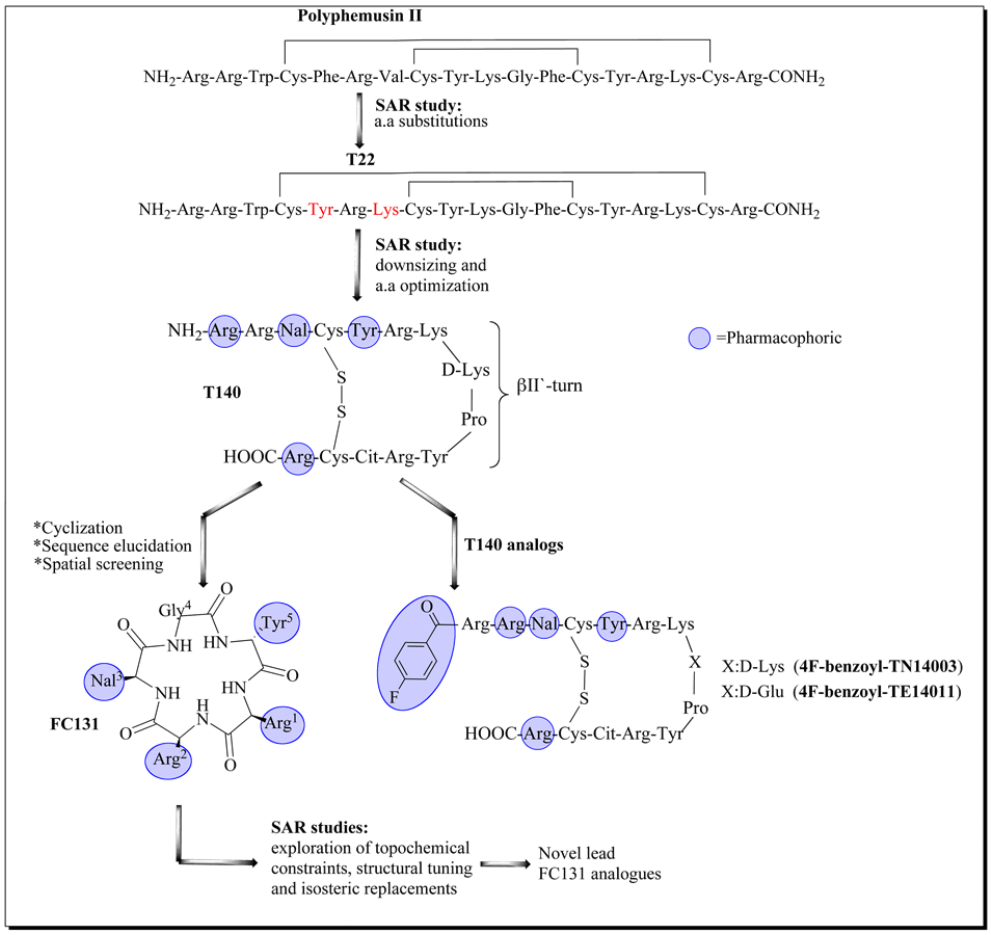



**Figure 16.** Development of Peptide-based CXCR4 antagonists from polyphemusin II.

### 7.2 Small-molecule CXCR4 antagonists

#### 7.2.1 Non-peptidic small-molecules

The non-peptide based CXCR4 antagonists which is the largest and most structurally diverse category of CXCR4 antagonists, comprises over 10 different chemical classes.[132, 133] An extensively studied representative in this category is the bicyclam AMD3100 (**1**, Figure 17), originally regarded as a highly potent and selective inhibitor in HIV fusion and uncoating.[153] However, in subsequent clinical trials it was indicated that AMD3100 could additionally mobilize various hematopoietic cells, while its overall efficacy in affecting disease activity in HIV-1 patients was considered low.[154, 155] AMD3100 exhibits furthermore, poor oral bioavailability mainly due to the increased positive charge (+2; in each cyclam ring at physiological pH). Accordingly, the development of AMD3100 into drug for anti-HIV application was discontinued, but it has since 2008 been in the market (Plerixafor or Mozobil) as a drug for stem colony mobilization, and it is administered by subcutaneous injection.[156]

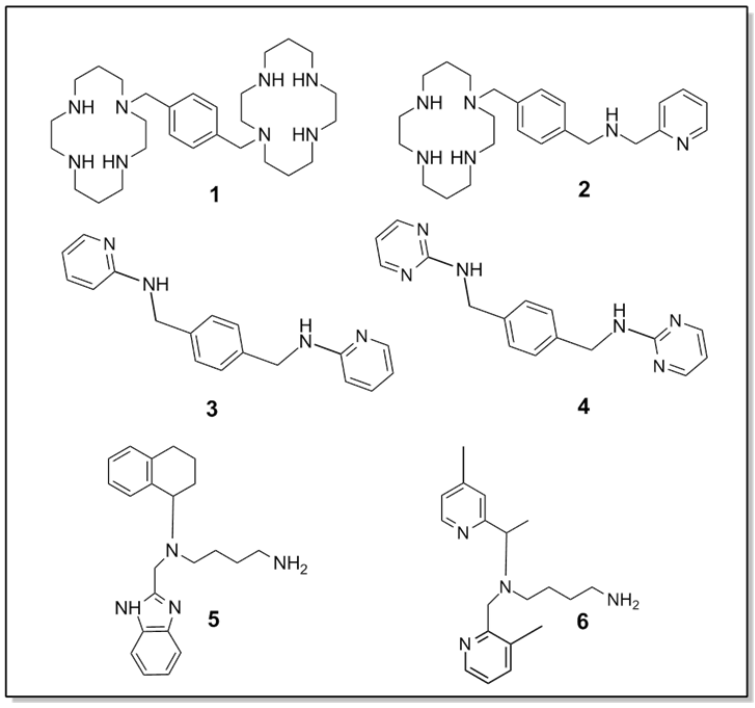



**Figure 17.** Non-peptide based CXCR4 antagonists.

In subsequent attempts to improve oral bioavailability, compound AMD3465 (**2**, Figure 17) was developed by substitution of one of the cyclam rings in the precursor (AMD3100) with a (pyridin-2-ylmethyl)amino moiety.[157] This monocyclam derivative was found to be a 10-fold more potent CXCR4 antagonist than AMD3100,[130] although oral bioavailability was still low. Furthermore, replacements of both cyclam moieties of AMD3100 with heteroaromatic moieties led, through rational design and analysis[158] to the

discovery of potent compounds WZ811 (**3**) and MSX-122 (**4**) (Figure 17). The two latter compounds were found to block CXCR4 at subnanomolar concentrations but failed to exhibit a good pharmacokinetic profile.[159, 160] The non-cyclam AMD-analogue, -11070[161] (**5**, Figure 17), is a potent orally bioavailable CXCR4 antagonist[162] shown to work in a synergistic manner with other HIV-inhibitors such as reverse transcriptase- and protease inhibitors. Due to liver histology changes and high risk of hepatotoxicity found in animal studies, AMD11070 is currently on hold for further development.[163] A number of derivatives of AMD11070 have been recently reported[164, 165] with the potential for further development (e.g., compound **6**, Figure 17).

#### 7.2.2 Small-molecule peptide mimetics

In an attempt to develop linear CXCR4 antagonists based on the T140-pharmacophore groups (Figure 16), thus involving Arg, Nal, Tyr and the 4-fluorobenzoyl moiety, a series of small-molecule linear CXCR4 antagonists were reported by Tamamura *et al*.[166] (exemplified by compounds **7** and **8** in Figure 18). These compounds had however lower potency than FC131, suggesting that the conformational restriction of the cyclic backbone of FC131 is essential for potency.

Subsequently, Ueda *et al*.[167] used a constrained and rigid indole template to incorporate pharmacophoric side-chains (surrogates of Arg and Nal groups of FC131) in a scaffold-hopping approach (see Chapter 3). Accordingly, compounds with a non-peptidic template were obtained (exemplified by compound **9** in Figure 18) and despite their lower potency, these ligands can serve as useful leads for further optimization. Interestingly, the only other scaffolding approach to develop tripeptidomimetic CXCR4 antagonists based on key side chains of FC131, included a scaffold ring synthetically derived from diketopiperazine mimetics (compound **10**, Figure 18).[168] This attempt resulted however in very low to no activity in comparison to the reference ligand FC131.

KRH-1636 (**11**, Figure 18) is another low molecular weight and selectively potent inhibitor of CXCR4 for X4 HIV-strains.[169] This ligand was previously considered to mimic the tripeptide Arg-Arg-Nal fragment of FC131,[170] and it constitutes an important prototype for design of linear peptidomimetics as CXCR4 antagonists (as exemplified by compound **12**[171] in Figure 18). Furthermore, KRH-1636 (**11**) was found to block HIV replication *in vivo* in

SCID mouse model while an intra-duodenal administration in rats resulted in high bioavailability suggesting that the compound might be orally bioavailable. However, a more conventional oral pharmacokinetic study in rats has not been reported. An alkyl amino analogue of KRH-1636 (KRH2731•5HCl; structure not disclosed yet) has high bioavailability (37% through oral administration in rats) and possesses potent CXCR4 antagonistic activity.[171] A key challenge for effective therapeutic effect and development of promising candidates into drugs remains the achievement of good oral bioavailability. As already mentioned however, peptide mimetics are expected to possess improved pharmacokinetic and pharmacodynamics traits (including good oral activity)[172, 173] and therefore, are considered as more useful targets for the drug discovery process.

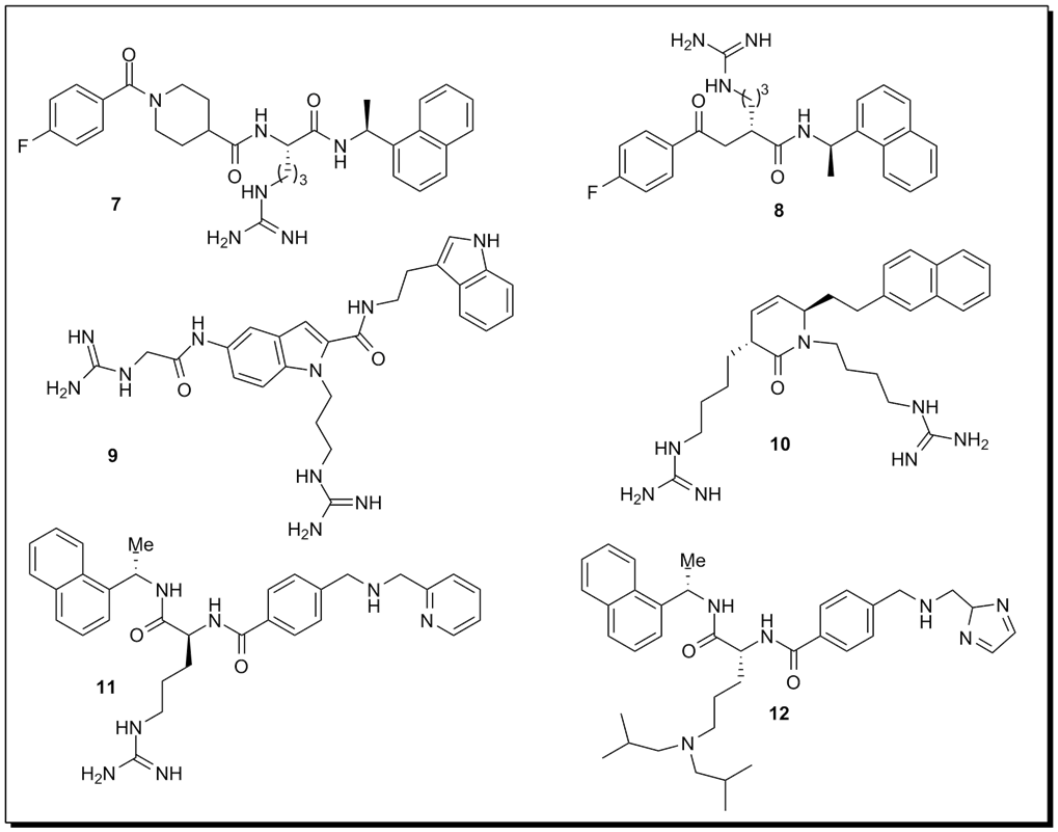



**Figure 18.** Examples of small-molecule peptide mimetics as CXCR4 antagonists.

## 8. AIMS

The main aim of this project was to rationally develop tripeptidomimetic CXCR4 antagonists based on the existing and generated knowledge about the SAR and pharmacophore for the lead cyclopentapeptide CXCR4 antagonist FC131. An additional aim was to determine whether the known tripeptide-like CXCR4 antagonist KRH-1636 is a mimic of the cyclopentapeptide FC131, *i.e.* if the two compounds bind to the receptor in the same way.

*Specific objectives*

- To investigate SAR for the aromatic positions 3 (2-Nal[3]) and -5 (D-Tyr[5]) in the cyclopentapeptide CXCR4 antagonist FC131 (PAPER I).

- To investigate SAR of simplified analogues based on the Arg-Arg-2-Nal tripeptide fragment (PAPER II).

- To design and synthesize a novel class of scaffold-based tripeptidomimetics (PAPER II).

- To determine the binding mode of the known tripeptidomimetic CXCR4 antagonist KRH-1636 (PAPER III).

## 9. STRUCTURE-ACTIVITY RELATIONSHIP STUDIES OF THE LEAD CYCLOPENTAPEPTIDE CXCR4 ANTAGONIST FC131 (PAPER I)

### 9.1 Background

Several research groups,[145, 148, 152] have reported experimental data investigating the SAR in substituted analogues of Arg[1] and Arg[2] (positions 1 and 2) in FC131 (Figure 19). Accordingly, the collective data have shown that both the structural identity and positive charge of the Arg[2] side-chain are indispensable pharmacophoric features. Position 1 (Arg[1]) is found to be generally more "forgiving" toward structural modifications, while a positively ionizable group is favored for activity; suggesting a potential salt-bridge interaction of Arg[1] with a negatively charged receptor site. Contrary to position 2 however, the H-bond potential (guanidino or amino), the stereochemistry (L or D), and the spacer properties (length, degree of flexibility/rigidity) were not found as critical.[145] Gly in position 4 was first utilized[143] with the intention to serve as a linker (linking the T140 pharmacophoric residues, Figure 16) during the construction and screening of libraries to discover FC131. Reported SAR data[143, 150] showed however that replacements of Gly led consistently to reduced potency compared to the parent ligand FC131. It is therefore suggested that the inherent flexibility and small size of Gly (due to the absence of side-chain), enables conformations that are energetically unfavorable for other amino acid substitutions.

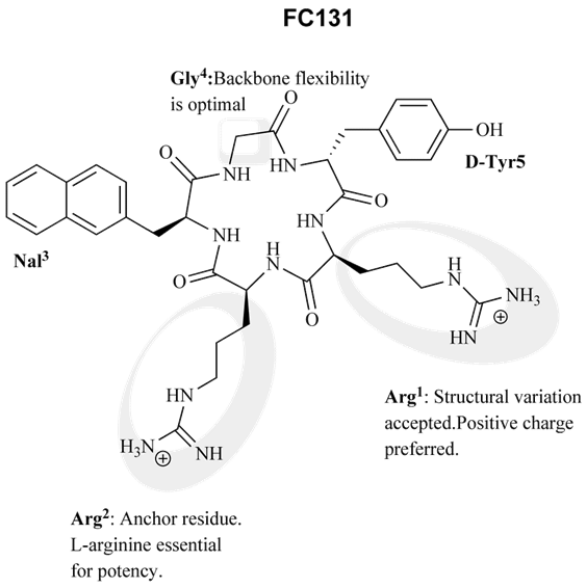



**Figure 19.** Schematic representation of previous SAR findings for positions 1, 2 and 4 in FC131.

However, existing SAR data for the two aromatic residues 2-Nal[3] and D-Tyr[5] have been less informative. In a previously reported Ala-scan of FC131, the Ala[3] and D-Ala[5] analogues were both classified as inactive;[152] however, the roles and relative importance of the 2-Nal[3] and D-Tyr[5] side chains have been unclear, leading to some ambiguity in

pharmacophore definitions.[167, 170]

## 9.2 Design

### 9.2.1 Position 3 (2-Nal$^3$)

The limited existing SAR data for position 3 (2-Nal$^3$) of FC131 showed that substitution with D-2-Nal (inverse stereochemistry) resulted in reduced activity (by 25-fold),[143] while *N*-methylation (*N*-Me-2-Nal$^3$), and substitution with a conformationally constricted tryptophan (Trp) derivative (Tricyclic) resulted in low, and no activity respectively.[150, 152] However, substitutions with a Trp-residue and a sulphur-containing Trp-analogue which are structurally similar to Nal, resulted in very good activity.[152]

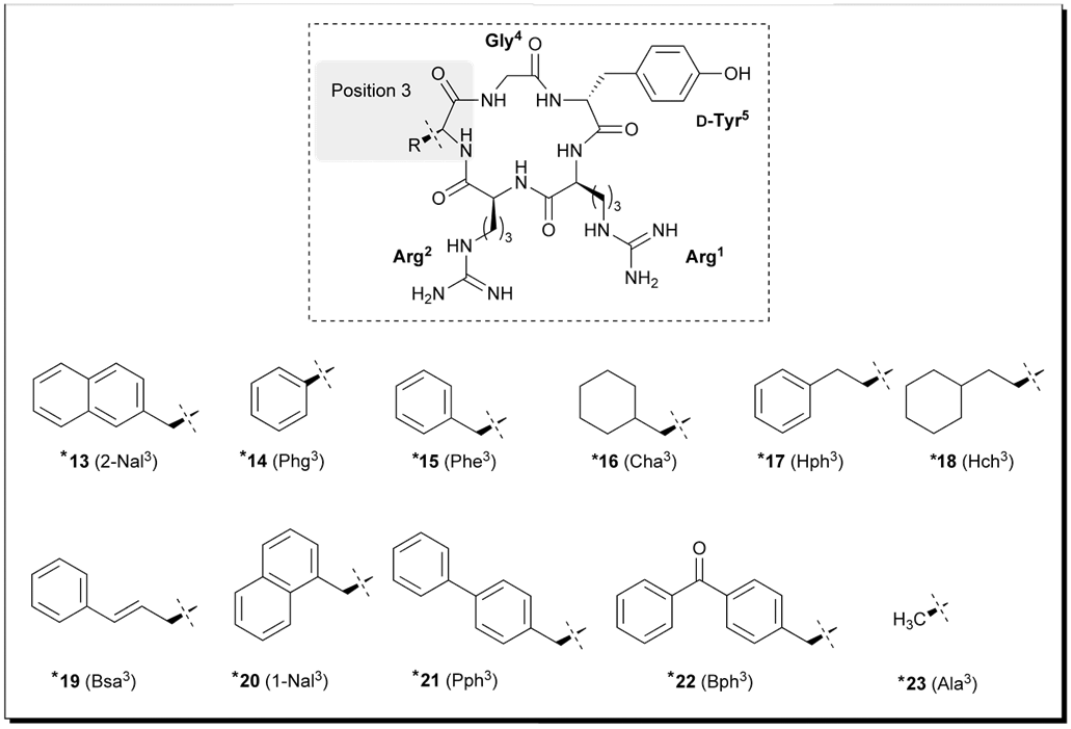



**Figure 20.** Structures of the lead cyclopentapeptide **13**, and the synthesized Xaa$^3$ analogues **14-23.**

***** Compounds **13-23** have been prepared by other group members.[144]

Thus, the existing SAR on position 3 are mostly directed on backbone effects (*N*-methylation, constriction) and stereochemistry, without further investigations on significantly different aromatic side chains than the 2-Nal ring.

Accordingly, we probed the position 3 by preparing a series of analogues with aromatic and aliphatic side residues of different size and shape affording a compound series with small (**14–16**), medium (**17–20**), and large (**21, 22**) side chains (Figure 20). The known Ala[3] analogue **23** was also included as reference.

#### 9.2.2 Position 5 (D-Tyr[5])

A number of Xaa[5] substituted analogues of FC131 have been reported,[149, 150, 152, 174] and the collective literature data indicate that both the size of D-Tyr side chain, the D-configuration, and the 4-hydroxyl group are ideal for potency of FC131. However, no conformational SAR study probing the rotameric state (in topographical χ-space) of D-Tyr[5] side-chain in the receptor-bound conformation of FC131 has been reported before.

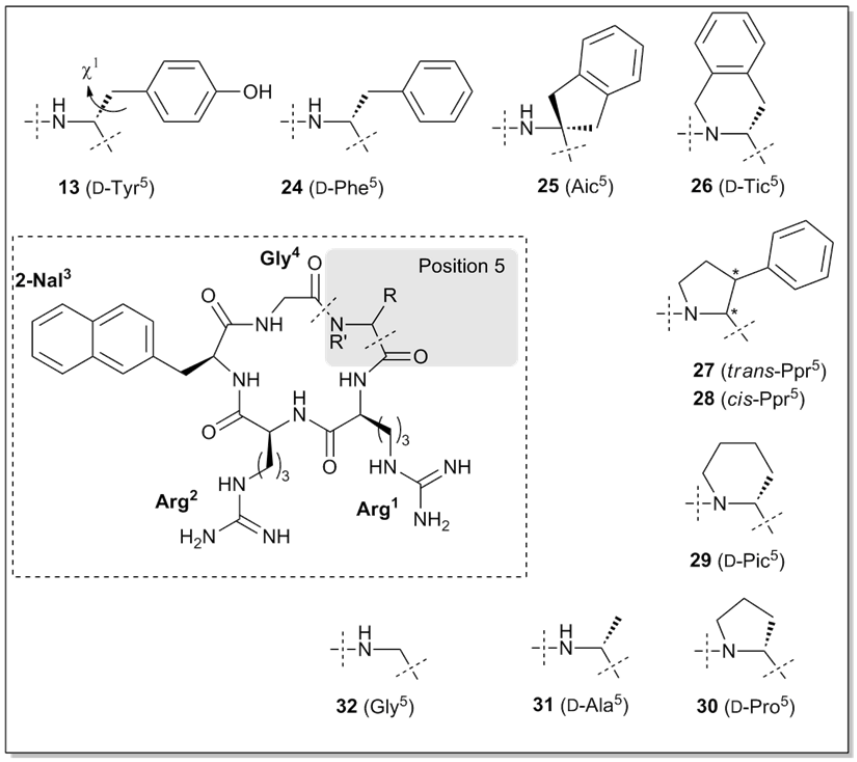



**Figure 21.** Structures of the lead cyclopentapeptide **13** (shown with the $\chi^1$ angle, and the synthesized Xaa[5] analogues **24-32**.

The $\chi^1$-torsional angle can adopt three accessible low-energy staggered conformations (gauche (-), gauche (+), and trans, see Chapter 3-Figure 4). It is assumed that upon ligand-

receptor interaction, the D-Tyr$^5$ $\chi^1$-angle will adopt one of these three possible conformations. Clearly, the side chain conformations are critical to molecular recognition.

Therefore, in order to determine the 3D rotamer state of D-Tyr$^5$ in $\chi$-space, *i.e.* the $\chi^1$-torsional angle and orientation of the side-chain, we constructed a series of cyclopentapeptide analogues with conformationally constrained position 5 as depicted in Figure 21. The rationale behind our approach lies in the use of constrained amino acids with a particular $\chi^1$-angle in order to identify a potent analogue, which at the same time can reveal the right side-chain orientation of D-Tyr$^5$.

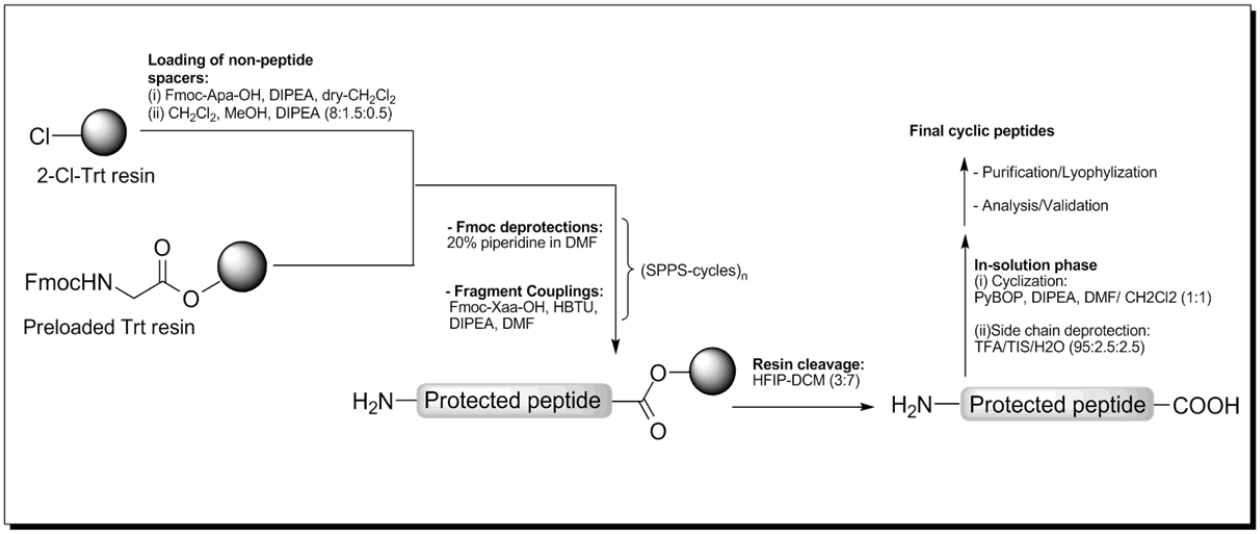
*Conformational probes introduced in position 5.* Based on the fact that the D-Phe$^5$ analogue was only 2-fold less potent than FC131 we used D-Phe mimetics (compounds **25-28**, Figure 21) as conformational probes instead of the corresponding D-Tyr mimetics which were not commercially available; thus, D-Phe$^5$ (compound **24**) was used as reference in our assay. For compound **25**, the achiral indane-constrained Aic$^5$ was used to link $C^{\alpha}$ to $C^{\gamma}$**.** Aic has been previously used to probe the orientation of aromatic amino acid side-chains,[175-177] and the $\chi^1$-angle is found to prefer the gauche (-) (-60$^0$), and trans (180$^0$) torsional angles. In compound **26**, $C^{\gamma}$ was linked to $N^{\alpha}$ using D-Tic, which was originally shown by Hruby et al.[178] to limit the conformations to gauche (-) or gauche (+), excluding trans. For compounds **27** and **28**, we originally intended to use D-Ppr (a proline chimera) in order to adopt the trans and one of the gauche conformations (depending on which diastereomer is used) although, trans is favoured over gauche.[53] The enantiomerically pure (3S)-D-Ppr and (3R)-D-Ppr were not commercially available, and therefore racemic trans-Ppr (containing (3S)-D-Ppr and (3R)-L-Ppr) and cis-Ppr (containing (3R)-D-Ppr and (3S)-L-Ppr) were used as building blocks. Due to the relatively low activity of these analogues (see sub-section 9.4.2) no attempt was done to separate or synthesize them as pure (3S)-D-Ppr and (3R)-D-Ppr.

*Backbone effects in position 5.* Furthermore, we proceeded to isolate the backbone effects imposed by compounds **26-28,** by preparing analogues **29** and **30** as well as the known D-Ala$^5$ analogue (compound **31**). Lastly, the Gly$^5$ analogue was prepared (**32**) in order to investigate the absence of constraints in position 5.

## 9.2 Chemistry

### 9.2.1 Synthesis of cyclopentapeptides and macrocyclic compounds (PAPERS I & II)

All cyclic pentapeptides and the macrocyclic compounds (see section 10.1.2, PAPER II) were prepared by a combination of solid- and solution-phase synthesis as depicted in Scheme 1. Synthesis of the linear pentapeptide precursors was carried out by standard Fmoc-based solid-phase peptide synthesis (SPPS) using a trityl resin preloaded with Fmoc-Gly, based on an optimized protocol and procedures by Chan.[179] For the linear precursors of macrocyclic compounds (carrying a non-Gly linker residue), the 2-Cl-Trt resin was used for the initial loading of the spacer. In short, DIPEA with HBTU were used to aid the coupling of each amino acid, and 20% piperidine/DMF was used for $N^{\alpha}$-Fmoc deprotections. The side-chain protected peptide was selectively cleaved from the resin with HFIP in DCM. Head-to-tail cyclization was achieved in dilute solution (DMF:DCM, 1:1) using PyBOP, and DIPEA, followed by side-chain deprotection with a cleavage cocktail of TFA:TIS:$H_2O$. All crude products were precipitated in cold ether, purified by RP-HPLC and lyophilized. The ligand structures were characterized by high-resolution mass spectrometry (HRMS), NMR, and all final products showed over 95% purity as determined by analytical RP-HPLC.




**Scheme 1.** Overall synthetic strategy for cyclopentapeptides and macrocyclic compounds (for macrocycle structures see section 10.1.2).

### 9.3 Biological (functional assay) procedure (PAPERS I-III)

The antagonistic potency of the compounds in the present project was determined in a functional assay measuring inhibition of CXCL12-induced activation of human CXCR4 transiently expressed in COS-7 cells. The $IP_3$-assay makes it possible to test antagonism in the presence of constant concentration of the agonist (CXCL12 in our case). Shortly, COS-7 cells are transfected with receptor cDNA and chimeric G protein $G\alpha_{\Delta 6qi4myr}$, which turns the normal $G\alpha_i$-coupled signal into the $G\alpha_q$ signal; thus, the phospholipase C pathway is triggered and the receptor activation can be measured as PI-turnover.[180] The $IP_3$-assay relies on the incorporation of tritiated [$^3$H] myo-inositol into the cells for radiolabeling of the PIP2 turnover product $IP_3$ ($^3$H-$IP_3$). Based on the method used to quantify the radiolabeled $IP_3$, two different assays were used:

*SPA-PI turnover assay*. Following cell lysis, the extracts are mixed directly with yttrium silicate (Ysi) SPA beads. A radiolabeled-$IP_3$ molecule binds on a functional group on the scintillant material (Ysi), and the isotopes ($^3$H) emit β-rays, which in turn stimulate the fluoromicrosphere to emit light. The rate of photons in unit time (CPM) is measured on a scintillation counter.

*Traditional $IP_3$-assay.* Alternatively**,** a different assay (the "traditional" $IP_3$-assay which includes an anion-exchange chromatography step) was used, which was found to show the same result. The extracts are filtered and purified on an anion exchange resin. After addition of a liquid scintillation cocktail, *γ*-radiation is counted in a Beckman Coulter counter LS6500.

### 9.4 SAR

#### 9.4.1 SAR for position 3 (2-Nal[3]) of FC131

Compounds with relatively small side-chains in position 3 (**14-16** and **23**) had very low activity (>100 μM) suggesting that the smaller size side-chains prevent the formation of optimal ligand interactions in the receptor binding pocket. The lack of activity in Phe[3] analogue **15** shows further, that the main contribution comes from the distal aromatic ring of 2-Nal[3] (**13**). Compounds with the largest size side-chains in position 3 (**21, 22**) showed also very low activity (>100 μM) suggesting that a possible steric-hindrance effect from the bulky side chains could disturb the anchoring of the position 3 in the receptor binding site.

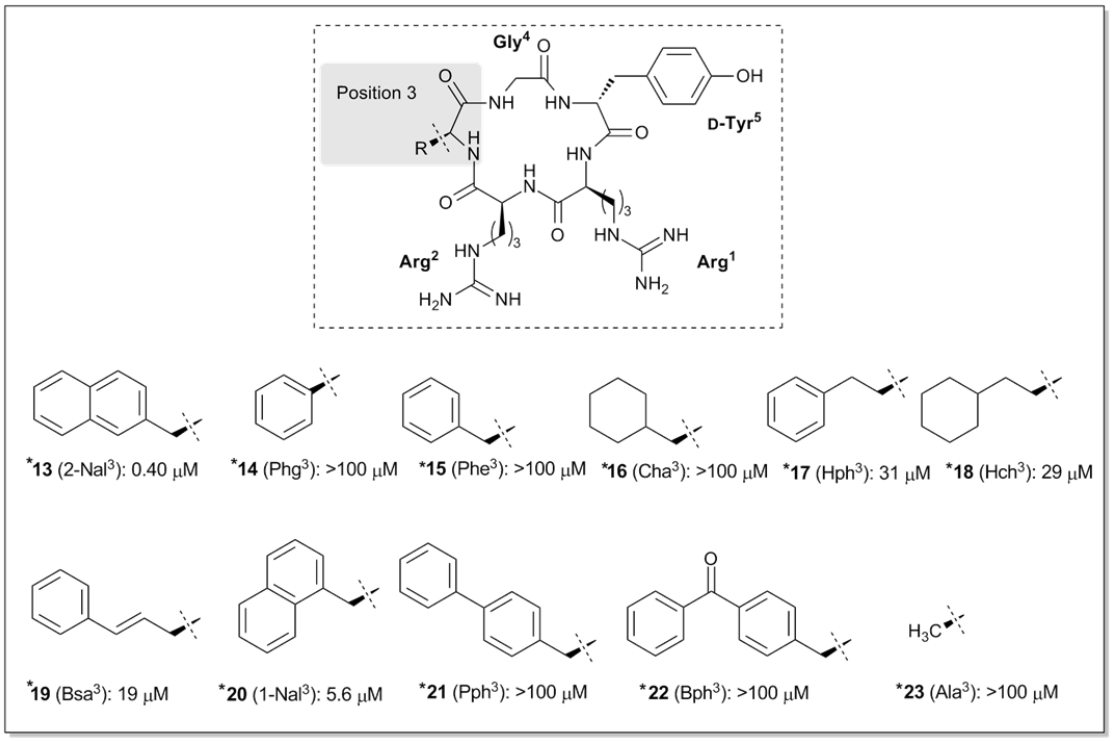



**Figure 22.** Structures and antagonistic potencies ($EC_{50}$) of the lead cyclopentapeptide **13**, and the synthesized Xaa[3] analogues **14-23**.

***** Compounds **13-23** have been prepared by other group members.[144]

However, compounds with a medium size substituent (**17-20**) displayed activity <100 μM (**13** > **20** > **19, 18, 17**). Clearly, the size of 2-Nal[3] (**13**) is most closely related to the medium size side-chain analogues (**17-20**). The most potent compound was the isomeric 1-Nal[3] analogue **20** (5.6 μM) and the 14-fold reduction in potency with respect to **13**, shows that the geometry of the 2-Nal side chain is favored over that of the 1-Nal side chain.

In sum, our SAR study on position 3 (2-Nal[3]) of FC131 shows that the 2-Nal group remains the optimal residue in this position and the distal aromatic ring of 2-Nal is critical in order to maintain potency.

#### 9.4.2 SAR for position 5 (D-Tyr[5]) of FC131

All the constrained analogues **25-29** (Figure 23) showed moderate to low activity indicating that the conformational restriction in the mobility of the position 5 (side-chain) is not well accommodated in the binding site of CXCR4. The overall effects of the structural alterations appear to overweigh the positive effect of constraining $\chi^1$, resulting in unfavorable receptor interactions.

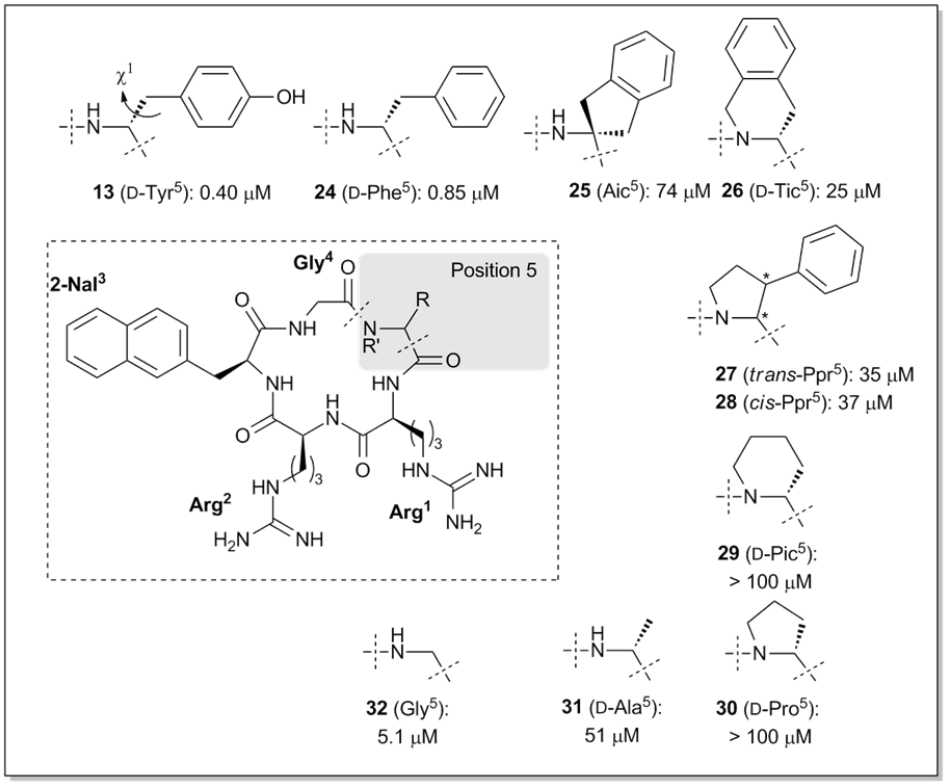



**Figure 23.** Structures and antagonistic potencies ($EC_{50}$) of the lead cyclopentapeptide **13**, and the synthesized Xaa[5] analogues **24-32**.

The increased steric volume in the side-chains of **25-29** might also not be compatible with the receptor-binding pocket. Furthermore, Aic and D-Tic in compounds **25** and **26** respectively, constrain the $\chi^2$ angle (see Figure 4) in addition to $\chi^1$, which again affects the

planarity of the phenyl ring. Also, the $N^{\alpha}$ alkylation (compounds **26-28**) removes the H-bond donor ability of the amide.

The four constrained analogues **25-28** appear not only to constrain the local conformation of the side-chain, but also to affect the peptide backbone conformation. Accordingly, α,α-disubstituted amino acids (**25**) are known to induce a helical backbone conformation,[181] while cyclic amino acids based on D-Pic and D-Pro (**26-28**) restrict the φ backbone torsion and promote trans/cis isomerization of the preceding amide bond.[182] The isolated effect on the backbone was tested in analogues **29**-**31**. While the restriction in **29** and **30** led to low activities, compound **31** showed moderate activity. This confirms that the backbone effects are partly responsible for the relatively low potency of compounds **26-28**. Surprisingly, the simplified and flexible Gly$^5$ analogue was found a 10-fold more potent than **31** and only a 6-fold less potent relative to **24**, which indicates that the reduced size and increased conformational flexibility of Gly$^5$ partly compensate for the side chain removal.

Overall, the D-Tyr$^5$ side-chain appears not to be an essential pharmacophoric element for cyclopentapeptide CXCR4 antagonists.

### 9.5 Rationalization of SAR: binding model

#### 9.5.1 Molecular modeling

*Ligand preparation.* The cyclopentapeptide ligands **13** (FC131), **24** (D-Phe$^5$ analogue), and **25** (Gly$^5$ analogue) were built in Maestro [183] using our previously proposed bioactive backbone conformation for cyclopentapeptide CXCR4 antagonists[170] as input structure.

*Protein preparation.* The X-ray crystal structure of CXCR4 bound to the 16-mer peptide antagonist CVX15 (PDB code 3OE0)[75] was prepared and optimized with the Protein Preparation Wizard workflow[184] as previously described[145] (see also sub-section 12.4.1).

*Docking simulations.* The three ligands were docked to the prepared CXCR4 structure using the induced fit docking workflow[185] which accounts for the conformational flexibility of both ligand and receptor using our optimized protocol.[145] Inspection of the generated poses (top 10 poses within an energy window of 30 kcal/mol) for each ligand resulted in the identification of the common binding mode for ligands **13**, **24**, and **25**.

### 9.5.2 Proposed binding mode for the lead cyclopentapeptide FC131

To rationalize the SAR data for aromatic positions 3 and -5 in cyclopentapeptide ligands, compounds **13**, **24** and **25** were docked to the X-ray structure of CXCR4 as described in the preceeding sub-section (9.5.1).

The proposed binding model for FC131 is shown in Figure 24. In comparison, the three docked ligands differing in position 5 (**13**: D-Tyr$^5$; **24**: D-Phe$^5$; **25**: Gly$^5$), had only minor differences among their top scoring poses. A superimposition of the three ligand conformations is depicted in Figure 25, indicating an overlap in the binding modes for the three ligands. Accordingly, the 2-Nal$^3$ side-chain is packed in a well-defined hydrophobic subpocket mainly composed of residues in TMH 5 (Figure 24). The restrictions of this subpocket would explain the reduced potency of the Xaa$^3$ analogues **14-23** (Figure 22). The D-Tyr$^5$ side-chain in position 5 is oriented toward TMH 1, and interacts with residues in the extracellular *N*-terminal fragment of CXCR4 (Figure 24). A cation-π

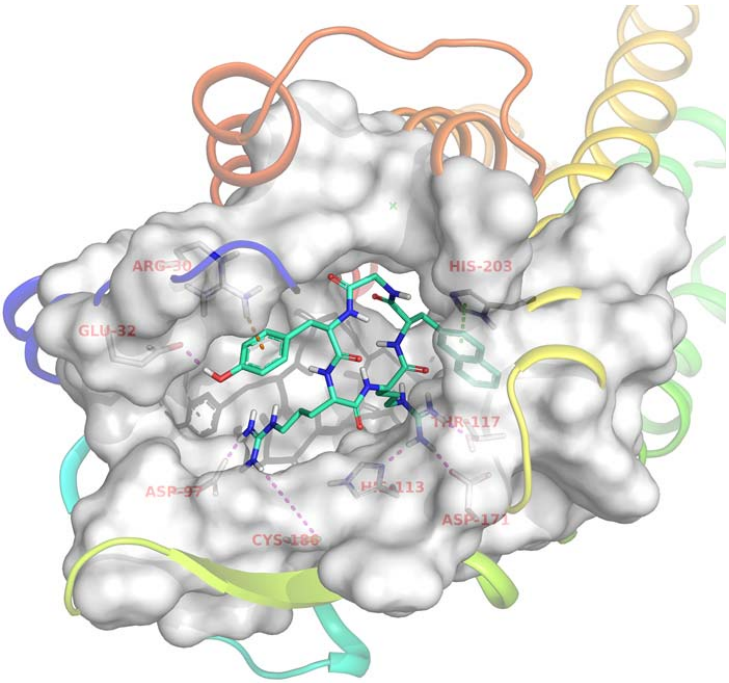



**Figure 24:** Proposed binding mode for the lead FC131. Ligand in light blue sticks, binding pocket as grey surface and the TMHs as colored ribbons.

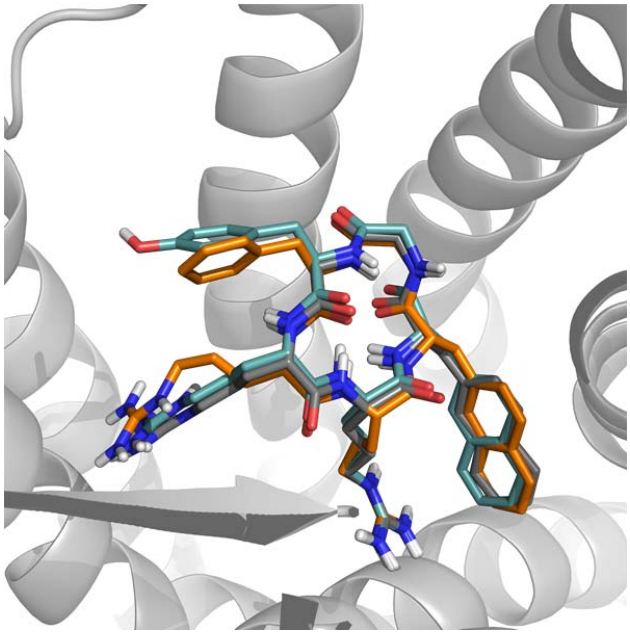


**Figure 25:** Overlay of the three conformations derived from docking: FC131 in light blue sticks, D-Phe$^5$ in orange sticks and Gly$^5$ in grey sticks. TMHs as grey ribbons.

interaction is seen between the guanidino group of Arg$^{30}$ and the phenyl rings of D-Tyr$^{5}$ and D-Phe$^{5}$, while the 4-hydroxyl group forms an additional H-bond with Glu$^{32}$, which would explain the higher activity of D-Tyr$^{5}$- in comparison to the D-Phe$^{5}$ analogue. The phenyl rings of **13** and **24** are solvent exposed, located in a relatively open region, and the lack of a defined subpocket for Xaa$^{5}$ would therefore explain why the D-Tyr$^{5}$ side-chain can be removed without a dramatic loss of potency (**25**: Gly$^{5}$). In sum, SAR data, supported by molecular modeling, indicate that the Arg$^{2}$ and 2-Nal$^{3}$ side-chains are buried in the receptor, while the side chains of D-Tyr$^{5}$ and Arg$^{1}$ are partly solvent exposed.

## 10. BACKBONE DISSECTION AND MACROCYCLIZATION (PAPER II)

### 10.1 Design

#### 10.1.1 Linear compounds

Our SAR findings on the aromatic positions of FC131 (Chapter 9) clearly showed that the naphthyl group in position 3 (L-2-Nal[3]) is indispensable for activity. The lack of a defined binding pocket for D-Tyr[5] in line with previous findings,[150] imply further that the activity of the cyclopentapeptides mainly resides in the remaining Arg[1]-Arg[2]-2-Nal[3] tripeptide fragment. In an effort to further study the tripeptide fragment, based on present and previous SAR (see also section 9.1) the naphthyl group and the L-Arg[2] (position 2) which both were found very sensitive to structural modifications,[145, 152] were therefore kept constant.

However, position 1 (Arg[1]) has been shown to be relatively tolerant to structural modifications, both with respect to stereochemistry (L- or D-arginine) and the chemical nature of the side chain.[145, 148] Moreover, incorporation of a D-amino acids in an all-L sequence during cyclization have a turn-inducing effect,[186] enhancing the macrolactamization reaction for ring-closure. In the following work, we therefore decided to focus on the D-Arg[1] epimers, using the lead cyclopentapeptide **33** (Figure 26) as a starting point.

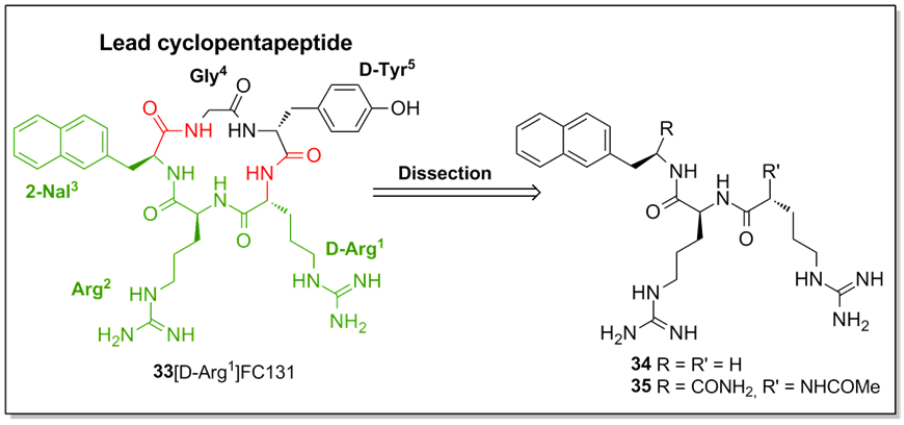



**Figure 26.** Linear analogues **34** and **35** and reference cyclopentapeptide compound **33**.

From a drug design perspective, we first applied a minimalistic approach aiming to isolate the essential structural elements of the tripeptidic pharmacophore, and the reference ligand **33** was therefore dissected to the linear

compound **34** (Figure 26). Furthermore, in order to probe the contribution of the terminal amide groups, compound **35** was further included in the study.

#### 10.1.2 Macrocyclic compounds

As part of our efforts to design small-molecule peptidomimetics as CXCR4 antagonists, we focused on the Arg[1]-Arg[2]-2-Nal[3] motif, using the D-Arg[1][FC131] analogue (**33**) as a reference compound (see also sub-section 10.1.1). Starting by dissection of the cyclopentapeptide structure we concluded (based on SAR, sub-section 10.3.1), that a further strategy should aim to reduce the undesired flexibility inherent in the linear analogues. Thus, herein we reintroduced cyclic constraints (macrocycles) in a systematic manner (Figure 27).

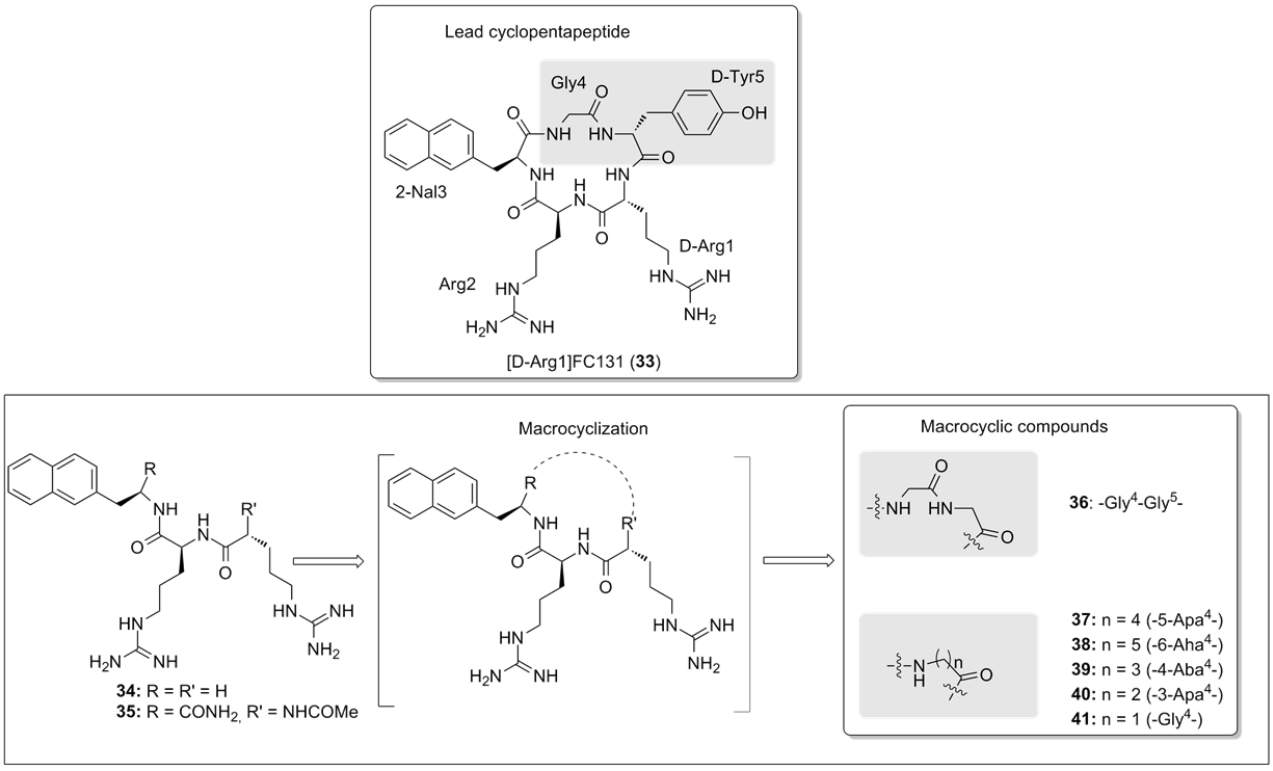



**Figure 27.** Macrocyclization strategy and synthesized macrocyclic analogues **36-41**.

Macrocyclization *i.e.* the introduction of a global cyclic constraint, is often used with the aim to force linear peptidic structures bearing pharmacophoric groups, into a bioactive conformation with increased binding affinity for the receptor. From this point of view, macrocyclization is an attractive strategy (often leading to increased receptor affinity and

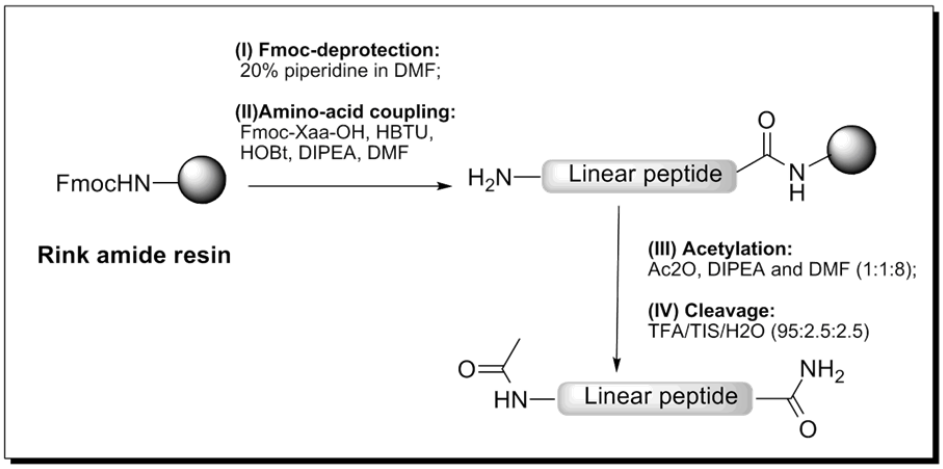
specificity)[187] when designing bioactive compounds, and the term "macrocyclic peptidomimetics" has been used to describe such compounds.[188] The nature of the macrocyclic constraint may vary, but typically the macrocyclic peptidomimetics are part peptide and part non-peptide, *i.e.* built on a hybrid template.

Thus, our objective is to force the D-Arg[1]-Arg[2]-2-Nal[3] motif into a more restricted conformation either by using a Gly-Gly dipeptide spacer to give a simplified cyclopentapeptide (compound **36**), or by using 5-aminopentanoic acid (5-Apa), giving the same ring size (compound **37**, Figure 27). Moreover, the importance of ring size (ring expansion/contraction strategy) was investigated by using ω-amino carboxylic acid spacers of different length (compounds **38-41**, Figure 27).

## 10.2 Chemistry

### 10.2.1 Synthesis of *N*-acetylated D-Arg[1]-Arg[2]-2-Nal[3] tripeptide amide

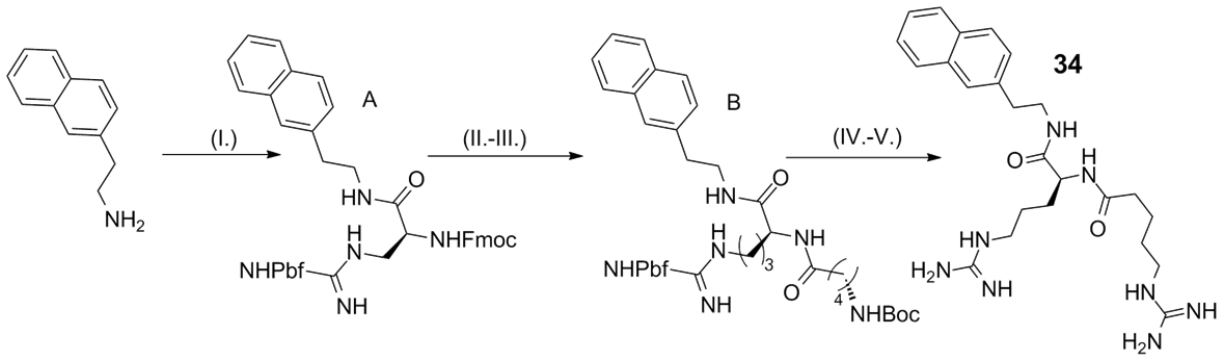
The linear tripeptide analogue **35** (Figure 26) was prepared by SPPS on an Fmoc-NH-Rink amide resin affording *N*-terminal acetylated tripeptide resin upon completion (Scheme 2). Cleavage from resin gave the *C*-terminal amidated tripeptide which was purified (RP-HPLC), and lyophilized to final product.




**Scheme 2.** Synthesis of linear compound **35.**

**10.2.2 Synthesis of linear compound 34**

In solution preparation of linear compound **34** was performed by means of a five-step synthesis as depicted in Scheme 3. Both HBTU and HATU were employed as activating reagents for the amide coupling reactions in presence of a tertiary base (DIPEA), and 2-ethanolamine was used for accelerated Fmoc-group deprotection. Treatment of unit B with TFA-cleavage cocktail removed both the guanidine protection (Pbf) and the *tert*-butyl group. Guanylation of the free amino group (of unit B) was completed with no presence of diguanylated byproducts detected during monitoring of the reaction. Intermediate products were purified by flash chromatography, and the final crude was purified by RP-HPLC and lyophilized to afford compound **34**.




**Scheme 3.** Reagents and conditions: (I.) HBTU, DIPEA, Fmoc-Arg(Pbf)-OH, dry DMF; (II.) 2-ethanolamine, DMF; (III.) HATU, DIPEA, N-Boc-5-aminopentanoic acid, DMF (IV.) TFA/TIS/$H_2O$ (95:2.5:2.5); (V.) 1H-pyrazole-1-carboxamidine hydrochloride, DIPEA, DMF.

The experimental procedure for the synthesis of Gly-containing cyclopentapeptide ligands (**33**, **36** and **41**) and macrocyclic analogues (**37-40**) was described in section 9.2.

### 10.3 Biological Activity: Linear and macrocyclic analogues

The antagonistic potency of the compounds included in the study was determined in a functional assay as described in section 9.3.

#### 10.3.1 SAR for linear analogues

In the attempt to determine a minimal structural motif for biological activity in CXCR4, the linear analogue **34** (Figure 28), with reduced peptide backbone was prepared. Compound **34** displayed moderate activity with a 112-fold lower potency than the parent ligand **33** reflecting the extensive downsize of the cyclopentapeptide structure as well as the increased flexibility. However, considering the extensive peptide dissection to derive **34**, the outcome (58 μM, Figure 28) represents a rather decent activity profile to confirm the importance of the three functionalities. Similar linear tripeptidic CXCR4 antagonists based on the Arg-Arg-Nal motif were reported before,[166, 189] having much the same activity as **34,** in comparison to the parent cyclopentapeptide.

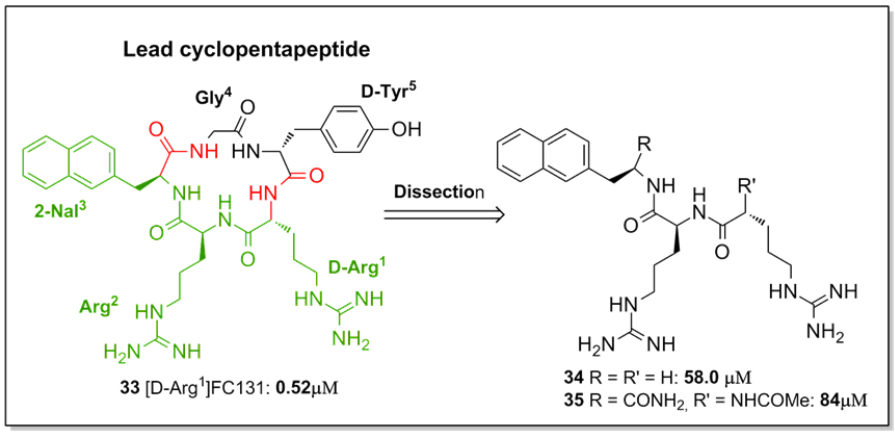



**Figure 28.** Parent ligand **33** and linear analogues **34** and **35** with $EC_{50}$ values.

Compound **35**, was less active than **34**, showing that the terminal amide groups are not contributing favorably to activity.

An apparent issue with the increased flexibility of such linear analogues is the negative contribution to the binding affinity for the receptor.[187] Thus, from a drug design perspective, a further optimization strategy, should aim to the introduction of more conformational restriction (see 10.1.2).

### 10.3.2 SAR for macrocyclic compounds

The Gly[4]-Gly[5] dipeptide linker-analogue **36** displayed moderate activity ($EC_{50}$ = 44 $\mu$M) with 2-fold increase in potency relative to the linear analogue **35** (see Figure 28), while simplification of **36** by replacement of Gly[4]-Gly[5] with the flexible 5-Apa[4] hydrocarbon spacer (**37,** same ring size n=4) resulted in significantly reduced potency (>100 $\mu$M). Accordingly, the contribution from the Gly[4]-Gly[5] amide bond in compound **36** is favorable to activity, either by a geometrical effect or through direct binding interactions. For the ring expansion/contraction strategy, analogues **38-40** with ring expansion (n=5), and ring contraction (n=3 and n=2) showed lower activity than **36** (>100 $\mu$M). However, the activity trend was changed in the case of the most constrained analogue **41** (n=1), that was equipotent to the linear analogue **34** (Figure 28).

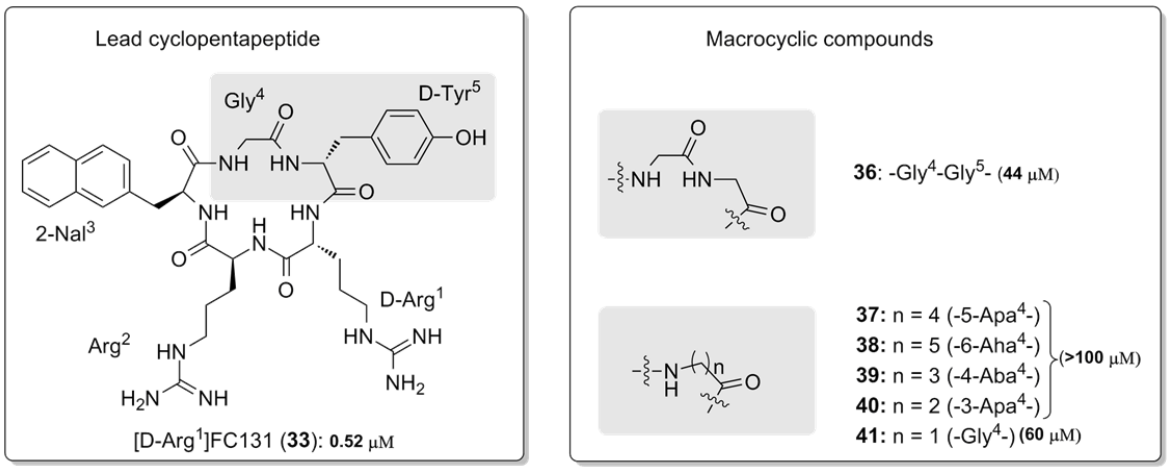



**Figure 29.** Parent ligand **33** and macrocyclic compounds **36-41** with $EC_{50}$-values.

Even if compound **41** displays same activity profile as the flexible linear analogue **34**, in a drug design context, the physicochemical properties of the former molecule are significantly improved through cyclization while the same moderate activity is retained. Overall, the implication is that ring contraction and hence, conformational restriction, could be a key strategy in the search for a more rigid scaffold to accommodate the pharmacophoric side-chains (see chapter 11).

## 11. DESIGN & SYNTHESIS OF SCAFFOLD-BASED TRIPEPTIDOMIMETICS (PAPER II)

### 11.1 Design

#### 11.1.1 Background

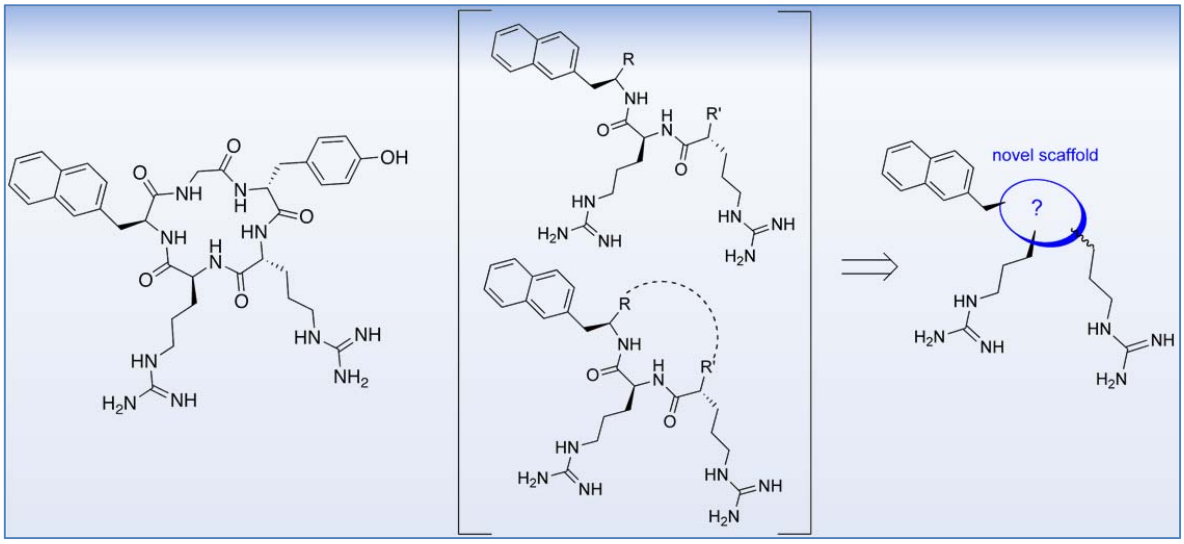



**Figure 30.** Background work suggesting further optimization through a novel rigidified scaffold to incorporate the three key side chains.

In a systematic approach using the D-Arg[1][FC131] epimer cyclopentapeptide as a reference, we initially dissected the cyclopentapeptide structure and linear tripeptidic analogues were obtained with moderate activity (sub-section 10.3.1). We further reintroduced cyclic constraints to reduce flexibility in the linear analogues, and successfully identified that activity was retained in the smallest ring-size analogue (**41**, 10.3.2).

A further consideration overall in our strategy to rationally design peptidomimetics was our intention to retain an important characteristic of the reference cyclopentapeptides, which is the mimicry of peptide turns[190] (see also section 3.2). Turns appear to be a common ligand recognition element for peptides and GPCRs.[191] Encouraged by the activity of **41** (see 10.3.2) we proceeded therefore to identify a more constrained scaffold capable of mimicking the turn-structure of the parent cyclopentapeptide **33**; thereby, presenting the side chains and backbone of the D-Arg[1]-Arg[2]-2-Nal[3] fragment in the required 3D-

orientation (topographical peptidomimetics).[172] Overall, we desired a more constrained and rigid skeleton which could potentially lead to more drug-like candidates.

#### 11.1.2 Prototype bicyclic tripeptidomimetics

In a pharmacophore-based approach, design of such turn-mimetics relies on the determined bioactive conformation of the parent (peptidic) ligand. We therefore employed our previously reported a 3D pharmacophore model that describes the spatial arrangement of the pharmacophoric side chains as well as the bioactive conformation of the cyclopentapeptide backbone.[170] We wanted to maintain the spatial orientation of the cyclopentapeptide pharmacophore by using a template which would not extend beyond the boundaries of the peptide. For this purpose we chose the tripeptide-derived 3,6,8-trisubstituted bicyclic structure **A** (Figure 31A) to replace the peptide backbone.

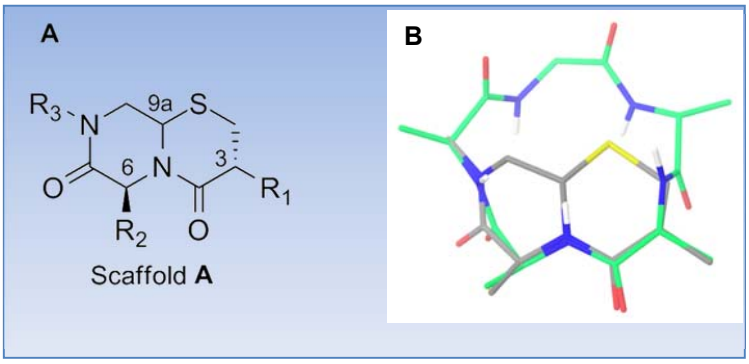



**Figure 31. A.** Structure of scaffold **A**; **B.** Superimposition of a low-energy conformation of **A** (grey carbon atoms) and the bioactive backbone conformation of the cyclopentapeptide CXCR4 antagonists (green carbon atoms) as defined by our 3D pharmacophore model.[170]

A structural comparison of low energy conformations of **A** with our 3D pharmacophore model indicated that side chains in scaffold **A** were oriented in a similar way as the parent cyclopentapeptide (Figure 31B).

Two of the three stereocenters of Scaffold **A** (Figure 32) can be synthetically predefined (see also Chemistry), and therefore two diastereomeric scaffolds (**A** and **A'** Figure 32A) are considered. An overlay of low-energy conformations of the two scaffolds (shown in Figure 32B) indicated almost identical conformations with respect to the orientation of the two amide bonds and the three sidechains (Figure 32B).

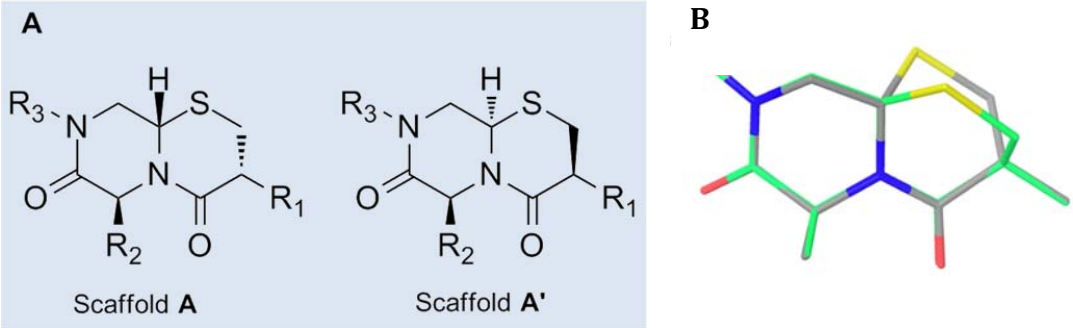



**Figure 32. A.** Structures of the diastereomeric scaffolds **A** and **A'; B.** Superimposition of low-energy conformations of **A** (grey carbon atoms) and **A'** (green carbon atoms).

Novel, scaffold-based tripeptidomimetic CXCR4 antagonists were therefore pursued based on **A** and **A'** bicyclic templates. As illustrated in Figure 33, two different pairs of compounds were targeted. The first pair (**42** and **43**) was based on use of cysteine and glycine to construct the $R^1$ side chain, resulting in an amide bond in $R^1$. The other pair (**44** and **45**) contained the same arginine $R^1$ side chain as the parent cyclopentapeptide.

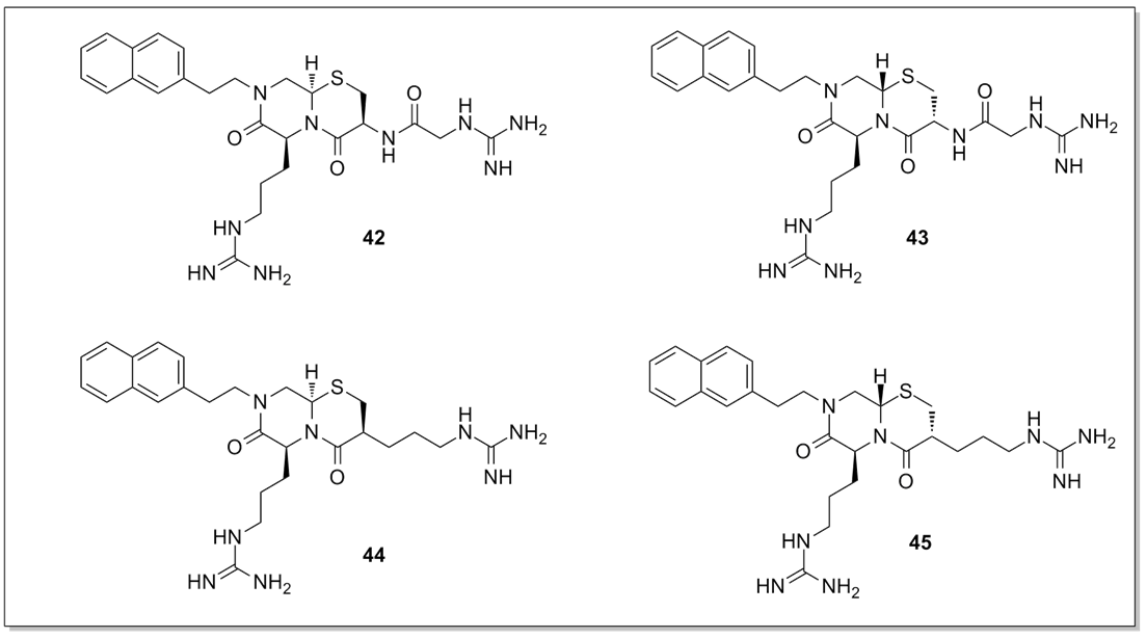
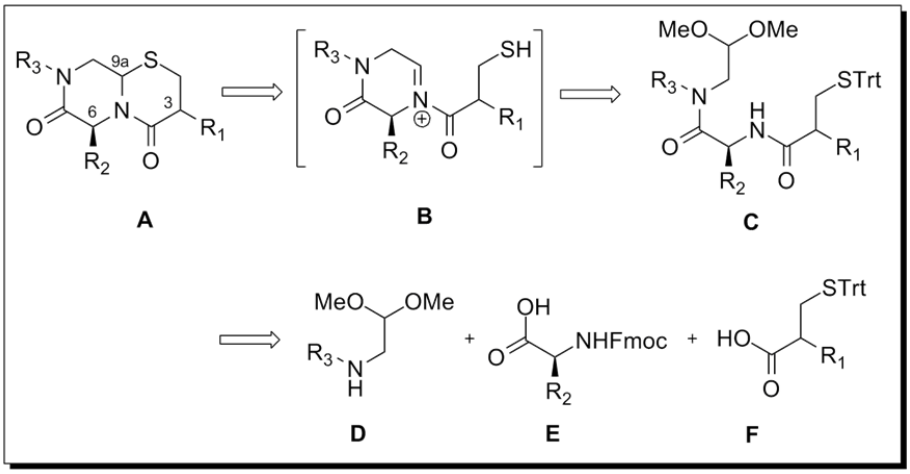



**Figure 33.** Targeted prototype molecules.

### 11.2 Chemistry

A retrosynthetic analysis of scaffold A is depicted in Scheme 4. Accordingly, building unit C is prepared through amide bond condensations of units D-F in a stepwise manner. A critical step is the TFA treatment of the acetal unit C whereby, a transient aldehyde is formed which undergoes condensation with adjacent amide nitrogen at the backbone to form the *N*-acyliminium ion intermediate B. Furthermore, nucleophilic attack from the deprotected thiol group leads to formation of a second ring, generating thus, a new stereocenter (the third) in the target bicycle A.[192, 193]




**Scheme 4.** Retrosynthetic approach to obtain bicycle A. The three stereocenters are numbered by carbon-atom.

*Stereoselective cyclization:* In a study by Grimes et al.,[194] it was concluded that the cyclization step occurs stereoselectively; the configuration at the bridgehead (C-9a) depends on the configuration at C-3, and in the absence of a $R^1$ substituent on the configuration at C-6. Accordingly, the $R_1$ substituent (at C-3) has an influence on the stereodirection of cyclization with *si* attack favored, giving a *cis* relationship between $H_{9a}$ and $H_3$ (Figure 34). In the absence of $R^1$ substitution at C-3, the substituent at C-6 ($R^2$) exerts some effect on the stereoselectivity, and the predominating effect favors the result of si attack with *trans* relationship between $H_6$ and $H_{9a}$.[194] The assigned stereochemistry by NMR on the obtained bicycles (**42-45** including unpublished data) was in agreement with this proposed outcome.

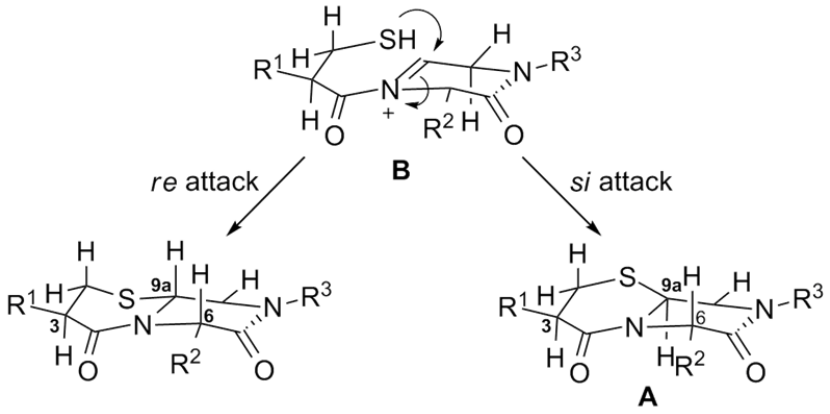
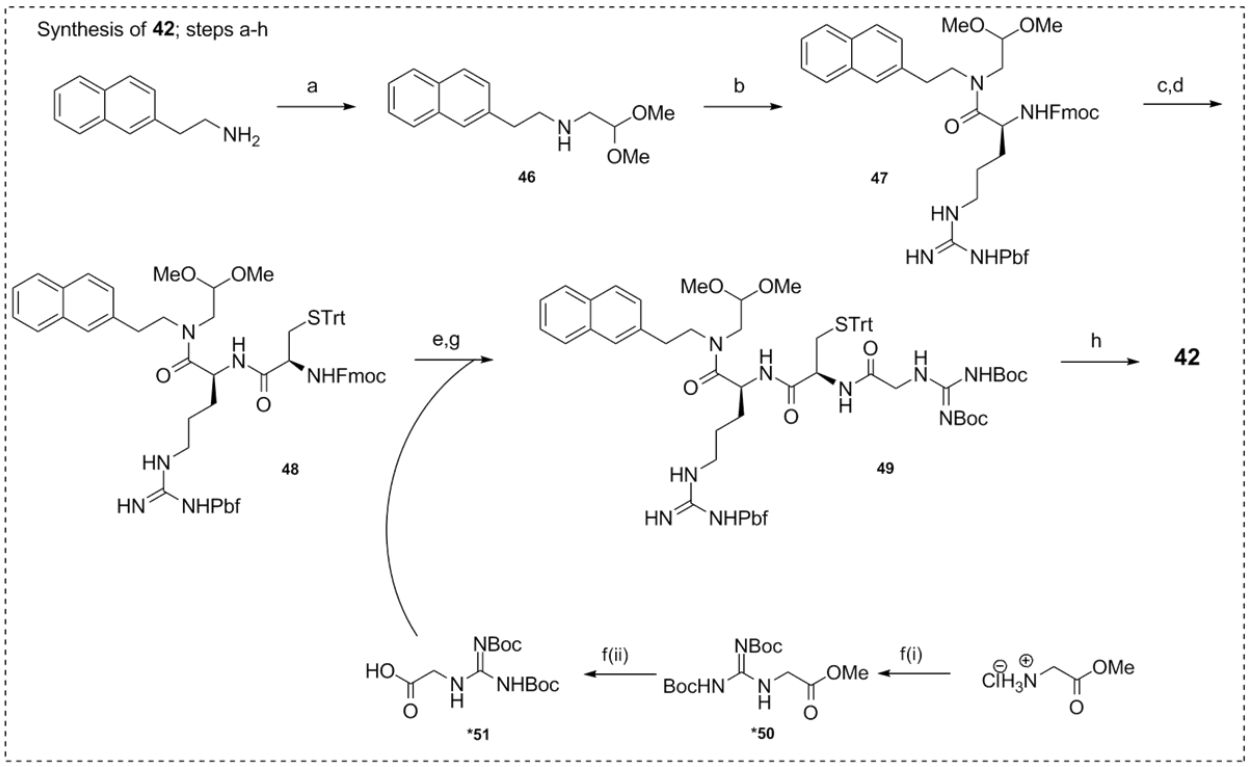



**Figure 34.** Proposed mechanism for cyclization.[194]

Unit D (Scheme 4) was commonly prepared through *N*-alkylation of the aromatic ethylamine using bromoacetaldehyde dimethyl acetal as depicted in Scheme 5. Moreover, a protected Arg residue was used as unit E in the synthesis of the prototypes **42-45**. The synthetic strategy employed in the preparation of unit F varied according to the targeted $R^1$-side chain structure.

*Synthesis of analogues **42** and **43***. For analogues **42** and **43** (see Figure 33 for structures), Fmoc-Cys(Trt)-OH (L-configuration for **42,** and D-configuration for **43**) was used in conjunction with a guanidinylated glycin (**51,** Scheme 5) for the incorporation of $R^1$-side chain (for illustration, see synthesis of **42** in Scheme 5). **51** was prepared by guanidinylation of glycin methyl ester hydrochloride using *N,N*-di-Boc-1H-pyrazole-1-carboxamidine followed by hydrolysis of the methyl ester of the resulting **50** using LiOH in a mixture of water and acetone.[195, 196] In brief, synthesis of analogue **42** (Scheme 5) was achieved through the alkylation of 2-(naphthalene-2-yl)ethanamine with bromoacetaldehyde dimethyl acetal in refluxing THF to give secondary amine **46**. This amine was in turn coupled with Fmoc-protected arginine to give **47**. Further Fmoc-deprotection and coupling with appropriately protected L-cysteine gave **48**, which was submitted to another Fmoc-deprotection, and then coupled with carboxylic acid **51** to give the linear precursor **49**. Treatment with TFA, thioanisole and water facilitated global deprotection leading to formation of the acyliminium ion intermediate that after

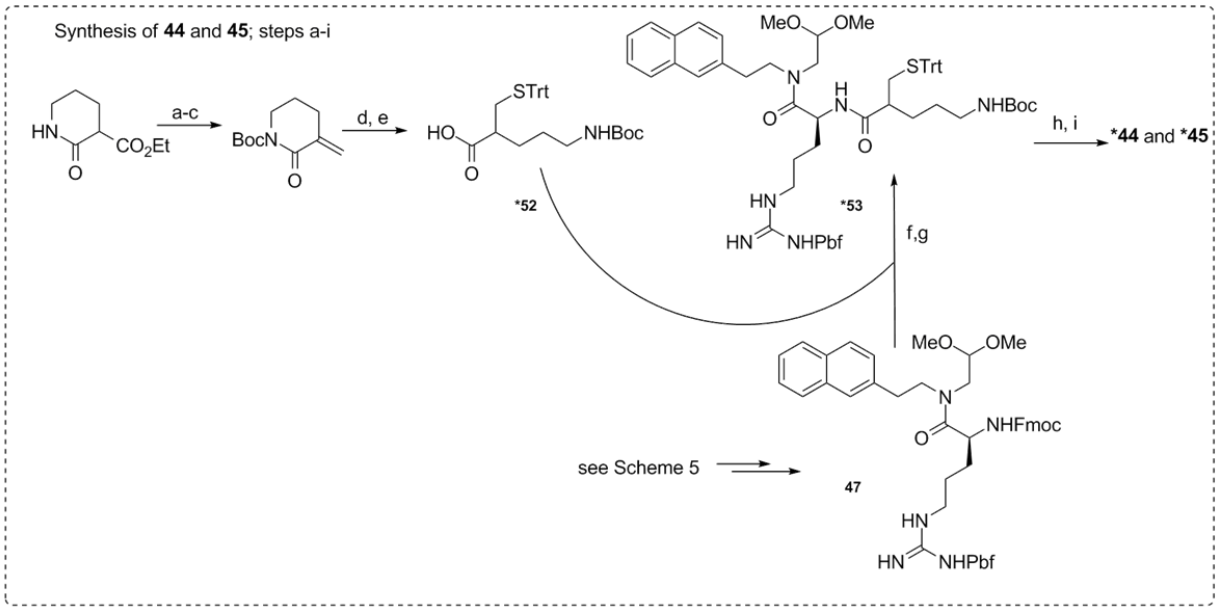
nucleophilic attack by the thiol gave target compound **42**. Similarly compound **43** was prepared using D-cysteine to prepare the intermediate product **48**. Upon deprotection and cyclization, the configuration of the newly formed stereocenters of **42** and **43** (see Figure 33) were determined using the 2D $^{1}$H ROESY experiment. The known configurations of C-6 (*S* for both **42** and **43**) and C-3 (*R* for **42** and S for **43**) were used as prerequisites for determination of the configuration of C-9a (*S* for **42** and *R* for **43**).




**Scheme 5.** *Reagents and conditions:* **Synthesis of 42:** (a) $BrCH_2CH(OMe)_2$, THF, reflux; (b) Fmoc-Arg(Pbf)-OH, HATU, DIPEA, DMF; (c) $Et_2HN$, $CH_2Cl_2$; (d) Fmoc-D-Cys(Trt)-OH, HATU, DIPEA, $CH_2Cl_2$; (e) $Et_2HN$, $CH_2Cl_2$; (fi) *N*,*N*-di-Boc-1*H*-pyrazole-1-carboxamidine, DIEA, DMF; (fii) LiOH, $H_2O$/acetone; (g) HATU, DIPEA, $CH_2Cl_2$; (h) TFA/thianisole/$H_2O$.

***** Building units **50**, **51** were prepared by Pernille Rasmussen.[197]

*Synthesis of analogues **44** and **45**.* For analogues **44** and **45** unit F (**52**) was afforded as a racemic carboxylic acid mixture after a multistep synthetic strategy as depicted in Scheme 6. Fmoc deprotection of **47** (see Scheme 5) and subsequent coupling with **52** gave **53** as inseparable mixture of diastereoisomers. After deprotection of **53** and cyclization (as previously described for **42**) the amine group ($R^1$ side chain) was guanidinylated to give **44** and **45**, which were separated upon purification by RP-HPLC. The configuration at the formed stereocenters was determined using the 2D $^1$H ROESY experiment (for **44**: C-3 (R) C-9a (S) C-6 (S) and for **45:** C-3 (S) C-9a (R) C-6 (S)).




**Scheme 6.** *Reagents and conditions:* **Synthesis of 44 and 45:** (a) $CaCl_2$, $NaBH_4$, MeOH, 0 °C to r.t.; (b) DCC, CuI, toluene, 110 °C; (c) $Boc_2O$, $Et_3N$, DMAP, $CH_2Cl_2$/DMF; (d) $Ph_3CSH$, $Et_3N$, $CH_2Cl_2$; (e) 1M aq. LiOH, THF; (f) $Et_2HN$, $CH_2Cl_2$; (b) HBTU, DIPEA, $CH_2Cl_2$; (h) TFA/thioanisole/$H_2O$; (i) 1*H*-pyrazole-1-carboxamidine hydrochloride, DIPEA, DMF.

*****Compounds **44**, **45** were prepared by Erik A. Berg.[198]

Intermediate products were purified by flash column chromatography and final products were purified by RP-HPLC.

**11.3 Biological Activity**

The antagonistic potency of the compounds included in the study was determined in a functional assay as described in sub-section 9.3.

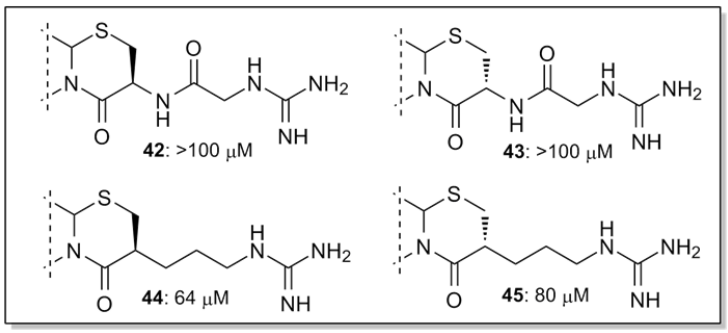



**Figure 35.** Prototype bicyclic tripeptidomimetics displayed with the $R_1$-sidechain; $EC_{50}$ values shown.

Compounds **42** and **43** exhibited low to no activity (>100 $\mu$M), and the lower activity than **44** and **45** can be therefore attributed to the peptide bond introduced in $R^1$ side chain. The planarity of the amide bond in combination with a conformationally constrained $R^1$ side chain had an unfavorable effect on the biological activity of **42** and **43**. Clearly, keeping the structural and topographical identity of the key side chains intact (as in compounds **44** and **45**), appeared to be more beneficial for biological activity.

Furthermore, compounds **44** and **45** which have opposite stereochemistry in two of the three chiral centers ((3R, 9aS, 6S and 3S, 9aS, 6S respectively), were found to be almost equipotent. This suggests that the D-/L- configuration of $R^1$ sidechain does not contribute significantly to the receptor binding conformation of these ligands; which is further supported by the comparison of low-energy conformations of the two diastereomeric scaffolds displaying their side chains in similar orientations (Figure 32). Overall, the scaffold-based tripeptidomimetics **44** and **45** are significantly less potent than the parent cyclopentapeptides (FC131 and D-Arg$^1$[FC131]) that they are based on; considering however, the rather extensive structural change through this scaffold-hop, the obtained analogs represent promising `hits´.

In two earlier attempts to develop scaffold-based peptidomimetic CXCR4 antagonists (see sub-section 7.2.2, Figure 18, compounds **9**, **10**),[167, 168] a drop in activity for initial compounds compared to the highly optimized FC131 was also found. This further reflects the general complexity of the initial "scaffold hop" for prototype compounds. Further studies into the optimal chain length for the three substituents and the nature of the

pharmacophoric side chains, is currently in progress (see section 11.4). These studies should provide further insight into the potential of the scaffold reported herein, and SAR data will be reported in due course.

### 11.4 Additional compounds

Based on the collective knowledge on the pharmacophore and SAR for the cyclic pentapeptide analogues in conjunction with the novel hits represented by the prototype compounds **44** and **45** (Figure 33), additional bicyclic analogues have been designed, and synthesized in the course of the present project and considerable time and effort has been devoted to this work. Presently, 11 new analogues (including diastereomers) have been prepared and purified; there are now in various stages with respect to structure determination (NMR) and biological evaluation. The design, synthesis, characterization and biological activity of these compounds will eventually be repeated as part of a separate paper; consequently, structures and experimental data are not disclosed here.

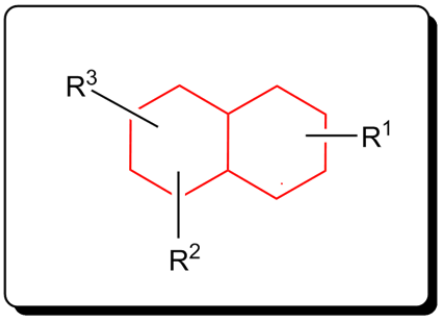



**Figure 36**. Further structural tuning and optimization.

## 12. BINDING MODE FOR THE TRIPEPTIDOMIMETIC CXCR4 ANTAGONIST KRH-1636 (PAPER III)

### 12.1 Design and SAR

After its discovery[169] in 2003, very little is known about the prototype CXCR4 antagonist KRH-1636. In order to investigate the importance of the different functionalities of this potent antagonist (compound **54**, Figure 37), we first carried out a limited SAR study. A selection of structurally modified analogues were synthesized targeting the three functionalities ($R^{1-3}$ positions as depicted in Figure 37), and along with **54**, analogues **55-59** were prepared.

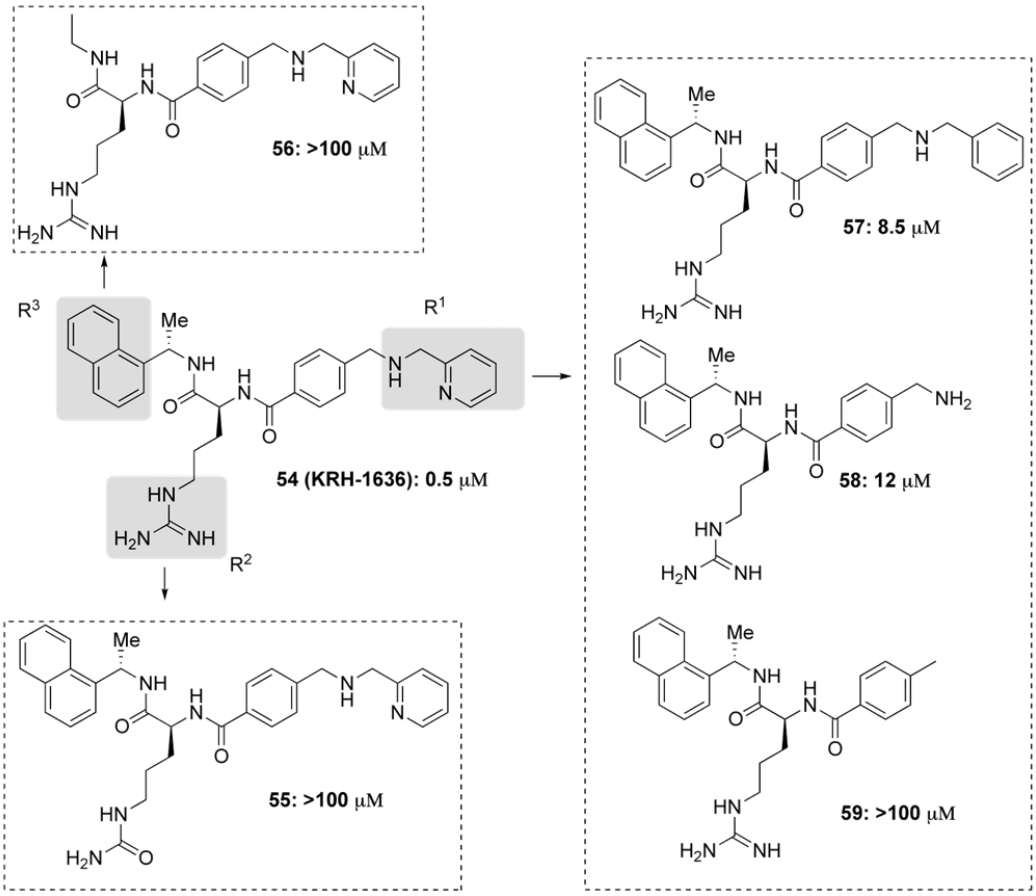



**Figure 37.** Overview on design strategy for preparation of KRH-1636 (**54**) and analogues **55-59**; [$R^1$-analogues **57-59**; $R^2$-analogue **55**; $R^3$ analogue **56**].

Focusing initially on the $R^1$-position of KRH-1636, the *N*-(pyridine-2-ylmethyl) moiety was gradually removed in analogues **55-59**. Main properties of interest for this functionality

($R^1$) are the positive charge of the dialkylated amine at physiological pH, and the potential of pyridine ring interactions with the receptor residues. In analogues **55**, the guanidine was substituted with a neutral isostere urea group, hence neutralizing the charge, and in **56** the aromatic group of KRH-1636 was removed.

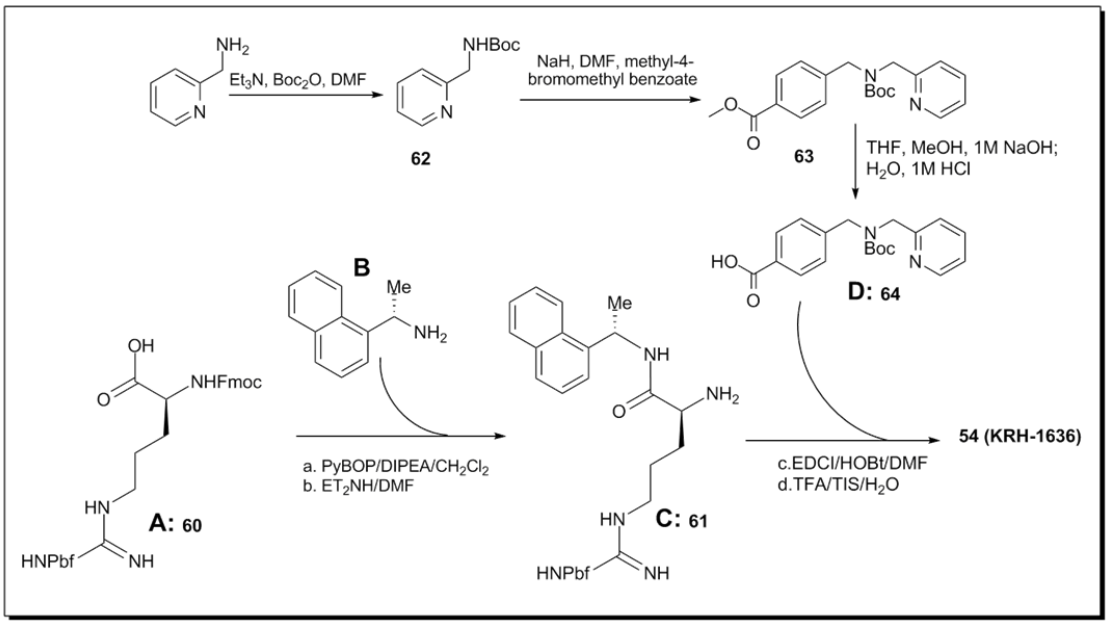
The CXCR4 antagonistic potency of **54-59** along with the lead compound KRH-1636 was assessed in a functional assay as described in section 9.3. Analogue **57,** displayed lower potency (by 17-fold) compared to **54** indicating the importance of the pyridine nitrogen. The pyridine ring was further removed affording a free amine in analogue **58**, which showed somewhat lower potency than **57** revealing thus, the role of the pyridine-2-ylmethyl group. Moreover, removal of the amino group in **59** had a detrimental effect on activity, confirming the charged secondary amine group as indispensable part of the KRH-1636 pharmacophore. A significant loss in activity compared to parent ligand **54** was furthermore found for both compounds **55** and **56**, with modifications in $R^2$- and $R^3$-positions, respectively; thus, the positive charge of the guanidine group and the aromatic moiety (naphthyl) are essential pharmacophoric elements.

In short, the results from this initial SAR study suggested that the aromatic group, the Arg, and the secondary amine in $R^1$-position could be considered as highly essential for the activity of KRH-1636, while the pyridine ring certainly contributes to the higher potency.

**12.2 Chemistry**

KRH-1636 (**54**), and analogues **55-59** (see Figure 37 for structures) were prepared based on modified procedures previously reported for the synthesis of KRH-1636[199] (Scheme 7). For KRH-1636, unit **C** was prepared by coupling of **60** (unit **A:** Fmoc-Arg-OH) with (*S*)-1-(1-naphthyl) ethylamine (Unit **B**) using PyBOP as the coupling reagent, followed by treatment of **61** with $Et_2NH$ in DMF for Fmoc-deprotection to give free amine.

Unit **D** (**64**) was prepared by *N*-alkylation of Boc-protected 2-(aminomethyl) pyridine **62** (Scheme 7) with methyl 4-bromomethyl benzoate, to give carboxylic acid (**64**) after hydrolysis of the methyl ester (**63**). Coupling of the two halves was facilitated by EDCI/HOBt to give KRH-1636 after removal of the Boc- and Pbf protecting groups by treatment with TFA/TIS/$H_2O$ cleavage cocktail.




**Scheme 7.** Synthesis of KRH-1636.

For the preparation of analogues **58** and **59**, the left hand side (unit **C**) was coupled with N-Boc 4-(aminomethyl)benzoic acid and 4-methylbenzoic acid, respectively. Analogue **57** was prepared by *N*-alkylation of **58** with benzyl bromide using $K_2CO_3$ as base and DMF as the solvent. The $R^2$ analogue **55** was prepared using Fmoc-protected citrulline as the starting material (unit **A**), while the $R^3$ analogue **56** was prepared by coupling of ethyl amine (unit **B**) to Fmoc-Arg-OH to afford the unit **C**.

After standard workup, all intermediate products were purified by flash column chromatography on silica gel, and final products were purified by RP-HPLC and lyophilized.

### 12.3 Site-directed mutagenesis

Mutagenesis studies were carried out in order to experimentally determine key binding interactions of KRH-1636 with the CXCR4 receptor. Residues previously found to be involved in the binding modes of FC131,[145, 200, 201] as well as cyclam-based CXCR4 antagonists (AMD3100 and analogues)[202-204] were selected, and a library of 24 CXCR4-mutants was constructed. Most of the mutations included a replacement with Ala; however, certain acidic residues such as Asp were alternatively subjected to isosteric substitutions

with the structurally similar and uncharged Asn residue. Selective substitutions with Trp were also done in order to introduce more bulky side-chains.

#### 12.3.1 Functional assay

*Method.* The antagonistic potency of the compounds was determined in a functional assay (SPA-PI turnover assay and/or the "traditional" $IP_3$-assay) measuring inhibition of CXCL12-induced activation of wt- and mutant-CXCR4 receptors transiently expressed in COS-7 cells. The methods were described in section 9.3.

*Results.* The mutants (Table 1) were initially tested in a functional assay with primary focus to determine the ability of KRH-1636 (**54**) to inhibit CXCL12-induced activation. The KRH-1636 antagonistic potency was significantly reduced upon mutations of TMH 3 residue $His^{113}$, TMH 6 residue $Asp^{262}$ and TMH 7 residue $His^{281}$ (Table 1). Moreover, the potency of KRH-1636 was also reduced upon mutations of $Tyr^{45}$, $Asp^{171}$, and $Gln^{200}$ in TMH 1, -4 and -5, respectively and hydrophobic residues $Trp^{252}$ and $Ile^{259}$ in TMH 6 and $Ile^{284}$ in TMH 7. Substitutions of the bulky $Trp^{94}$ residue and the acidic $Asp^{97}$ residue in TM-2 to Ala (W94A and D97A) led to increased potency (>2-fold).

In order to probe which parts of the ligand interact with the different receptor parts, the $R^1$ analogs **57** and **58** (Figure 35) were also included in the mutagenesis study to further investigate the molecular interactions of the structurally modified $R^1$-side chain for the three analogs with CXCR4.

Interpretation of the functional-SDM data (Table 1) for all three ligands indicated furthermore, that mutations in TMHs 1-4 and ECL-2 affected the potency of the ligands in a similar manner, while several of the mutations in TMHs 5-7 affected the three ligands in a distinctive manner. In sum, the overall trend identified in the functional experiments for analogues **54**, **57** and **58** (Table 1) suggests that the structurally distinct $R^1$ group contacts TMHs 5-7.

The CXCR4-mutants H113A-, D171N-, D262N- and H281A-CXCR4 were consistently reported to be involved in the binding modes of other prototype CXCR4 antagonists such as the AMD-series analogues -3100 and -3465.[203-205] (see Figure 17). The effect of these mutations on the potency of KRH-1636 compared to wt-CXCR4, is depicted in Figure 38.

Although, the majority of the mutations did not affect the ability of CXCL12 to activate the receptor (see Table 1; similar potencies to the wt-CXCR4 receptor were observed), the E288A, D97A, D187A, W94A and Y116A mutants reduced the potency of CXCL12 to activate the receptor by an 8-14 fold reduction (Table 1).

Substitutions of residues $Glu^{288}$ (TMH 7), $Asp^{187}$ (ECL-2) and $Asp^{97}$ (TMH 2) to alanine (E288A-, D187A- and D97A-CXCR4 mutants), were previously reported to affect the binding of the endogenous ligand CXCL12.[75, 202, 206] Moreover, $Tyr^{116}$ and $Trp^{94}$ along with the three aforementioned residues are implicated as part of "the site two" of "the two-site model" for CXCL12 binding to CXCR4 receptor[75, 207] (see also section 5.2.2). Based on these findings, a potential limitation in using CXCL12 to probe the effects of mutation on those residues in the functional assay is therefore suggested.

**Table 1. Functional data**. $F_{mut}$ indicates the difference between the potency on wt- and mutant CXCR4. Red indicates $F_{mut}$ >25, orange indicates $F_{mut}$ from 10-25, while yellow indicates $F_{mut}$ from 5-10 and green indicates $F_{mut}$ <0.5.

| | Residue | | CXCL12 | | KRH-1636 | | Analog 57 | | Analog 58 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Position[a] | Number | $EC_{50}$ *nM* | $F_{mut}$ | $EC_{50}$ *μM* | $F_{mut}$ | $EC_{50}$ *μM* | $F_{mut}$ | $EC_{50}$ *μM* | $F_{mut}$ |
| **wt** | | **CXCR4** | 1,5 | 1,0 | 0,50 | 1,0 | 8,5 | 1,0 | 12 | 1,0 |
| TM-1 | I:07/1.39 | Y45A | 3,4 | 2,3 | 4,0 | 8,0 | 56 | 6,6 | 181 | 15 |
| TM-2 | II:20/2.60 | **W94A*** | 18 | 13 | 0,06 | 0,12 | 0,6 | 0,07 | 0,64 | 0,05 |
| | II:23/2.63 | **D97A*** | 20 | 14 | 0,21 | 0,41 | 0,54 | 0,06 | 1,4 | 0,12 |
| TM-3 | III:05/3.29 | **H113A*** | 0,84 | 0,58 | 13 | 26 | > 100 | | > 100 | |
| | III:08/3.32 | **Y116A*** | | | | | | | | |
| | III:09/3.33 | T117A | 1,7 | 1,17 | 0,17 | 0,34 | 0,87 | 0,10 | 1,4 | 0,12 |
| TM-4 | IV:20/4.60 | **D171N*** | 3,2 | 2,2 | 2,7 | 5,3 | 68 | 8,1 | 231 | 20 |
| ECL-2 | Cys+1 | D187A | 13 | 9 | 0,35 | 0,69 | 4,1 | 0,49 | 5,4 | 0,46 |
| | Cys+2 | R188A | 0,53 | 0,36 | 0,86 | 1,7 | 8,4 | 0,99 | 25 | 2,1 |
| | Cys+3 | F189A | 1,8 | 1,2 | 0,13 | 0,26 | 0,74 | 0,09 | 1,0 | 0,09 |
| | Cys+4 | Y190A | 1,2 | 0,82 | 0,27 | 0,53 | 12 | 1,4 | 12 | 1,01 |
| TM-5 | V:01/5.35 | V196A | 1,4 | 0,96 | 1,4 | 2,8 | 4,2 | 0,50 | 13 | 1,1 |
| | V:04/5.38 | F199A | 1,4 | 0,98 | 0,69 | 1,4 | | | | |
| | V:05/5.39 | Q200A | 1,4 | 0,9 | 3,7 | 7,4 | 17 | 2,0 | 250 | 21 |
| | V:05/5.39 | Q200W | 1,8 | 1,2 | 10 | 21 | 14 | 1,6 | 85 | 7,2 |
| | V:08/5.42 | H203A | 1,4 | 1,0 | 0,71 | 1,4 | | | | |
| TM-6 | VI:13/6.48 | W252A | 0,78 | 0,53 | 3,8 | 7,5 | 27 | 3,2 | 173 | 15 |
| | VI:16/6.51 | Y255A | 1,1 | 0,77 | 1,7 | 3,3 | | | | |
| | VI:20/6.55 | I259A | 2,1 | 1,4 | 2,6 | 5,1 | | | | |
| | VI:20/6.55 | I259W | 1,3 | 0,91 | 1,5 | 2,9 | | | | |
| | VI:23/6.58 | **D262N*** | 5,8 | 4,0 | 13 | 26 | 27 | 3,2 | 11 | 0,9 |
| TM-7 | VII:-02/7.32 | **H281A*** | 1,8 | 1,2 | 31 | 61 | 53 | 6,2 | 155 | 13 |
| | VII:02/7.35 | I284A | 2,3 | 1,6 | 3,4 | 6,7 | 34 | 4,0 | 306 | 26 |
| | VII:06/7.39 | **E288A*** | 14 | 9,6 | 1,7 | 3,4 | | | | |

* Also tested in binding assay. The data are presented as the mean of at least three independent experiments.

[a] The position of each residue given is based on the generic numbering system proposed by Baldwin and modified by Schwartz[208] followed by the Ballesteros/Weinstein numbering system.[209]

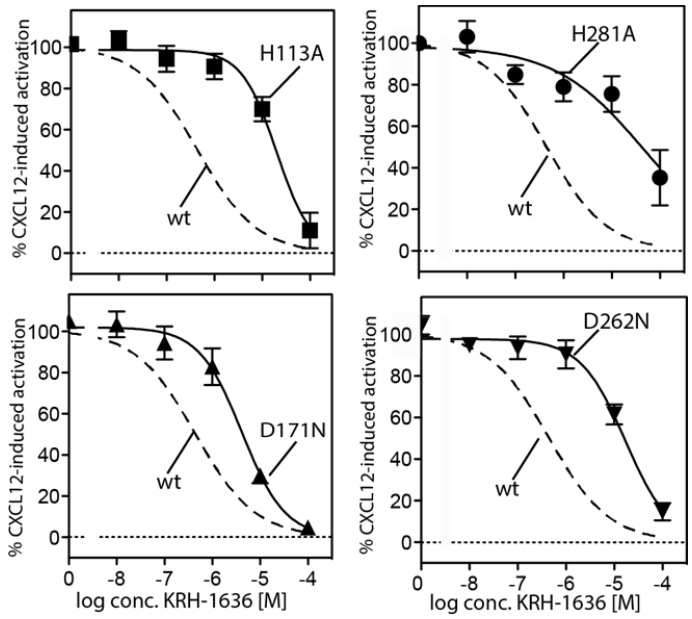



**Figure 38.** Effects of H113A, D171N, H281A and D262N on the inhibition of CXCL12-induced CXCR4 activation by KRH-1636.

**12.3.2 Binding assay**

*Method.* Competition binding experiments were carried out with [$^{125}$I]-12G5 used as radioligand, on transiently transfected COS-7 cells. The affinity was determined based on the ability of the tested compounds to displace radiolabeled monoclonal antibody [$^{125}$I]-12G5. Nonspecific binding was determined in the presence of unlabeled 12G5.

*Results.* A competitive binding assay was conducted to address the issue of reduced activity of CXCL12 in the five (W94A, E288A, D97A, D187A, Y116A)-CXCR4 mutants, and the monoclonal antibody 12G5 was used to assess the ligand affinity. In addition to these five mutants, a number of mutations found to be important during the functional assay were also included (Table 2). All mutants had close to wild-type receptor affinities for 12G5 suggesting that they did not affect the receptor cell surface expression.

Despite the strong antagonistic potency measured in the functional assay, KRH-1636 displayed lower binding affinity in wt-CXCR4 measured against $^{125}$I-12G5 (Figure 39). This further suggests that this compound and the two derivatives bind allosterically to the

receptor as compared to 12G5. The two $R^1$ analogs **57** and **58** displayed furthermore even lower affinity (Figure 39) in comparison to KRH-1636 for the wt-CXCR4, and were therefore rendered unfit for further binding experiments.

The affinity of KRH-1636 determined against $^{125}$I-12G5, was strongly affected in cells expressing the TMH 3 mutants H113A and Y116A, TMH 4 mutant D171N, TMH 5 mutant D262N, and the two mutations in TMH 7, H281A and E288A. Accordingly, the aforementioned mutations reduced the $IC_{50}$-values 8- to >12-fold in respect to the wt-CXCR4 (Table 2). Furthermore, KRH-1636 could not displace 12G5 in the Y116A-mutant. In line with the functional data on antagonistic potency, residues His$^{113}$, Asp$^{171}$, Asp$^{262}$ and His$^{281}$ were also found to affect the affinity of KRH-1636 for the receptor. In addition, residues in TMH 3 and TMH 7, Tyr$^{116}$ and Glu$^{288}$ respectively, appear to be key interaction points for KRH-1636 in CXCR4.

**Table 2. Binding assay**. $F_{mut}$ indicates the difference between the affinity on wt- and mutant CXCR4. Red indicates $F_{mut}$ >10, orange indicates $F_{mut}$ from 0-10. The data are presented as the mean of at least three independent experiments.

| | Residue | | 12G5 | | KRH-1636 | |
|---|---|---|---|---|---|---|
| | | | $IC_{50}$ *(nM)* | *Fmut* | $IC_{50}$ *(μM)* | *Fmut* |
| wt | | CXCR4 | 4,7 | 1,0 | 8,0 | 1,0 |
| TM-2 | II:20/2.60 | W94A | 1,9 | 0,40 | 1,7 | 0,21 |
| | II:23/2.63 | D97A | 9,4 | 2,0 | 8,5 | 1,1 |
| TM-3 | III:05/3.29 | H113A | 2,7 | 0,6 | > 100 | >12 |
| | III:08/3.32 | Y116A | 8,7 | 1,9 | | |
| TM-4 | IV:20/4.60 | D171N | 2,2 | 0,47 | 67 | 8,3 |
| ECL-2 | Cys+1 | D187A | 16 | 3,4 | 26 | 3,2 |
| TM-6 | VI:23/6.58 | D262N | 3,7 | 0,79 | > 100 | >12 |
| TM-7 | VII:-02/7.32 | H281A | 2,4 | 0,52 | 63 | 7,8 |
| | VII:06/7.39 | E288A | 2,1 | 0,45 | > 100 | >12 |

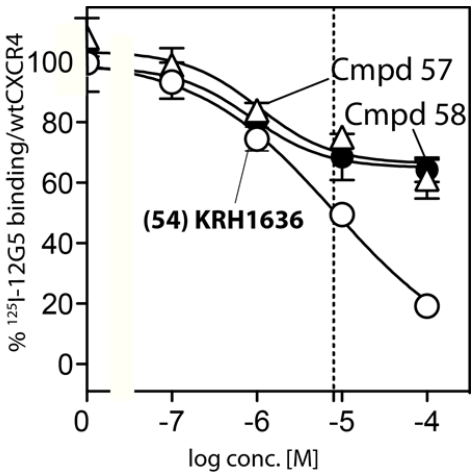



**Figure 39.** Binding affinity of KRH-1636 and $R^1$ analogues **57** and **58** in competition with [$^{125}$I]-12G5 as radioligand in wt-CXCR4,

### 12.4 Computational Modeling

In line with the experimental data, we carried out a prediction of the binding mode for KRH-1636 using the X-ray crystal structure of CXCR4.[75] The ligand was docked in the prepared protein structure using the induced-fit docking protocol.[185] All molecular modeling calculations were performed using the Maestro software packages (Schrödinger, U.S).[183]

#### 12.4.1 Procedures

*Protein preparation.* The X-ray structure of human CXCR4 bound to the 16-mer peptide antagonist CVX15 (PDB code 3OE0)[75] was imported and prepared with the Protein Preparation Wizard workflow.[184] Using the preprosessing tool, bond orders were assigned, missing hydrogen atoms added, missing disulfide bonds created, termini capped, and all water molecules at a distance greater than 5 Å from the ligand were deleted. Subsequently, the structures were refined using the H-bond assignment tool by which, protonation states (pH 7.4, PROPKA) for ionizable residues were optimized, hydrogens of altered species were minimized, and water orientations were sampled. After automated optimization, the $Asp^{262}$ was assigned correct protonation state using the interactive optimizer. The structure was then subjected to a restrained minimization (root-mean square deviation (RMSD): 0.30 Å for heavy atoms) in order to relieve strain.

*Induced-fit docking.* We chose the induced-fit docking methodology which comprises Glide to account for ligand flexibility, and Prime to account for receptor flexibility (*i.e.* side chain rotamers). The binding cavity was first defined with the enclosing box (26Å length) centered on $Asp^{187}$. The standard precision (SP) scoring function[210] was applied in the initial docking stage allowing the ligand to be docked flexibly with a softened energy potential[211]. Next, Prime was employed to sample the receptor degrees of freedom within 5.0 Å of the ligand, and a minimization of the ligand-protein complex is performed. The last stage includes a redocking of the ligand (in Glide) employing an extra precision (XP) scoring function[212] with a hard potential energy function. In the initial docking step, the number of poses to be saved was set to 50, while in the redocking step the default value (20 poses) was chosen within an energy window of 30 kcal/mol.

### 12.4.2 Derived binding model

Visual inspection of the 15 generated poses resulted in the identification of a pose that was in agreement with the experimental data. In the proposed binding model for KRH-1636 (Figure 40), the $R^1$-side chain adopts a bending conformation around the secondary amino group; the pyridine and phenyl rings lie in the same plane, and the $R^1$-side chain is globally oriented toward the TMH 6-7 region. The TMH 7 residue $His^{281}$ forms an aromatic π-π stacking interaction with the pyridine ring, and the TMH 6 residue $Asp^{262}$ is involved in bimodal interactions with the charged secondary amine of the $R^1$-side chain in KRH-1636. Furthermore, a water-mediated interaction of $Glu^{288}$ with the ligand backbone, an interaction of $Gln^{200}$ with $Asp^{262}$, and a ligand-backbone interaction with $Arg^{188}$ were detected.

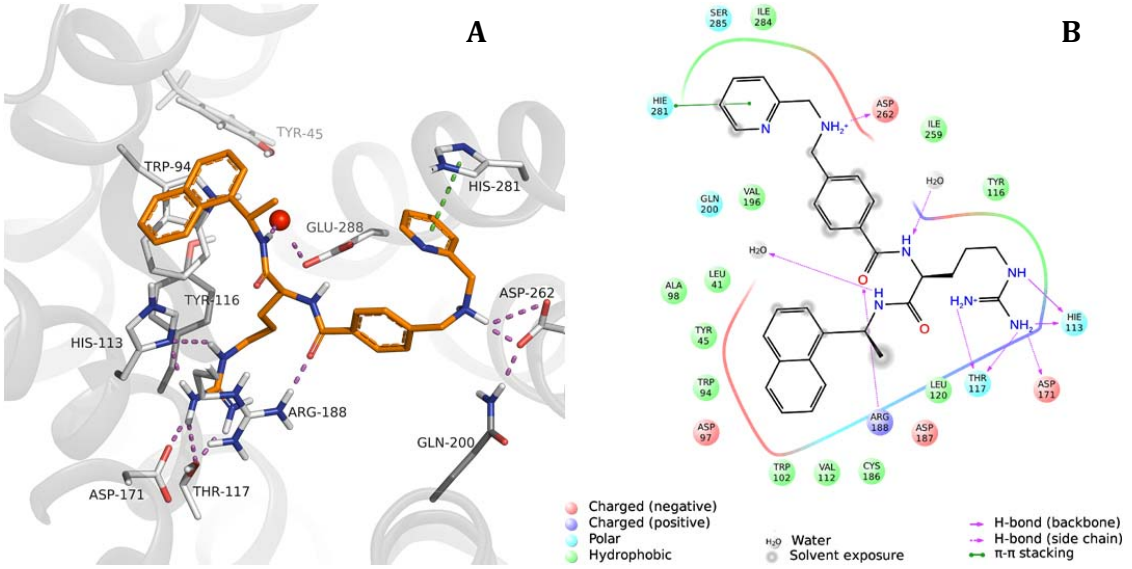



**Figure 40. A.** Binding conformation of KRH-1636 in the CXCR4 crystal structure as calculated by induced-fit docking methodology; KRH-1636 ligand is shown in orange sticks, receptor in white sticks and ribbons, and water molecule as red sphere**;** H-bonds are shown in magenta color, and π-π interactions in green color. Only main interacting residues are shown. **B.** A 2D-representation of the docking pose for KRH-1636.

**12.5 Rationalization of binding mode**

In summary, receptor mapping from the mutagenesis studies showed that both the potency and affinity of KRH-1636 were strongly dependent on the receptor residues $His^{113}$, $Asp^{171}$, $Asp^{262}$, and $His^{281}$. The antagonistic potency of KRH-1636 was also found to depend on $Tyr^{45}$ and $Gln^{200}$, while binding affinity was also strongly affected by mutation of $Tyr^{116}$ and $Glu^{288}$. The experimental data in conjunction with the derived binding model from the docking simulations suggest therefore that the L-Arg guanidino group of KRH-1636 forms polar interactions with TMH 3 and -4 residues $His^{113}$ and $Asp^{171}$ respectively, the *N*-pyridinylmethylene moiety is anchored by TMH 6 residue $Asp^{262}$ and TMH 7 residue $His^{281}$, while the naphthyl group is embedded in a hydrophobic region defined by the side-chains of $Tyr^{45}$, $Trp^{94}$, and $Tyr^{116}$ in the TMH 1-3 subpocket of the receptor.

Prior to the publication of the X-ray structure of CXCR4 in 2010,[75] an attempt to describe the binding mode of KRH-1636[213] suggested an interaction pattern involving receptor residues $Asp^{262}$, $Glu^{288}$ and $Asp^{171}$; which resembles the proposed binding mode of another potent CXCR4 antagonist, the AMD3100[202, 214] (compound **1,** Figure 16). However, the study[213] was merely based on homology modeling, using the bovine rhodopsin crystal structure as a template, which is a suboptimal representation of CXCR4,[215] without including any experimentally determined ligand-receptor interactions. Our own study provides first experimental evidence to describe the binding of KRH-1636 on CXCR4 based on SAR, site-directed mutagenesis, and molecular docking on the CXCR4 crystal structure suggesting thus, a more accurate binding model for this non-peptide antagonist.

Furthermore, it was recently suggested a binding mode for FC131 based on ligand structure- activity studies, SDM and *in silico* docking to the X-ray structure of CXCR4.[216] The proposed receptor interactions for the $Arg^2$ guanidino group in FC131 ($His^{113}$, $Thr^{117}$, and $Asp^{171}$) are consistent with the interactions of the $R^2$-guanidino group in KRH-1636 in our present study. Subsequently, a structural comparison of the reported binding mode for FC131[216] with the present binding model for KRH-1636 (see Figure 41) shows that the L-$Arg^2$ of FC131 superimposes relatively well with the Arg ($R^2$-position) of KRH-1636 but the side-chains for 2-napthyl group (FC131) and 1-Nal (KRH-1636) have reverse coordinates. Similarly, comparisons of the KRH-1636 model with the binding mode for the $Arg^1$-$Arg^2$-1-

Nal[3] fragment of the 16-mer peptide antagonist CVX15 in complex with CXCR4[75] (model comparisons not shown) indicated the same pattern to the comparison with FC131; the 1-naphthyl group (KRH-1636) faces the TMH 2 as compared to the TMH 5 orientation of the naphthyl groups in FC131 (Figure 41), and only the central arginine (Arg[2]) of CVX15 appears to overlap well with the guanidine group of KRH-1636.

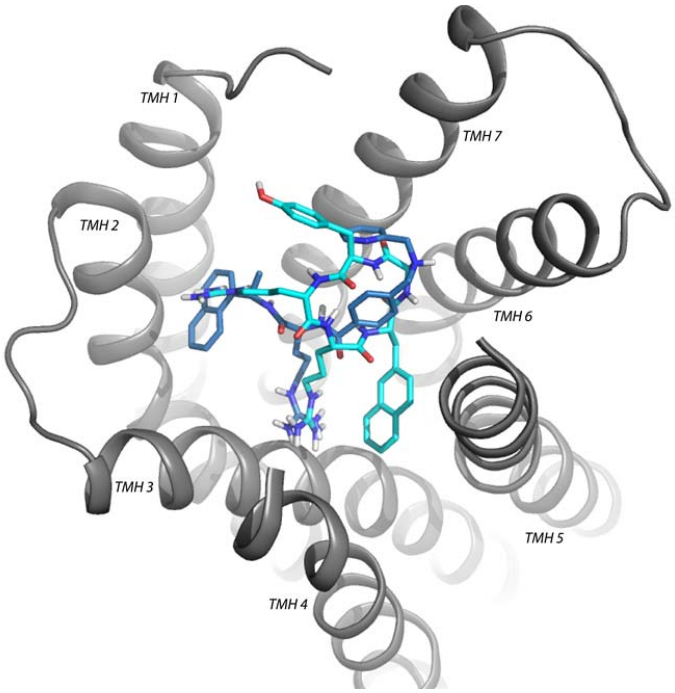



**Figure 41.** Superimposition of binding conformations of KRH-1636 and FC131 in the CXCR4 crystal structure; KRH-1636 in deep blue sticks and FC131 in light blue sticks; receptor in grey sticks and ribbons.

Moreover, our present binding model indicates that the $R^1$ side-chain of KRH-1636 shares common binding interactions with the previously reported binding mode of the potent CXCR4 antagonist and AMD-series analogue -3465. This monocyclam analogue (AMD3465, see Figure 16, compound **2**, sub-section 7.2.1) has the *N*-pyridinylmethylene group in common with KRH-1636, and was suggested in the same way as KRH-1636 to pick up an interaction with the TMH 7 residue His[281].[204, 205]

In conclusion, we contradict earlier suggestions[170] that KRH-1636 mimics the binding mode of the Arg[1]-Arg[2]-2-Nal[3] fragment of FC131, and we propose herein a binding mode for KRH-1636 where, the Arg-guanidino group ($R^2$ side-chain) overlaps with the binding modes of FC131 and CVX15, while the *N*-pyridinylmethylene moiety ($R^1$ side-chain) overlaps with the binding mode of the right-hand side of AMD-3465.

## 13. CONCLUSIONS AND FUTURE PERSPECTIVES

The following conclusions and future directions can be drawn from the current thesis:

- Our SAR data for the aromatic residues of the lead cyclopentapeptide and CXCR4 antagonist FC131 showed that the naphthyl group (2-Nal, position 3) is generally more important for activity than the phenol group in position 5 (D-Tyr$^5$). Molecular modeling indicated further, the lack of a defined binding pocket for D-Tyr$^5$ suggesting that this residue is dispensable for activity.

- Our collective SAR data suggested that tripeptidomimetics could be based on the Arg$^1$-Arg$^2$-2-Nal$^3$ fragment, and novel, scaffold-based tripeptidomimetics were investigated in this project.

- KRH-1636 has been previously assumed to exhibit a similar molecular recognition pattern as FC131. Subsequently, site-directed mutagenesis to map the ligand-receptor interactions in conjunction with molecular docking suggested that that the central arginine-guanidino group of KRH-1636 has overlapped receptor interactions with the Arg$^2$ of FC131; however, the naphthyl rings of the two ligands have reversed coordinates in the receptor binding site.

- From the existing and generated SAR data on FC131, peptidomimetic drug design could potentially be based on a minimal recognition motif, the dipeptide Arg$^2$-2-Nal$^3$ fragment, which could serve as a starting point for future design of dipeptidomimetics as CXCR4 antagonists.

- Biological testing of the prototype bicyclic tripeptidomimetics showed that they represent promising candidates for further structural tuning and optimization. The synthetic route applied to prepare the novel scaffold ligands, allows furthermore

synthetic access to a range of future target molecules. Further SAR studies are being pursued aiming to optimize the bicyclic-scaffold mimetics.

- Binding mode studies on KRH-1636 suggest that there are two alternative paths to future structure-based design of peptidomimetic CXCR4 antagonists based on FC131 and KRH-1636, respectively.

## 14. REFERENCES


1. Sharma, S., Singh, R. & Rana, S. Bioactive Peptides: A Review. *Int. J. BIO autom.* (15), 223-250.
2. Snyder, S.H. & Innis, R.B. Peptide neurotransmitters. *Annu. Rev. Biochem* **1979,** 48, 755-782.
3. Watanabe, T., Sato, K., Itoh, F., Wakabayashi, K., Shichiri, M. & Hirano, T. Endogenous bioactive peptides as potential biomarkers for atherosclerotic coronary heart disease. *Sensors* **2012,** 12, 4974-4985.
4. Labella, F.S., Geiger, J.D. & Glavin, G.B. Administered peptides inhibit the degradation of endogenous peptides. The dilemma of distinguishing direct from indirect effects. *Peptides* **1985,** 6, 645-660.
5. Thundimadathil, J. Cancer Treatment Using Peptides: Current Therapies and Future Prospects. *J. Amino Acids* **2012,** 2012.
6. Naider, F. & Anglister, J. Peptides in the treatment of AIDS. *Curr. Opin. Struct. Biol.* **2009,** 19, 473-482.
7. Von Recum, H.A. & Pokorski, J.K. Peptide and protein-based inhibitors of HIV-1 co-receptors. *Exp. Biol. Med.* **2013,** 238, 442-449.
8. Gordon, Y.J., Romanowski, E.G. & Mcdermott, A.M. A review of antimicrobial peptides and their therapeutic potential as anti-infective drugs. *Curr. Eye Res.* **2005,** 30, 505-515.
9. Flight, M.H. Drug Discovery: Structure-led design. *Nature* **2013,** 502, S50-S52.
10. Hruby, V.J. & Cai, M. Design of peptide and peptidomimetic ligands with novel pharmacological activity profiles. *Annu. Rev. Pharmacol. Toxicol.* **2013,** 53, 557-80.
11. Ganellin, C., Lindberg, P. & Mitscher, L. Glossary of terms used in medicinal chemistry. **1998**.
12. Purcell, R. & Lockey, R. Immunologic responses to therapeutic biologic agents. *J. Investig. Allergol. Clin. Immunol.* **2008,** 18, 335-342.
13. Bursavich, M.G. & Rich, D.H. Designing non-peptide peptidomimetics in the 21st century: inhibitors targeting conformational ensembles. *J. Med. Chem.* **2002,** 45, 541-558.
14. Oishi, S. & Fujii, N. Peptide and peptidomimetic ligands for CXC chemokine receptor 4 (CXCR4). *Org. Biomol. Chem.* **2012,** 10, 5720-31.
15. Oprea, T.I. Current trends in lead discovery: Are we looking for the appropriate properties? *J. Comput. Aided Mol. Des.* **2002,** 16, 325-334.
16. Lipinski, C. & Hopkins, A. Navigating chemical space for biology and medicine. *Nature* **2004,** 432, 855-861.
17. Van De Waterbeemd, H., Carter, R.E., Grassy, G., Kubinyi, H., Martin, Y.C., Tute, M.S. & Willett, P. Glossary of Terms Used in Computational Drug Design (IUPAC Recommendations 1997). *Annu. Rep. Med. Chem.* **1998,** 33, 397-409.
18. Chapman, T. Drug discovery: the leading edge. *Nature* **2004,** 430, 109-115.
19. Lipinski, C.A., Lombardo, F., Dominy, B.W. & Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **1997,** 23, 3-25.
20. Ajay, Walters, W.P. & Murcko, M.A. Can we learn to distinguish between "drug-like" and "nondrug-like" molecules? *J. Med. Chem.* **1998,** 41, 3314-3324.
21. Walters, W.P., Murcko, A.A. & Murcko, M.A. Recognizing molecules with drug-like properties. *Curr. Opin. Chem. Biol.* **1999,** 3, 384-387.

22. Lajiness, M.S., Vieth, M. & Erickson, J. Molecular properties that influence oral drug-like behavior. *Curr. Opin. Drug Discov. Devel.* **2004,** 7, 470-477.
23. Hruby, V.J. Designing peptide receptor agonists and antagonists. *Nat. Rev. Drug Discov.* **2002,** 1, 847-58.
24. Wermuth, C.-G., Ganellin, C.R., Lindberg, P. & Mitscher, L.A. Glossary of Terms Used in Medicinal Chemistry (IUPAC Recommendations 1997. *Annu. Rep. Med. Chem.* **1998,** 33, 385-395.
25. Acharya, C., Coop, A., Polli, J.E. & Mackerell Jr, A.D. Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. *Curr. Comput. Aided Drug Des.* **2011,** 7, 10.
26. Williams, R.J. Flexible drug molecules and dynamic receptors. *Angew. Chem. Int. Ed. Engl.* **1977,** 16, 766-77.
27. Humblet, C. & Marshall, G.R. Three-dimensional computer modeling as an aid to drug design. *Drug Dev. Res.* **1981,** 1, 409-434.
28. Burgen, A.S., Roberts, G.C. & Feeney, J. Binding of flexible ligands to macromolecules. *Nature* **1975,** 253, 753-5.
29. Hruby, V.J. & Balse, P.M. Conformational and topographical considerations in designing agonist peptidomimetics from peptide leads. *Curr. Med. Chem.* **2000,** 7, 945-70.
30. Hruby, V.J. Peptide science: exploring the use of chemical principles and interdisciplinary collaboration for understanding life processes. *J. Med. Chem.* **2003,** 46, 4215-4231.
31. Chen, L., K Morrow, J., T Tran, H., S Phatak, S., Du-Cuny, L. & Zhang, S. From laptop to benchtop to bedside: structure-based drug design on protein targets. *Curr. Pharm. Des.* **2012,** 18, 1217-1239.
32. Craik, D.J., Fairlie, D.P., Liras, S. & Price, D. The future of peptide-based drugs. *Chem. Biol. Drug Des.* **2013,** 81, 136-47.
33. Fiercepharma http://www.fiercepharma.com/special-reports/enbrel Accessed 16.09.2013.
34. Lipinski, C.A. Lead-and drug-like compounds: the rule-of-five revolution. *Drug Discov. Today Techno.* **2004,** 1, 337-341.
35. Mcgregor, D.P. Discovering and improving novel peptide therapeutics. *Curr. Opin. Pharmacol.* **2008,** 8, 616-619.
36. Hummel, G., Reineke, U. & Reimer, U. Translating peptides into small molecules. *Mol. Biosyst.* **2006,** 2, 499-508.
37. Loffet, A. Peptides as drugs: is there a market? *J. Pept. Sci.* **2002,** 8, 1-7.
38. Ladner, R.C., Sato, A.K., Gorzelany, J. & De Souza, M. Phage display-derived peptides as therapeutic alternatives to antibodies. *Drug Discov. Today* **2004,** 9, 525-529.
39. Latham, P.W. Therapeutic peptides revisited. *Nat. Biotechnol.* **1999,** 17, 755-758.
40. Sato, A.K., Viswanathan, M., Kent, R.B. & Wood, C.R. Therapeutic peptides: technological advances driving peptides into development. *Curr. Opin. Biotechnol.* **2006,** 17, 638-642.
41. Pichereau, C. & Allary, C. Therapeutic peptides under the spotlight. *Eur. Biopharm. Rev* **2005**, 88-91.
42. Bray, B.L. Large-scale manufacture of peptide therapeutics by chemical synthesis. *Nat. Rev. Drug Discov.* **2003,** 2, 587-93.
43. Merrifield, R.B. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **1963,** 85, 2149-2154.
44. Gupta, S. & Payne, J. Evaluation of the conformational propensities of peptide isosteres as a basis for selecting bioactive pseudopeptides. *J. Pept. Res.* **2001,** 58, 546-561.

45. Gillespie, P., Cicariello, J. & Olson, G.L. Conformational analysis of dipeptide mimetics. *Biopolymers* **1997,** 43, 191-217.
46. Giannis, A. & Kolter, T. Peptidomimetics for receptor ligands—discovery, development, and medical perspectives. *Angew. Chem. Int. Ed. Engl.* **1993,** 32, 1244-1267.
47. Hruby, V.J., Li, G., Haskell-Luevano, C. & Shenderovich, M. Design of peptides, proteins, and peptidomimetics in chi space. *Biopolymers* **1997,** 43, 219-66.
48. Du Vigneaud, V., Denning, G.S., Drabarek, S. & Chan, W. The Synthesis and Pharmacological Study of 4-Decarboxamido-oxytocin (4-α-Aminobutyric Acid-oxytocin) and 5-Decarboxamido-oxytocin (5-Alanine-oxytocin). *J. Biol. Chem.* **1964,** 239, 472-478.
49. Marshall, G.R. & Bosshard, H. Angiotensin II. Studies on the biologically active conformation. *Circ. Res.* **1972,** 31, Suppl 2: 143.
50. Toniolo, C. Conformationally restricted peptides through short-range cyclizations. *Int. J. Pept. Protein Res.* **1990,** 35, 287-300.
51. Spatola, A. Chemistry and Biochemistry of Amino Acids, Peptides, and Proteins Vol.7, 267–357 (ed. Weinstein, B.) (Marcel Dekker, New York, 1983)
52. Weide, T., Modlinger, A. & Kessler, H. Spatial screening for the identification of the bioactive conformation of integrin ligands (Springer, 2007).
53. Gibson, S.E., Guillo, N. & Tozer, M.J. Towards control of χ-space: conformationally constrained analogues of Phe, Tyr, Trp and His. *Tetrahedron* **1999,** 55, 585-615.
54. Cowell, S.M., Lee, Y.S., Cain, J.P. & Hruby, V.J. Exploring Ramachandran and chi space: conformationally constrained amino acids and peptides in the design of bioactive polypeptide ligands. *Curr. Med. Chem.* **2004,** 11, 2785-98.
55. Lloyd, D.G., Buenemann, C.L., Todorov, N.P., Manallack, D.T. & Dean, P.M. Scaffold hopping in de novo design. Ligand generation in the absence of receptor information. *J. Med. Chem.* **2004,** 47, 493-496.
56. Schneider, G. & Fechner, U. Computer-based de novo design of drug-like molecules. *Nat. Rev. Drug Discov.* **2005,** 4, 649-63.
57. Rester, U. Dock around the clock–current status of small molecule docking and scoring. *QSAR Comb. Sci.* **2006,** 25, 605-615.
58. Lewell, X.Q., Judd, D.B., Watson, S.P. & Hann, M.M. RECAP retrosynthetic combinatorial analysis procedure: A powerful new technique for identifying privileged molecular fragments with useful applications in combinatorial chemistry. *J. Chem. Inf. Comput. Sci.* **1998,** 38, 511-522.
59. Kato, Y., Itai, A. & Iitaka, Y. A novel method for superimposing molecules and receptor mapping. *Tetrahedron* **1987,** 43, 5229-5236.
60. Kato, Y., Inoue, A., Yamada, M., Tomioka, N. & Itai, A. Automatic superposition of drug molecules based on their common receptor site. *J. Comput. Aided Mol. Des.* **1992,** 6, 475-486.
61. Schneider, G., Neidhart, W., Giller, T. & Schmid, G. "Scaffold-Hopping" by Topological Pharmacophore Search: A Contribution to Virtual Screening. *Angew. Chem. Int. Ed. Engl.* **1999,** 38, 2894-2896.
62. Böhm, H.-J., Flohr, A. & Stahl, M. Scaffold hopping. *Drug Discov. Today Techno.* **2004,** 1, 217-224.
63. Bernstein, F.C., Koetzle, T.F., Williams, G.J., Meyer Jr, E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. & Tasumi, M. The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **1977,** 112, 535-542.

64. Wise, A., Gearing, K. & Rees, S. Target validation of G-protein coupled receptors. *Drug Discov. Today* **2002,** 7, 235-246.
65. Pierce, K.L., Premont, R.T. & Lefkowitz, R.J. Seven-transmembrane receptors. *Nat. Rev. Mol. Cell Biol.* **2002,** 3, 639-50.
66. Tyndall, J.D. & Sandilya, R. GPCR agonists and antagonists in the clinic. *Med. Chem.* **2005,** 1, 405-21.
67. Overington, J.P., Al-Lazikani, B. & Hopkins, A.L. How many drug targets are there? *Nat. Rev. Drug Discov.* **2006,** 5, 993-996.
68. Schiöth, H.B. & Fredriksson, R. The GRAFS classification system of G-protein coupled receptors in comparative perspective. *Gen. Comp. Endocrinol.* **2005,** 142, 94-101.
69. Archbold, J.K., Flanagan, J.U., Watkins, H.A., Gingell, J.J. & Hay, D.L. Structural insights into RAMP modification of secretin family G protein-coupled receptors: implications for drug development. *Trends Pharmacol. Sci.* **2011,** 32, 591-600.
70. Kniazeff, J., Prézeau, L., Rondard, P., Pin, J.-P. & Goudet, C. Dimers and beyond: The functional puzzles of class C GPCRs. *Pharmacol. Ther.* **2011,** 130, 9-25.
71. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., Trong, I.L., Teller, D.C., Okada, T. & Stenkamp, R.E. Crystal structure of rhodopsin: AG protein-coupled receptor. *Sci. Signal.* **2000,** 289, 739.
72. Huang, J., Chen, S., Zhang, J.J. & Huang, X.-Y. Crystal structure of oligomeric β1-adrenergic G protein–coupled receptors in ligand-free basal state. *Nat. Struct. Mol. Biol.* **2013,** 20, 419-425.
73. Cherezov, V., Rosenbaum, D.M., Hanson, M.A., Rasmussen, S.G., Thian, F.S., Kobilka, T.S., Choi, H.-J., Kuhn, P., Weis, W.I. & Kobilka, B.K. High-resolution crystal structure of an engineered human β2-adrenergic G protein–coupled receptor. *Science* **2007,** 318, 1258-1265.
74. Chien, E.Y., Liu, W., Zhao, Q., Katritch, V., Han, G.W., Hanson, M.A., Shi, L., Newman, A.H., Javitch, J.A. & Cherezov, V. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science* **2010,** 330, 1091-1095.
75. Wu, B., Chien, E.Y., Mol, C.D., Fenalti, G., Liu, W., Katritch, V., Abagyan, R., Brooun, A., Wells, P., Bi, F.C., Hamel, D.J., Kuhn, P., Handel, T.M., Cherezov, V. & Stevens, R.C. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **2010,** 330, 1066-71.
76. Granier, S., Manglik, A., Kruse, A.C., Kobilka, T.S., Thian, F.S., Weis, W.I. & Kobilka, B.K. Structure of the δ-opioid receptor bound to naltrindole. *Nature* **2012,** 485, 400-404.
77. White, J.F., Noinaj, N., Shibata, Y., Love, J., Kloss, B., Xu, F., Gvozdenovic-Jeremic, J., Shah, P., Shiloach, J. & Tate, C.G. Structure of the agonist-bound neurotensin receptor. *Nature* **2012**.
78. Seifert, R. & Wenzel-Seifert, K. Constitutive activity of G-protein-coupled receptors: cause of disease and common property of wild-type receptors. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2002,** 366, 381-416.
79. Kobilka, B.K. & Deupi, X. Conformational complexity of G-protein-coupled receptors. *Trends Pharmacol. Sci.* **2007,** 28, 397-406.
80. Ahuja, S. & Smith, S.O. Multiple switches in G protein-coupled receptor activation. *Trends Pharmacol. Sci.* **2009,** 30, 494-502.
81. Schertler, G.F. Signal transduction: the rhodopsin story continued. *Nature* **2008,** 453, 292-293.

82. Rosenbaum, D.M., Rasmussen, S.G. & Kobilka, B.K. The structure and function of G-protein-coupled receptors. *Nature* **2009,** 459, 356-363.
83. Park, J.H., Scheerer, P., Hofmann, K.P., Choe, H.-W. & Ernst, O.P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **2008,** 454, 183-187.
84. Scheerer, P., Park, J.H., Hildebrand, P.W., Kim, Y.J., Krauß, N., Choe, H.-W., Hofmann, K.P. & Ernst, O.P. Crystal structure of opsin in its G-protein-interacting conformation. *Nature* **2008,** 455, 497-502.
85. Surgand, J.S., Rodrigo, J., Kellenberger, E. & Rognan, D. A chemogenomic analysis of the transmembrane binding cavity of human G-protein-coupled receptors. *Proteins* **2006,** 62, 509-38.
86. Smit, M.J., Vischer, H.F., Bakker, R.A., Jongejan, A., Timmerman, H., Pardo, L. & Leurs, R. Pharmacogenomic and structural analysis of constitutive g protein-coupled receptor activity. *Annu. Rev. Pharmacol. Toxicol.* **2007,** 47, 53-87.
87. Schwartz, T.W., Frimurer, T.M., Holst, B., Rosenkilde, M.M. & Elling, C.E. Molecular mechanism of 7TM receptor activation-a global toggle switch model. *Annu. Rev. Pharmacol. Toxicol.* **2006,** 46, 481-519.
88. Kleinau, G., Claus, M., Jaeschke, H., Mueller, S., Neumann, S., Paschke, R. & Krause, G. Contacts between extracellular loop two and transmembrane helix six determine basal activity of the thyroid-stimulating hormone receptor. *J. Biol. Chem.* **2007,** 282, 518-525.
89. Cabrera-Vera, T.M., Vanhauwe, J., Thomas, T.O., Medkova, M., Preininger, A., Mazzoni, M.R. & Hamm, H.E. Insights into G protein structure, function, and regulation. *Endocr. Rev.* **2003,** 24, 765-781.
90. Thomsen, W., Frazer, J. & Unett, D. Functional assays for screening GPCR targets. *Curr. Opin. Biotechnol.* **2005,** 16, 655-665.
91. Hogaboam, C.M., Carpenter, K.J., Schuh, J.M., Proudfoot, A.A., Bridger, G. & Buckland, K.F. The therapeutic potential in targeting CCR5 and CXCR4 receptors in infectious and allergic pulmonary disease. *Pharmacol. Ther.* **2005,** 107, 314-28.
92. Horuk, R. Chemokine receptors. *Cytokine Growth Factor Rev.* **2001,** 12, 313-335.
93. Dammann, O. & Oshea, T. Chemokine/chemokine receptor nomenclature. *J. Interferon Cytokine Res.* **2002,** 22, 1067-1068.
94. Clore, G.M., Appella, E., Yamada, M., Matsushima, K. & Gronenborn, A.M. Three-dimensional structure of interleukin 8 in solution. *Biochemistry* **1990,** 29, 1689-1696.
95. Baldwin, E.T., Weber, I.T., St Charles, R., Xuan, J.C., Appella, E., Yamada, M., Matsushima, K., Edwards, B.F., Clore, G.M., Gronenborn, A.M. & Et Al. Crystal structure of interleukin 8: symbiosis of NMR and crystallography. *Proc. Natl. Acad. Sci. U. S. A.* **1991,** 88, 502-6.
96. Clark-Lewis, I., Schumacher, C., Baggiolini, M. & Moser, B. Structure-activity relationships of interleukin-8 determined using chemically synthesized analogs. Critical role of NH2-terminal residues and evidence for uncoupling of neutrophil chemotaxis, exocytosis, and receptor binding activities. *J. Biol. Chem.* **1991,** 266, 23128-23134.
97. Murphy, P.M. The molecular biology of leukocyte chemoattractant receptors. *Annu. Rev. Immunol.* **1994,** 12, 593-633.
98. Clark-Lewis, I., Dewald, B., Loetscher, M., Moser, B. & Baggiolini, M. Structural requirements for interleukin-8 function identified by design of analogs and CXC chemokine hybrids. *J. Biol. Chem.* **1994,** 269, 16075-16081.
99. Burrows, S.D., Doyle, M.L., Murphy, K.P., Franklin, S.G., White, J.R., Brooks, I., Mcnulty, D.E., Scott, M.O. & Knutson, J.R. Determination of the monomer-dimer

equilibrium of interleukin-8 reveals it is a monomer at physiological concentrations. *Biochemistry* **1994,** 33, 12741-12745.
100. Rajarathnam, K., Sykes, B.D., Kay, C.M., Dewald, B., Geiser, T., Baggiolini, M. & Clark-Lewis, I. Neutrophil activation by monomeric interleukin-8. *Science* **1994,** 264, 90-2.
101. Crump, M.P., Gong, J.H., Loetscher, P., Rajarathnam, K., Amara, A., Arenzana-Seisdedos, F., Virelizier, J.L., Baggiolini, M., Sykes, B.D. & Clark-Lewis, I. 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, 6996-7007.
102. Ottersbach, K., Mclean, J., Isaacs, N.W. & Graham, G.J. A310 helical turn is essential for the proliferation-inhibiting properties of macrophage inflammatory protein-1 alpha (CCL3). *Blood* **2006,** 107, 1284-1291.
103. Mantovani, A. The chemokine system: redundancy for robust outputs. *Immunol. Today* **1999,** 20, 254-257.
104. Onuffer, J.J. & Horuk, R. Chemokines, chemokine receptors and small-molecule antagonists: recent developments. *Trends Pharmacol. Sci.* **2002,** 23, 459-467.
105. Rajagopal, S., Kim, J., Ahn, S., Craig, S., Lam, C.M., Gerard, N.P., Gerard, C. & Lefkowitz, R.J. β-arrestin-but not G protein-mediated signaling by the "decoy" receptor CXCR7. *Proc. Natl. Acad. Sci.* **2010,** 107, 628-632.
106. O'hayre, M., Salanga, C., Handel, T. & Allen, S. Chemokines and cancer: migration, intracellular signalling and intercellular communication in the microenvironment. *Biochem. J* **2008,** 409, 635-649.
107. Mantovani, A., Bonecchi, R. & Locati, M. Tuning inflammation and immunity by chemokine sequestration: decoys and more. *Nat. Rev. Immunol.* **2006,** 6, 907-18.
108. Schwarz, M.K. & Wells, T.N.C. New therapeutics that modulate chemokine networks. *Nat. Rev. Drug Discov.* **2002,** 1, 347-358.
109. Terricabras, E., Benjamim, C. & Godessart, N. Drug discovery and chemokine receptor antagonists: eppur si muove! *Autoimmun Rev* **2004,** 3, 550-6.
110. Siciliano, S.J., Rollins, T.E., Demartino, J., Konteatis, Z., Malkowitz, L., Van Riper, G., Bondy, S., Rosen, H. & Springer, M.S. Two-site binding of C5a by its receptor: an alternative binding paradigm for G protein-coupled receptors. *Proc. Natl. Acad. Sci.* **1994,** 91, 1214-1218.
111. Wells, T., Power, C.A., Lusti-Narasimhan, M., Hoogewerf, A.J., Cooke, R.M., Chung, C., Peitsch, M. & Proudfoot, A. Selectivity and antagonism of chemokine receptors. *J. Leukoc. Biol.* **1996,** 59, 53-60.
112. Allen, S.J., Crown, S.E. & Handel, T.M. Chemokine: receptor structure, interactions, and antagonism. *Annu. Rev. Immunol.* **2007,** 25, 787-820.
113. Bleul, C.C., Farzan, M., Choe, H., Parolin, C., Clarklewis, I., Sodroski, J. & Springer, T.A. The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry. *Nature* **1996,** 382, 829-833.
114. Farrens, D.L., Altenbach, C., Yang, K., Hubbell, W.L. & Khorana, H.G. Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science* **1996,** 274, 768-770.
115. Veldkamp, C.T., Seibert, C., Peterson, F.C., De La Cruz, N.B., Haugner, J.C., 3rd, Basnet, H., Sakmar, T.P. & Volkman, B.F. Structural basis of CXCR4 sulfotyrosine recognition by the chemokine SDF-1/CXCL12. *Sci Signal* **2008,** 1, ra4.

116. Rosenkilde, M.M., Benned-Jensen, T., Frimurer, T.M. & Schwartz, T.W. The minor binding pocket: a major player in 7TM receptor activation. *Trends Pharmacol. Sci.* **2010,** 31, 567-574.
117. Berson, J.F., Long, D., Doranz, B.J., Rucker, J., Jirik, F.R. & Doms, R.W. A seven-transmembrane domain receptor involved in fusion and entry of T-cell-tropic human immunodeficiency virus type 1 strains. *J. Virol.* **1996,** 70, 6288-6295.
118. Chabot, D.J., Zhang, P.-F., Quinnan, G.V. & Broder, C.C. Mutagenesis of CXCR4 identifies important domains for human immunodeficiency virus type 1 X4 isolate envelope-mediated membrane fusion and virus entry and reveals cryptic coreceptor activity for R5 isolates. *J. Virol.* **1999,** 73, 6598-6609.
119. Buckley, C.D., Amft, N., Bradfield, P.F., Pilling, D., Ross, E., Arenzana-Seisdedos, F., Amara, A., Curnow, S.J., Lord, J.M. & Scheel-Toellner, D. Persistent induction of the chemokine receptor CXCR4 by TGF-β1 on synovial T cells contributes to their accumulation within the rheumatoid synovium. *J. Immunol.* **2000,** 165, 3423-3429.
120. Luster, A.D. Chemokines - Chemotactic cytokines that mediate inflammation. *N. Engl. J. Med.* **1998,** 338, 436-445.
121. Baggiolini, M. Chemokines and leukocyte traffic. *Nature* **1998,** 392, 565-568.
122. Zou, Y.-R., Kottmann, A.H., Kuroda, M., Taniuchi, I. & Littman, D.R. Function of the chemokine receptor CXCR4 in haematopoiesis and in cerebellar development. *Nature* **1998,** 393, 595-599.
123. Tarasova, N.I., Stauber, R.H. & Michejda, C.J. Spontaneous and ligand-induced trafficking of CXC-chemokine receptor 4. *J. Biol. Chem.* **1998,** 273, 15883-15886.
124. Nanki, T., Takada, K., Komano, Y., Morio, T., Kanegane, H., Nakajima, A., Lipsky, P.E. & Miyasaka, N. Chemokine receptor expression and functional effects of chemokines on B cells: implication in the pathogenesis of rheumatoid arthritis. *Arthritis Res. Ther.* **2009,** 11, R149.
125. Balkwill, F. Cancer and the chemokine network. *Nat. Rev. Cancer* **2004,** 4, 540-50.
126. Didigu, C.A. & Doms, R.W. Novel approaches to inhibit HIV entry. *Viruses* **2012,** 4, 309-324.
127. Proudfoot, A.E. Chemokine receptors: multifaceted therapeutic targets. *Nat. Rev. Immunol.* **2002,** 2, 106-15.
128. Bockaert, J. & Pin, J.P. Molecular tinkering of G protein-coupled receptors: an evolutionary success. *EMBO J.* **1999,** 18, 1723-9.
129. Verani, A. & Lusso, P. Chemokines as natural HIV antagonists. *Curr. Mol. Med.* **2002,** 2, 691-702.
130. Princen, K. & Schols, D. HIV chemokine receptor inhibitors as novel anti-HIV drugs. *Cytokine Growth Factor Rev.* **2005,** 16, 659-677.
131. Scholten, D., Canals, M., Maussang, D., Roumen, L., Smit, M., Wijtmans, M., De Graaf, C., Vischer, H. & Leurs, R. Pharmacological modulation of chemokine receptor function. *Br. J. Pharmacol.* **2012,** 165, 1617-1643.
132. Choi, W.-T., Duggineni, S., Xu, Y., Huang, Z. & An, J. Drug discovery research targeting the CXC chemokine receptor 4 (CXCR4). *J. Med. Chem.* **2011,** 55, 977-994.
133. Debnath, B., Xu, S., Grande, F., Garofalo, A. & Neamati, N. Small Molecule Inhibitors of CXCR4. *Theranostics* **2013,** 3, 47.
134. Tamamura, H., Omagari, A., Hiramatsu, K., Gotoh, K., Kanamoto, T., Xu, Y., Kodama, E., Matsuoka, M., Hattori, T. & Yamamoto, N. 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, 1897-1902.
135. Murakami, T., Nakajima, T., Koyanagi, Y., Tachibana, K., Fujii, N., Tamamura, H., Yoshida, N., Waki, M., Matsumoto, A. & Yoshie, O. A small molecule CXCR4 inhibitor that blocks T cell line–tropic HIV-1 infection. *J. Exp. Med.* **1997,** 186, 1389-1393.
136. Xu, Y., Tamamura, H., Arakaki, R., Nakashima, H., Zhang, X., Fujii, N., Uchiyama, T. & Hattori, T. Marked increase in anti-HIV activity, as well as inhibitory activity against HIV entry mediated by CXCR4, linked to enhancement of the binding ability of tachyplesin analogs to CXCR4. *AIDS Res. Hum. Retroviruses* **1999,** 15, 419-427.
137. Murakami, T., Zhang, T.-Y., Koyanagi, Y., Tanaka, Y., Kim, J., Suzuki, Y., Minoguchi, S., Tamamura, H., Waki, M. & Matsumoto, A. Inhibitory mechanism of the CXCR4 antagonist T22 against human immunodeficiency virus type 1 infection. *J. Virol.* **1999,** 73, 7489-7496.
138. Tamamura, H., Sugioka, M., Odagaki, Y., Omagari, A., Kan, Y., Oishi, S., Nakashima, H., Yamamoto, N., Peiper, S.C. & Hamanaka, N. 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, 359-362.
139. Tamamura, H., Hiramatsu, K., Kusano, S., Terakubo, S., Yamamoto, N., Trent, J.O., Wang, Z., Peiper, S.C., Nakashima, H. & Otaka, A. Synthesis of potent CXCR4 inhibitors possessing low cytotoxicity and improved biostability based on T140 derivatives. *Org. Biomol. Chem.* **2003,** 1, 3656-3662.
140. Tamamura, H., Hiramatsu, K., Mizumoto, M., Ueda, S., Kusano, S., Terakubo, S., Akamatsu, M., Yamamoto, N., Trent, J.O. & Wang, Z. Enhancement of the T140-based pharmacophores leads to the development of more potent and bio-stable CXCR4 antagonists. *Org. Biomol. Chem.* **2003,** 1, 3663-3669.
141. Tamamura, H., Hiramatsu, K., Kusano A, S., Terakubo, S., Yamamoto, N., Trent, J.O., Wang, Z., Peiper, S.C., Nakashima, H., Otaka, A. & Fujii, N. Synthesis of potent CXCR4 inhibitors possessing low cytotoxicity and improved biostability based on T140 derivatives. *Org. Biomol. Chem.* **2003,** 1, 3656-62.
142. Tamamura, H., Omagari, A., Oishi, S., Kanamoto, T., Yamamoto, N., Peiper, S.C., Nakashima, H., Otaka, A. & Fujii, N. 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, 2633-2637.
143. Fujii, N., Oishi, S., Hiramatsu, K., Araki, T., Ueda, S., Tamamura, H., Otaka, A., Kusano, S., Terakubo, S., Nakashima, H., Broach, J.A., Trent, J.O., Wang, Z.X. & Peiper, S.C. Molecular-size reduction of a potent CXCR4-chemokine antagonist using orthogonal combination of conformation- and sequence-based libraries. *Angew. Chem. Int. Ed. Engl.* **2003,** 42, 3251-3.
144. Mungalpara, J., Zachariassen, Z.G., Thiele, S., Rosenkilde, M.M. & Vabeno, J. Structure-activity relationship studies of the aromatic positions in cyclopentapeptide CXCR4 antagonists. *Org. Biomol. Chem.* **2013**.
145. Mungalpara, J., Thiele, S., Eriksen, Ø., Eksteen, J., Rosenkilde, M.M. & Våbenø, J. Rational Design of Conformationally Constrained Cyclopentapeptide Antagonists for CXC Chemokine Receptor 4 (CXCR4). *J. Med. Chem.* **2012,** 55, 10287-10291.
146. Kobayashi, K., Oishi, S., Hayashi, R., Tomita, K., Kubo, T., Tanahara, N., Ohno, H., Yoshikawa, Y., Furuya, T. & Hoshino, M. Structure–Activity Relationship Study of a

CXC Chemokine Receptor Type 4 Antagonist, FC131, Using a Series of Alkene Dipeptide Isosteres. *J. Med. Chem.* **2012,** 55, 2746-2757.

147. Tamamura, H., Hiramatsu, K., Ueda, S., Wang, Z.X., Kusano, S., Terakubo, S., Trent, J.O., Peiper, S.C., Yamamoto, N., Nakashima, H., Otaka, A. & Fujii, N. 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.* **2005,** 48, 380-391.
148. Tamamura, H., Araki, T., Ueda, S., Wang, Z., Oishi, S., Esaka, A., Trent, J.O., Nakashima, H., Yamamoto, N., Peiper, S.C., Otaka, A. & Fujii, N. Identification of novel low molecular weight CXCR4 antagonists by structural tuning of cyclic tetrapeptide scaffolds. *J. Med. Chem.* **2005,** 48, 3280-9.
149. Tanaka, T., Nomura, W., Narumi, T., Esaka, A., Oishi, S., Ohashi, N., Itotani, K., Evans, B.J., Wang, Z. & Peiper, S.C. 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.* **2009,** 7, 3805-3809.
150. Ueda, S., Oishi, S., Wang, Z.X., Araki, T., Tamamura, H., Cluzeau, J., Ohno, H., Kusano, S., Nakashima, H., Trent, J.O., Peiper, S.C. & Fujii, N. Structure-activity relationships of cyclic peptide-based chemokine receptor CXCR4 antagonists: disclosing the importance of side-chain and backbone functionalities. *J. Med. Chem.* **2007,** 50, 192-8.
151. Tamamura, H., Mizumoto, M., Hiramatsu, K., Kusano, S., Terakubo, S., Yamamoto, N., Trent, J.O., Wang, Z., Peiper, S.C., Nakashima, H., Otaka, A. & Fujii, N. Topochemical exploration of potent compounds using retro-enantiomer libraries of cyclic pentapeptides. *Org. Biomol. Chem.* **2004,** 2, 1255-7.
152. Tamamura, H., Esaka, A., Ogawa, T., Araki, T., Ueda, S., Wang, Z., Trent, J.O., Tsutsumi, H., Masuno, H. & Nakashima, H. Structure–activity relationship studies on CXCR4 antagonists having cyclic pentapeptide scaffolds. *Org. Biomol. Chem.* **2005,** 3, 4392-4394.
153. De Clercq, E., Yamamoto, N., Pauwels, R., Baba, M., Schols, D., Nakashima, H., Balzarini, J., Debyser, Z., Murrer, B.A., Schwartz, D. & Et Al. Potent and selective inhibition of human immunodeficiency virus (HIV)-1 and HIV-2 replication by a class of bicyclams interacting with a viral uncoating event. *Proc. Natl. Acad. Sci. U. S. A.* **1992,** 89, 5286-90.
154. Hendrix, C.W., Collier, A.C., Lederman, M.M., Schols, D., Pollard, R.B., Brown, S., Jackson, J.B., Coombs, R.W., Glesby, M.J., Flexner, C.W., Bridger, G.J., Badel, K., Macfarland, R.T., Henson, G.W., Calandra, G. & Group, A.H.S. Safety, pharmacokinetics, and antiviral activity of AMD3100, a selective CXCR4 receptor inhibitor, in HIV-1 infection. *J. Acquir. Immune Defic. Syndr.* **2004,** 37, 1253-62.
155. Hendrix, C.W., Flexner, C., Macfarland, R.T., Giandomenico, C., Fuchs, E.J., Redpath, E., Bridger, G. & Henson, G.W. Pharmacokinetics and safety of AMD-3100, a novel antagonist of the CXCR-4 chemokine receptor, in human volunteers. *Antimicrob. Agents Chemother.* **2000,** 44, 1667-1673.
156. Flomenberg, N., Devine, S.M., Dipersio, J.F., Liesveld, J.L., Mccarty, J.M., Rowley, S.D., Vesole, D.H., Badel, K. & Calandra, G. The use of AMD3100 plus G-CSF for autologous hematopoietic progenitor cell mobilization is superior to G-CSF alone. *Blood* **2005,** 106, 1867-1874.

157. Bridger, G.J., Boehringer, E.M., Wang, Z., Schols, D., Skerlj, R.T. & Bogucki, D.E. Preparation of antiviral macrocyclic polyamines. WO2000002870A1, **2000**.
158. Zhu, A., Zhan, W., Liang, Z., Yoon, Y., Yang, H., Grossniklaus, H.E., Xu, J., Rojas, M., Lockwood, M. & Snyder, J.P. Dipyrimidine amines: a novel class of chemokine receptor type 4 antagonists with high specificity. *J. Med. Chem.* **2010,** 53, 8556-8568.
159. Yu, L., Cecil, J., Peng, S.-B., Schrementi, J., Kovacevic, S., Paul, D., Su, E.W. & Wang, J. Identification and expression of novel isoforms of human stromal cell-derived factor 1. *Gene* **2006,** 374, 174-179.
160. Fujii, N. & Tamamura, H. Peptide-lead CXCR4 antagonists with high anti-HIV activity. *Curr. Opin. Investig. Drugs* **2001,** 2, 1198-202.
161. Bridger, G., Skerlj, R., Kaller, A., Harwig, C., Bogucki, D., Wilson, T.R., Crawford, J., Mceachern, E.J., Atsma, B., Nan, S., Zhou, Y., Schols, D., Smith, C.D. & Di, F.R.M. Preparation of azolylmethylaminotetrahydroquinolines and related compounds as chemokine receptor binding agents. WO2002022600A2, **2002**.
162. Skerlj, R.T., Bridger, G.J., Kaller, A., Mceachern, E.J., Crawford, J.B., Zhou, Y., Atsma, B., Langille, J., Nan, S. & Veale, D. Discovery of Novel Small Molecule Orally Bioavailable C− X− C Chemokine Receptor 4 Antagonists That Are Potent Inhibitors of T-Tropic (X4) HIV-1 Replication. *J. Med. Chem.* **2010,** 53, 3376-3388.
163. Moyle, G., Dejesus, E., Boffito, M., Wong, R.S., Gibney, C., Badel, K., Macfarland, R., Calandra, G., Bridger, G. & Becker, S. Proof of activity with AMD11070, an orally bioavailable inhibitor of CXCR4-tropic HIV type 1. *Clin. Infect. Dis.* **2009,** 48, 798-805.
164. Skerlj, R., Bridger, G., Mceachern, E., Harwig, C., Smith, C., Wilson, T., Veale, D., Yee, H., Crawford, J. & Skupinska, K. Synthesis and SAR of novel CXCR4 antagonists that are potent inhibitors of T tropic (X4) HIV-1 replication. *Bioorg. Med. Chem. Lett.* **2011,** 21, 262-266.
165. Skerlj, R., Bridger, G., Mceachern, E., Harwig, C., Smith, C., Kaller, A., Veale, D., Yee, H., Skupinska, K. & Wauthy, R. Design of novel CXCR4 antagonists that are potent inhibitors of T-tropic (X4) HIV-1 replication. *Bioorg. Med. Chem. Lett.* **2011,** 21, 1414-1418.
166. Tamamura, H., Tsutsumi, H., Masuno, H., Mizokami, S., Hiramatsu, K., Wang, Z., Trent, J.O., Nakashima, H., Yamamoto, N., Peiper, S.C. & Fujii, N. Development of a linear type of low molecular weight CXCR4 antagonists based on T140 analogs. *Org. Biomol. Chem.* **2006,** 4, 2354-7.
167. Ueda, S., Kato, M., Inuki, S., Ohno, H., Evans, B., Wang, Z.X., Peiper, S.C., Izumi, K., Kodama, E., Matsuoka, M., Nagasawa, H., Oishi, S. & Fujii, N. Identification of novel non-peptide CXCR4 antagonists by ligand-based design approach. *Bioorg. Med. Chem. Lett.* **2008,** 18, 4124-9.
168. Niida, A., Tanigaki, H., Inokuchi, E., Sasaki, Y., Oishi, S., Ohno, H., Tamamura, H., Wang, Z., Peiper, S.C. & Kitaura, K. Stereoselective Synthesis of 3, 6-Disubstituted-3, 6-dihydropyridin-2-ones as Potential Diketopiperazine Mimetics Using Organocopper-Mediated a nti-SN2'Reactions and Their Use in the Preparation of Low-Molecule CXCR4 Antagonists. *J. Org. Chem.* **2006,** 71, 3942-3951.
169. Ichiyama, K., Yokoyama-Kumakura, S., Tanaka, Y., Tanaka, R., Hirose, K., Bannai, K., Edamatsu, T., Yanaka, M., Niitani, Y., Miyano-Kurosaki, N., Takaku, H., Koyanagi, Y. & Yamamoto, N. 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, 4185-90.

170. Våbeno, J., Nikiforovich, G.V. & Marshall, G.R. A minimalistic 3D pharmacophore model for cyclopentapeptide CXCR4 antagonists. *Biopolymers* **2006,** 84, 459-71.
171. Yamazaki, T., Saitou, A., Ono, M., Yokoyama, S., Bannai, K., Hirose, K. & Yanaka, M. Preparation of amino acid amide derivatives as antagonists of chemokine CXCR4 receptor. WO2003029218A1, **2003**.
172. Ripka, A.S. & Rich, D.H. Peptidomimetic design. *Curr. Opin. Chem. Biol.* **1998,** 2, 441-452.
173. Wiley, R.A. & Rich, D.H. Peptidomimetics derived from natural products. *Med. Res. Rev.* **1993,** 13, 327-384.
174. Tanaka, T., Tsutsumi, H., Nomura, W., Tanabe, Y., Ohashi, N., Esaka, A., Ochiai, C., Sato, J., Itotani, K. & Murakami, T. Structure-activity relationship study of CXCR4 antagonists bearing the cyclic pentapeptide scaffold: identification of the new pharmacophore. *Org. Biomol. Chem.* **2008,** 6, 4374-4377.
175. Mosberg, H.I. & Kroona, H.B. Incorporation of a novel conformationally restricted tyrosine analog into a cyclic,. delta. opioid receptor selective tetrapeptide (JOM-13) enhances. delta. receptor binding affinity and selectivity. *J. Med. Chem.* **1992,** 35, 4498-4500.
176. Casimir, J., Iterbeke, K., Van Den Nest, W., Tourwe, D., Trescol-Biémont, M.C., Rabourdin-Combe, C., Dumortier, H., Muller, S., Gerlier, D. & Paris, J. Conformational restriction of the Tyr53 side-chain in the decapeptide HEL [52-61]: effects on binding to MHC-II I-Ak molecule and TCR recognition. *J. Pept. Res.* **2000,** 56, 398-408.
177. Ruzza, P., Cesaro, L., Tourwé, D., Calderan, A., Biondi, B., Maes, V., Menegazzo, I., Osler, A., Rubini, C. & Guiotto, A. Spatial conformation and topography of the tyrosine aromatic ring in substrate recognition by protein tyrosine kinases. *J. Med. Chem.* **2006,** 49, 1916-1924.
178. Hruby, V.J., Al-Obeidi, F. & Kazmierski, W. Emerging approaches in the molecular design of receptor-selective peptide ligands: conformational, topographical and dynamic considerations. *Biochem. J.* **1990,** 268, 249-62.
179. Chan, W.C. & White, P.D. Fmoc solid phase peptide synthesis (Oxford University Press, 2000).
180. Heydorn, A., Ward, R.J., Jorgensen, R., Rosenkilde, M.M., Frimurer, T.M., Milligan, G. & Kostenis, E. Identification of a novel site within G protein α subunits important for specificity of receptor-G protein interaction. *Mol. Pharmacol.* **2004,** 66, 250-259.
181. Toniolo, C., Bonora, G.M., Bavoso, A., Benedetti, E., Di Blasio, B., Pavone, V. & Pedone, C. Preferred conformations of peptides containing α, α-disubstituted α-amino acids. *Biopolymers* **1983,** 22, 205-215.
182. Weiss, M.S., Jabs, A. & Hilgenfeld, R. Peptide bonds revisited. *Nat. Struct. Mol. Biol.* **1998,** 5, 676-676.
183. Schrödinger http://www.schrodinger.com/.
184. Schrödinger Suite 2012 Protein Preparation Wizard; Epik Version 2.2, S., Llc, New York, Ny, 2012; Impact Version 5.7, Schrödinger, Llc, New York, Ny, 2012; Prime Version 2.3, Schrödinger, Llc, New York, Ny, 2012.
185. Schrödinger Suite 2012 Induced Fit Docking Protocol; Glide Version 5.8, S., Llc, New York, Ny, 2012; Prime Version 3.1, Schrödinger, Llc, New York, Ny, 2012.
186. White, C.J. & Yudin, A.K. Contemporary strategies for peptide macrocyclization. *Nat Chem* **2011,** 3, 509-24.

187. Ruben, A.J., Kiso, Y. & Freire, E. Overcoming roadblocks in lead optimization: a thermodynamic perspective. *Chem. Biol. Drug Des.* **2006,** 67, 2-4.
188. Fairlie, D.P., Abbenante, G. & March, D.R. Macrocyclic Peptidomimetics - Forcing Peptides into Bioactive Conformations. *Curr. Med. Chem.* **1995,** 2, 654-686.
189. Narumi, T., Tanaka, T., Hashimoto, C., Nomura, W., Aikawa, H., Sohma, A., Itotani, K., Kawamata, M., Murakami, T. & Yamamoto, N. Pharmacophore-based small molecule CXCR4 ligands. *Bioorg. Med. Chem. Lett.* **2012,** 22, 4169-4172.
190. Stradley, S.J., Rizo, J., Bruch, M.D., Stroup, A.N. & Gierasch, L.M. 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 β-turns. *Biopolymers* **1990,** 29, 263-287.
191. Ruiz-Gomez, G., Tyndall, J.D., Pfeiffer, B., Abbenante, G. & Fairlie, D.P. Update 1 of: Over one hundred peptide-activated G protein-coupled receptors recognize ligands with turn structure. *Chem. Rev.* **2010,** 110, PR1-PR41.
192. Vojkovský, T., Weichsel, A. & Pátek, M. Solid-phase synthesis of heterocycles containing an 1-acyl-3-oxopiperazine skeleton. *J. Org. Chem.* **1998,** 63, 3162-3163.
193. Kohn, W.D. & Zhang, L. Solid-phase synthesis of peptide–heterocycle hybrids containing a tripeptide-derived 6, 6-fused bicyclic subunit. *Tetrahedron Lett.* **2001,** 42, 4453-4457.
194. Grimes Jr, J.H., Zheng, W. & Kohn, W.D. Diastereoselectivity in the solid-phase synthesis of peptide heterocycle hybrids. *Tetrahedron Lett.* **2004,** 45, 6333-6336.
195. Allen, M.J., Raines, R.T. & Kiessling, L.L. Contrast agents for magnetic resonance imaging synthesized with ring-opening metathesis polymerization. *J. Am. Chem. Soc.* **2006,** 128, 6534-6535.
196. Balakrishnan, S., Zhao, C. & Zondlo, N.J. Convergent and Stereospecific Synthesis of Molecules Containing α-Functionalized Guanidiniums via α-Guanidino Acids. *J. Org. Chem.* **2007,** 72, 9834-9837.
197. Rasmussen, P., Synthetic studies towards peptidomimetic CXCR4 antagonists, *Master thesis,* University of Bergen.
198. Berg, E. A., Synthetic studies towards bicyclic antagonists, *Master thesis,* University of Bergen.
199. Yanaka, M., Yamazaki, T., Bannai, K. & Hirose, K. CXCR4-antagonistic drugs composed of nitrogen-containing compound. US 0157818, August 12, **2004**.
200. 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.* **2012,** 22, 2146-2150.
201. Demmer, O., Dijkgraaf, I., Schumacher, U., Marinelli, L., Cosconati, S., Gourni, E., Wester, H.J. & Kessler, H. Design, synthesis, and functionalization of dimeric peptides targeting chemokine receptor CXCR4. *J. Med. Chem.* **2011,** 54, 7648-62.
202. Gerlach, L.O., Skerlj, R.T., Bridger, G.J. & Schwartz, T.W. Molecular interactions of cyclam and bicyclam non-peptide antagonists with the CXCR4 chemokine receptor. *J. Biol. Chem.* **2001,** 276, 14153-14160.
203. Rosenkilde, M.M., Gerlach, L.O., Jakobsen, J.S., Skerlj, R.T., Bridger, G.J. & Schwartz, T.W. Molecular mechanism of AMD3100 antagonism in the CXCR4 receptor. *J. Biol. Chem.* **2004,** 279, 3033-3041.
204. Rosenkilde, M.M., Gerlach, L.O., Hatse, S., Skerlj, R.T., Schols, D., Bridger, G.J. & Schwartz, T.W. Molecular mechanism of action of monocyclam versus bicyclam non-

peptide antagonists in the CXCR4 chemokine receptor. *J. Biol. Chem.* **2007,** 282, 27354-27365.
205. Wong, R.S., Bodart, V., Metz, M., Labrecque, J., Bridger, G. & Fricker, S.P. Comparison of the potential multiple binding modes of bicyclam, monocylam, and noncyclam small-molecule CXC chemokine receptor 4 inhibitors. *Mol. Pharmacol.* **2008,** 74, 1485-1495.
206. Brelot, A., Heveker, N., Montes, M. & Alizon, M. Identification of residues of CXCR4 critical for human immunodeficiency virus coreceptor and chemokine receptor activities. *J. Biol. Chem.* **2000,** 275, 23736-23744.
207. Truax, V.M., Zhao, H.Y., Katzman, B.M., Prosser, A.R., Alcaraz, A.A., Saindane, M.T., Howard, R.B., Culver, D., Arrendale, R.F., Gruddanti, P.R., Evers, T.J., Natchus, M.G., Snyder, J.P., Liotta, D.C. & Wilson, L.J. Discovery of Tetrahydroisoquinoline-Based CXCR4 Antagonists. *ACS Med. Chem. Lett.* **2013,** 4, 1025-1030.
208. Generic numbering system proposed by Baldwin and modified by Schwartz (Baldwin, E. T.; Weber, I. T.; St Charles, R.; Xuan, J. C.; Appella, E.; Yamada, M.; Matsushima, K.; Edwards, B. F.; Clore, G. M.; Gronenborn, A. M.; et al. Crystal structure of interleukin 8: symbiosis of NMR and crystallography. *Proc Natl Acad Sci U S A* **1991**, 88, 502-6; Schwartz, T. W. Locating ligand-binding sites in 7TM receptors by protein engineering. *Curr Opin Biotechnol* **1994**, 5, 434-444.).
209. Ballesteros, J.A. & Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci.* **1995,** 25, 366-428.
210. Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., Repasky, M.P., Knoll, E.H., Shelley, M. & Perry, J.K. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **2004,** 47, 1739-1749.
211. Sherman, W., Day, T., Jacobson, M.P., Friesner, R.A. & Farid, R. Novel procedure for modeling ligand/receptor induced fit effects. *J. Med. Chem.* **2006,** 49, 534-553.
212. Friesner, R.A., Murphy, R.B., Repasky, M.P., Frye, L.L., Greenwood, J.R., Halgren, T.A., Sanschagrin, P.C. & Mainz, D.T. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J. Med. Chem.* **2006,** 49, 6177-6196.
213. Kawatkar, S.P., Yan, M., Gevariya, H., Lim, M.Y., Eisold, S., Zhu, X., Huang, Z. & An, J. Computational analysis of the structural mechanism of inhibition of chemokine receptor CXCR4 by small molecule antagonists. *Exp. Biol. Med.* **2011,** 236, 844-850.
214. Wong, R.S.Y., Bodart, V., Metz, M., Labrecque, J., Bridger, G. & Fricker, S.P. Comparison of the potential multiple binding modes of bicyclam, monocylam, and noncyclam small-molecule CXC chemokine receptor 4 inhibitors. *Mol. Pharmacol.* **2008,** 74, 1485-1495.
215. Planesas, J.M., Pérez-Nueno, V.I., Borrell, J.I. & Teixidó, J. Impact of the CXCR4 structure on docking-based virtual screening of HIV entry inhibitors. *J. Mol. Graphics Modell.* **2012**.
216. Thiele, S.; Mungalpara, J.; Steen, A.; Rosenkilde, M. M.; Våbenø, J. Determination of the binding mode for the cyclopentapeptide CXCR4 antagonist FC131 using a dual approach of ligand modifications and receptor mutagenesis. *Submitted*.
217. Delano, W.L. The PyMOL molecular graphics system. 2002.

218. Horuk, R. Chemokine receptors and HIV-1: the fusion of two major research fields. *Immunol. Today* **1999**, 20, 89-94.
219. Luster, A.D. Chemokines - Chemotactic cytokines that mediate inflammation. *N. Engl. J. Med.* **1998**, 338, 436-445.

# PAPER I

# PAPER II

# PAPER III