http://dx.doi.org/10.5012/bkcs.2012.33.2.717

Prediction of Binding Mode between Chemokine Receptor CCR2 and Its Known Antagonists using Ligand Supported Homology Modeling

Jong-Hoon Kim,[†,‡] Jee Woong Lim,[‡] Seung-Woo Lee,[‡] Kyoungrak Kim,[‡] and Kyoung Tai No[†,*]

[†]*Department of Biotechnology, Yonsei University, Seoul 120-749, Korea.* [*]*E-mail: ktno@yonsei.ac.kr*
[‡]*Research and Development Center, YangJi Chemical, Suwon 443-766, Korea*
Received August 16, 2011, Accepted December 12, 2011

Key Words : GPCR, Ligand supported homology modeling, CCR2, Docking evaluation, Consensus scoring, Binding mode

Chemokines play a crucial role in the trafficking of leukocytes in the body through the binding to their related receptors. Chemokine Receptor 2 (CCR2) belongs to G-Protein Coupled Receptor (GPCR) family and expressed in monocytes, immature dendritic cells, activated T lymphocytes, basophils, and endothelial and vascular smooth-muscle cells. CCR2 binds several chemokines; CCL2, CCL7, CCL8, and CCL13.[1] CCR2 and its ligand have been implicated in the pathophysiology of a number of diseases, including rheumatoid arthritis, multiple sclerosis, atherosclerosis, organ transplant rejection, and insulin resistance.[2] Knowledge of the structural basis on CCR2-ligand interaction could help facilitate the design of novel CCR2 antagonists. We modeled and predicted the binding sites of widely known CCR2 antagonists, TAK-779[3] and Teijin-lead[4] (Fig. 1), using ligand supported homology modeling method.

Homology modeling predicts the three-dimensional structure of a given protein sequence based on its alignment to reference proteins of known three-dimensional structure (so-called templates). It is known as the most successful technique for predicting the three-dimensional protein structure.[5] In conventional homology modeling method, the ligand molecules are not considered in the process of protein model building. After building the protein model, the ligand molecules are docked into the protein model. These approaches often showed the consequence that residue side chains involved in ligand binding are inappropriately modeled. The time-consuming re-modeling of the binding site or docking is required to generate sensible side-chain and ligand orientations. In the case of GPCR, the conventional homology modeling approaches for obtaining the binding site models are more challenging because the number of the available templates is very limited and the sequence identity to the templates is very low. The available templates for GPCR are the solved three-dimensional structures of bovine rhodopsin,[6] β_2-adrenergic receptor (β_2AR),[7] turkey β_1-adrenergic receptor,[8] and human adenosine A_{2A} receptor.[9] The homology modeling from those limited number of templates may result in uncertainty not only regarding backbone positioning but also about side chain conformations. One approach to solve this 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. Another approach is to use dynamics and ligand data as a restraint. This method also has a possibility of poor quality.[10] Thus, the development of modified homology modeling methods for constructing a plausible ligand binding mode is an important theme in GPCR-ligand modeling research.

The incorporation of ligand information from the early stage of homology modeling has been studied as an alternative strategy to the conventional homology modeling. This approach was first developed and termed as *ligand supported homology modeling* by Klebe and coworkers.[11] Recently, the improved and diversified methods have been published as ligand-steered homology modeling[12] or homology-modeling protein-ligand interactions.[13] The results presented reasonable binding site models. We recently reported the protocol and evaluation for the proprietary CCR2 antagonist by ligand-supported homology modeling.[14] We here employed the protocol to predict the binding mode for known CCR2 antagonists, TAK-779 and Teijin-lead.

The method in this study for structure modeling of CCR2 and its antagonists is summarized in Figure 2. In the first place of homology modeling, we retained the bound ligand, carazolol, from the crystal structure of β_2AR. Then, carazolol was replaced by the various conformations of the target

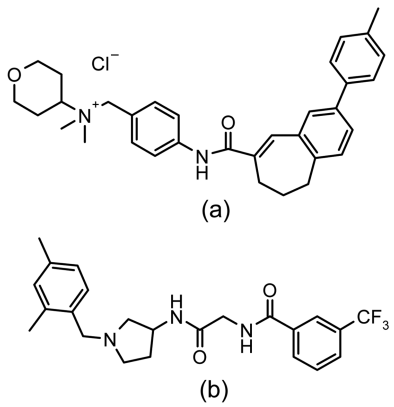

Figure 1. Chemical structures of TAK-779 (a) and Teijin-lead (b). The reported IC_{50} for TAK-779 binding to CCR2 is 27 nM.[3] The reported IC_{50} for Teijin-lead binding to CCR2 is 54 nM.[4]

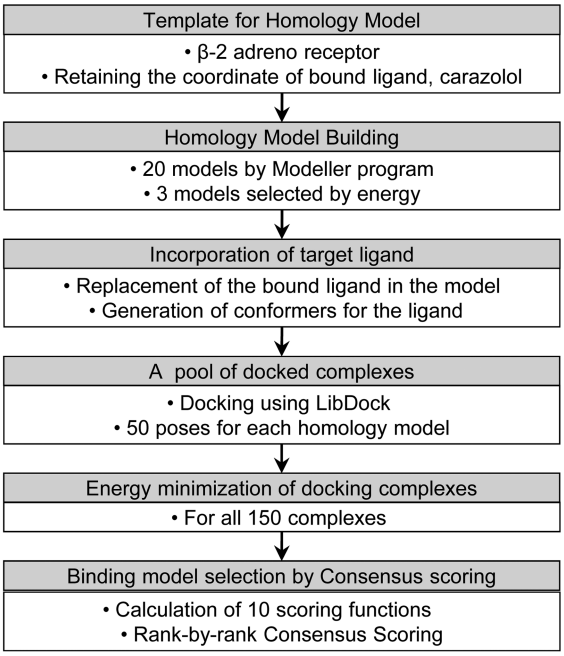

Figure 2. The overall scheme for ligand supported homology modeling and finding the best binding model.

ligand. Docking complexes were optimized by energy minimization and constitute the pool of binding modes. Consensus scoring of various docking scores was employed to identify the best docking mode.

The automatic sequence alignment for the entire CCR2 sequences was performed against two GPCR sequences, bovine rhodopsin (PDB code 1F88)[15] and β_2AR (PDB code 2RH1)[7] using the clustal W program in Discovery Studio (DS) version 2.1 package (Accelrys, San Diego, CA). Then, manual modification was performed to ensure no gaps occurred in the transmembrane (TM) helices of CCR2, in comparison to the β_2AR structure.

The automated homology modeling was carried out using the program MODELLER[16] in the DS software package. The bound ligand in the crystal structure of β_2AR, carazolol, was copied to maintain a binding cavity in the space surrounded by the 7TM helices. This incorporation of the ligand formed ligand supported strategy at the initial stage of modeling. Twenty models were generated and sorted by probability density functions (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. Although these three models had similar backbone structures within their respective 7TM regions, the orientations of side chains are varied. These variations provided the flexibility of the models in the early stages of generating the docking pose pool.

The generation of the conformers for TAK-779 and Teijin-lead was carried out using the BEST algorithm in the Catalyst program (Accelrys). The BEST method uses a poling algorithm[17] to generate a diverse set of low-energy conformations. The maximum number of conformers was set to 255 with a 20 kcal/mol energy window from the lowest energy. The conformers from the ligand were docked into the three homology models using the LibDock program in DS. 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. The selection of the residue Glu291 was grounded from the reported results of site-directed mutagenesis studies.[18,19] Several residues in the upper helices region of CCR2 have been shown to participate in the binding of small molecule ligands. Among those residues, the interaction between the acidic side chain Glu291 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.[20] Consequently, we tried to track and maintain this interaction during the docking calculation process. In LibDock calculation, the number of saving poses was assigned to 50 for each protein model. Since three protein models from homology study were tried, a pool comprising 150 different docking poses for each ligand were obtained.

All 150 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 15 Å radius sphere from the center of the docked conformer. All atoms in the sphere, including atoms from each ligand, were set to engage in the minimization calculation. CHARMm force field and Smart Minimizer algorithm in DS were employed. The dielectric constant was set to 4.0, which was suggested as a well-working protein dielectric constant.[21] A maximum of 1,000 iterations were allowed. This step constituted the pool of possible binding modes for each ligand.

To select the best binding mode from the pool of 150 binding-complexes, the consensus scoring method of rank-by-rank strategy[22] was used. The docking scores were evaluated for the 150 binding complexes using ten scoring functions implemented through the DS program, including LigScore1, LigScore2, PLP1, PLP2, Jain, PMF01, PMF04, Ludi1, Ludi2, and Ludi3. The ranks of the scores for each scoring function were calculated. Then, the average ranks from the ranks of ten scoring functions, the rank-by-rank, were determined for each binding mode. The top ranked CCR2-lignad complex was selected to the binding site model (Table 1). The application of consensus scoring could improve performance by compensating for the deficiencies of each of the scoring functions.[23]

Table 1 shows the rank-by-rank consensus scoring result with the top ranked model for TAK-779 and Teijin-lead. We selected these as the binding model and checked their robustness. We evaluated whether the protein model maintain 7TM helical structure by the analysis with PROCHEK.[24] The protein structures in the selected binding model of TAK-779 and Teijin-lead are acceptable in terms of backbone arrangement.

Considering the apparent flexibility of GPCRs and the

Table 1. Consensus scoring results by Rank-by-Rank. Among the 150 binding site models, the top models for TAK-779 and Teijin-lead are shown

Top ranked model	Rank by each scoring function										Rank-by-Rank
	LigScore1	LigScore2	PLP1	PLP2	Jain	PMF 01	PMF 04	Ludi1	Ludi2	Ludi3	
TAK-779	2	4	53	31	4	7	8	2	1	3	11.5
Teijin-lead	19	5	1	1	5	7	11	14	19	15	9.7

Table 2. Fold change in binding IC_{50} of TAK-779 and Teijin lead for Mutant CCR2, relative to wild type CCR2.[26] The values were calculated from the site-directed mutagenesis data by Berkout *et al.*[18]

Mutant	Fold Change in IC_{50}	
	TAK-779	Teijin-lead
Y49F	4	0.9
Y120A	15	5.4
H121A	2.6	7.6
Q288A	3.3	0.9
E291Q	30	Inactive
T292A	11	22.9

diversity in ligand types, binding pockets, and the binding mode, the quality of GPCR homology models strongly depends on its consistency with experimental data.[25] We evaluated the obtained binding models through the comparison with the reported experimental results regarding CCR2-ligand interaction. Table 2 shows the summarization of interacting residues and their effects obtained from the site-directed mutagenesis and the binding study of TAK-779 and Teijin-lead by other group.[18,26] The most crucial point is that the interaction with Glu291 is inevitable. The predicted binding sites for TAK-779 and Teijin-lead obtained from this study are shown in Figure 3 and Figure 4, respectively. The carboxylate of Glu291 showed hydrogen bonding with amide in both compounds. Tyr120 showed helping this interaction by the interaction with Glu291 and the direct contact with aromatic ring in TAK-779 and Teijin-lead. This stacking interaction is more definite in TAK-779 binding site. Berkhout *et al.* presented the interaction model between the quaternary nitrogen in TAK-779 and the carboxylate in Glu291 using the homology model from rhodopsin template.[18] However, Hall *et al.* reported the model in which the amide hydrogen in TAK-779 interacts with Glu291.[19] In this study, Tyr49 showed the hydrogen bonding with the oxygen in the tetrahydro-pyranyl moiety in TAK-779 binding site model. This interaction was also observed in the model by Hall *et al.* Table 2 shows Tyr49 has significant influence on

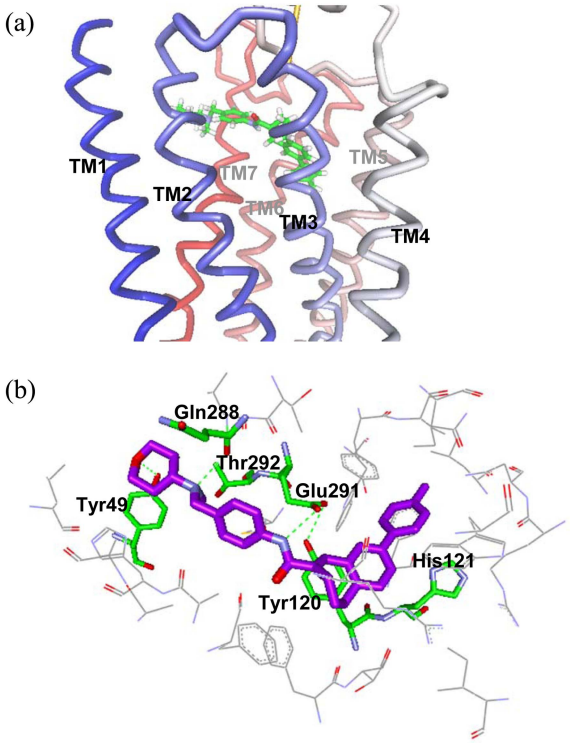

Figure 4. The homology model structure of CCR2 with TAK-779 binding site. (a) Side view of the binding site. TAK-779 is in green. (b) Display of the interacting residues with TAK-779 (purple). The residues in Table 2 are displayed in green. Dashed green lines indicate hydrogen bonding.

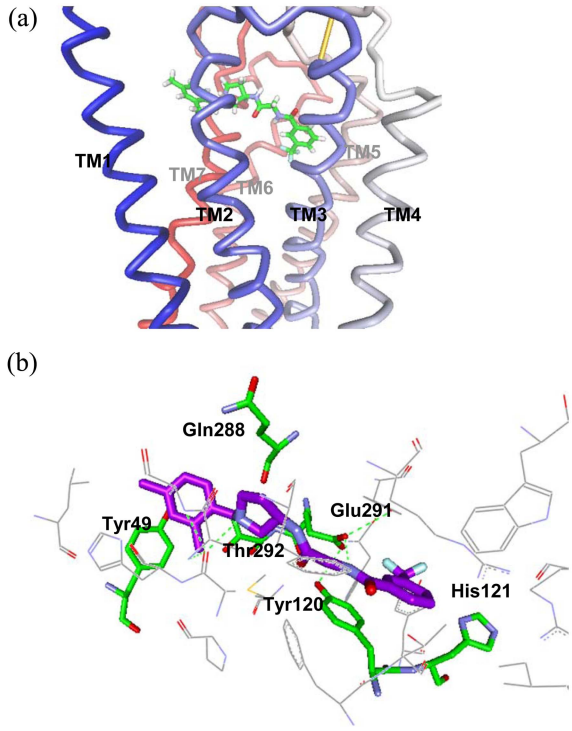

Figure 5. The homology model structure of CCR2 with Teijin-lead binding site. (a) Side view of the binding site. Teijin-lead is in green. (b) Display of the interacting residues with Teijin-lead (purple). The residues in Table 2 are displayed in green. Dashed green lines indicate hydrogen bonding.

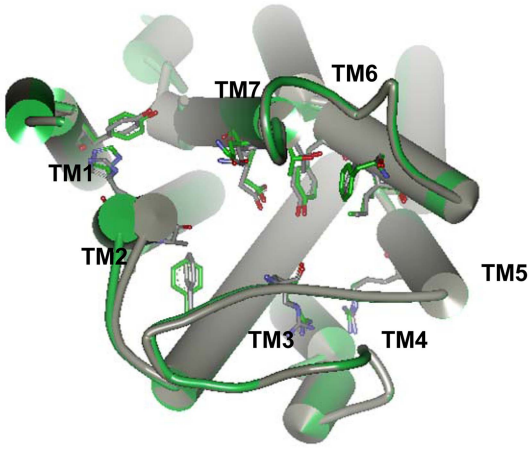

Figure 6. The comparison of CCR2 in the binding models. Teijin-lead binding model is displayed in gray and TAK-779 binding model is in green.

TAK-779 binding. The model from this study, having hydrogen bonding between amide in the ligand and Glu291 and Tyr49 interaction, provides plausible elucidation on these experimental observations. Gln288 showed closer contact in TAK-779 binding site than in Teijin-lead binding site, while His121 closely contacted with aromatic ring in Teijin-lead binding site. Thr292 showed close contact with both compounds. Although high fold change in T292A mutant is not clearly explained by our binding site models, the interactions by Gln288, His121, and Thr292, also supported the experimental data in Table 2. Figure 5 shows the comparison of CCR2 models between Teijin-lead binding and TAK-779 binding. TM2 in TAK-779 binding model is slightly shifted and the several side chains show a few differences in the conformation. However, the overall structure of CCR2 has consistency in the binding of both compounds.

In summary, ligand-supported homology models of chemokine receptor CCR2 were built based on the crystal structure of human β_2-adrenergic receptor and the binding site model of the known antagonists, TAK-779 and Teijin-lead, were constructed by consensus scoring strategy from the binding sites pool. The binding pockets for these ligands were characterized and compared with the reported site-directed mutagenesis studies. Our modeling studies indicate that the binding pockets for CCR2 antagonists are located in the upper transmembrane (TM) part near Glu291. The binding site models are consistent with experimental data, thus could be beneficial to novel ligand design.

References

1. Murphy, P. M.; Baggiolini, M.; Charo, I. F.; Hbert, C. A.; Horuk, R.; Matsushima, K.; Miller, L. H.; Oppenheim, J. J.; Power, C. A. *Pharmacol. Rev.* **2000**, *52*, 145.
2. Viola, A.; Luster, A. D. *Annu. Rev. Pharmacol. Toxicol.* **2008**, *48*, 171.
3. Baba, M.; Nishimura, O.; Kanzaki, N.; Okamoto, M.; Sawada, H.; Iizawa, Y.; Shiraishi, M.; Aramaki, Y.; Okonogi, K.; Ogawa, Y.; Meguro, K.; Fujino, M. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 5698.
4. Moree, W. J.; Kataoka, K.-i.; Ramirez-Weinhouse, M. M.; Shiota, T.; Imai, M.; Tsutsumi, T.; Sudo, M.; Endo, N.; Muroga, Y.; Hada, T.; Fanning, D.; Saunders, J.; Kato, Y.; Myers, P. L.; Tarby, C. M. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 1869.
5. Moult, J.; Fidelis, K.; Kryshtafovych, A.; Rost, B.; Hubbard, T.; Tramontano, A. *Proteins* **2007**, *69*, 3.
6. 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.; Yamamoto, M.; Miyano, M. *Science* **2000**, *289*, 739.
7. Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G. F.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K.; Stevens, R. C. *Science* **2007**, *318*, 1258.
8. Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G. W.; Tate, C. G.; Schertler, G. F. X. *Nature* **2008**, *454*, 486.
9. Jaakola, V.-P.; Griffith, M. T.; Hanson, M. A.; Cherezov, V.; Chien, E. Y. T.; Lane, J. R.; Ijzerman, A. P.; Stevens, R. C. *Science* **2008**, *322*, 1211.
10. Todd, A. E.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. *Abstr. Pap. Am. Chem. Soc.* **2005**, *230*, U2898.
11. Evers, A.; Gohlke, H.; Klebe, G. *J. Mol. Biol.* **2003**, *334*, 327.
12. Cavasotto, C. N.; Orry, A. J. W.; Murgolo, N. J.; Czarniecki, M. F.; Kocsi, S. A.; Hawes, B. E.; O'Neill, K. A.; Hine, H.; Burton, M. S.; Voigt, J. H.; Abagyan, R. A.; Bayne, M. L.; Monsma, F. J. *J. Med. Chem.* **2008**, *51*, 581.
13. Dalton, J. A. R.; Jackson, R. M. *J. Mol. Biol.* **2010**, *399*, 645.
14. Kim, J.-H.; Lim, J.; Lee, S.-W.; Kim, K.; No, K. *J. Mol. Model.* **2011**, *17*, 2707.
15. Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. *Science* **2000**, *289*, 739.
16. Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779.
17. Smellie, A.; Teig, S. L.; Towbin, P. *J. Comput. Chem.* **1995**, *16*, 171.
18. Berkhout, T. A.; Blaney, F. E.; Bridges, A. M.; Cooper, D. G.; Forbes, I. T.; Gribble, A. D.; Groot, P. H. E.; Hardy, A.; Ife, R. J.; Kaur, R.; Moores, K. E.; Shillito, H.; Willetts, J.; Witherington, J. *J. Med. Chem.* **2003**, *46*, 4070.
19. Hall, S. E.; Mao, A.; Nicolaidou, V.; Finelli, M.; Wise, E. L.; Nedjai, B.; Kanjanapangka, J.; Harirchian, P.; Chen, D.; Selchau, V.; Ribeiro, S.; Schyler, S.; Pease, J. E.; Horuk, R.; Vaidehi, N. *Mol. Pharmacol.* **2009**, *75*, 1325.
20. Rosenkilde, M. M.; Schwartz, T. W. *Curr. Top. Med. Chem.* **2006**, *6*, 1319.
21. Schutz, C. N.; Warshel, A. *Proteins* **2001**, *44*, 400.
22. Wang, R.; Wang, S. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1422.
23. Yang, J.-M.; Chen, Y.-F.; Shen, T.-W.; Kristal, B. S.; Hsu, D. F. *J. Chem. Inf. Model.* **2005**, *45*, 1134.
24. Laskowski, R. A.; MacArthur, M. W.; Moss, D.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283.
25. Graaf, C. d.; Rognan, D. *Curr. Pharm. Des.* **2009**, *15*, 4026.
26. Carter, P. H.; Tebben, A. J. In *Methods Enzymol.*; Handel, T. H., Hamel, D. J., Eds., 2009; Vol. 461, p 249