

THESIS FOR THE DEGREE MASTER OF PHARMACY

INVESTIGATIONS OF BINDING MODE FOR THE
CYCLOPENTAPEPTIDE CXCR4 ANTAGONIST FC131
BY INDUCED FIT DOCKING

By

Natnael Abera



Section of Medicinal Chemistry

Department of Pharmacy

University of Tromsø

May 2009

Acknowledgements

This work was done at the department of Pharmacy, Section of Medicinal chemistry, University of Tromsø, Norway from October 2008 to May 2009. I want to express my thanks to my supervisor, Jon Våbenø for his indescribable help and guidance throughout this work. His support, inspiration and motivation right from the beginning of this work have been tremendous; many thanks to Jon Våbenø.

I want to thank the members of Rehoboth Baptist Church for their prayer and support throughout my studies.

A special thank to Dad and Mom for their love and constant support, thank you so much for helping me get through this.

I also want to thank Kiya for her understanding and unwavering support throughout my study.

I finally want to thank God, my Lord; you have been my home and great provider.

Natnael Abera

Tromsø, May 2009

Abstract

The GPCR CXCR4 is a chemokine receptor that by activation of its natural ligand SDF-1 α is involved in the pathology of several diseases like cancer metastasis, leukemia cell progression and rheumatoid arthritis. The finding that CXCR4 plays a critical role for HIV-1 entry into T cells prompts additional motivation for the design of CXCR4 inhibitor. The establishment of the possible binding mode(s) for the cyclopentapeptide FC131 is decisive for the development of such inhibitor. Induced fit docking, which allowed flexibility for both the ligand and receptor structure, was used to generate ligand-receptor complexes. The resulting poses were compared based on their XP score and two ligand binding modes were suggested. In addition to this, mutational analysis on three CXCR4 residues which are believed to be important for HIV infection of T cells was performed.

Table of Contents

1	INTRODUCTION.....	9
1.1	THE CXCR4 RECEPTOR	9
1.2	CXCR4 ANTAGONISTS FROM THE LITERATURE.....	10
1.3	DESIGN OF PEPTIDOMIMETIC CXCR4 ANTAGONISTS	12
1.4	COMPUTER AIDED DRUG DESIGN.....	12
1.4.1	<i>Homology modeling</i>	12
1.4.2	<i>Docking Methods</i>	13
1.5	OBJECTIVE OF THE THESIS	14
2	METHODS.....	15
2.1	GENERAL	15
2.2	HOMOLOGY MODELING OF THE CXCR4 TRANSMEMBRANE HELIX (TMH) BUNDLE.....	15
2.2.1	<i>Query (CXCR4) sequence</i>	15
2.2.2	<i>Template structure</i>	15
2.2.3	<i>Sequence alignment</i>	15
2.2.4	<i>Building the CXCR4 structure model</i>	16
2.2.5	<i>Refining the modeled structure</i>	16
2.2.6	<i>The final receptor model</i>	17
2.3	INDUCED FIT DOCKING	17
2.3.1	<i>Preparing the ligand structure</i>	17
2.3.2	<i>Preparing the receptor model(s)</i>	18
2.3.3	<i>Induced Fit Docking protocol</i>	18
3	RESULTS AND DISCUSSION	21
3.1	HOMOLOGY MODEL	21
3.1.1	<i>Choice of template</i>	21
3.1.2	<i>Alignment</i>	21
3.1.3	<i>The final receptor model</i>	22
3.2	DOCKING RESULTS.....	24
3.3	POSSIBLE BINDING MODE(S) OF FC131	33
4	CONCLUSIONS.....	35
5	REFERENCES.....	37

1 Introduction

1.1 The CXCR4 receptor

Chemokine receptors are integral membrane proteins which serve as specific binding sites for chemokines, which constitute a large family of chemotactic cytokines. The binding of chemokines to their specific chemokine receptors mediates diverse biological processes like angiogenesis, hematopoiesis, organogenesis and leukocyte trafficking under homeostatic and inflammatory conditions.¹

The chemokine receptor CXCR4 (receptor code 2.1:CHK:4:CXCR4)² is an alpha-helical 7 transmembrane (7TM) rhodopsin-like receptor which belongs to the large superfamily of receptor proteins called G protein-coupled receptors (GPCRs).

The natural ligand for CXCR4 is stromal cell-derived factor (SDF-1 α), also known as chemokine ligand 12 (CXCL12), is a small cytokine which belongs to the CXC chemokines.³ CXCL12 is a strong chemotactic for lymphocytes and also controls many other important biological processes like stem cell movement, development of neurons, angiogenesis and activation of leukocytes.⁴

Beside the CD4 primary cellular receptor, CXCR4 was after a decade long search identified as the major coreceptor for human immunodeficiency virus type 1 (HIV-1) fusion and entry into T-cells.⁵

Furthermore the CXCR4 receptor has been shown to be involved in many other diseases like cancer metastasis, leukemia cell progression and rheumatoid arthritis.⁶

The involvement of CXCR4 in several different diseases has made it a potential therapeutic target for cancer treatment and inhibition of HIV-1 entry.

1.2 CXCR4 antagonists from the literature

Several types of CXCR4 antagonists have over the past decade been shown to exhibit potent and selective anti-HIV activity. The CXCR4 antagonists AMD3100, KRH-1636, T140 and FC131 are some of the agents in this class. The basic center in these molecules is one of the common features they have in common.

The bicyclam AMD3100 is a symmetric compound consisting of two monocyclam rings connected by an aromatic 1,4-phenylenebis(methylene)-linker (Fig.1.1). AMD3100 demonstrates anti-HIV effect by inhibiting the CXCR4 co-receptor and blocking HIV-1 membrane fusion and entry to the host cell.⁷

Mutational studies has revealed that Asp171, Asp262 and Glu288 are key interaction points for AMD3100⁸. It suggested that Asp171 interacts with one of the bicyclam rings and the other ring is sandwiched between Asp262 and Glu288.⁹

AMD070 is another small molecule which is believed to exhibit an anti-HIV activity through CXCR4 inhibition (Fig. 1.1). AMD070 is a derivate of AMD3100 and shows oral bioavailability.¹⁰

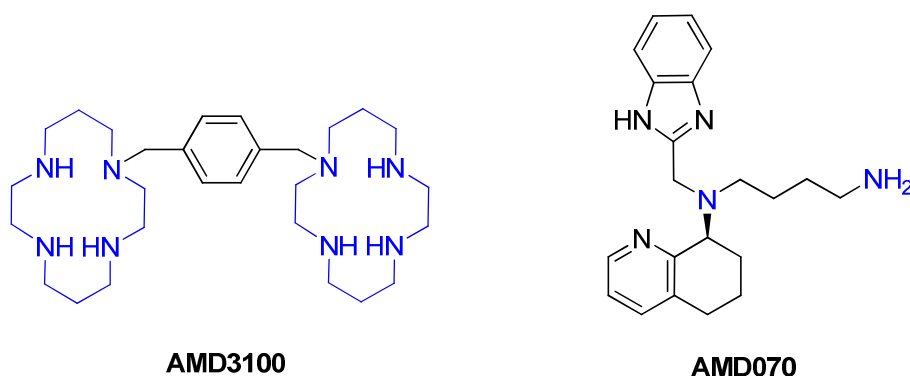


Fig. 1.1: Structure of AMD3100 and AMD070. The basic center of the structures is marked with blue color.

KRH-1636 (Kureha Chemical Industries) (Fig.1.2) is a relatively new low-molecular weight nonpeptide compound having a potent anti-HIV activity and reasonable bioavailability both *in vivo* and *in vitro*. KRH-1636 selectively hinders HIV-1 infection by inhibiting viral entry and membrane fusion to the host cell through CXCR4 coreceptor.¹¹

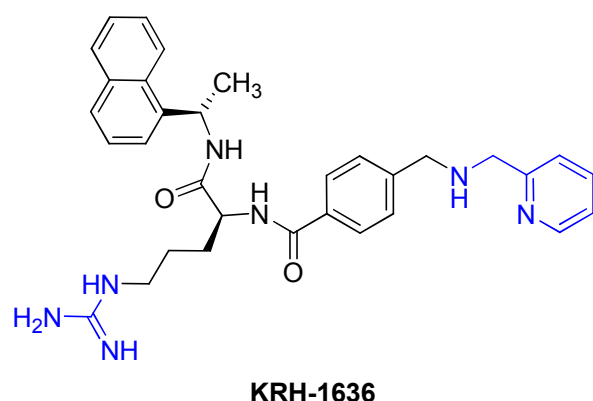


Fig. 1.2: Structure of KRH-1636. Part of the structure which is colored with blue is the basic center of the molecule.

The cyclopentapeptide (CPP) FC131 [c(Gly¹-D-Tyr²-Arg³-Arg⁴-Nal⁵); Nal is 2-naphthylalanine] (Fig.1.3), developed by molecular size reduction of another CXCR4 antagonist, the 14-residue peptide T140 [Arg¹-Arg²-Nal³-c(Cys⁴-Tyr⁵-Arg⁶-Lys⁷-D-Lys⁸-Pro⁹- Tyr¹⁰-Arg¹¹-Cit¹²-Cys¹³)-Arg¹⁴] (Fig.1.3) has also shown to be a potent CXCR4 antagonist. The size reduction (from T140 to FC131) was done based on the identification of the four bioactive amino acid residues Arg², Nal³, Tyr⁵, and Arg¹⁴ of the T140 molecule.¹²

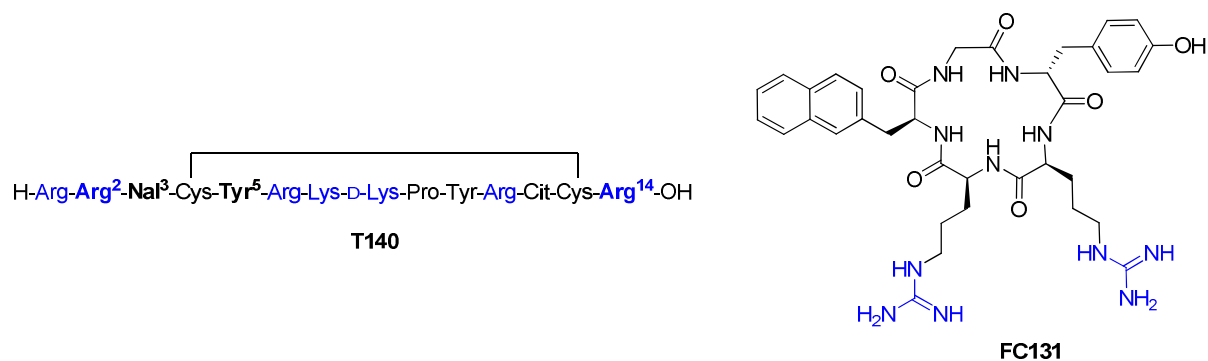


Fig. 1.3: Structures of T140 and the downsized peptide FC131 where Nal = L-3-(2 naphthyl) alanine and Cit = L-citrulline. The basic center of the molecules is colored with blue.

Side chain substitution at the different positions of FC131 resulted in analogs with varying affinity to the CXCR4 receptor. According to SAR studies, FC131 still remains among the molecules with the highest affinity to the CXCR4 receptor.¹³

1.3 Design of peptidomimetic CXCR4 antagonists

Limited stability and oral bioavailability of peptides is a major drawback for the development into drugs.¹⁴

The peptide mimetic design principles^{15,16,17} converts or modifies peptide antagonists into non-peptide or peptidomimetic compounds. A peptidomimetic molecule mimics the biological activity of the peptide, in addition to providing metabolic stability and oral bioavailability.^{18,19}

The small molecule CPP FC131, with a well defined 3D pharmacophore and established binding mode represents a good starting point for the development of peptidomimetic compounds.²⁰⁻²¹

1.4 Computer aided drug design

The ideal situation for computer-aided drug design is that the detailed 3D molecular structure of the drug target and preferably also the 3D structure of the ligand-receptor complex is known.

Though sequencing of the human genome has created the possibility of identifying many unknown proteins, the 3D structure of most membrane proteins, including CXCR4 is still unknown.²²

However, computer based homology modeling and docking programs have provided us the opportunity to predict the 3D structure of such proteins and also perform ligand docking, as described below.

1.4.1 Homology modeling

Homology modeling is a method of constructing and predicting an atomic resolution model of the target protein from its amino acid sequence based on an experimentally determined 3D structure of a related homologous protein called the template protein.

The method of homology modeling of membrane proteins is a relatively new and immature because of the few number of experimentally resolved membrane protein structures.²³

The four basic steps in homology modeling are: (1) identifying the template structure sequence, (2) aligning the query sequence with the template structure sequence, (3) building the model structure of the query based on the information from the template structure and (4) evaluating the predicted model.²³

Homology modeling is therefore a useful methodology in predicting undetermined protein structures like the CXCR4 receptor.²⁴

1.4.2 Docking Methods

As the number of proteins with a known three-dimensional structure is increasing, the need for computational docking techniques, which involves the prediction of a ligand's conformation and orientation in the target's binding site, has grown rapidly particularly in the biotechnology and pharmaceutical industries.^{25,26}

Rigid receptor docking, which uses fixed receptor sites derived from high-resolution crystal structures, is one of the most widely used methods; however, this method often fails because many receptor-ligand interactions cause ligand-induced conformational changes. Rigid receptor docking is therefore not a suitable method where ligand-induced receptor/enzyme conformational change is relevant.²⁷

Because of the high demand of flexibility, most docking methods assume the protein to be rigid while many other docking programs consider the ligand to be flexible.^{28,24}

Induced fit docking (IFD) delivered by Schrodinger integrates two powerful programs called Prime (protein structure prediction) and Glide (rigid receptor docking), and is one of the most advanced docking programs which is especially useful when conformational change on both the receptor and ligand structure is induced during ligand receptor binding.²⁹

1.5 Objective of the thesis

The main goal of this project is to identify a plausible binding mode for the cyclopentapeptide CXCR4 antagonist FC131.

This will be done by induced fit docking of FC131 to a receptor model of CXCR4, which will be based on the newly published X-ray structure of the human β_2 -adrenergic receptor.³⁰

Specifically, the focus will be on the three acidic residues (Asp171, Asp262 and Glu288) that have been shown to be important for binding of AMD3100 to CXCR4.

2 Methods

2.1 General

All calculations were done using commercially available software from Schrödinger³¹ (2008 Suite) on a Dell Precision 390 N-Series workstation. Specifically, the following Schrödinger modules were used: MacroModel³² for ligand minimization, Prime³³ for homology modelling, and the Induced Fit Docking protocol³⁴ for ligand docking.

2.2 Homology modeling of the CXCR4 Transmembrane helix (TMH) bundle

The 3D structure model of the CXCR4 receptor was generated by using Schrödinger's Prime module, which consists of the following 5 steps.

2.2.1 Query (CXCR4) sequence

The amino acid sequence of the human CXCR4 receptor (entry number P61073) was obtained from ExpASy (**Expert Protein Analysis System**) Proteomics Server³⁵ and imported in FASTA format (a text based format where a single letter code is used to represent amino acid sequences).

2.2.2 Template structure

The high resolution crystal structure sequence of the human β_2 -adrenergic receptor (PDB-ID 2RH1_A), was selected as the template structure.

2.2.3 Sequence alignment

The "Align GPCR" option which is a specially designed program capable of aligning GPCRs by identifying transmembrane helices and fingerprint (X.50)* matching was

* Ballesteros-Weinstein numbering is used throughout the text as superscripts to the protein numbering. Within each helix is a single, most conserved residue among the class A GPCRs. This residue is designated X.50, where X is the number of the transmembrane helix. All other residues on that helix are numbered relative to this conserved position.

used to suggest the pair-wise alignment between the template and the query sequences. The suggested pair-wise alignment was manually edited in order to remove gaps in TMHs and achieve overlap between fingerprint residues of GPCRs.³⁶

2.2.4 Building the CXCR4 structure model

The “Build structure” screen was used to predict the CXCR4 structure, omit structural discontinuities of more than 20 residues and optimize the side chains of the receptor model.

2.2.5 Refining the modeled structure

The termini and loops of the originally predicted receptor structure were removed in order to obtain only the 7 TMHs.

The helical boundary (start-end) residues of the 7 TMHs of the predicted CXCR4 receptor was defined based on the secondary structure assignment (SSA) of the template β_2 -adrenergic structure and were assigned as follows; [TMH1: Phe36-Tyr65 (the first and last residues respectively); TMH2: Met72-Asn101; TMH3: Gly105-Val139; TMH4: Leu150-Ile173; TMH5: Val197-Ser229; TMH6: Gln233-Phe264; and TMH7: His281-Leu301].

The remaining TMH bundle for CXCR4 was subjected, by using the protein preparation wizard, to addition of hydrogen atoms and capping of terminals. During cap termini, the program removes formal charges in the backbone structure by capping the N- and C-termini of the helixes with ACE (N-acetyl) and NMA (N-methyl amide) groups, respectively.

The Ramachandran plot revealed that the dihedral angle, phi (ϕ), of the edited 7 TMH segment residues (Phe36, Met72, Gly105, Leu150, Val197, Gln233 and His281) was changed to 180° during capping. The angles were therefore adjusted to the value they had in the originally predicted receptor model (Table 1).

Table 2.1: The dihedral values after phi (ϕ) adjusted.

Residue	Phe36	Met72	Gly105	Leu150	Val197	Gln233	His281
phi (ϕ)	-61.1	-59.3	84.5	-58.8	-120	-41.6	-58.8

The side chain conformations of the phi (ϕ) adjusted TMH bundle was finally optimized to its energetically most stable position resulting in the final TMH bundle of the CXCR4 receptor.

2.2.6 The final receptor model

The conformations of the two amino acid residues, Asp171 and Asp262, were changed based on visual inspection of the originally predicted receptor structure.

Rotating Asp171 and Asp262 by a 180° from their original conformation was primarily meant to increase the exposure of these residues, and mainly of Asp171, to the binding pocket(s) of the receptor structure resulting in the hereafter called the wtCXCR4 receptor.

2.3 Induced fit docking

2.3.1 Preparing the ligand structure

The starting conformation of the ligand molecule, FC131[c(Gly¹-D-Tyr²-Arg³-Arg⁴-Nal⁵), was generated from the proposed bioactive conformation of Ala³FC131 c(Gly¹-D-Tyr²-Ala³-Arg⁴-Nal⁵).²⁰

The residue Ala³ of Ala³FC131 was mutated to Arg³ resulting in the CPP FC131 (Gly¹-D-Tyr²-Arg³-Arg⁴-Nal⁵). The Macro Model program was used to minimize the energy of the ligand molecule. FC131 was built with positive charge on both Arginines (Arg³ and Arg⁴).

2.3.2 *Preparing the receptor model(s)*

The following two receptor categories were prepared and used in this project; (1) the originally predicted wild type (wtCXCR4) receptor, and (2) the mutant-receptor where the residues Asp262 and Glu288 of the wtCXCR4 receptor were mutated, one at a time, to Asn262 and Gln288, respectively to perform mutagenesis analysis.

2.3.3 *Induced Fit Docking protocol*

The default settings of Schrodinger's IFD protocol were used in all the jobs except the values specified below.

The receptor binding site represented by the energy grids of a cubic box, defined with 46 Å and 14 Å for the outer and inner cubic boxes respectively, was used in all the jobs and was centered at residue number 116 (Tyr116), which represents the approximate center of the TMH bundle.

In the docking jobs with constraints, the hydrogen bond acceptor atoms were selected from the workspace and applied in the respective jobs.

Step 1. Initial glide docking:

During this stage, Glide generated 100 ligand-receptor complexes called poses.

Step 2. Prime induced fit:

The receptor side chains of every pose within 5 Å distance (default setting) of the ligand molecule were optimized.

Step 3. Glide redocking:

At this stage, each ligand structure within 30.0 Kcal/mol from the lowest energy pose and the top 20 resulting structure complexes from step 2 was redocked, using Glide XP, into their corresponding low energy receptor structures.

The IFD protocol was used to perform the following seven jobs described below.

FC131 docking to the wild type (wtCXCR4) receptor:

1. IFD_FC131_wtCXCR4

FC131 docking to the wtCXCR4 receptor with constraint at Asp171, Asp262 and Glu288 respectively:

2. IFD_FC131_wtCXCR4_const171

3. IFD_FC131_wtCXCR4_const262

4. IFD_FC131_wtCXCR4_const288

FC131 docking to the wtCXCR4 receptor with Asp262 and Glu288 mutated to Asn262 and Gln288 with constraint at the respective residues:

5. IFD_FC131_wtCXCR4_Asp-262-Asn_const262

6. IFD_FC131_wtCXCR4_Glu288-Gln_const288

3 Results and discussion

3.1 Homology model

3.1.1 Choice of template

Since sequence similarity usually implies significant structural resemblance, the sequence of a related homologous template protein can be used in the prediction of the structure of the query sequence.³⁷ Selecting the right homolog template sequence is therefore a critical step toward the production of a good query structure.

A chemokine receptor with highest possible resolution would have been an ideal template structure for CXCR4. However, this is not available, and instead we had to choose between the more distantly related GPCR receptors with known X-ray structure; these are Rhodopsin receptor (with resolution between 2.2-4.2 Å), β_2 -Adrenergic receptor (resolution between 2.4-3.4 Å), β_1 -Adrenergic receptor (2.7 Å) and Adenosine A_{2A} receptor (resolution of 2.6 Å)³⁸; where the structures are from Bovine/squid, Human, Turkey and Human respectively.

The newly published 7 transmembrane crystal structure of human β_2 -adrenergic GPCR receptor, PDB code 2RH1_A and resolution of 2.4 Å was chosen based on its high resolution, the diffusible ligand (in contrast to Rhodopsin) and that the template is from the same species (human) as the query sequence.

3.1.2 Alignment

Literature data from GPCR family alignment shows that, the conserved fingerprint residues within the 7 transmembrane helices are: helix I [Gly (1.49) and Asn (1.50)], helix II [Leu (2.46) and Asp (2.50)], helix III [(Cys (3.25) and Asp (3.49), Arg (3.50), Tyr (3.51)], helix IV [Trp (4.50) and Pro (4.59)], helix V [Pro (5.50) and Tyr (5.58)], helix VI [Phe (6.44), Trp (6.48) and Pro (6.50)], and helix VII [Asn (7.49), Pro (7.50), and Tyr (7.53) of the NPXXY motif]³⁶, where the most conserved residues in each transmembrane helices of the rhodopsin-like GPCRs are known as the X.50 residues. In the β_2 -adrenergic receptor, these are Asn51, Asp79, Arg131, Trp158, Pro211, Pro288, and Pro323.

The result of the pairwise alignment between the query and template sequences (Figure 3.1) showed that the conserved residues of the query and template sequence overlapped and were in good agreement with earlier sequence analysis of GPCRs³⁶.

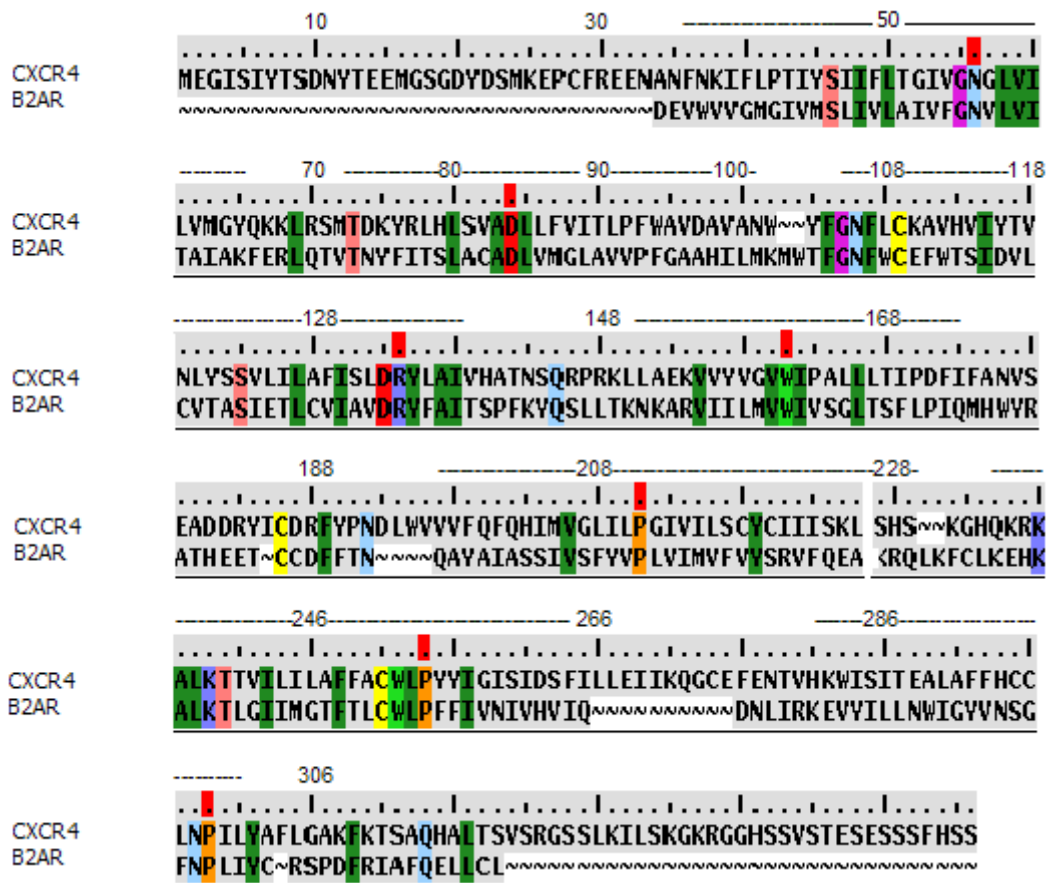


Figure 3.1: Alignment of CXCR4 and the β_2 -adrenergic receptor sequences. The dashed line over the alignment sequences shows the helical regions of the CXCR4. The most conserved residues (X.50) are marked with a red colored rectangle over them.

3.1.3 The final receptor model

The Ramachandran plot (Figure 3.2) for the final receptor structure shows that the dihedral angles for most of the TMH residues of wtCXCR4 are in the typical region for alpha-helix residues (-60° , -60°).

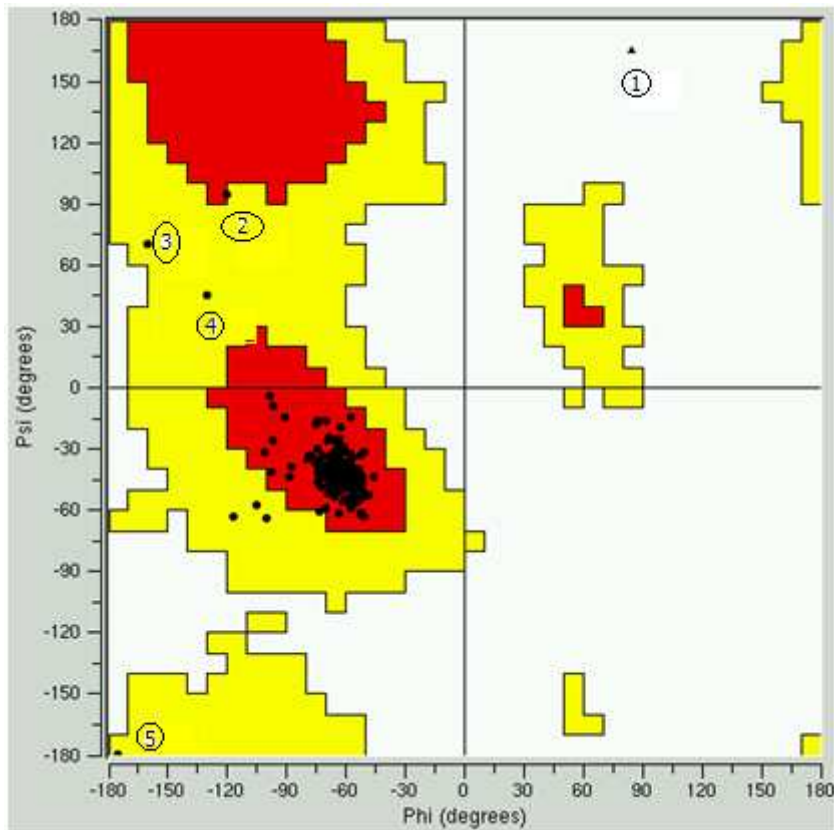


Figure 3.2: Ramachandran plot for wtCXCR4, where the numbering represents residues (1) Gly105, (2) Val197, (3) Ser229, (4) Ala100 and (5) Asn101.

Gly is less restricted because of its side chain (more flexible), and is fully allowed for Gly105 to be at the position shown; for the other residues, the deviation shown is believed to cause by “break/bend” at the ends of the helices.

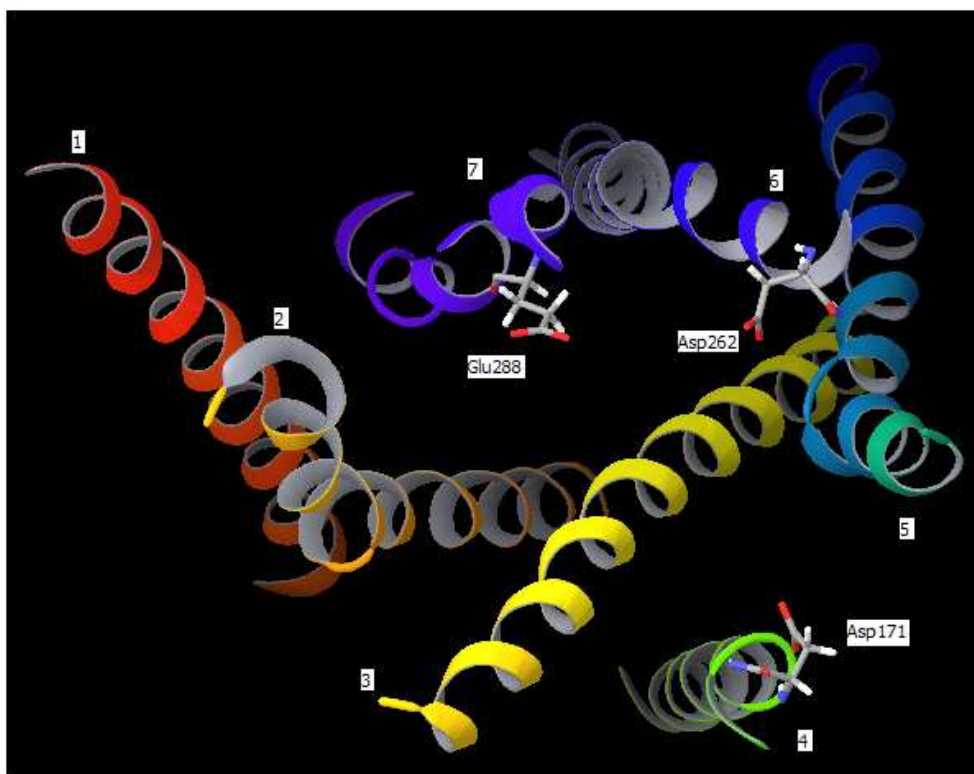


Figure 3.3: 7 TMH bundle of wtCXCR4 receptor showing Asp171 (TMH4), Asp262 (TMH6) and Glu288 (TMH7).

As shown in figure 3.3, the 7 TMHs of wtCXCR4 are organized in such a way that the residues Asp262, Glu288 are placed near to each other, whereas Asp171 is placed at a distance from these two residues. In addition to this, Asp171 is shielded from the core of the TMH bundle by transmembrane helix number 3.

3.2 Docking results

Favorable interactions were anticipated between the important pharmacophoric groups (Figure 3.3), D-Tyr², Arg³, Arg⁴ and Nal⁵, of FC131 and the contact residues of the CXCR4 receptor.

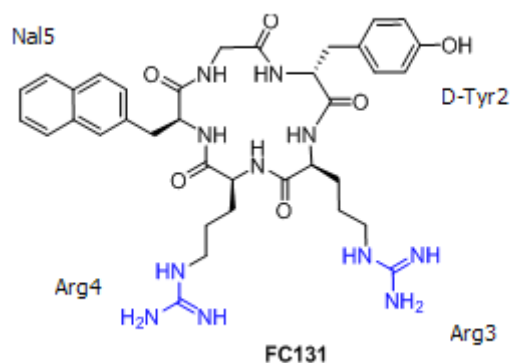


Figure 3.4: The structure of FC131. D-Tyr², Arg³, Arg⁴ and Nal⁵ are the important pharmacophoric groups. The basic center of FC131 is marked with blue color.

Since Arg³ and Arg⁴ of FC131 are positively charged, it is especially expected that these side chains of FC131 interact with the negatively charged residues (Asp171, Asp262 and Glu288) in the extracellular part of the CXCR4 TMH bundle; whereas the hydrophobic side chains D-Tyr² and Nal⁵ make hydrogen/hydrophobic- and hydrophobic bonding, respectively.

The extra precision (XP) scoring function in Glide is designed to identify ligand poses that would be expected to have unfavorable energies. Only active compounds will have available poses that avoid these penalties and also receive favorable scores for appropriate hydrophobic contacts between the receptor and the ligand, hydrogen-bonding interactions, and so on. The main purpose of the XP method is to weed out false positives and to provide a better correlation between good poses and good scores. The XP score was therefore used rank the poses and the three pose from each job. The three poses with the best XP score measured in Kcal/mol are reported for each job (job 1-job 7).

Table 3.1: FC131 docking to the wtCXCR4 receptor without constraint (job 1)

Ligand side-chains	Contact residues of the receptor structure		
	Pose 1	Pose 2	Pose 3
Arg ⁴	Val99/His281	Ala98	Glu288
Arg ³	Asp262	Asp262	Ala98
Tyr ²	**	Gln200	**
Nal ⁵	Phe292/Ala95/Phe49	Val99/Tyr45/Pro92	Ile48/Ala289/Val96
XP score	-13.82	-12.95	-11.95

** The ligand side chain is not involved in interaction.

From the result of job 1 (Table 3.1) and (Figure 3.5), we can clearly see that Asp262 and Glu288 interacted with the ligand, whereas Asp171 was not involved in binding to FC131 for the top three poses.

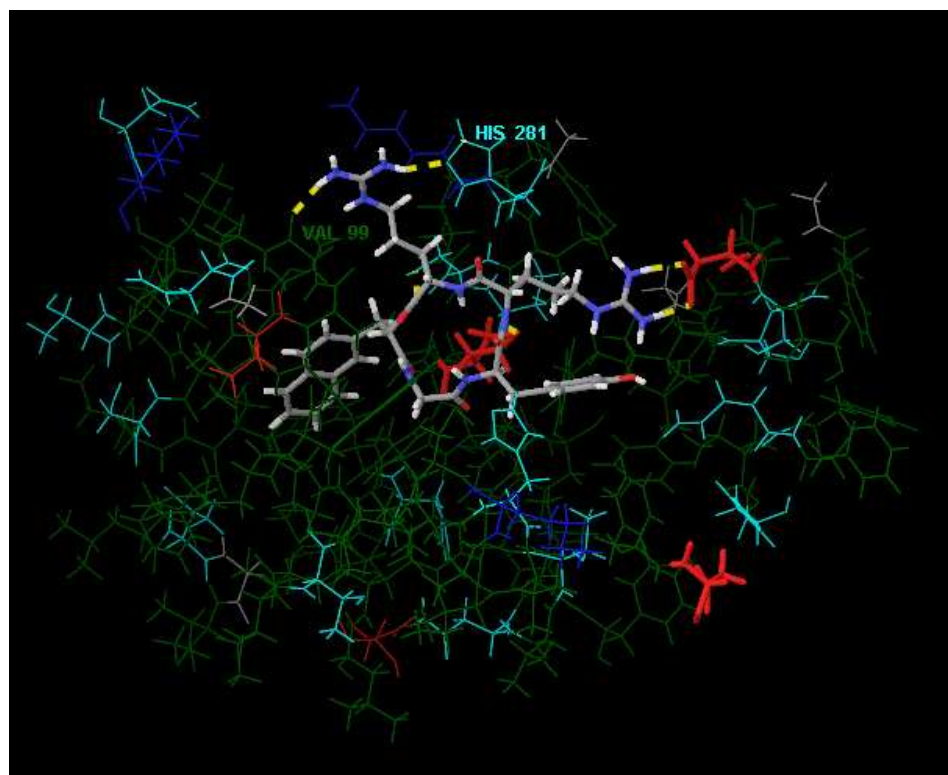


Figure 3.5: Figure showing the binding mode of the best scored pose form job 1. Arg³ and Arg⁴ are shown to involve in binding to the receptor.

The purpose of job 2 (Table 3.2), with constraint on Asp171, was therefore to see if the ligand molecule is able to bind to this particular residue and also analyze the consequence on the XP score when Asp171 is forced to participate in binding.

Table 3.2: FC131 docking to the wtCXCR4 receptor with constraint on Asp171 (job 2)

Ligand side-chains	Contact residues of the receptor structure		
	Pose 1	Pose 2	Pose 3
Arg ⁴	Asp171/Tyr121	Asp171/Tyr121	Glu288
Arg ³	Asp171	Asp171	Asp171/Tyr121
Tyr ²	**	Asp262	Asp171
Nal ⁵	Tyr256/Tyr255/Leu120/Tyr116	Tyr256/Tyr255/Leu120/Tyr116	Val99/Val98
XP score	-10.90	-10.79	-8.63

** The ligand side chain is not involved in interaction.

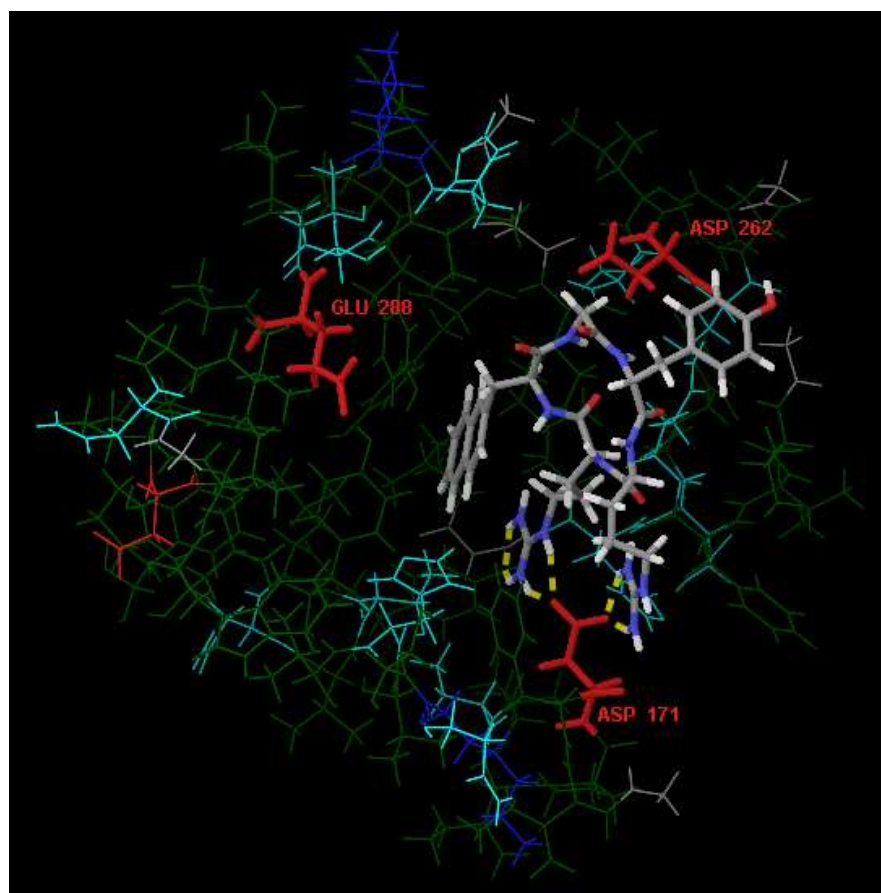


Figure 3.6: Figure showing the binding mode of the best scored pose form job 2. Both Arg3 and Arg4 are bounded to Asp171.

As shown in table 3.2, interaction of Asp171 with the ligand molecule is possible; but the XP score has decreased by about 3 Kcal/mol, when compared with the XP score of job 1 (Table 3.1).

This decrease in the XP score corresponds to a 100-fold reduction in affinity, and indicates that, interaction of the ligand FC131 with Asp171 is unfavorable. This result is also in good agreement with what an earlier binding mode study of FC131 to the CXCR4 receptor has suggested.²¹

Table 3.3: FC131 docking to the wtCXCR4 receptor with constraint on Asp262 (job 3)

Ligand Side-chains	Contact residues of the receptor structure		
	Pose 1	Pose 2	Pose 3
Arg ⁴	His281/Val99	**	Ala98/Ser285
Arg ³	Asp262	Asp262/Thr287	Asp262
Tyr ²	**	**	Gln200
Nal ⁵	Phe292/Ala95/Phe49	Val99/Tyr45/Pro92	Ile48/Ala289/Val96
XP score	-13.89	-13.45	-12.96

** The ligand side chain is not involved in interaction.

The result from job 3 (Table 3.3) indicates the involvement of Asp262 in binding of Arg³, but the XP score from this job is very similar to the XP score from job 1 (Table 3.1).

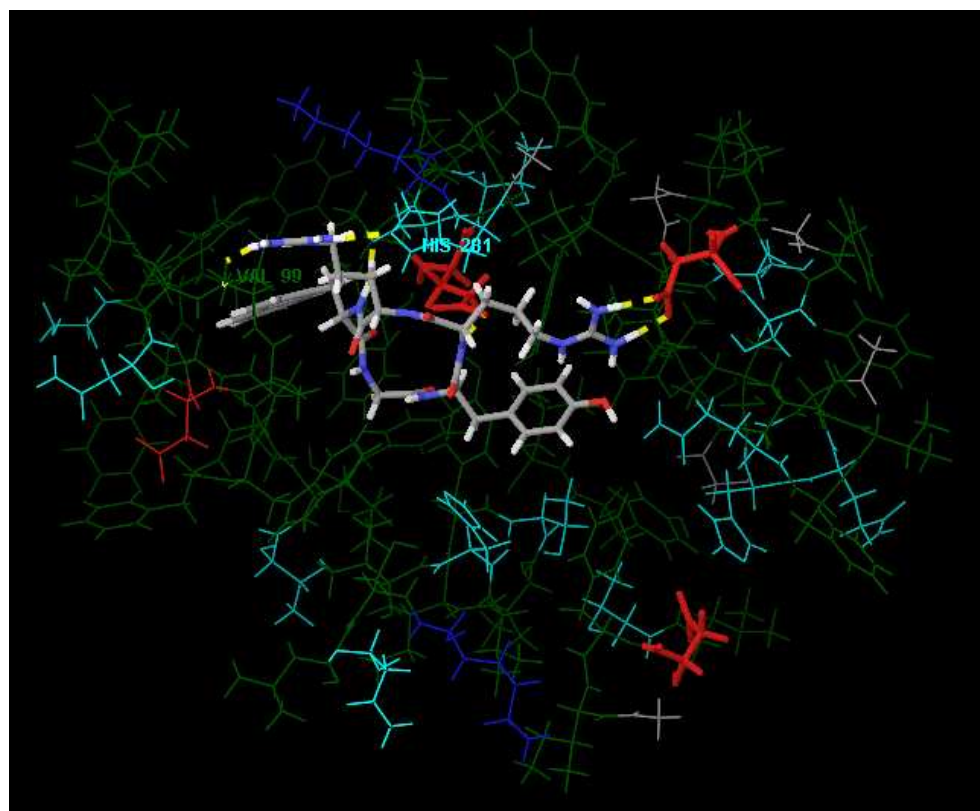


Figure 3.7: Figure showing the binding mode of the best scored pose form job 3. Picture showing the Arg³-Asp262 interaction as described in table 3.3

Table 3.4: FC131 docking to the wtCXCR4 receptor with constraint on Glu288 (job 4)

Ligand side-chains	Contact residues of the receptor structure		
	Pose 1	Pose 2	Pose 3
Arg ⁴	Glu288/Ser285	Glu288/Ile285	Glu288/His281
Arg ³	Asp262	Asp262	Asp262
Tyr ²	**	Asp171	Asp171
Nal ⁵	Val112,Leu91, Tyr116,Trp94	Ala95/Ala98 Phe292/His113	Leu120/Tyr255/Tyr256/His113 Thr117/Ile204/Tyr116/Phe172 Tyr121/Phe292
XP score	-13.02	-12.98	-12.92

** The ligand side chain is not involved in interaction.

Just like job 3 (Table 3.3), the result from job 4 (Table 3.4) does not show a significant change in the XP score when compared to job 1 (Table 3.1). The result from this job (Table 3.4) indicates the involvement of Glu288 in binding of Arg⁴.

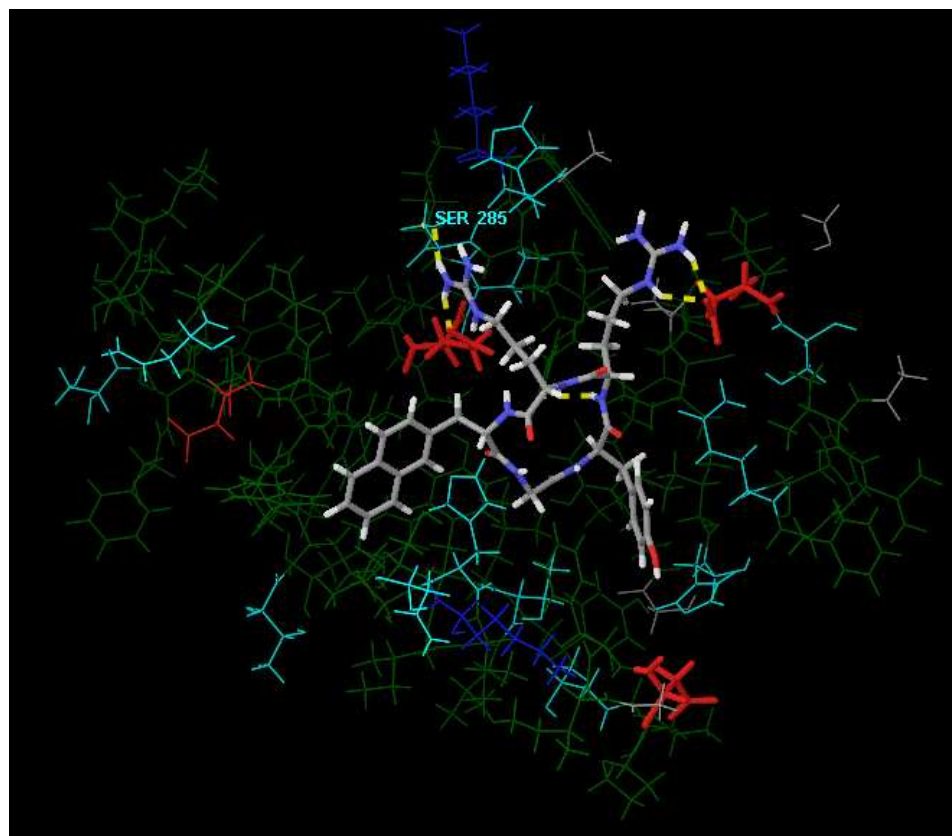


Figure 3.8: Figure showing the binding mode of the best scored pose form job 4. Picture showing the Arg⁴-Glu288 interaction as described in table 3.4

Since modeling (docking) is a theoretical approach which helps us predict the reality, the results from this work was taught to be supported by experimental data which could help us approve or disapprove the theoretical result.

Site directed mutagenesis (SDM) is a technique which is widely used to perform such kinds of analysis. Mutation is created at a defined site in a molecule containing the desired change and the outcome of this mutation will be analyzed.

The ligand molecule, FC131, was therefore sent to Copenhagen for SDM-studies, but the result which was expected to come at the beginning of this work, is still unfinished. The aim was to see if it was possible to correlate the experimental data with the docking data.

Based on the following two mutational data, the effect of mutating Asp262Asn and Glu288Gln was examined in job 5 and job 6 respectively.

(1) Mutational studies done by substituting Asp262 to Asn, has been shown to affect (decrease) HIV coreceptor activity³⁹. It was also reported that mutation of Asp262 to Ala262 showed to reduce the anti-HIV activity of T140.⁴⁰

(2) Glu288Gln⁴¹ and Glu288Ala⁹ mutations has been shown to reduce signaling of SDF-1 α significantly.

Mutational data has confirmed the importance of Asp171 for binding of HIV gp120⁴², and binding mode analysis studies showed the involvement of Asp171 in binding with T140⁴⁰ (which FC131 is derived from) and AMD3100.⁸ However, mutating Asp171 was not necessary since binding of Asp171 to FC131 is shown to be unfavorable (job 2 or Table 3.2).

Table 3.5: FC131 docking to Asp262Asn mutated CXCR4 with constraint on Asn262 (job 5)

Ligand side-chains	Contact residues of the receptor structure		
	Pose 1	Pose 2	Pose 3
Arg ⁴	Glu288/Tyr255	Glu288/His113	Glu288
Arg ³	Asn262/Gln200	Asn262	Asn262/Gln200
Tyr ²	**	**	Lys110
XP score	-9.55	-9.24	-9.21

** The ligand side chain is not involved in interaction.

The result from job 5 (Table 3.5), with the best XP score of -9.55, shows a reduction in the XP score when compared to the result from job 1 (Table 3.1) and job 3 (Table 3.3).

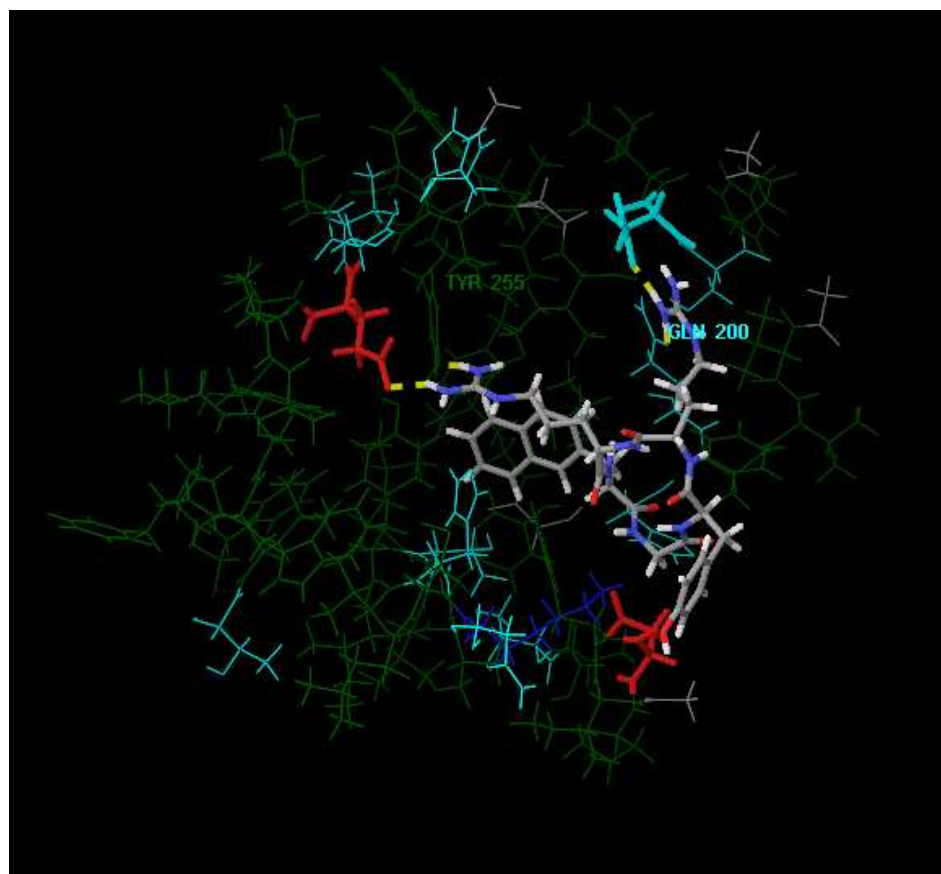


Figure 3.9: Figure showing the binding mode of the best scored pose from job 5. As shown in the picture and also table 3.5, Arg³ binds to Asn262 and Arg⁴ to Glu288.

A positive contribution of Asp262 to the ligand interaction was not suggested in a previous study.²¹ However, the decrease in the XP score when Asp262 is mutated to Asn262, shows that Asp262 contributes positively to the ligand-receptor interaction and seems important for the binding of the ligand molecule to the CXCR4 receptor.

Table 3.6: FC131 docking to Glu288Gln mutated CXCR4 with constraint on Gln288 (job 6)

Ligand side-chains	Contact residues of the receptor structure		
	Pose 1	Pose 2	Pose 3
Arg ⁴	Asp262	Asp262/Gln200	Asp262
Arg ³	**	Asp262	His113
Tyr ²	Ala98	**	**
XP score	-14.09	-13.26	-13.04

** The ligand side chain is not involved in interaction.

The result from job 6 (Table 3.6) shows a small increase in the XP score when compared to job 1 (Table 3.1). We saw from job 4 (Table 3.4) that Glu288 is involved in binding of FC131, and the affinity is therefore expected to decrease (by about 100-1000 fold) when this residue is mutated to Gln288.

The observed result from job 6 (Table 3.6), however, showed no/a small increase in affinity of the ligand when Glu288 is mutated to Gln. This is in contrast to what one would expect, and if this also was observed experimentally, it would be concluded that Glu288 is not involved in binding. This shows that the result from job 6 (Table 3.6) is a false negative which leads to a wrong conclusion.

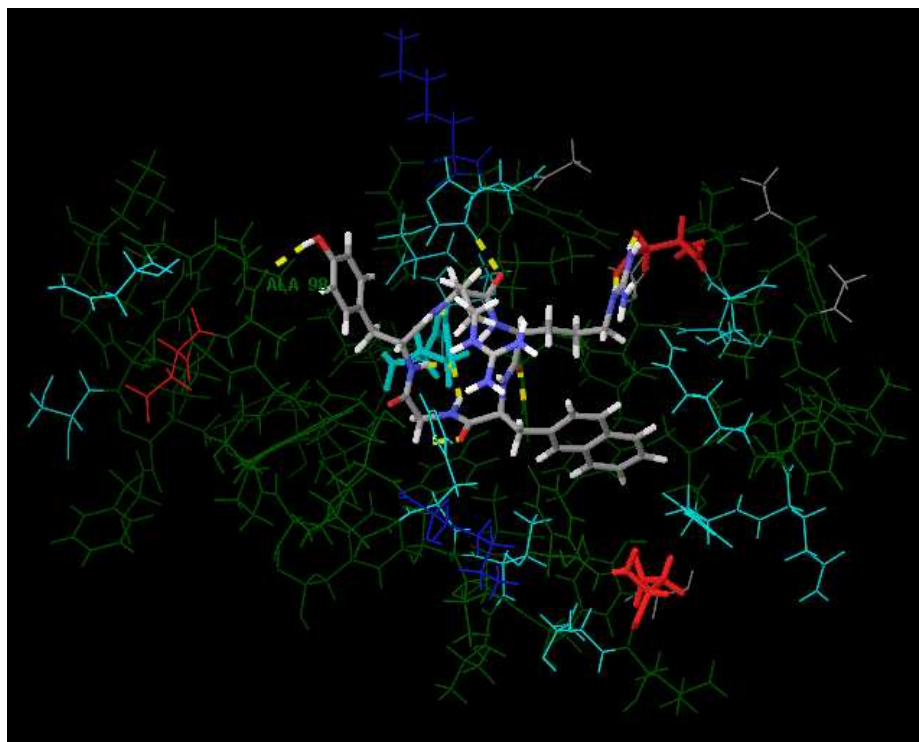


Figure 3.10: Figure showing the binding mode of the best scored pose from job 6. The ligand is rotated by a 180° directing Arg³ to Glu288 and Arg⁴ to Asp262.

Figure 3.10 shows that Gln288 binds to the backbone of FC131 and that the binding fashion of the ligand molecule is rotated by about 180° when we compare it with the binding fashion from job 4 (Figure 3.8).

3.3 Possible binding mode(s) of FC131

Candidate binding mode for FC131, based on the four different docking jobs, (job 1, job 3 and job 4) plus one additional job (docking to wtCXCR4 with constraint on Asp262 and Glu288), which the result is not shown here, were used in analyzing the binding mode of FC131 to the CXCR4 receptor.

The following two ligand binding modes (Table 3.7) were suggested by analyzing the contact residues of the three best poses from each job.

Table 3.7 Contact residues of the representative binding modes.

Pharmacophoric Groups	Binding mode 1	Binding mode 2
Arg ⁴	Ala98, Val99, His281, Ser285	His281, Ile284, Ser285, Glu288
Arg ³	Asp262, Thr287	Asp262
Tyr ²	Gln200	Asp171
Nal ⁵	Phe292, Ala95, Phe49, Val99 Tyr445 Pro92, Ile48, Ala289, Val96	Tyr116, Phe298, Cys109, Pro92, Leu91, Ala95, Val112, Ala98, Trp94

The positioning of the ligand molecule is almost “the same” in both of the suggested binding modes, i.e. the side chains of FC131 are directed to the same receptor residue (Tyr² to Asp171, Arg³ to Asp262 and Arg⁴ to Glu288) in both cases.

The main difference of these two binding modes is the placing of the backbone which is positioned near Glu288 in one case (binding mode 1) and near Asp262 (binding mode 2) in the other case.

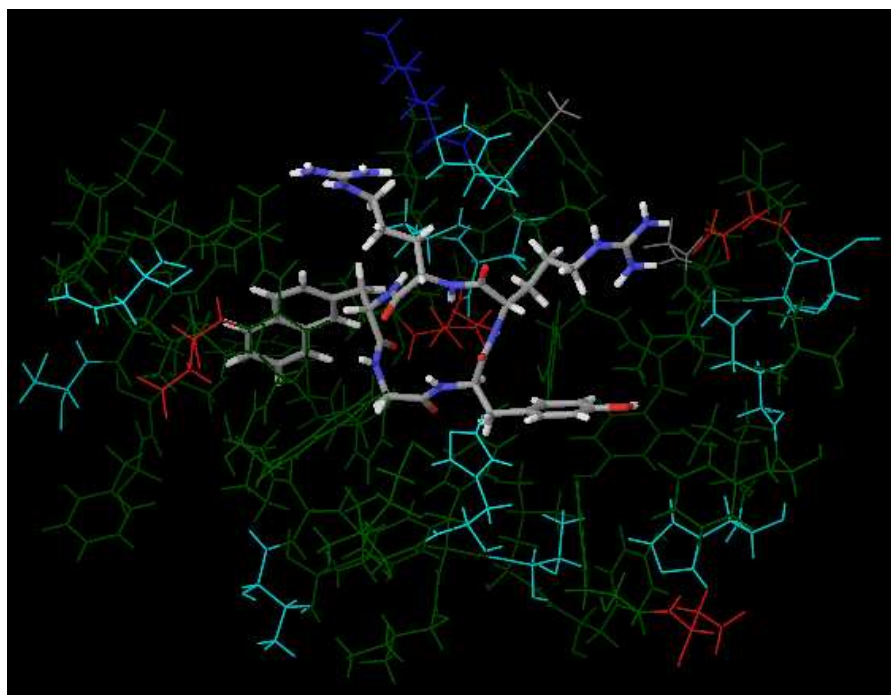


Figure 3.11: Binding mode 1 of FC131. The backbone is placed near Glu288.

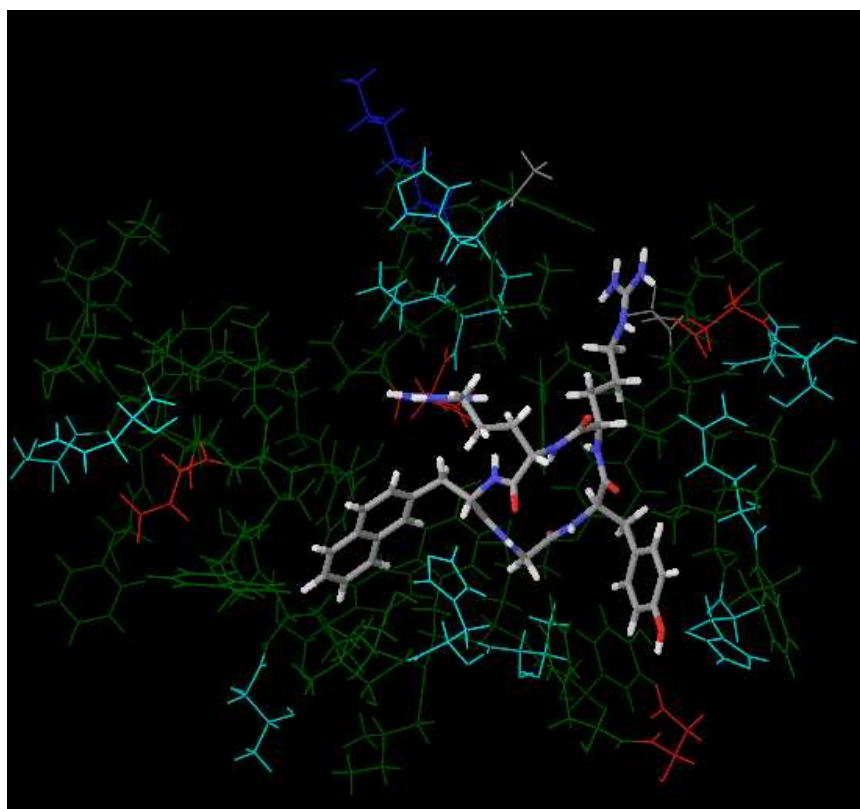


Figure 3.12: Binding mode 2 of FC131. The backbone is placed near Asp262.

4 Conclusions

Based on induced fit docking of FC131 to a homology modeled 7TMH CXCR4 receptor, the residues Asp262 and Glu288 seems to be involved in ligand binding by interacting with Arg³ and Arg⁴, respectively. Asp171, as shown from this present study, does not seem to be involved in binding to FC131.

When docking to the Asp262Asn and Glu288Gln mutants, a reduction in affinity was observed for Asp262Asn, which would indicate that this residue is important for ligand-receptor interaction. For Glu288Gln, a small increase in affinity was observed, which would suggest that Glu288 is not involved in binding. However, since the involvement of Glu288 was already established, the docking result for Glu288Gln must be considered as a false negative. This indicates that the docking protocol used in the present study may not be suited for determining the effects of receptor mutations.

Two possible binding modes of the ligand molecule are suggested. These two binding modes of the ligand molecule have similar binding fashion but the ligands are positioned in two different “binding pockets” of the receptor structure.

The induced fit docking program used in this work allows flexibility of both the ligand and receptor structure which is important when ligand-induced conformational changes are relevant; this is one of the strongest side in the present study.

The advantages of using the human β_2 -adrenergic receptor as template structure, when compared to templates that have been used in earlier studies (rhodopsin), is the diffusible ligand (in contrast to rhodopsin), the high crystal resolution and the origin (human). However, the template is not a chemokine receptor and this is one of the main drawbacks in this work.

Another weakness in this study is the removal the extra- and intracellular loops; which can have importance for the binding site conformation and/or ligand receptor binding.

5 References

1. Charo, I. F.; Ransohoff, R. M., The many roles of chemokines and chemokine receptors in inflammation. *N Engl J Med* **2006**, 354, (6), 610-21.
2. <http://www.iuphar-db.org/GPCR/ReceptorDisplayForward?receptorID=2216>.
3. Balabanian, K.; Lagane, B.; Infantino, S.; Chow, K. Y.; Harriague, J.; Moepps, B.; Arenzana-Seisdedos, F.; Thelen, M.; Bachelier, F., The chemokine SDF-1/CXCL12 binds to and signals through the orphan receptor RDC1 in T lymphocytes. *J Biol Chem* **2005**, 280, (42), 35760-6.
4. Burns, J. M.; Summers, B. C.; Wang, Y.; Melikian, A.; Berahovich, R.; Miao, Z.; Penfold, M. E.; Sunshine, M. J.; Littman, D. R.; Kuo, C. J.; Wei, K.; McMaster, B. E.; Wright, K.; Howard, M. C.; Schall, T. J., A novel chemokine receptor for SDF-1 and I-TAC involved in cell survival, cell adhesion, and tumor development. *J Exp Med* **2006**, 203, (9), 2201-13.
5. Feng, Y.; Broder, C. C.; Kennedy, P. E.; Berger, E. A., HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor. *Science* **1996**, 272, (5263), 872-7.
6. Tamamura, H.; Tsutsumi, H.; Fujii, N., The chemokine receptor CXCR4 as a therapeutic target for several diseases. *Mini Rev Med Chem* **2006**, 6, (9), 989-95.
7. De Clercq, E.; Yamamoto, N.; Pauwels, R.; Balzarini, J.; Witvrouw, M.; De Vreese, K.; Debyser, Z.; Rosenwirth, B.; Peichl, P.; Datema, R.; et al., Highly potent and selective inhibition of human immunodeficiency virus by the bicyclam derivative JM3100. *Antimicrob Agents Chemother* **1994**, 38, (4), 668-74.
8. 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, (17), 14153-60.
9. 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: transfer of binding site to the CXCR3 receptor. *J Biol Chem* **2004**, 279, (4), 3033-41.
10. De Clercq, E., The AMD3100 story: The path to the discovery of a stem cell mobilizer (Mozobil). *Biochem Pharmacol* **2009**, 77, (11), 1655-64.
11. 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, (7), 4185-90.
12. Tamamura, H.; Otaka, A.; Fujii, N., Development of anti-HIV agents targeting dynamic supramolecular mechanism: entry and fusion inhibitors based on CXCR4/CCR5 antagonists and gp41-C34-remodeling peptides. *Curr HIV Res* **2005**, 3, (4), 289-301.
13. 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, (9), 3280-9.
14. Princen, K.; Schols, D., HIV chemokine receptor inhibitors as novel anti-HIV drugs. *Cytokine Growth Factor Rev* **2005**, 16, (6), 659-77.
15. 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, (2), 249-62.
16. Hruby, V. J.; Li, G.; Haskell-Luevano, C.; Shenderovich, M., Design of peptides, proteins, and peptidomimetics in chi space. *Biopolymers* **1997**, 43, (3), 219-66.
17. al-Obeidi, F.; Hruby, V. J.; Sawyer, T. K., Peptide and peptidomimetic libraries. Molecular diversity and drug design. *Mol Biotechnol* **1998**, 9, (3), 205-23.
18. Hartley, O.; Gaertner, H.; Wilken, J.; Thompson, D.; Fish, R.; Ramos, A.; Pastore, C.; Dufour, B.; Cerini, F.; Melotti, A.; Heveker, N.; Picard, L.; Alizon, M.; Mosier, D.; Kent, S.; Offord, R., Medicinal chemistry applied to a synthetic protein: development of highly potent HIV entry inhibitors. *Proc Natl Acad Sci U S A* **2004**, 101, (47), 16460-5.

19. Kazmierski, W. M.; Kenakin, T. P.; Gudmundsson, K. S., Peptide, peptidomimetic and small-molecule drug discovery targeting HIV-1 host-cell attachment and entry through gp120, gp41, CCR5 and CXCR4. *Chem Biol Drug Des* **2006**, *67*, (1), 13-26.
20. Vabeno, J.; Nikiforovich, G. V.; Marshall, G. R., A minimalistic 3D pharmacophore model for cyclopentapeptide CXCR4 antagonists. *Biopolymers* **2006**, *84*, (5), 459-71.
21. Vabeno, J.; Nikiforovich, G. V.; Marshall, G. R., Insight into the binding mode for cyclopentapeptide antagonists of the CXCR4 receptor. *Chem Biol Drug Des* **2006**, *67*, (5), 346-54.
22. Griffith, R.; Luu, T. T.; Garner, J.; Keller, P. A., Combining structure-based drug design and pharmacophores. *J Mol Graph Model* **2005**, *23*, (5), 439-46.
23. Yuzlenko, O.; Kiec-Kononowicz, K., Molecular modeling of A1 and A2A adenosine receptors: comparison of rhodopsin- and beta2-adrenergic-based homology models through the docking studies. *J Comput Chem* **2009**, *30*, (1), 14-32.
24. Taft, C. A.; Da Silva, V. B.; Da Silva, C. H., Current topics in computer-aided drug design. *J Pharm Sci* **2008**, *97*, (3), 1089-98.
25. Westbrook, J.; Feng, Z.; Chen, L.; Yang, H.; Berman, H. M., The Protein Data Bank and structural genomics. *Nucleic Acids Res* **2003**, *31*, (1), 489-91.
26. Bajorath, J., Integration of virtual and high-throughput screening. *Nat Rev Drug Discov* **2002**, *1*, (11), 882-94.
27. Anderson, A. C.; O'Neil, R. H.; Surti, T. S.; Stroud, R. M., Approaches to solving the rigid receptor problem by identifying a minimal set of flexible residues during ligand docking. *Chem Biol* **2001**, *8*, (5), 445-57.
28. Sousa, S. F.; Fernandes, P. A.; Ramos, M. J., Protein-ligand docking: current status and future challenges. *Proteins* **2006**, *65*, (1), 15-26.
29. 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*, (2), 534-53.
30. 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.; Stevens, R. C., High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science* **2007**, *318*, (5854), 1258-65.
31. <http://www.schrodinger.com>.
32. MacroModel, version 9.6, Schrödinger, LLC, New York, NY, 2008.
33. Prime, version 2.0, Schrödinger, LLC, New York, NY, 2008.
34. Schrödinger Suite 2008 Induced Fit Docking protocol; Glide version 5.0, Schrödinger, LLC, New York, NY, 2005; Prime version 1.7, Schrödinger, LLC, New York, NY, 2005.
35. <http://www.expasy.org>.
36. Mirzadegan, T.; Benko, G.; Filipek, S.; Palczewski, K., Sequence analyses of G-protein-coupled receptors: similarities to rhodopsin. *Biochemistry* **2003**, *42*, (10), 2759-67.
37. Marti-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A., Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct* **2000**, *29*, 291-325.
38. Hanson, M. A.; Stevens, R. C., Discovery of new GPCR biology: one receptor structure at a time. *Structure* **2009**, *17*, (1), 8-14.
39. Hatse, S.; Princen, K.; Gerlach, L. O.; Bridger, G.; Henson, G.; De Clercq, E.; Schwartz, T. W.; Schols, D., Mutation of Asp(171) and Asp(262) of the chemokine receptor CXCR4 impairs its coreceptor function for human immunodeficiency virus-1 entry and abrogates the antagonistic activity of AMD3100. *Mol Pharmacol* **2001**, *60*, (1), 164-73.
40. Trent, J. O.; Wang, Z. X.; Murray, J. L.; Shao, W.; Tamamura, H.; Fujii, N.; Peiper, S. C., Lipid bilayer simulations of CXCR4 with inverse agonists and weak partial agonists. *J Biol Chem* **2003**, *278*, (47), 47136-44.
41. 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*, (31), 23736-44.

42. Tian, S.; Choi, W. T.; Liu, D.; Pesavento, J.; Wang, Y.; An, J.; Sodroski, J. G.; Huang, Z., Distinct functional sites for human immunodeficiency virus type 1 and stromal cell-derived factor 1alpha on CXCR4 transmembrane helical domains. *J Virol* **2005**, 79, (20), 12667-73.