J Mol Model (2011) 17:2707–2716
DOI 10.1007/s00894-010-0943-x

SHORT COMMENT

Ligand supported homology modeling and docking evaluation of CCR2: docked pose selection by consensus scoring

Jong-Hoon Kim · Jee Woong Lim · Seung-Woo Lee · Kyoungrak Kim · Kyoung Tai No

Received: 29 August 2010 / Accepted: 20 December 2010 / Published online: 7 January 2011

Abstract Chemokine receptor 2 (CCR2) is a G-protein coupled receptor (GPCR) and a crucial target for various inflammatory and autoimmune diseases. The structure based antagonists design for many GPCRs, including CCR2, is restricted by the lack of an experimental three dimensional structure. Homology modeling is widely used for the study of GPCR-ligand binding. Since there is substantial diversity for the ligand binding pocket and binding modes among GPCRs, the receptor-ligand binding mode predictions should be derived from homology modeling with supported ligand information. Thus, we modeled the binding of our proprietary CCR2 antagonist using ligand supported homology modeling followed by consensus scoring the docking evaluation based on all modeled binding sites. The protein-ligand model was then validated by visual inspection of receptor-ligand interaction for consistency of published site-directed mutagenesis data and virtual screening a decoy compound database. This model was able to successfully identify active compounds within the decoy database. Finally, additional hit compounds were identified through a docking-based virtual screening of a commercial database, followed by a biological assay to validate CCR2 inhibitory activity. Thus, this procedure can be employed to screen a large database of compounds to identify new CCR2 antagonists.

Keywords CCR2 · Consensus scoring · Docking evaluation · GPCR · Ligand supported homology modeling · Virtual screening

J.-H. Kim · K. T. No (✉)
Department of Biotechnology, Yonsei University,
Seoul 120-749, Republic of Korea
e-mail: ktno@yonsei.ac.kr

J.-H. Kim
e-mail: jonghoon@yonsei.ac.kr

J.-H. Kim · J. W. Lim · S.-W. Lee · K. Kim
Research and Development Center, YangJi Chemicals,
Suwon 443-766, Republic of Korea

J. W. Lim
Department of Chemistry, Sejong University,
Seoul 143-747, Republic of Korea

K. T. No
Bioinformatics and Molecular Design Research Center,
Seoul 120-749, Republic of Korea

Introduction

Chemokines play a crucial role in the trafficking of leukocytes in the body through the binding to their cognate receptors. Chemokine receptors belong to the superfamily of G-protein-coupled receptors (GPCRs). CCR2 is one of the chemokine receptors and is expressed by a variety of cell types, including monocytes, immature dendritic cells, activated T lymphocytes, basophils, and endothelial and vascular smooth-muscle cells. CCR2 binds several chemokines, including CCL2, CCL7, CCL8, and CCL13 [1, 2]. Among the chemokines that bind CCR2, monocyte chemoattractant protein-1 (MCP-1 or CCL2) is upregulated in many autoimmune and inflammatory conditions [3]. CCR2 and MCP-1 have been implicated in the pathophysiology of several inflammatory and auto-immune diseases, including rheumatoid arthritis [4], multiple sclerosis [5], atherosclerosis [6], organ transplant rejection [7], and insulin resistance [8]. Therefore, blocking the binding of MCP-1 to CCR2 may have therapeutic potential in treating chronic inflammatory conditions, and the discovery and development of small molecule CCR2 antagonists has been regarded as an important goal.

A number of CCR2 antagonists have been reported during the past decade and are structurally diverse (Fig. 1). Among the various scaffolds reported, the glycinamide-linked series, such as compounds 1, 3, and 6 (Fig. 1), have been explored by several companies [9, 10]. Recently, we synthesized new Glycinamide-linked compounds (Fig. 2) that have antagonistic activity against CCR2 [11]. Compound 7, shown in Fig. 2, displayed good binding affinity and cellular functional (chemotaxis) inhibition, as shown. We hereby report the docking study of our lead molecule, compound 7, using ligand supported homology modeling of CCR2.

Structure-based ligand modeling of GPCRs has been hampered by the difficulty in solving X-ray crystallographic structure of GPCRs. Thus, structural information about GPCRs is often inferred by de novo structure prediction strategies [12–14] or by homology modeling from the solved GPCR structures. For example, in 2000 bovine rhodopsin was the first GPCR to have an atomic structure determined [15]. This structure has been used as a template protein for homology modeling of many GPCRs. The recently published crystal structure of the engineered human β_2 adrenergic receptor (β_2AR) [16, 17] provided alternative structural information to bovine rhodopsin. In contrast to the rhodopsin crystal structure, the β_2AR crystal structure has a larger and more open space in the upper helices region and contains a reversibly bound ligand [16]. Because of this feature, we used the structure of β_2AR as the template for a ligand supported homology modeling study.

A number of small molecule ligands for GPCRs are known to bind within the upper region of the protein, along the seven TM helices. Since homology models of GPCRs are built on the same β_2AR template, they reveal common TM arrangements. Therefore, the difference in ligand binding sites across GPCRs mainly comes from side chain orientation. Thus, refinement of the side chain orientation and conformation is crucial in the process of homology modeling for GPCRs. One method is to optimize the side chains through mechanics and/or dynamics in an empty pocket; however, this method may be inaccurate, due to possible disruption of the binding site. Recently, it has been suggested that ligand-supported homology modeling [18, 19], ligand-steered homology modeling [20], and homology-modeling protein-ligand interactions [21] can reduce the uncertainty in binding site modeling results. These methods introduce the ligand in the early stage of homology modeling and retain the ligand during the optimization procedure. Thus, it is important to incorporate bound ligand information in the process of modeling GPCR binding pockets, if known ligands are available. Considering the apparent flexibility of GPCRs and the diversity in ligand types, binding pockets, and binding modes, GPCR models should be tailor-made, using all possible experimental data to guide their construction and validation during the different modeling steps [22].

Based on these the ligand-supported strategies, we used a protocol in which the ligand was incorporated throughout the modeling process. We kept the ligand from the β_2AR crystal structure, carazolol, during the initial homology model building of CCR2. Then compound 7 was replaced during the process of building a pool of the docking poses. We generated the pool of ligand-receptor complexes using LibDock [23, 24] docking to several homology models obtained from the structure of β_2AR and refined each complex through energy minimization of the region surrounding the bound ligand.

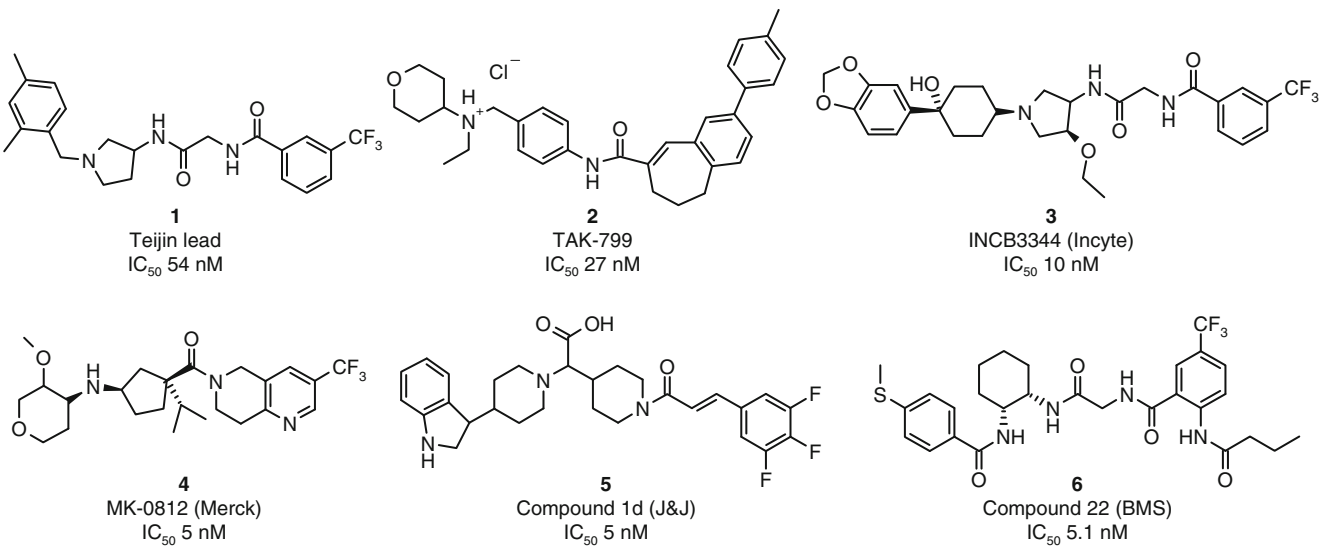


Fig. 1 Various CCR2 antagonists

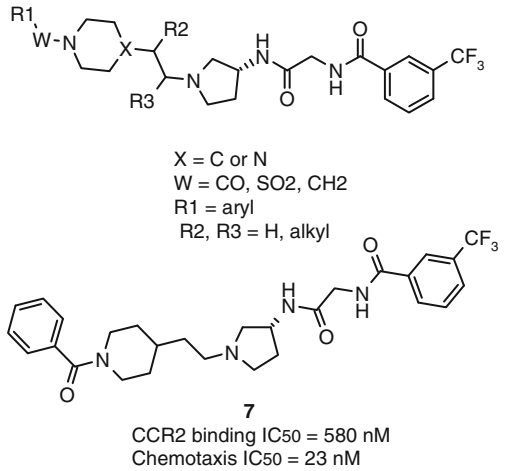


Fig. 2 4-Piperidinoethyl and piperazinoethyl pyrrolidine analogs have been synthesized. The structure and in vitro activities of the lead molecule, compound 7, are shown

Another challenging task in developing a docking model is selecting receptor-ligand binding mode from the docking pose pool. To do this we employed a consensus scoring method. The top ranked model from the consensus scoring, using 10 scoring functions, was assumed to be the binding mode of compound 7. This binding mode was then validated by small scale virtual screening of a decoy database and a database that included the active compounds. Through this procedure, the docking pose of compound 7 in the CCR2 binding pocket was successfully proposed. In addition, we applied this docking model to a virtual screening of a commercial library and identified a hit compound.

Materials and methods

Sequence alignment between CCR2b and human β_2 adrenergic receptor

The sequence of human chemokine receptor CCR2b was obtained from the NCBI database, and the crystal structure of human β_2AR (PDB code 2RH1) and rhodopsin (PDB code 1F88) were obtained from the PDB database. Modeling was performed using modules bundled in the Discovery Studio (DS) version 2.1 package (Accelrys, San Diego, CA).

The multiple sequence alignment of β_2AR, rhodopsin, and CCR2b was generated by an alignment protocol based on the clustal W program in DS. The initial sequence alignment was carried out using automatic sequence alignment using a BLOSUM 30 pair-wise alignment scoring matrix and default values in DS. Then the alignment was modified to maintain the conservation of key residues and establish no gap in the helices region by comparing β_2AR and CCR2b sequences.

Building homology models

Homology modeling was carried out using the program MODELLER [25] in the DS software package. During the homology modeling calculation, the ligand (carazolol) in the crystal structure of β_2AR was copied to maintain a cavity in the space surrounded by the seven TM helices. The disulfide bond between Cys113 in Helix III and Cys190 in Extracellular loop II (ECL2) was included and all of the intracellular and extracellular loops were generated. Twenty models were generated and sorted by PDF total energy. Three models having the lowest energy were selected and refined by the energy minimization with a fixed backbone constraint. The CHARMm force field and conjugate gradient algorithms with the maximum of 500 iterations were applied. The models were evaluated by PROCHECK [26].

Building a pool of ligand-docked receptors

Ligand files were prepared using the SD file format and then each ligand was ionized at the physiological pH range of 7.3 - 7.5 using the module in DS. The generation of ligand conformers was carried out using the BEST algorithm in the Catalyst program (Accelrys). The BEST method uses a poling algorithm [27] to generate a diverse set of low-energy conformations. The maximum number of conformers was set to 255 with a 20 kcal mol^{-1} energy window from the lowest energy. The conformers from the ligand were docked into the three homology models using the LibDock program. The binding site sphere for LibDock calculation was defined to be centered on a point located halfway between the Glu291 acidic carbon and the Tyr120 phenolic oxygen, and had a radius of 13 Å. Hydrophilic and lipophilic hot spots were generated with a density of 250. The number of saving poses was assigned to 50 for each protein model. Since three protein models from homology study were tried, 150 different docking poses for the ligand were obtained and analyzed. The rest of the docking parameters were set to default values.

Optimization of docking poses

Docking complexes were energy-minimized using ligand minimization protocols within the DS software package. The flexible residues in the receptor were defined to include any residues within a 15Å radius sphere from the center of the docked conformer. All atoms in the ligand were also included in the minimization calculation. CHARMm force field and smart minimizer algorithm in DS were employed. The dielectric constant was set to 4.0 and a maximum of 1000 iterations were allowed. This minimization calculation was applied to all 150 ligand-CCR2 binding complexes.

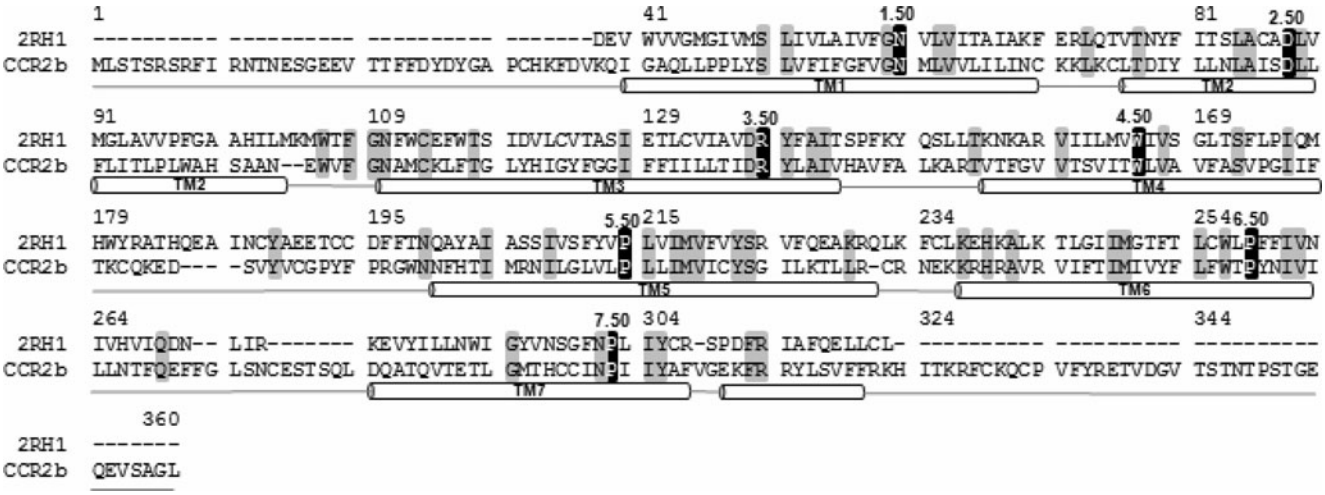


Fig. 3 Sequence alignment of human CCR2b and human β_2AR (PDB code 2RH1). The Ballesteros-Weinstein numbering [38] system was adopted to mark all the conserved amino acid residues in each helix and these residues are shadowed in black. Identical residues between CCR2b and β_2AR are shadowed in gray. Residue numbering corresponds to the CCR2b protein sequence. The secondary structure of the CCR2 receptor, based on β_2AR crystal structure, is indicated out below the sequence

Analysis of docking poses

After minimization calculations, docking scores were evaluated for the 150 binding complexes using scoring functions implemented through the DS program, including LigScore1 [28], LigScore2 [28], PLP1 [29], PLP2 [29], Jain [30], PMF01 [31], PMF04 [31], LUDI1 [32], LUDI2 [32], and LUDI3 [32]. Consensus scoring using a rank-by-rank strategy [33] was also evaluated using these ten scoring functions. All of the binding models are ranked by the average ranks predicted for each of the scoring functions. For example, if a binding model ranks no. 3 according to scoring function A and ranks no. 5 according to scoring function B, then its average rank will be (3+5)/2=4. The top ranked ligand-protein complex from the consensus scoring method was selected as the best binding site model. To confirm the validity of protein coordinates, the ligand was removed from the complex and the resulting structure was evaluated by PROCHECK.

Preparation of the validation set

The active compound database, which contained compounds having CCR2 antagonistic activity, was constructed from the in-house synthesized proprietary compounds. The active database includes 27 compounds whose IC_{50} against CCR2 binding are below 5 μM. The potentially inactive compound database was constructed from the Comprehensive Medicinal Chemistry (CMC) database (Symyx Technologies, Inc., Sunnyvale, CA). This database consisted of 1206 molecules that have MW ranging from 400 to 550. A SD file for the databases was prepared by having the compounds ionized in the physiological pH range of 7.3 - 7.5 using the module in DS.

Validation of the binding site model using virtual screening of small molecule databases

LigandFit [34], a docking program in DS, was used for the docking calculations. The LigandFit docking procedure

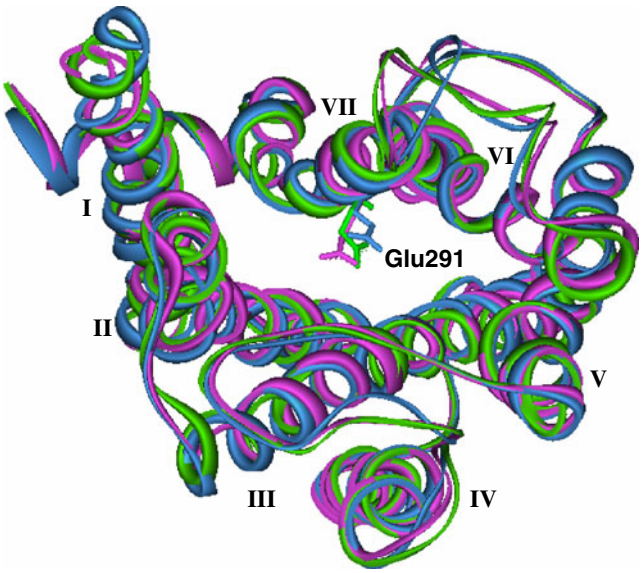


Fig. 4 The three initial models derived by PDF energy in MODELLER from 20 homology models. Model A is shown in green, B in blue, and C in purple. The side chain of Glu291 (one of the residues that interacts with the ligand) from each model clearly shows the variation in the orientation

Table 1 Data from the Ramachandran plot of the three initial CCR2 homology models

Initial model	MFR[a]		AAR[b]		GAR[c]		Subtotal[d]	Number of Gly and Pro	Total[e]
	n	%	n	%	n	%			
Model A	217	84.4	35	13.6	5	1.9	257	26	283
Model B	215	83.7	37	14.4	5	1.9	257	26	283
Model C	218	84.8	35	13.6	4	1.6	257	26	283

[a] The most favored regions

[b] The additional allowed regions

[c] The generously allowed regions

[d] The total number of residues evaluated except glycines and prolines

[e] The total number of amino acids in the receptor models

consists of two parts: 1) definition of cavity as the active site for docking; 2) docking ligands into the site. The volume of the docked ligand was defined as the binding site cavity. The calculations were performed with the set of default parameters. One docked conformation for each molecule was collected. Docking scores per each molecule were evaluated using ten scoring functions in DS. Enrichment factors were calculated at 10% of the ranked database using the following equation [35].

$$EF^{10\%} = \left(Hits_{sampled} / N_{sampled}\right) \div \left(Hits_{total} / N_{total}\right).$$

Virtual screening of a commercial small molecule database

The binding site model of compound 7 was virtually screened using a chemokine focused library (ChemDiv, Inc., San Diego, CA). LigandFit was used for the docking calculation and set to retain the top ranked conformations. The docking poses were evaluated using the 10 scoring functions. Selected compounds were tested using a CCR2 ligand binding inhibition assay (below).

Biological assay

Membranes from the stable HEK293-EBNA cell line expressing the human chemokine CCR2b receptor (PerkinElmer Life and Analytical Sciences, Boston, USA) were used. For the CCR2b receptor binding assay, cell membranes (8 μg/well), 0.03 nM [^{125}I]-MCP-1 (PerkinElmer) and appropriate concentrations of the test compounds were added to 0.25 ml of 25 mM HEPES (pH 7.2) buffer containing 2 mM $CaCl_2$, 1 mM $MgCl_2$ and 0.2% BSA. The mixture was incubated for 60 min at 27 °C, and the reaction was terminated by rapid filtration using a cell harvester (Inotech, Switzerland) through Filtermat A GF/C glass fiber filter presoaked in 0.3% polyethylenimine. The filter was covered with Melti-Lex, sealed in a sample bag followed by drying in a microwave oven, and bound radioactivity was determined by scintillation counting using MicroBeta Plus (Wallac, Turku,

Table 2 Consensus scoring results by rank-by-rank. Among the 150 binding site models, the top 10 models sorted by rank-by-rank are shown. Mode 1, 3, 7, 8, 9, and 10 models originated from homology model A, and the others from B. Mode 1 ranked the best and was selected to the best binding model

Binding site model	Rank by each scoring function										Rank-by-rank
	Lig score1	Lig score2	PLP1	PLP2	Jain	PMF01	PMF04	LUDI1	LUDI2	LUDI3	
Mode 1	10	6	31	9	14	4	9	1	1	3	8.8
Mode 2	3	2	20	17	6	27	11	13	9	18	12.6
Mode 3	13	1	34	14	21	20	23	9	12	6	15.3
Mode 4	9	3	6	4	13	51	55	14	11	2	16.8
Mode 5	5	13	1	1	3	94	69	4	2	25	21.7
Mode 6	16	17	4	3	2	69	67	10	8	31	22.7
Mode 7	8	4	33	42	30	21	15	23	29	28	23.3
Mode 8	1	5	9	12	16	59	57	29	34	27	24.9
Mode 9	17	16	54	49	10	11	4	34	42	15	25.2
Mode 10	25	21	48	47	12	9	1	32	44	14	25.3

Table 3 Data from the Ramachandran plot of CCR2 docked with compound 7

Model	MFR[a]		AAR[b]		GAR[c]		Subtotal[d]	Number of Gly and Pro	Total[e]
	n	%	n	%	n	%			
Docked pose	221	86.0	31	12.1	5	1.9	257	26	283

[a] The most favored regions

[b] The additional allowed regions

[c] The generously allowed regions

[d] The total number of residues evaluated except glycines and prolines

[e] The total number of amino acids in the receptor models

Finland). Nonspecific binding was determined in the presence of 0.05 μM human MCP-1. Competition binding studies were carried out with 10 μM test compounds run in duplicate tubes.

Results and discussion

Homology modeling of CCR2

Homology modeling of the CCR2 receptor based on the structure of β_2AR at 2.4 Å (PDB code 2RH1) includes the following steps: 1) automatic sequence alignment; 2) manual modification of the alignment; 3) generation of homology models using the ligand-bound template; and 4) relaxation of strain by the energy minimization.

The automatic sequence alignment for the entire CCR2 sequences was performed against two GPCR sequences, bovine rhodopsin and β_2AR. Then, manual modification was performed to ensure no gaps occurred in the TM helices of CCR2, in comparison to the β_2AR structure. The positions of all conserved amino acids found in most GPCRs were maintained according to the β_2AR sequence. The final alignment is shown in Fig. 3.

Twenty homology models of CCR2 were initially built based on the structure of the β_2AR-carazolol complex using MODELLER, step 3. Three homology models in the lowest probability density functions (PDF) energy were selected for the next stage, step 4. After removing the carazolol ligand from the three models, energy minimization with a fixed backbone coordinate was performed to relieve the structural strain of side chains that occurs due to the presence of non-conserved residues during homology modeling. Although these three models had similar backbone structures within their respective seven TM regions, the orientations of side chains are varied (Fig. 4). These variations provided the flexibility of the models in the early stages of generating the docking pose pool. Analysis of the models using PROCHECK is shown in Table 1. All three models are acceptable for use in docking studies.

Table 4 Fold change in binding IC_{50} of known CCR2 antagonists for Mutant CCR2, relative to wild type CCR2. These data are come from the references, 36 and 42

Mutant	Fold change in IC_{50}	
	Teijin lead	TAK-799
Y49F	0.9	4
Y120A	5.4	15
H121A	7.6	2.6
D284A	1.9	1.8
Q288A	0.9	3.3
E291Q	Inactive	30
T292A	22.9	11

Generation of a pool of ligand-binding poses through ligand docking

Experimental evidence of ligand-receptor interactions are obtained usually from the results of site-directed mutagen-

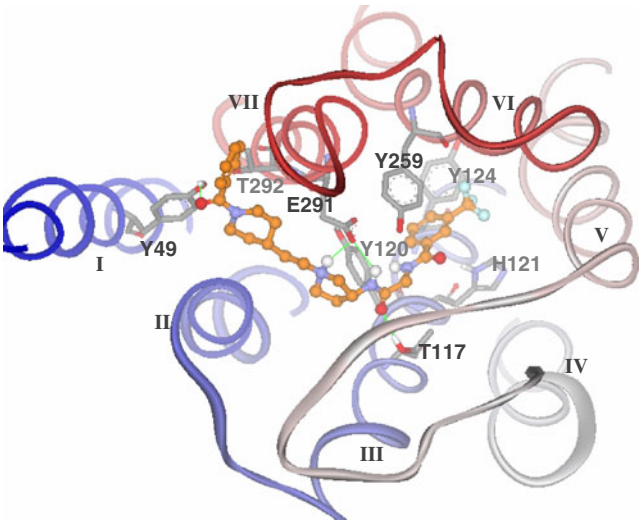


Fig. 5 Protein-ligand interactions in the binding site model. Several interacting residues previously shown by site-directed mutagenesis to be important for ligand interactions [42] are displayed. Hydrogen bonds are indicated by green lines

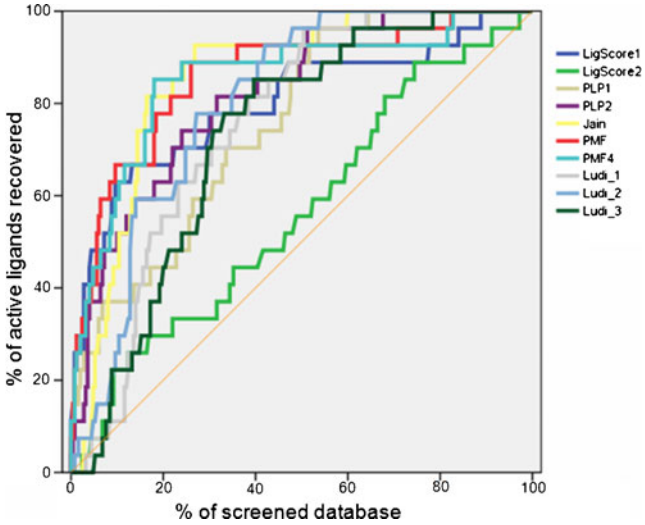


Fig. 6 Fraction of known CCR2 antagonists recovered as a function of the ranked database. Colored bold lines indicate the results from each scoring function and thin line indicates random

esis studies. Several residues in the upper helices region of CCR2 have been shown to participate in the binding of small molecule ligands [36, 37]. Among those residues, the interaction between the acidic side chain Glu291$^{7.39}$(residues are numbered according to the Ballesteros-Weinstein nomenclature [38]) in TM helix VII and the basic amine of the ligand has been considered critical in CCR2, as well as in other chemokine receptor family members [39]. Consequently, we tried to track and maintain this interaction during the docking calculation process. Therefore, the binding site sphere for LibDock docking calculation was defined to have the center on the acidic oxygen in Glu291. During docking, the flexible conformations of the ligand were generated and applied. The diverse conformers of compound 7 generated by the Catalyst program were docked into each of three CCR2 models. The top 50 ranked conformers, which were evaluated and ranked by their docking score in LibDock, were saved. Thus, an initial pool of 150 binding site models was obtained from the docking between the flexible ligands and the rigid protein.

To refine the docking models, the flexibility of the protein should be considered. Protein flexibility was generated through energy minimization. All atoms in both ligand and protein within a sphere of 15Å radius from the center of the docked ligand were included in the energy minimization procedure. Optimization of the complexes through energy minimization resulted in the final pool of 150 ligand-receptor complexes.

Selection of the binding site model

Ligand incorporation into the binding site model may reduce the uncertainty in side chain orientations. However, uncertainty in the models remained because they are based upon a homology model. Thus, inferring the final binding site model from a binding pose pool is challenging. We employed the consensus scoring method to infer the best model. The consensus scoring method is generally employed for virtual screening to improve performance by compensating for the deficiencies of each of the scoring functions [40, 41]. We thought models selected by consensus scoring compensate for bias that comes from the exact calculation technique. Further, this method is able to accurately describe the interactions between the GPCR and the antagonist, and provides a high predictability value for structure based screening.

There are several strategies for consensus scoring, including rank-by-rank, rank-by-number, and rank-by-vote of multiple scoring functions [33]. We used the rank-by-rank strategy. This strategy uses relative ranks rather than absolute binding affinities for ranking binding models. Ten kinds of scoring functions were used to evaluate all 150 binding models in the pool. Table 2 shows the top 10 ranked models using the rank-by-rank strategy. Binding mode 1, derived from the initial model A, scored best and was selected to be the binding site model. The binding site model was analyzed by PROCHECK. Table 3 shows that the final binding site model was improved in the number of the residues in the most favored regions (MFR) compared to the initial model in Table 1. This means the energy minimization refinement stabilized the secondary structure of the receptor.

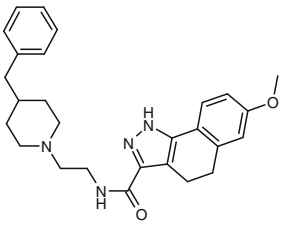
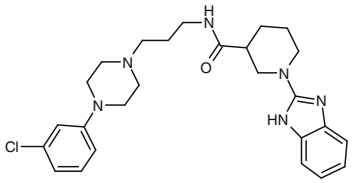
Validation of the binding site model

Table 4 shows interacting residues obtained from the mutagenesis studies using two antagonists in Fig. 1 [36, 42]. The interaction with Glu291 is inevitable and the interactions with Tyr120, His121, and Thr292 are significant for the binding of antagonists. The interaction with Tyr49 or Gln288 seems to depend on the type of antagonists. Figure 5 shows that binding mode 1 well described the interaction between the ammonium moiety of

Table 5 Enrichment factors (EF) for virtual screening of the validation set. EF values for each scoring function were calculated at 10% of the ranked database

	Ligscore1	Ligscore2	PLP1	PLP2	Jain	PMF	PMF04	Ludi1	Ludi2	Ludi3
$EF^{10\%}$	4.8	1.8	3.7	4.0	3.3	5.1	4.4	1.1	1.8	2.2

Table 6 Hit compounds and percent inhibition of CCR2 binding

Compound ID	Structure	% inhibition @50μM
G365-0350		54
G796-2120		27

the ligand and the Glu291 carboxylate in the helix VII. The docked compound 7 is stabilized through hydrogen bonds with Glu291, Tyr49, and Thr117. Tyr120 supports the interaction between the ligand and Glu291 through a hydrogen bond with Glu291. His121 interacts with trifluoromethyl-substituted phenyl ring of compound 7 and comprise a hydrophobic pocket with Tyr124 and Tyr259. Thr292 interacts with both piperidinyl and phenyl rings in compound 7.

Virtual screening with a decoy database and an active database was performed to evaluate the derived binding model. The active 27 compounds used were our in-house synthesized CCR2 antagonists [11] and have a binding affinity IC_{50} below 5 μM. The putatively inactive compounds that comprise a decoy database were selected from the CMC database, and had molecular weight ranging from 400 to 550. The range was designed to have molecular weights similar to the in-house CCR2 antagonists. The CMC database includes the structures of each of the marketed drugs. These drugs were developed using targets different from CCR2 and can therefore be considered putatively inactive compounds with respect to CCR2 antagonism.

Virtual screening was performed using the LigandFit docking algorithm. Ten scoring functions were employed for the evaluation. The fraction of the active ligands recovered as a function of the ranked database is shown in Fig 6. The ranked database contained only 250 molecules, which resulted in docking poses, from a total of 1206 molecules in the decoy database. The enrichment factor analysis based on the ranked database is shown in Table 5. Of the 10 scoring functions, the binding site model showed better enrichment for LigScore1, PLP1, PLP2, PMF, and PMF04. Thus, if the whole decoy database is considered, the ability of the model to discriminate binders and nonbinders can be improved. The validation results showed that the selected model possess good predictability and is suitable for structure based molecule screening.

The binding model selected by consensus scoring showed consistency with interactions inferred from experimental data and predictability toward screening for active compounds. This model is not the best scored model for each scoring function (Table 2). However, the averaged result from 10 scoring functions produced good model and implied an alternative strategy for GPCR homology model without dynamics calculation.

Virtual screening of a commercial database

LigandFit docking screening of the ChemDiv database, which consists of 23,066 compounds, was performed. The top 10 ranked compounds from each scoring function were examined by visual inspection for a rational docking pose and commercial availability. Fourteen of the commercially available compounds were tested in a biological screen for the ability to inhibit CCR2 binding. The two compounds exhibited >25% inhibition at a concentration of 50 μM

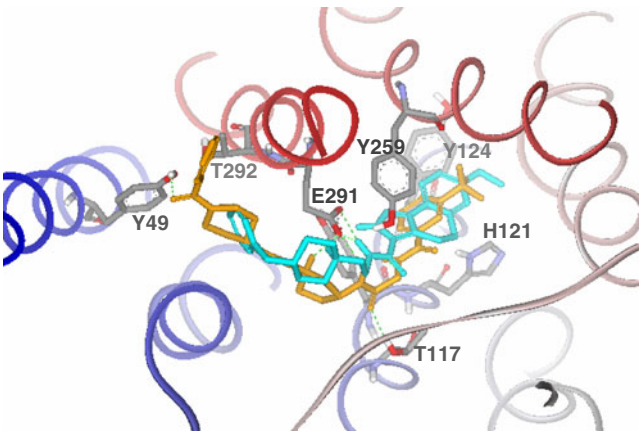


Fig. 7 Comparison of docked poses between compound 7 (orange) and hit compound G365-0350 (cyan). Hydrogen bonds are indicated by green dashed lines

(Table 6). Although the IC_{50} value was not measured, G365-0350, which showed 54% inhibition at 50 μM, was identified as a hit. The comparison of the docking poses between compound 7 and G365-0350 are illustrated in Fig. 7. While G365-0350 has hydrogen bonds only with Glu291, compound 7 has additional hydrogen bonds with Tyr49 and Thr117, and G365-0350 does not fully occupy the binding site pocket modeled for compound 7.

Conclusions

Binding site models of CCR2 based on the proprietary compound 7 were built by ligand supported homology modeling using the atomic structure of the human β_2 adrenergic receptor. A pool of diverse binding poses was initially constructed by allowing the flexible conformations of the ligand. Then, the binding poses in the pool were refined through the energy minimization, in which the flexibility of the protein was taken into consideration. Consensus scoring of docking evaluation functions was used to identify the best binding site model from the pool. Considering the apparent flexibility of GPCRs and the diversity in ligand types, binding pockets, and the binding modes, the quality and applicability of GPCR models strongly depends on its consistency with experimental data, more than on the exact modeling techniques employed or the prospective predictability of the model for structure based virtual screening [22]. Thus, the model presented here fit these criteria, as it fit well with reported site-directed mutagenesis data and the enrichment study using small scale virtual screening with databases consisting of active and inactive compounds showed the model possessed good predictability. The binding model derived from ligand supported homology modeling into a pool of binding modes and selection by consensus scoring has usefulness, although it has the possible limitation to be not the most stabilized model in terms of free energy.

Finally, the model was used to perform a virtual screen for potential CCR2 antagonists, which was followed by using a biological assay to identify lead compound hits. This method entails a powerful filter to screen for CCR2 antagonists from large molecule databases.

References

1. Charo IF, Myers SJ, Herman A, Franci C, Connolly AJ, Coughlin SR (1994) Molecular cloning and functional expression of two monocyte chemoattractant protein 1 receptors reveals alternative splicing of the carboxyl-terminal tails. Proc Nat Acad Sci USA 91:2752–2756
2. Murphy PM, Baggiolini M, Charo IF, Hebert CA, Horuk R, Matsushima K, Miller LH, Oppenheim JJ, Power CA (2000) International union of pharmacology. XXII. Nomenclature for chemokine receptors. Pharmacol Rev 52:145–176
3. Daly C, Rollins BJ (2003) Monocyte chemoattractant protein-1 (CCL2) in inflammatory disease and adaptive immunity: Therapeutic opportunities and controversies. Microcirculation 10:247
4. Szekanecz Z, Szucs G, Szanto S, Koch AE (2006) Chemokines in rheumatic diseases. Curr Drug Targets 7:91–102
5. Charo IF, Ransohoff RM (2006) The many roles of chemokines and chemokine receptors in inflammation. N Engl J Med 354:610–621
6. Libby P (2002) Inflammation in atherosclerosis. Nature 420:868–874
7. Hancock WW, Gao W, Faia KL, Csizmadia V (2000) Chemokines and their receptors in allograft rejection. Curr Opin Immunol 12:511–516
8. Kamei N, Tobe K, Suzuki R, Ohsugi M, Watanabe T, Kubota N, Ohtsuka-Kowatari N, Kumagai K, Sakamoto K, Kobayashi M, Yamauchi T, Ueki K, Oishi Y, Nishimura S, Manabe I, Hashimoto H, Ohnishi Y, Ogata H, Tokuyama K, Tsunoda M, Ide T, Murakami K, Nagai R, Kadowaki T (2006) Overexpression of monocyte chemoattractant protein-1 in adipose tissues causes macrophage recruitment and insulin resistance. J Biol Chem 281:26602–26614
9. Moree WJ, K-i K, Ramirez-Weinhouse MM, Shiota T, Imai M, Tsutsumi T, Sudo M, Endo N, Muroga Y, Hada T, Fanning D, Saunders J, Kato Y, Myers PL, Tarby CM (2008) Potent antagonists of the ccr2b receptor. Part 3: SAR of the 3-aminopyrrolidine series. Bioorg Med Chem Lett 18:1869–1873
10. Moree WJ, K-i K, Ramirez-Weinhouse MM, Shiota T, Imai M, Sudo M, Tsutsumi T, Endo N, Muroga Y, Hada T, Tanaka H, Morita T, Greene J, Barnum D, Saunders J, Kato Y, Myers PL, Tarby CM (2004) Small molecule antagonists of the CCR2b receptor. Part 2: Discovery process and initial structure-activity relationships of diamine derivatives. Bioorg Med Chem Lett 14:5413–5416
11. Lim JW, Oh Y, Kim J-H, Oak M-H, Na Y, Lee J-O, Lee S-W, Cho H, Park W-K, Choi G, Kang J (2010) Synthesis and biological evaluation of 3-aminopyrrolidine derivatives as CC chemokine receptor 2 antagonists. Bioorg Med Chem Lett 20:2099–2102
12. Shacham S, Marantz Y, Bar-Haim S, Kalid O, Warshaviak D, Avisar N, Inbal B, Heifetz A, Fichman M, Topf M, Naor Z, Noiman S, Becker OM (2004) Predict modeling and in-silico screening for G-protein coupled receptors. Proteins 57:51–86
13. Trabanino RJ, Hall SE, Vaidehi N, Floriano WB, Kam VWT, Goddard WA 3rd (2004) First principles predictions of the structure and function of G-protein-coupled receptors: Validation for bovine rhodopsin. Biophys J 86:1904–1921
14. Vaidehi N, Floriano WB, Trabanino R, Hall SE, Freddolino P, Choi EJ, Zamanakos G, Iii WAG (2002) Prediction of structure and function of G protein-coupled receptors. Proc Nat Acad Sci USA 99:12622–12627
15. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, Trong IL, Teller DC, Okada T, Stenkamp RE, Yamamoto M, Miyano M (2000) Crystal structure of rhodopsin: a G protein-coupled receptor. Science 289:739–745
16. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, Choi H-J, Kuhn P, Weis WI, Kobilka BK, Stevens RC (2007) High-resolution crystal structure of an engineered human β_2-adrenergic G protein coupled receptor. Science 318:1258–1265
17. Rasmussen SGF, Choi H-J, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, Burghammer M, Ratnala VRP, Sanishvili R, Fischetti RF, Schertler GFX, Weis WI, Kobilka BK (2007) Crystal structure of the human β_2 adrenergic G-protein-coupled receptor. Nature 450:383–387

18. Evers A, Klebe G (2004) Ligand-supported homology modeling of G-protein-coupled receptor sites: Models sufficient for successful virtual screening13. Angew Chem Int Edit 43:248–251
19. Evers A, Gohlke H, Klebe G (2003) Ligand-supported homology modelling of protein binding-sites using knowledge-based potentials. J Mol Biol 334:327–345
20. Cavasotto CN, Orry AJW, Murgolo NJ, Czarniecki MF, Kocsi SA, Hawes BE, O'Neill KA, Hine H, Burton MS, Voigt JH, Abagyan RA, Bayne ML, Monsma FJ (2008) Discovery of novel chemotypes to a G-protein-coupled receptor through ligand-steered homology modeling and structure-based virtual screening. J Med Chem 51:581–588
21. Dalton JAR, Jackson RM (2010) Homology-modelling protein-ligand interactions: Allowing for ligand-induced conformational change. J Mol Biol 399:645–661
22. Cd G, Rognan D (2009) Customizing G protein-coupled receptor models for structure-based virtual screening. Curr Pharm Des 15:4026–4048
23. Diller DJ, Merz KM (2001) High throughput docking for library design and library prioritization. Proteins 43:113–124
24. Rao SN, Head MS, Kulkarni A, LaLonde JM (2007) Validation studies of the site-directed docking program Libdock. J Chem Inf Model 47:2159–2171
25. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815
26. Laskowski RA, MacArthur MW, Moss D, Thornton JM (1993) PROCHEK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291
27. Smellie A, Teig SL, Towbin P (1995) Poling: Promoting conformational variation. J Comput Chem 16:171–187
28. Krammer A, Kirchhoff PD, Jiang X, Venkatachalam CM, Waldman M (2005) Ligscore: a novel scoring function for predicting binding affinities. J Mol Graph Model 23:395–407
29. Willett P, Barnard JM, Downs GM (1998) Chemical similarity searching. J Chem Inf Comput Sci 38:983–996
30. Muegge I, Martin YC (1999) A general and fast scoring function for protein − ligand interactions: a simplified potential approach. J Med Chem 42:791–804
31. Gehlhaar DK, Verkhivker GM, Rejto PA, Sherman CJ, Fogel DR, Fogel LJ, Freer ST (1995) Molecular recognition of the inhibitor AG-1343 by HIV-1 protease: conformationally flexible docking by evolutionary programming. Chem Biol 2:317–324
32. Böhm HJ (1994) The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. J Comput Aid Mol Des 8:243–256
33. Wang R, Wang S (2001) How does consensus scoring work for virtual library screening? An idealized computer experiment. J Chem Inf Comput Sci 41:1422–1426
34. Venkatachalam CM, Jiang X, Oldfield T, Waldman M (2003) Ligandfit: a novel method for the shape-directed rapid docking of ligands to protein active sites. J Mol Graph Model 21:289–307
35. Pearlman DA, Charifson PS (2001) Improved scoring of ligand−protein interactions using owfeg free energy grids. J Med Chem 44:502–511
36. Berkhout TA, Blaney FE, Bridges AM, Cooper DG, Forbes IT, Gribble AD, Groot PHE, Hardy A, Ife RJ, Kaur R, Moores KE, Shillito H, Willetts J, Witherington J (2003) CCR2: characterization of the antagonist binding site from a combined receptor modeling/mutagenesis approach. J Med Chem 46:4070–4086
37. Hall SE, Mao A, Nicolaidou V, Finelli M, Wise EL, Nedjai B, Kanjanapangka J, Harirchian P, Chen D, Selchau V, Ribeiro S, Schyler S, Pease JE, Horuk R, Vaidehi N (2009) Elucidation of binding sites of dual antagonists in the human chemokine receptors CCR2 and CCR5. Mol Pharmacol 75:1325–1336
38. Ballesteros JA, Weinstein H (1995) Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations of G protein-coupled receptors. Methods Neurosci 25:366–428
39. Rosenkilde MM, Schwartz TW (2006) GluVII:06 - a highly conserved and selective anchor point for non-peptide ligands in chemokine receptors. Curr Top Med Chem 6:1319–1333
40. Yang J-M, Chen Y-F, Shen T-W, Kristal BS, Hsu DF (2005) Consensus scoring criteria for improving enrichment in virtual screening. J Chem Inf Model 45:1134–1146
41. Wang R, Lu Y, Wang S (2003) Comparative evaluation of 11 scoring functions for molecular docking. J Med Chem 46:2287–2303
42. Carter PH, Tebben AJ (2009) The use of receptor homology modeling to facilitate the design of selective chemokine receptor antagonists. In: Handel TH, Hamel DJ (eds) Method in enzymology, vol 461. Academic, San Diego, pp 249–279