

Molecular docking and 3D QSAR studies on 1-amino-2-phenyl-4-(piperidin-1-yl)-butanes based on the structural modeling of human CCR5 receptor

Yong Xu, Hong Liu,* Chunying Niu, Cheng Luo, Xiaomin Luo, Jianhua Shen, Kaixian Chen and Hualiang Jiang*

Drug Discovery and Design Center, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 201203, China

Received 22 July 2004; revised 31 August 2004; accepted 31 August 2004
Available online 18 September 2004

Abstract—In the present study, we have used an approach combining protein structure modeling, molecular dynamics (MD) simulation, automated docking, and 3D QSAR analyses to investigate the detailed interactions of CCR5 with their antagonists. Homology modeling and MD simulation were used to build the 3D model of CCR5 receptor based on the high-resolution X-ray structure of bovine rhodopsin. A series of 64 CCR5 antagonists, 1-amino-2-phenyl-4-(piperidin-1-yl)-butanes, were docked into the putative binding site of the 3D model of CCR5 using the docking method, and the probable interaction model between CCR5 and the antagonists were obtained. The predicted binding affinities of the antagonists to CCR5 correlate well with the antagonist activities, and the interaction model could be used to explain many mutagenesis results. All these indicate that the 3D model of antagonist–CCR5 interaction is reliable. Based on the binding conformations and their alignment inside the binding pocket of CCR5, three-dimensional structure–activity relationship (3D QSAR) analyses were performed on these antagonists using comparative molecular field analysis (CoMFA) and comparative molecular similarity analysis (CoMSIA) methods. Both CoMFA and CoMSIA provide statistically valid models with good correlation and predictive power. The $q^2(r_{\text{cross}}^2)$ values are 0.568 and 0.587 for CoMFA and CoMSIA, respectively. The predictive ability of these models was validated by six compounds that were not included in the training set. Mapping these models back to the topology of the active site of CCR5 leads to a better understanding of antagonist–CCR5 interaction. These results suggest that the 3D model of CCR5 can be used in structure-based drug design and the 3D QSAR models provide clear guidelines and accurate activity predictions for novel antagonist design.

© 2004 Elsevier Ltd. All rights reserved.

1. Introduction

Human immunodeficiency virus type 1 (HIV-1) infection, which eventually leads to the acquired immunodeficiency syndrome (AIDS), remains to be a lethal disease threatening human's health.¹ To inhibit HIV-1 replication and slow the destruction of the immune system, current therapies utilize a combination of protease and reverse transcriptase inhibitors.² While suppression of viral replication through combination therapy delays progression to AIDS, the virus is not eradicated and the immune system eventually succumbs to infection.^{3–5}

Concerns about the long side effects of protease inhibitors and the increasing transmission of resistant variants emphasize the need to identify new classes of drugs, which are able to suppress HIV-1 replication efficiently.

The recently discovered chemokine receptor CCR5, a primary co-receptor essential for HIV recognition and entry into cell, has been identified as a potential new target for the treatment of HIV-1 infection.⁶ The CCR5 receptor belongs to the rhodopsin family of G-protein-coupled receptors (GPCR) characterized by seven trans-membrane motif.⁷ Chemokines MIP-1 α , MIP-1 β , and RANTES are the natural ligands of CCR5.⁸ Individuals homozygous or heterozygous for a 32-base pair deletion in the gene for CCR5 exhibit either resistance to HIV-1 infection or delayed disease progression.^{9,10} These observations provided remarkable evidence that functional inhibition of CCR5 receptor can be highly

Keywords: CCR5; Homology modeling; Docking; 3D QSAR; CoMFA; CoMSIA.

*Corresponding authors. Tel.: +86 21 50807188; fax: +86 21 50807088; e-mail: hliang@mail.shcnc.ac.cn

protective against HIV-1 infection. Therefore, great efforts have been initiated to identify suitable CCR5 antagonist as potential anti-HIV-1 therapeutic agents.

To date, many chemically diverse small molecule CCR5 antagonists have been reported through patents and scientific literatures.^{11–25} For examples, TAK-779 was the first CCR5 antagonist developed in Takeda Chemical Industries in Japan.^{24,25} Merck Research Laboratories reported a series of novel CCR5 antagonists as anti-HIV-1 agents derived from 1-amino-2-phenyl-4-(piperidin-1-yl)-butane.^{19,20,26,27} Correlating the physico-chemical properties or structural features of compounds with their biological activities is believed to gain an insight into the interaction mechanism of CCR5 receptor to antagonists, providing useful clues for designing new anti-HIV-1 drugs. On the basis of rough alignments of a series of CCR5 antagonists using atomic RMS (root mean square) fit and rigid body field fit techniques, preliminary 3D QSAR studies on the putative active conformation were performed by Song et al.²⁸ The 3D QSAR study based on simple alignment did not give detailed information of how antagonists interact with CCR5 and what interaction features between the receptor and antagonists are important for biological activity. A molecular modeling-guided mutagenesis study²⁹ gave a relatively clear image of CCR5–antagonists interactions; this study, however, did not revealed a precise interaction feature since the interaction model of receptor–ligand was derived from a manual docking based on a pharmacophore. However, structural information of GPCR is difficult to obtain using experimental techniques because of enormous difficulties in the preparation of samples suitable for subsequent X-ray or NMR determinations. Nevertheless, site-directed mutagenesis and other binding experiments have still given us information important for its interaction with antagonists. These available experimental data can be used to guide the modeling of CCR5. In the present study, we used an approach combining homology modeling, molecular dynamics simulation, and automated docking methods to investigate the antagonist–CCR5 interaction mechanism and the structure–activity relationship of the antagonists.

Firstly, taking the crystal structure of bovine rhodopsin (PDB code of 1F88 at 2.80 Å resolution)³⁰ as a template, the homology modeling and molecular dynamics simulation techniques were employed to build three-dimensional (3D) model of the human CCR5 receptor. Afterwards, a series of 64 1-amino-2-phenyl-4-(piperidin-1-yl)-butane analogues were docked into the putative binding site of the 3D model of CCR5. The binding conformations of these antagonists and their alignment in the active site of the receptor were used to construct 3D QSAR models, which can be further applied in activity prediction at a faster speed.^{31,32} The 3D QSAR models were constructed by using approaches of comparative molecular field analysis (CoMFA)³³ and comparative molecular similarity indices analysis (CoMSIA).³⁴ The aims of the present research are (i) demonstrate the common binding model of 1-amino-2-phenyl-4-(piperidin-1-yl)-butane compounds with

CCR5; (ii) elucidate the structural features associate with the chemical modifications and explain the SAR data for CCR5 antagonists; and (iii) obtain stable and predictive QSAR models involving the main intermolecular interactions between antagonists and CCR5, which can be used in rapidly and accurately predicting the activities of new designed antagonists.

2. Computational details

2.1. Molecular structures and optimization

A series of 64 molecules were employed in the present study, their chemical structures are shown in Table 1.^{17–20} The initial structures of these compounds were built on the basis of the structure of compound **2** and energetically minimized using Tripos force field with Gasteiger–Hückel charges.³⁵ The N-protonated forms of the molecules, which are the prevalent species at physiological pH, were used in the calculations. All of the calculations were performed on a SGI O2 workstation, and a SGI 3200 server using the Sybyl 6.8 program.³⁶

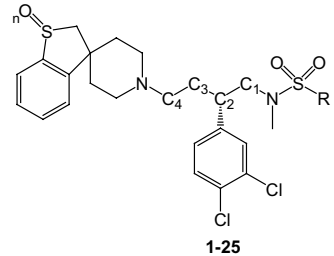
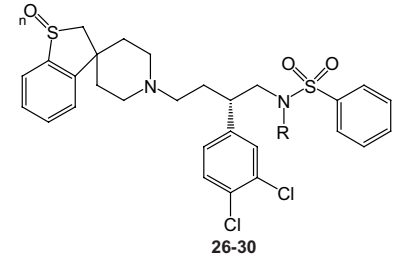
2.2. Structural modeling of CCR5

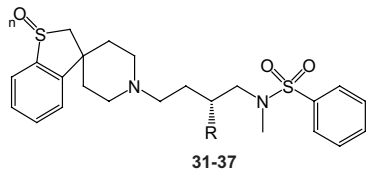
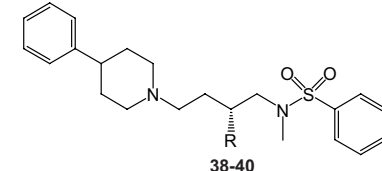
2.2.1. Modeling α -helix bundle. The sequence of the 352 amino acids of human CCR5 receptor was taken from the SwissProt database (entry P51681; <http://us.expasy.org/sprot/>). Sequence analysis and conserved-residues identification were carried out among chemokine receptors and other rhodopsin-like GPCRs. Using the crystal structure of bovine rhodopsin at 2.80 Å resolution (PDB entry 1F88)³⁰ as a template, the Homology module of InsightII³⁷ and the ClustalW algorithm³⁸ were applied in sequence alignment, and the Blosum scoring matrix³⁹ was employed to obtain the best-fit alignment. The best alignment was selected according to not only the value of the alignment score, but also the reciprocal positions of conserved residues. In this way, the trans-membrane (TM) segments were identified and transformed into α -helices.

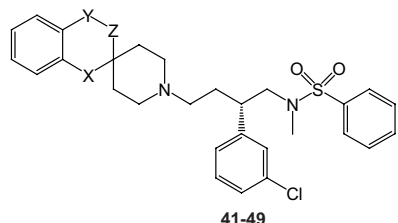
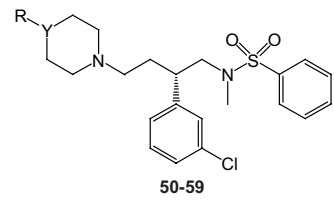
2.2.2. Modeling extracellular and intracellular regions. The FASTA program⁴⁰ was used to identify sequence homologies through the in-house database⁴¹ containing 700 loops and proteins with medium to high sequence identity. ClustalW³⁸ was then used to determine the fragments that had higher homology with the extracellular and intracellular regions of CCR5. The reasonable fragment conformation was chosen from the top 10 candidates that had the lowest root mean square (RMS) values and considerable geometrical compatibility. The conserved disulfide bond between residues Cys101 at the beginning of TM3 and Cys178 in the middle of extracellular loop 2 (EL2) was also created and kept as a constraint in the geometric optimization.

The resultant structure of CCR5 was optimized using molecular mechanics method with the following parameters: a distance-dependent dielectric constant of 5.0; nonbonded cutoff 8 Å, Amber force field and Kollman-all-atom charges; and conjugate gradient minimization

Table 1. Chemical structures of 1-amino-2-phenyl-4-(piperidin-1-yl)-butane analogues and their antagonist activities^a

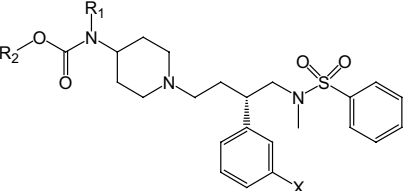
 <p>1-25</p>				 <p>26-30</p>			
No	<i>n</i>	R	pIC ₅₀ ^b	No	<i>n</i>	R	pIC ₅₀ ^b
1	0	Phenyl	6.00	16*	2	4-NO ₂ -Phenyl	7.22
2	1	Phenyl	7.46	17	1	4-MeO-Phenyl	7.40
3	2	Phenyl	7.00	18	1	4-Ph-Phenyl	7.40
4	1	2-Thienyl	7.22	19	1	Naphth-1-yl	6.44
5	2	2-Thienyl	6.92	20	1	Naphth-2-yl	7.22
6*	1	NMe ₂	6.47	21	1	Indan-5-yl	7.16
7	1	Benzyl	6.23	22	1	Pyridin-3-yl	7.00
8	1	Methyl	6.07	23	1	Quinolin-8-yl	7.05
9	1	<i>n</i> -Octyl	5.85	24	1	Quinolin-3-yl	6.92
10	1	Cyclopentyl	7.00	25	1	1-Me-Imidazol-4-yl	6.47
11	1	Cyclohexyl	7.00	26*	0	H	5.80
12	1	2-Cl-Phenyl	7.10	27	0	Et	5.00
13	1	3-Cl-Phenyl	7.16	28	1	H	6.64
14	1	4-Cl-Phenyl	7.40	29	1	Et	6.89
15	2	3-NO ₂ -Phenyl	6.82	30	2	Et	6.80

 <p>31-37</p>				 <p>38-40</p>			
31	0	(<i>S</i>)-3-Cl-Phenyl	6.57	36*	1	(<i>S</i>)-3,4-OCH ₂ O-Phenyl	6.70
32	1	(<i>S</i>)-3-Cl-Phenyl	8.00	37	2	(<i>S</i>)-3,4-OCH ₂ O-Phenyl	7.10
33	2	(<i>S</i>)-3-Cl-Phenyl	7.82	38		(<i>S</i>)-3-Cl-Phenyl	7.52
34	1	(<i>S</i>)-4-Cl-Phenyl	6.57	39		(<i>S</i>)-4-F-Phenyl	6.00
35	1	(<i>S</i>)-4-F-Phenyl	6.24	40		(<i>S</i>)-3,4-OCH ₂ O-Phenyl	6.52

 <p>41-49</p>				 <p>50-59</p>			
X	Y-Z			Y	R		
41	-CH ₂ CH ₂ -	6.75	50	-N-	C ₆ H ₅	6.16	
42	-NHCH ₂ -	7.30	51	-N-	2-MeC ₆ H ₄	5.62	
43	-C(O)CH ₂ -	8.30	52	-CH-	2-MeC ₆ H ₄	6.40	
44	-C(O)NH-	7.35	53	-CH-	2-MeOC ₆ H ₄	7.16	
45	-C(O)N(Me)-	7.00	54	-CH-	3-CF ₃ C ₆ H ₄	6.92	
46*	-C(O)NHCH ₂ -	7.46	55	-CH-	4-ClC ₆ H ₄	6.70	
47	-N(MeSO ₂)CH ₂ -	7.46	56*	-CH-	4-FC ₆ H ₄	7.60	
48	-CH(OH)CH ₂ -	7.00	57	-CH-	C ₆ H ₅ CH ₂	6.60	
49	-O-	6.58	58	-CH-	C ₆ H ₅ CH ₂ CH ₂	7.19	
			59	-CH-	C ₆ H ₅ CH ₂ CH ₂ CH ₂	8.30	

(continued on next page)

Table 1 (continued)

							
60-64							
	<u>R₁</u>	<u>R₂</u>			<u>R₁</u>	<u>R₂</u>	
60	<i>n</i> -Pr	4-NO ₂ -C ₆ H ₄ -CH ₂	8.82	63	<i>n</i> -Pr	4-NH ₂ COC ₆ H ₄ -CH ₂	8.70
61	Allyl	4-NO ₂ -C ₆ H ₄ -CH ₂	8.70	64	<i>n</i> -Pr	4-NH ₂ COC ₆ H ₄ -CH ₂	8.52
62	<i>n</i> -Pr	3-NH ₂ COC ₆ H ₄ -CH ₂	9.10				

The kind of atomic charges were assigned as Kollman-all-atom for CCR5 and Gasteiger–Marsili³⁵ for the ligands.

Finally, the docked complexes of ligand–receptor were selected according to the criteria of interacting energy combined with geometrical matching quality. These complexes were used as the starting conformation for further energetic minimization and geometrical optimization before the final binding models were achieved.

2.4. Binding free energy prediction

Typically, three binding energy terms used in the previous versions of AutoDock⁵⁰ were included in the score function; the van der Waals interaction represented as a Lennard-Jones 12–6 dispersion/repulsion term, the hydrogen bonding represented as a directional 12–10 term, and the Coulombic electrostatic potential. Therefore, the binding energy of 1-amino-2-phenyl-4-(piperidin-1-yl)-butane analogues to CCR5 could be simply described as the electrostatic, van der Waals and hydrogen-bonding interaction energy, respectively. On the basis of the traditional molecular force field model of interaction energy, a new score function at the level of binding free energy was derived and adopted in the version of AutoDock 3.0.3.^{47,48} Not only the restriction of internal rotors, the global rotation, and the translation were modeled depending on the number of torsion angles of the ligand but also the desolvation upon binding and the hydrophobic effect (solvent entropy changes at solute–solvent interfaces) were calculated. The total binding free energy was empirically calibrated based on the above-stated terms and a set of coefficient factors.⁴⁸ Thus, the new score function was sufficient to rank the antagonist in the different levels of binding affinities. The total binding free energy between 1-amino-2-phenyl-4-(piperidin-1-yl)-butane analogues and CCR5 was calculated using the new scoring function encoded in AutoDock 3.0.3.^{47,48}

2.5. 3D QSAR studies

To more fully explore the specific contributions of electrostatic, steric, and hydrophobic effects for these 1-amino-2-phenyl-4-(piperidin-1-yl)-butane analogues binding to CCR5, CoMFA,³³ and CoMSIA³⁴ studies were performed for these antagonists based on the conformational alignment predicted from the molecular docking.

2.5.1. CoMFA. Usually, steric and electronic field energies were probed using an sp^3 carbon atom and a +1 net charge atom, respectively. Steric and electrostatic interactions were calculated using the Tripos force field⁵¹ with a distance-dependent dielectric constant at all intersections in a regularly spaced (2 Å) grid. The minimum σ (column filtering) was set to 2.0 kcal/mol to improve the signal-to-noise ratio by omitting those lattice points whose energy variation was below this threshold. A cutoff of 30 kcal/mol was adopted, and the regression analysis was carried out using the full cross-validated partial least squares (PLS) method (leave

one out) with CoMFA standard options for scaling of variables. The final model (non-cross-validated conventional analysis) was developed with the optimum number of components equal to that yielding the highest q^2 .

2.5.2. CoMSIA. In this study, three physicochemical properties, namely steric, electrostatic, and hydrophobic fields, have been evaluated. The steric contribution was reflected by the third power of the atomic radii of the atoms. Electrostatic properties were introduced as atomic charges resulted from molecular docking. An atom-based hydrophobicity was assigned according to the parameterization developed by Ghose et al.⁵² The lattice dimensions were selected with a sufficiently large margin (>4 Å) to enclose all aligned molecules. Singularities were avoided at atomic positions in CoMSIA fields because a Gaussian-type distance dependence of the physicochemical properties was adopted, thus no arbitrary cutoffs were required. In general, similarity indices $A_{F,k}$ between the compounds of interest were computed by placing a probe atom at the intersections of the lattice points using Eq. 1,

$$A_{F,k}^q(j) = - \sum_{i=1}^n W_{\text{probe},k} W_{ik} e^{-\alpha r_{iq}^2} \quad (1)$$

where q represents a grid point; i is summation index over all atoms of the molecule j under computation; W_{ik} is the actual value of physicochemical property k of atom i ; $W_{\text{probe},k}$ is the value of the probe atom. In present study, similarity indices were computed using a probe atom ($W_{\text{probe},k}$) with charge +1, radius 1 Å, hydrophobicity +1, and attenuation factor α of 0.3 for the Gaussian-type distance. The statistical evaluation for the CoMSIA analyses was performed in the same way as described for CoMFA.

3. Results and discussion

3.1. 3D model of CCR5 receptor

The X-ray crystal structure of rhodopsin is the only one experimentally determined structure for GPCRs.³⁰ However, recently, great success has been achieved in the structure prediction of GPCRs based on the crystal structure of bovine rhodopsin,^{53–56} and even the homolog models of GPCRs have been used successfully in virtual screening of chemical databases for drug discovery.⁵⁶ Accordingly, the crystal structure of bovine rhodopsin is an appreciated template for modeling the 3D structures of GPCR family proteins. The primary 3D model of CCR5 resulted from the structural modeling is shown in Figure 1. PROCHECK⁴² statistics indicated that ~94% residues in the CCR5 model are in either the most favored or in the additionally allowed regions of the Ramachandran plot. The overall main chain and side chain parameters, as evaluated by PROCHECK, are all very favorable.

The 3D model of CCR5 was optimized alternately using molecular mechanics and molecular dynamics (see Section 2.3). Two kinds of interacting networks are

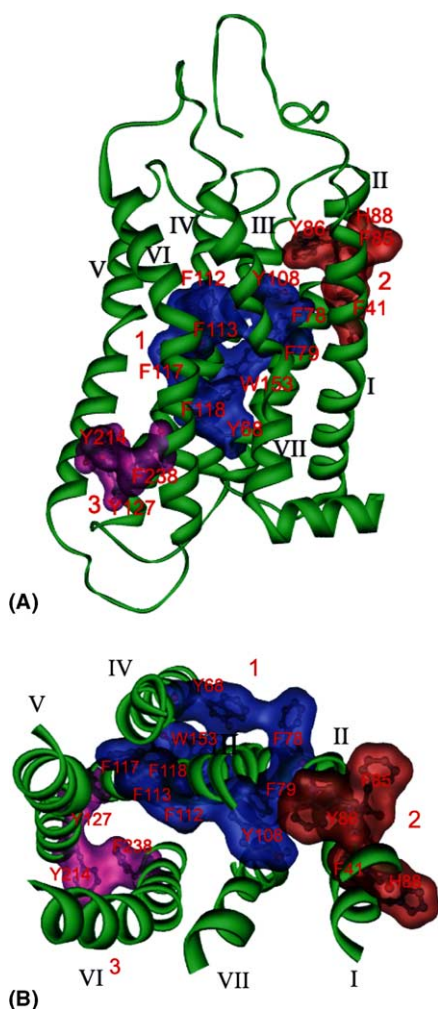


Figure 1. Ribbon representation of the 3D model for CCR5 receptor. Intramolecular hydrophobic interactions represented as three aromatic clusters are also shown: (1) cluster among TM2, TM3, and TM4; (2) cluster between TM1 and TM2; and (3) cluster among TM3, TM5, and TM6. (A) Side view. (B) Top view from the intracellular side (without loops, C- and N-terminal for clear visual of TMs). This image was rendered in the POV-Ray program.⁵⁷

observed from the optimized 3D model of CCR5, that is, aromatic residue clusters (Fig. 1) and a hydrogen-bonding network (Table 2). For the first network, aromatic residues are observed to assemble three clusters, maintaining the geometry of the TMs by the favorable stacking interactions as in known membrane proteins.^{58,59} One aromatic cluster (cluster 1 in Fig. 1) locates at the middle of TM2, TM3, and TM4, being composed of residues Tyr68, Phe78, Phe79, Tyr108, Trp158, Phe112, Phe113, Phe117, and Phe118; these aromatic interactions make TM2, TM3, and TM4 congregate tightly. Another aromatic cluster (cluster 2 in Fig. 1) locates at the upside of TM1 and TM2 formed by the side chains of Phe41, Phe85, Trp86, and His88. The third aromatic cluster (cluster 3 in Fig. 1), formed by residues Tyr127, Tyr214, and Phe238, packs TM3 with TM5 and TM6 through van der Waals interactions. The second network is formed by hydrogen bonds between residues conserved across the chemokines receptor subtypes (Table 2). All the hydrogen bonds along

with above mentioned hydrophobic interactions enable the helices conformation more stable.

Sequence alignment of bovine rhodopsin with CCR5 is shown in Figure 2A, which indicates that residues composing the trans-membrane regions of GPCRs are relatively conserved, the trans-membrane regions of CCR5 are located at the similar positions of rhodopsin. The structural superposition of the CCR5 model to the X-ray crystal structure of bovine rhodopsin is shown in Figure 2B and C, indicating again the conservation of the trans-membrane regions, the root mean square deviation (RMSD) for superposition the C α atoms of the seven trans-membrane helices of CCR5 with that of bovine rhodopsin is ~ 1.0 Å. Moreover, similar to the structure of rhodopsin, Cys178 of EL2 forms a conserved disulfide bond with Cys101 of TM3, which makes EL2 adopt a conformation that covers the putative binding pocket.

3.2. Interactions between antagonists and CCR5 binding pocket

The automated molecular docking may produce several options of binding conformation for each antagonist. The conformation corresponding to the lowest binding energy with CCR5 was selected as the most possible binding conformation. Thus the most probable binding conformations of the 58 compounds in the training set and the six compounds in the testing set were obtained. The 3D binding model for compound 2 is shown in Figure 3A and B. The molecular superposition of the binding conformation for these 64 compounds extracted from the optimized antagonist–CCR5 complexes is shown in Figure 4. To illustrate the interaction mechanism, compound 2 was selected for more detailed analysis.

The aromatic rings of the ligands occupy the hydrophobic clusters of CCR5 (Table 3). Tyr108 and Phe112 in TM3 and Trp248 and Tyr251 in TM6 assemble one hydrophobic cluster, which accommodates the *N*-methyl-*N*-phenylsulfonylamino moiety of the ligands. C-2 phenyl fits into the hydrophobic cluster consisted of Ser179 in EL2, Tyr251 in TM6, and Met279 in TM7. The 4-substituted piperidine moiety resides in the cluster lined by Tyr37 in TM1, Trp86, Ala87, and Ala90 in TM2, and Leu104 and Tyr108 in TM3. Most of the site-directed mutagenesis on binding site of CCR5 have a direct effect on antagonist binding, and then resulting in decrease or even loss of activation of the receptor.^{25,29,44,45} Tyr108 in TM3 is a very important binding site in different GPCRs for neurotransmitters and peptides. Tyr108Ala mutation strongly reduced binding affinity 14–375-fold for different ligands.²⁹ In our present model, the Tyr108 side chain is positioned in the face-to-edge orientation with the indole ring of Trp86. This type of aromatic–aromatic interaction has been described as stabilizing a protein structure. The conformational changes induced in TM2 by Tyr108Ala mutation would affect the orientation of the side chain of Trp86. Similarly, Trp86Ala mutation would also abolish the aromatic–aromatic interaction, and conse-

Table 2. Hydrogen bonds (beyond the backbone) formed by residues of TMs

Location	Donor		Acceptor		Distance (Å)
	Residue	Group	Group	Residue	
TM1–TM7	Tyr37	OH	O ^{γ1}	Thr284	2.59
TM1–TM1	Asn48	N ^{δ2} H	O=C<	Gly44	2.58
TM1–TM7	Asn48	N ^{δ2} H	O=C<	Cys290	2.67
TM1–TM1	Asn57	N ^{δ2} H	O=C<	Asn57	2.78
TM2–TM4	Tyr68	OH	O=C<	Thr148	2.57
TM2–TM4	Asn71	N ^{δ2} H	O ^γ	Ser149	2.56
TM2–TM3	Trp86	N ^{ε1} H	O=C<	Met100	2.59
TM2–TM1	His88	N ^{δ1} H	O=C<	Pro34	2.56
TM3–TM3	Arg126	N ^{η1} H	O ^{δ1}	Asp125	2.42
TM4–TM2	Trp153	N ^{ε1} H	O ^{δ1}	Asn71	2.59
TM5–TM4	Lys197	N ^ζ H	O=C<	Phe158	2.67
TM5–TM5	Lys197	N ^ζ H	O ^{ε1}	Gln194	2.44
TM5–TM3	Cys213	S ^γ H	O=C<	Thr123	2.84
TM6–TM3	Arg230	N ^{η1} H	O ^{δ1}	Asp125	2.53
TM6–TM3	Arg230	N ^{η2} H	O=C<	Asp125	2.53
TM6–TM3	Arg230	N ^{η2} H	O ^{δ1}	Asp125	2.46
TM6–TM7	Arg232	N ^ε H	O=C<	Val300	3.09
TM6–TM7	Arg232	N ^{η1} H	O ^{δ2}	Glu302	2.36
TM6–TM3	Tyr244	OH	O=C<	Phe112	2.52
TM6–TM6	Trp248	N ^{ε1} H	OH	Tyr244	2.61
TM7–TM7	Gln277	NH	O ^{δ1}	Asp276	2.51
TM7–TM2	Cys290	S ^γ H	O=C<	Asp76	2.81
TM7–TM1	Cys291	S ^γ H	O=C<	Phe43	2.69
TM7–TM2	Asn293	N ^{δ2} H	O ^{δ2}	Asp76	2.54
TM7–TM7	Asn293	N ^{δ2} H	O=C<	His289	2.46
TM7–TM2	Tyr297	OH	O ^{δ1}	Asp66	2.45
TM7–TM7	Arg305	N ^{η1} H	O=C<	Ala298	2.52
TM7–TM7	Arg305	N ^{η1} H	O=C<	Phe299	2.47

quently affects the binding affinity of ligands. In fact, it has been demonstrated that Trp86Ala mutation reduced the binding affinities for the Merck compounds to CCR5.²⁹ In addition, Tyr251Ala mutation was also found to reduce the binding affinities of the ligands.²⁹ Our interaction model indicated that this mutation may decrease the hydrophobic interaction of the aromatic rings at C-1 and C-2 positions of compound **2** to CCR5.

Our model revealed that hydrogen bonding is an important interaction between the antagonists and CCR5. There are three hydrogen bonds formed between compound **2** and CCR5 (Fig. 3B). These hydrogen bonds act as anchors, intensely determining the binding orientations of the ligands in the binding pocket of CCR5. The basic piperidine nitrogen atom of compound **2** acts as a donor to form hydrogen bonds with the carboxyl oxygen atom of Glu283 and hydroxy oxygen atom of Thr177 (Fig. 3B). Glu283 is the only acidic residue in the extracellular end of the seven trans-membrane bundles. The binding models indicate that almost all of the compounds in this study adopt binding conformations that may form hydrogen bonds between the basic nitrogen of the compounds and the carboxyl oxygen atom of Glu283 in TM7. This result is in agreement with the mutagenesis data that a significantly reduced binding affinity was observed for the Merck compounds (reduced about 1000 times) and TAK779 upon Glu283Ala mutation.^{25,43} Quaternization of piperidine nitrogen seemed to be essential for the interaction between the

ligands and CCR5 as well as for the orientation of the 4-phenyl substitution. The sulfonamide oxygen of **2** acts as an acceptor to form a hydrogen bond with the amide nitrogen atom of Met287, indicating the sulfonamide may be a key pharmacophore for antagonist activity of the antagonists. This is in good agreement with the experimental result that compounds without sulfonamide such as the amides or substituted benzamide analogues are inactive.¹⁷

As mentioned above, the putative binding site is a pocket surrounded by TM1, TM2, TM3, TM6, and TM7. This cavity is very similar to the cavity proposed by others.^{25,29,53,61} The consistency between the 3D model of ligand–receptor complex and the experimental mutagenesis results indicates the reasonability of the modeled structure, and the binding pocket may act as a starting point for structure-based ligand design. In addition, our ligand–receptor interaction model revealed the importance of Thr177 and Met287 to the ligand binding, restricting the antagonists in their proper locations. So far, no site-directed mutagenesis analyses have been conducted on these two residues, which will be a direction for further functional study of CCR5.

3.3. The correlation between binding free energy and antagonist activity

Table 4 lists the predicted binding free energies of the 1-amino-2-phenyl-4-(piperidin-1-yl)-butane analogues to CCR5. Satisfied that the 3D structures of CCR5–antagonist

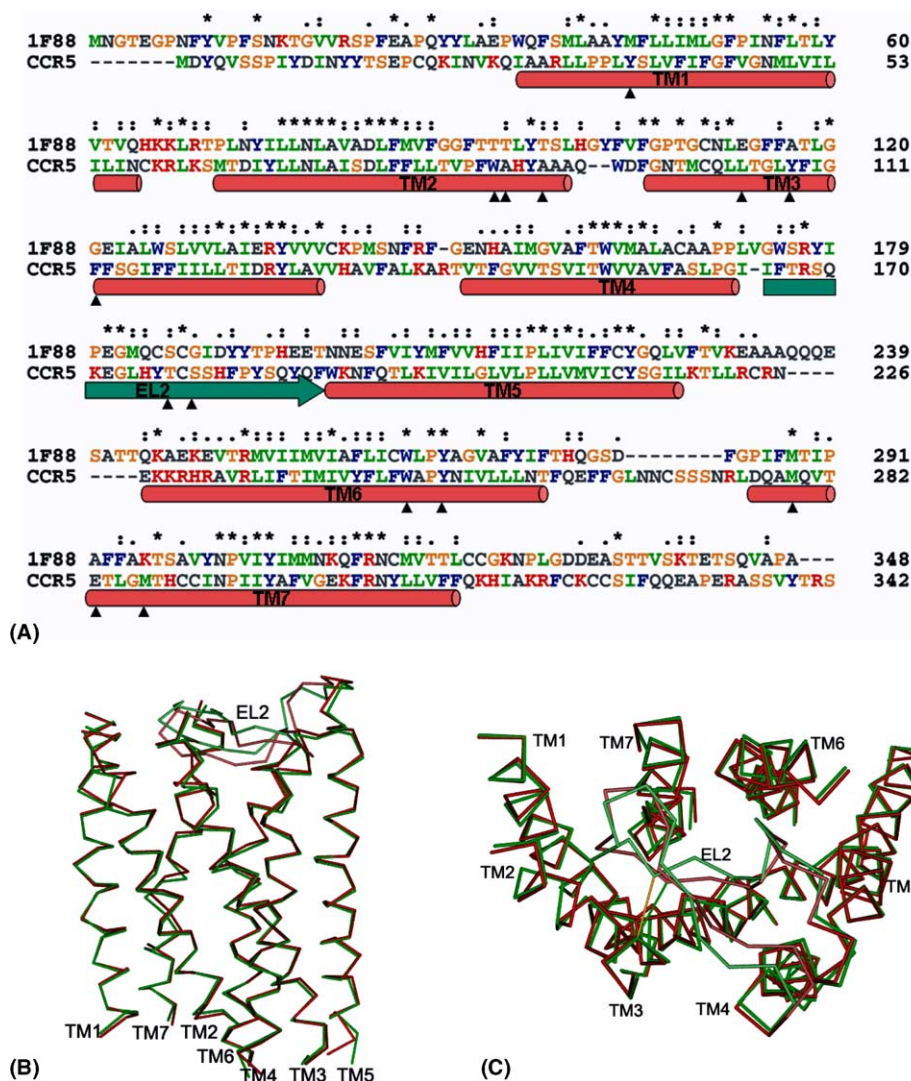


Figure 2. (A) Sequence alignment of CCR5 with rhodopsin (1F88) generated by CLUSTALW (1.81). In the sequences, an asterisk (*) indicates an identical or conserved residue, a colon (:) indicates a conserved substitution, a stop (.) indicates a semi-conserved substitution, a triangle (▲) indicates a residue formed the binding pocket. (B) Side view and (C) top view of superposition of the 3D model of CCR5 (green) with the X-ray crystal structure of rhodopsin (red), only seven TMs and EL2 regions are shown for clarity.

complexes were indeed reasonable, we then performed a linear regression analysis to explore whether the antagonist potencies of the 1-amino-2-phenyl-4-(piperidin-1-yl)-butane analogues could be correlated with the energetic parameters. By using experimentally determined IC_{50} values, we calculated the regression equation for the antagonist activities, $-\log IC_{50}$ values, using the total binding free energies, ΔG s, as the sole descriptor variable. A good correlation was found between the antagonist activities and the predicted binding energies (Eq. 2), and this relationship is graphically shown in Figure 5.

$$-\log IC_{50} = -2.163 - 0.417 \times \Delta G$$

$$(n = 58, r^2 = 0.591, F = 80.770, s = 0.517) \quad (2)$$

Based on the binding free energies and their correlation with the antagonist activities, we can give a more quantitative explanation to the structure–activity relationship

of the antagonistic mechanism for these antagonists. It is obvious that there would be about a 2.4 kcal/mol difference in binding free energy if there is one order of magnitude for numerical difference in the antagonist potency (IC_{50}). The trend of change in the binding free energy is in agreement with that of change in the antagonist activity.

The sophisticated method for binding free energy calculation is the free energy perturbation (FEP) approach.^{62,63} However, the FEP approach is time-consuming; thus, it cannot be extended to predict the binding affinities for a large set of molecules. Automated molecular docking can identify the binding conformation and predict the binding affinity very quickly; therefore, it can be applied in constructing the prediction model for a series of molecules in a tolerable time, as indicated above that the antagonist potency correlates well with the AutoDock predicted binding free energy (Eq. 2 and Fig. 5). This relationship suggests that those

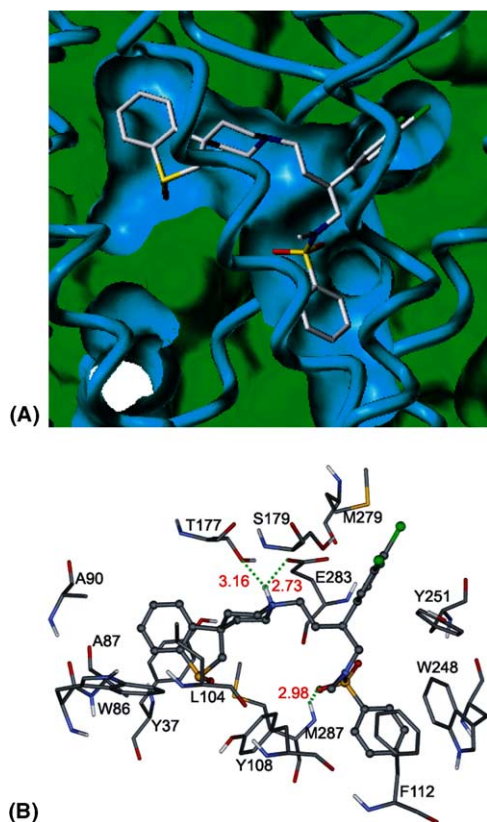


Figure 3. (A) Close-up view of the putative binding pocket of CCR5 bounded with antagonist **2**. This image was generated using the MOLCAD program encoded in Sybyl 6.8. (B) Three-dimensional structural model of the main interactions between compound **2** and the putative CCR5 binding pocket.

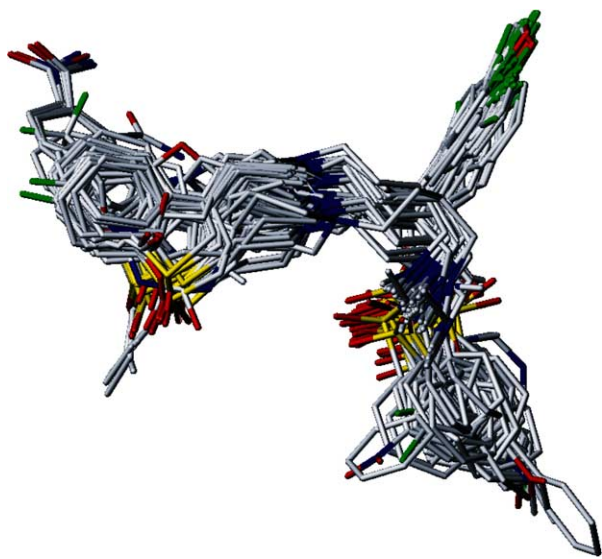


Figure 4. Superposition of 64 compounds based on probable binding conformations with CCR5.

potential CCR5 antagonists exhibiting stronger binding free energies using this paradigm would therefore be expected to have greater efficacy toward antagonist action.

3.4. 3D QSAR models

3.4.1. CoMFA. Although CoMFA is not able to appropriately describe all aspects of the binding forces, being based principally on standard steric and electrostatic molecular fields to model ligand–protein interaction, it is still a widely used tool for the study of QSAR at the 3D level. The major objective of CoMFA analysis for 1-amino-2-phenyl-4-(piperidin-1-yl)-butane analogues is to find the best predictive model within the system. PLS analysis results based on least squares fit are listed in Table 5, which shows that all of the statistical indexes are reasonably high. As listed in Table 5, CoMFA models with cross-validated $r^2_{\text{cross}}(q^2)$ of 0.568 for four components was obtained based on the binding conformations and their alignment in the active site of CCR5. The non-cross-validated PLS analysis was repeated with the optimum number of components, as determined by the cross-validated analysis, to give an r^2 of 0.947, $F = 237.484$, and the estimated standard error of 0.191. The values indicated a good conventional statistical correlation, and the CoMFA model had a fair predictive ability. The predicted antagonist activities of these 58 compounds are listed in Table 6 and also shown in Figure 6A, indicating that the fitting power is rational and potent and the predictive ability is satisfactory.

3.4.2. CoMSIA. CoMSIA analysis results are also summarized in Table 5. A CoMSIA model with an $r^2_{\text{cross}}(q^2)$ value of 0.587 for six components and a conventional r^2 of 0.967 for 1-amino-2-phenyl-4-(piperidin-1-yl)-butanes was obtained. These data demonstrate that the CoMSIA model is also fairly predictive, and the predicted antagonist potencies of these 58 compounds are listed in Table 6 and also shown in Figure 6B. The high value of the conventional r^2 relating to three different descriptor variables (steric, electrostatic, and hydrophobic) illustrates that these variables are necessary to describe the interaction mode of 1-amino-2-phenyl-4-(piperidin-1-yl)-butane analogues with the CCR5 receptor, as well as the field properties around the antagonists. CoMSIA analysis revealed that the steric, electrostatic, and hydrophobic field distributions are 16.6%, 41.5%, and 41.9% (Table 5), respectively, indicating that electrostatic and hydrophobic interactions equally dominate the binding between 1-amino-2-phenyl-4-(piperidin-1-yl)-butane analogues and CCR5. This is in agreement with the binding model of the antagonists to CCR5 (Fig. 3). Furthermore, the CoMFA steric field can be seen as a comprehensive contribution of the pure steric and hydrophobic effects because the steric contribution of CoMFA is 58.7% (Table 5), which is approximately equal to the summation of the steric and hydrophobic contributions of CoMSIA (58.5%). The CoMSIA model indicates that the fitting power is rational and potent and the predictive ability is satisfactory.

3.4.3. Validation of the 3D QSAR models. To validate our 3D QSAR models, six antagonists (star-denoted in Table 6) that were not included in generating CoMFA and CoMSIA models were selected as testing compounds. The results are simultaneously shown in Table 6 (star labeled), Figure 6A and B (in red star symbols).

Table 3. Hydrophobic contacts between compound **2** and CCR5*

CCR5 residue	Atom type	Ligand atom type	Distance (Å)	CCR5 residue	Atom type	Ligand atom type	Distance (Å)
Met287	C ^ε	C3	3.84	Cys178	C ^β	C1	3.90
Gly286	C	C41	3.70	Cys178	C	C6	3.54
Gly286	C	C15	3.51	Cys178	C	C1	3.75
Gly286	C	S12	3.86	Cys178	C ^α	C1	3.65
Gly286	C ^α	C56	3.41	Thr177	C	C1	3.87
Gly286	C ^α	C15	3.39	Thr177	C ^α	C49	3.76
Gly286	C ^α	S12	3.58	Tyr176	C ^β	C50	3.73
Glu283	C ^δ	C21	3.81	Tyr176	C	C50	3.54
Glu283	C ^δ	C7	3.05	Tyr176	C	C49	3.46
Glu283	C ^δ	C4	3.75	Phe112	C ^β	C43	3.58
Glu283	C ^γ	C7	3.85	Phe112	C ^β	C42	3.44
Glu283	C ^γ	C4	3.57	Tyr108	C ^{δ2}	C40	3.28
Glu283	C	C4	3.87	Tyr108	C ^γ	C40	3.81
Met279	S ^δ	C20	3.32	Tyr108	C ^β	C40	3.42
Met279	S ^δ	C19	3.54	Leu104	C ^{δ1}	C26	3.70
Met279	C ^γ	C20	3.85	Leu104	C ^{δ1}	C53	3.65
Tyr251	C ^{ε1}	C21	3.71	Leu104	C ^{δ1}	C49	3.58
Tyr251	C ^{ε1}	C20	3.71	Leu104	C ^{δ1}	C25	3.31
Tyr251	C ^{ε1}	C19	3.46	Leu104	C ^{δ1}	C1	3.10
Tyr251	C ^{ε1}	C18	3.17	Leu104	C ^{δ1}	C2	3.62
Tyr251	C ^{ε1}	C17	3.17	Leu104	C ^β	C26	3.68
Tyr251	C ^{ε1}	C16	3.47	Leu104	C ^β	C1	3.87
Tyr251	C ^{δ1}	C21	3.08	Ala90	C ^β	C51	3.57
Tyr251	C ^{δ1}	C20	3.28	Ala90	C ^β	C50	3.73
Tyr251	C ^{δ1}	C19	3.57	Ala87	C ^β	C52	3.21
Tyr251	C ^{δ1}	C18	3.67	Ala87	C ^α	C52	3.12
Tyr251	C ^{δ1}	C17	3.48	Ala87	C ^α	C51	3.39
Tyr251	C ^{δ1}	C16	3.20	Trp86	C ^{η2}	S54	3.45
Tyr251	C ^{δ1}	C10	3.62	Trp86	C ^{η2}	C53	3.77
Trp248	C ^{ε3}	C56	3.90	Trp86	C ^{ε3}	S54	3.30
Trp248	C ^{ε3}	C56	3.68	Trp86	C ^{ε3}	C53	3.39
Trp248	C ^{ε3}	C55	3.65	Trp86	C ^{ε3}	C52	3.15
Trp248	C ^{ε2}	C55	3.52	Trp86	C ^{ε3}	C52	3.60
Trp248	C ^{ε2}	C43	3.78	Val83	C ^{γ1}	S54	3.63
Trp248	C ^{δ2}	C56	3.89	Tyr37	C ^{ε2}	C53	3.85
Trp248	C ^{δ2}	C55	3.34	Tyr37	C ^{ε2}	C52	3.44
Trp248	C ^γ	C55	3.69	Tyr37	C ^{ε2}	C51	3.51
Ser179	C ^β	C18	3.85	Tyr37	C ^{δ2}	C52	3.68

* Calculated using the LIGPLOT program.⁶⁰

The predicted activities values are in good agreement with the experimental data in a statistically tolerable

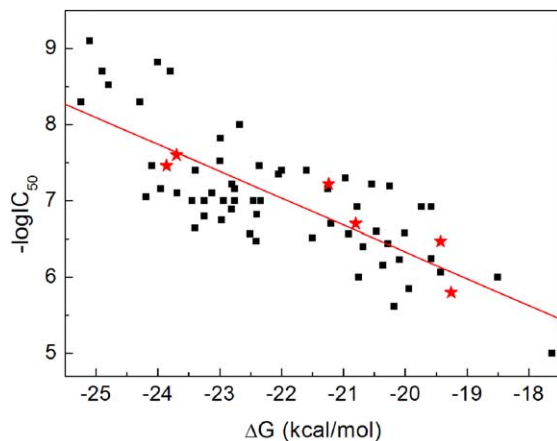


Figure 5. Correction between the binding free energies (ΔG s) of 1-amino-2-phenyl-4-(piperidin-1-yl)-butane compounds with CCR5 and the experimental activities ($-\log IC_{50}$ s). ■, compounds of the training set; ★, compounds of the testing set.

error range, $r^2 = 0.920$ and 0.686 for CoMFA and CoMSIA models, respectively. To investigate the structural differences of binding modes between the testing set of compounds and the training set, automated molecular docking was performed for the testing set using the same method as that of the training set. The testing results for the six antagonists indicate that the CoMFA and CoMSIA models can be further used in new antagonists design for CCR5.

3.4.4. Mapping of CoMFA and CoMSIA model onto CCR5 binding site. The QSAR produced by CoMFA, with its hundreds or thousands of terms, was usually represented as 3D ‘coefficient contour’. It shows regions where variations of steric or electrostatic nature in the structural features of the different molecules contained in the training set lead to increases or decreases in the activity. The CoMFA steric and electrostatic fields for the analysis based on the alignments of the binding conformations are presented as contour plots in Figure 7. To aid in visualization, compound **2** is displayed in the maps. In general, the colored polyhedra in the map surrounded all lattice points where the QSAR strongly

Table 4. Predicted binding free energies versus experimental activities ($-\log IC_{50}$ s) of the CCR5 antagonists

No	ΔG (kcal/mol)	$-\log IC_{50}$	No	ΔG (kcal/mol)	$-\log IC_{50}$
1	-20.75	6.00	33	-22.99	7.82
2	-24.10	7.46	34	-22.51	6.57
3	-23.25	7.00	35	-19.58	6.24
4	-20.54	7.22	36*	-20.80	6.70
5	-19.74	6.92	37	-23.69	7.10
6*	-19.43	6.47	38	-23.00	7.52
7	-20.09	6.23	39	-18.51	6.00
8	-19.43	6.07	40	-21.50	6.52
9	-19.94	5.85	41	-22.97	6.75
10	-22.46	7.00	42	-20.97	7.30
11	-22.34	7.00	43	-25.24	8.30
12	-23.13	7.10	44	-22.05	7.35
13	-22.76	7.16	45	-22.94	7.00
14	-23.39	7.40	46*	-23.86	7.46
15	-22.40	6.82	47	-22.36	7.46
16*	-21.24	7.22	48	-23.45	7.00
17	-22.00	7.40	49	-20.01	6.58
18	-21.60	7.40	50	-20.36	6.16
19	-20.28	6.44	51	-20.18	5.62
20	-22.80	7.22	52	-20.68	6.40
21	-23.95	7.16	53	-21.25	7.16
22	-22.76	7.00	54	-19.58	6.92
23	-24.19	7.05	55	-21.20	6.70
24	-20.78	6.92	56*	-23.70	7.60
25	-22.41	6.47	57	-20.47	6.60
26*	-19.26	5.80	58	-20.25	7.19
27	-17.63	5.00	59	-24.29	8.30
28	-23.40	6.64	60	-24.00	8.82
29	-22.81	6.89	61	-24.90	8.70
30	-23.25	6.80	62	-25.10	9.10
31	-20.92	6.57	63	-23.80	8.70
32	-22.68	8.00	64	-24.80	8.52

* Antagonists of the testing set.

Table 5. Statistical indexes of CoMFA and CoMSIA models based on 58 1-amino-2-phenyl-4-(piperidin-1-yl)-butane analogues binding conformations

	Cross-validated		Conventional		
	$q^2(r^2_{\text{cross}})$	Optimal comp.	r^2	s	F
CoMFA	0.568	4	0.947	0.191	237.484
CoMSIA	0.587	6	0.967	0.158	251.677
	Field distribution (%)				
	Electrostatic	Steric	Hydrophobic		
CoMFA	0.413	0.587			
CoMSIA	0.415	0.166	0.419		

associated changes in the compounds' field values with changes in biological potency. The CoMFA contour plots show green-colored regions where increased steric is associated with enhanced activity and yellow-colored regions where increased steric bulk is associated with decreased activity (Fig. 7A).

A big region of green contour near the terminal of 4-substituted piperidine moiety suggests that more bulky substituents in these positions will significantly improve the biological activities. For example, compound **60–64** exhibited a considerable gain in binding affinity due to the increased bulky substituents. As men-

tioned above, 4-substituted piperidine binds to the hydrophobic cluster lined by Tyr37 in TM1, Trp86, Ala87, and Ala90 in TM2, and Leu104 and Tyr108 in TM3. Around this hydrophobic cluster, there is enough space left to allow bulky group interact with CCR5 residues. The yellow polyhedron near 4-position substitutions on the C-2 phenyl indicates that increased steric bulk is unfavorable for the antagonist activities. This may rationalize why the antagonist activity of compounds have the order of **37** > **36** > **34** > **35**. This is quite complementary with the model structure of antagonist-CCR5 complex. The yellow polyhedron under the *N*-(methyl)-*N*-(phenylsulfonyl)-amino moiety may bring

Table 6. Predicted activities (PA) from CoMFA and CoMSIA models compared with the experimental activities (EA) and the residuals (δ)

No	pIC ₅₀	CoMFA		CoMSIA		No	pIC ₅₀	CoMFA		CoMSIA	
		PA	δ	PA	δ			PA	δ	PA	δ
1	6.00	6.12	−0.12	5.94	0.06	33	7.82	7.58	0.24	7.80	0.02
2	7.46	7.15	0.30	7.11	0.35	34	6.57	6.65	−0.08	6.99	−0.42
3	7.00	7.06	−0.06	6.98	0.03	35	6.24	6.48	−0.24	6.46	−0.22
4	7.22	6.89	0.33	7.19	0.04	36*	6.70	6.75	−0.05	7.02	−0.32
5	6.92	6.81	0.11	6.90	0.02	37	7.10	7.13	−0.03	7.08	0.02
6*	6.47	6.84	−0.37	6.78	−0.31	38	7.52	7.10	0.43	7.32	0.20
7	6.23	6.29	−0.06	6.01	0.22	39	6.00	6.22	−0.22	6.06	−0.06
8	6.07	5.85	0.22	5.87	0.20	40	6.52	6.43	0.10	6.54	−0.01
9	5.85	5.92	−0.07	5.85	0.00	41	6.75	6.77	−0.02	6.79	−0.04
10	7.00	6.84	0.16	6.95	0.05	42	7.30	7.20	0.10	7.36	−0.06
11	7.00	6.96	0.04	6.99	0.01	43	8.30	8.19	0.11	8.33	−0.03
12	7.10	7.28	−0.18	7.17	−0.07	44	7.35	7.30	0.05	7.15	0.20
13	7.16	7.02	0.14	7.07	0.09	45	7.00	7.16	−0.16	7.15	−0.15
14	7.40	7.41	−0.02	7.19	0.21	46*	7.46	7.45	0.01	7.02	0.44
15	6.82	6.63	0.19	6.82	0.01	47	7.46	7.31	0.15	7.53	−0.07
16*	7.22	7.54	−0.32	6.81	0.41	48	7.00	6.59	0.41	6.90	0.10
17	7.40	7.45	−0.06	7.37	0.03	49	6.58	6.90	−0.31	6.90	−0.31
18	7.40	7.61	−0.21	7.35	0.05	50	6.16	6.22	−0.06	6.25	−0.09
19	6.44	6.74	−0.30	6.81	−0.37	51	5.62	5.80	−0.18	5.54	0.08
20	7.22	7.42	−0.20	7.31	−0.09	52	6.40	6.22	0.18	6.23	0.17
21	7.16	7.21	−0.05	7.13	0.03	53	7.16	6.98	0.18	7.03	0.13
22	7.00	7.11	−0.11	7.14	−0.14	54	6.92	6.65	0.27	6.69	0.24
23	7.05	6.85	0.20	6.97	0.08	55	6.70	6.50	0.20	6.84	−0.14
24	6.92	7.05	−0.13	7.01	−0.09	56*	7.60	7.47	0.13	7.68	−0.08
25	6.47	6.86	−0.39	6.67	−0.20	57	6.60	6.75	−0.14	6.59	0.01
26*	5.80	6.26	−0.46	6.31	−0.51	58	7.19	7.36	−0.17	7.21	−0.03
27	5.00	5.09	−0.09	5.07	−0.07	59	8.30	8.41	−0.11	8.36	−0.06
28	6.64	6.73	−0.09	6.76	−0.12	60	8.82	8.71	0.11	8.77	0.06
29	6.89	6.99	−0.11	6.89	0.00	61	8.70	8.78	−0.08	8.67	0.03
30	6.80	6.87	−0.07	6.93	−0.14	62	9.10	8.91	0.19	8.95	0.15
31	6.57	6.57	0.00	6.43	0.14	63	8.70	8.69	0.01	8.72	−0.02
32	8.00	8.05	−0.05	7.89	0.11	64	8.52	8.78	−0.26	8.63	−0.11

* Antagonists of the testing set.

steric clash of these antagonists with Leu104 and Tyr108. This is consistent with the observation that the N-ethyl analogues (**27**, **29**, and **30**) are less potent than corresponding N-alkyl analogues (**1**, **2**, and **3**).

Regions where increase positive charge is favorable for antagonist activity are indicated in blue, while regions where increased negative charge is favorable for activity are indicated in red (Fig. 7B). The blue contours near 4-position substitutions on the C-2 phenyl indicate that positively charged substituents may increase the antagonist activity. Near the sulfonamide group, there are one red polyhedron region on one side and one blue polyhedron region on the other side. This information gave a powerful argument that the sulfonamide is a critical pharmacophore for this scaffold since the related amides and all substituted benzamide analogues were essentially inactive. The sulfonamide group plays a critical role in assuring the phenyl moiety insert into one of the CCR5 hydrophobic pocket formed by Tyr108, Phe112, Trp248, Tyr251, and His 289.

The steric and electrostatic fields of CoMSIA, as shown in Figure 7C and D, are generally in accordance with the field distribution of CoMFA maps (Fig. 7A and B). Besides the structural features already mentioned in the

CoMFA steric field analysis, there is a big yellow polyhedron at the center of the image that corresponds to the second extracellular loop (EL2) region. Obviously, it is very reasonable because that the binding pocket is partly covered by the second extracellular loop. Some of the EL2 residues, such as Thr177 and Ser179, take part in the binding of the antagonists with the CCR5 receptor and restrict the location of antagonists. The green contours near the sulfoxide group at the lower left of Figure 7C indicate that bulky substituents in these positions will significantly improve the biological activities. This is in agreement with that sulfoxide derivative was preferred and the corresponding sulfones were usually equipotent or slightly less active while the sulfides were much poorer or inactive antagonists.¹⁷ The blue contours (Fig. 7B) near 4-position substitutions on the C-2 phenyl in CoMFA electrostatic field and the red contours (Fig. 7D) near 3-position substitutions on the C-2 phenyl in CoMSIA indicate that electron-rich chlorine substitution at the 3-position appeared to be preferred.

Furthermore, the hydrophobic analysis of CoMSIA (Fig. 7E), based on the atomic hydrophobicity distribution, could demonstrate more clearly the hydrophobic interactions between the 1-amino-2-phenyl-4-(piperidin-1-yl)-butane analogues with CCR5. For the

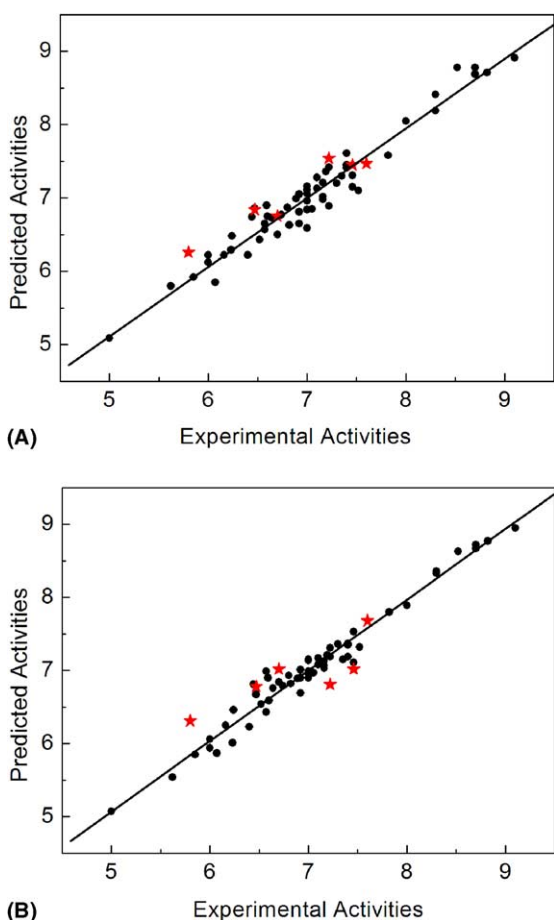


Figure 6. Correlation between predicted activities (PA) by CoMFA (A) and CoMSIA (B) models and the experimental activities of 1-amino-2-phenyl-4-(piperidin-1-yl)-butane analogues on CCR5 (●, compounds of the training set, $r^2 = 0.947$ and 0.967 for two QSAR models, respectively; ★, compounds of the testing set, $r^2 = 0.920$ and 0.686 for the two QSAR models, respectively).

hydrophobic maps of CoMSIA, contour plots show magenta-colored regions where increased hydrophobic interaction is associated with enhanced activity and white-colored regions where increased hydrophilic interaction is associated with increased activity. As shown in Figure 7E, the magenta-colored polyhedron at 4-position substitutions on the C-2 phenyl indicates that these structural moieties interact with the side chains of residues at the binding site of CCR5 through hydrophobic interaction. While the white-colored polyhedron at 3-position substitutions on the C-2 phenyl illustrates that adding hydrophobic group at this position would be detrimental to antagonist activity. This information farther validate that electron-rich chlorine substitution at the 3-position will increase activity. The big white-colored polyhedron near the sulfoxide group demonstrated again that sulfoxide is a better pharmacophore element than sulfones and sulfides.¹⁷ The magenta-colored polyhedra on the upper-left corner, on the upper-right corner and on the lower-right corner of the image suggest that these regions would prefer substituents with hydrophobic group. The residues around magenta-colored polyhedra are all hydrophobic. The consistency between the CoMSIA field distributions and the structural topo-

logical properties of the CCR5 binding site demonstrates the reasonability of the CoMSIA results.

Combining the CoMFA and CoMSIA contour maps with the 3D structural topology of the CCR5 binding site, several insights into the binding of 1-amino-2-phenyl-4-(piperidin-1-yl)-butane analogues with CCR5 can be readily observed. Not only does the field property coincide perfectly with the environmental characteristics of the binding pocket but also indication for some further structural modification of this kind of compounds could be found. Most of the amino acids around the 1-amino-2-phenyl-4-(piperidin-1-yl)-butanes in the binding pocket are hydrophobic in nature (Fig. 3B). These colored polyhedra of CoMFA and CoMSIA located in the cavity of the binding pocket are direct indexes for the kinds and magnitude of the substituents in the process of CCR5 antagonist synthesis.

4. Conclusions

Homology modeling and molecular dynamics simulations have been applied to build up the 3D model of the CCR5 receptor. From the resultant model, we have predicted the binding conformations of 1-amino-2-phenyl-4-(piperidin-1-yl)-butane compounds to CCR5 employing the LGA algorithm of AutoDock. The modeling results provide a satisfactory explanation for the binding between the antagonists and CCR5, and the predicted ligand–receptor interaction model is in agreement with the experimental mutagenesis results. Additionally, the binding free energies of these compounds calculated by LGA algorithm based on the ligand–receptor interaction models correlate very well with the reported antagonist potencies against CCR5. All these indicate the reasonability of the modeling results. On the basis of the binding conformations of these compounds, we have developed stable and predictive 3D QSAR models with acceptable r^2_{cross} values using CoMFA and CoMSIA techniques, and these models could be mapped back to the 3D topology of the binding site of the CCR5 receptor.

Understanding protein–ligand interactions is essential for designing novel synthetic candidates, while those interactions are difficult to describe. Structure-based design is focused on the elucidation of ligand–protein interactions but does not always lead to predictive models. On the other hand, 3D QSAR results based on CoMFA allow focus on those regions, where steric, electronic, or hydrophobic effects play a dominant role in ligand–receptor interactions. These models are usually built using alignment rules, which are not always similar to the bioactive conformation. In this study, we successfully combined with two approaches: the complex 3D model of 1-amino-2-phenyl-4-(piperidin-1-yl)-butane antagonists with CCR5 was derived by AutoDock 3.0 and the predictive 3D QSAR models were derived by using alignment conformations extracted directly from the 3D models of the ligand–protein complexes. This leads to a better understanding of important ligand–receptor interactions and thus provides guidelines for

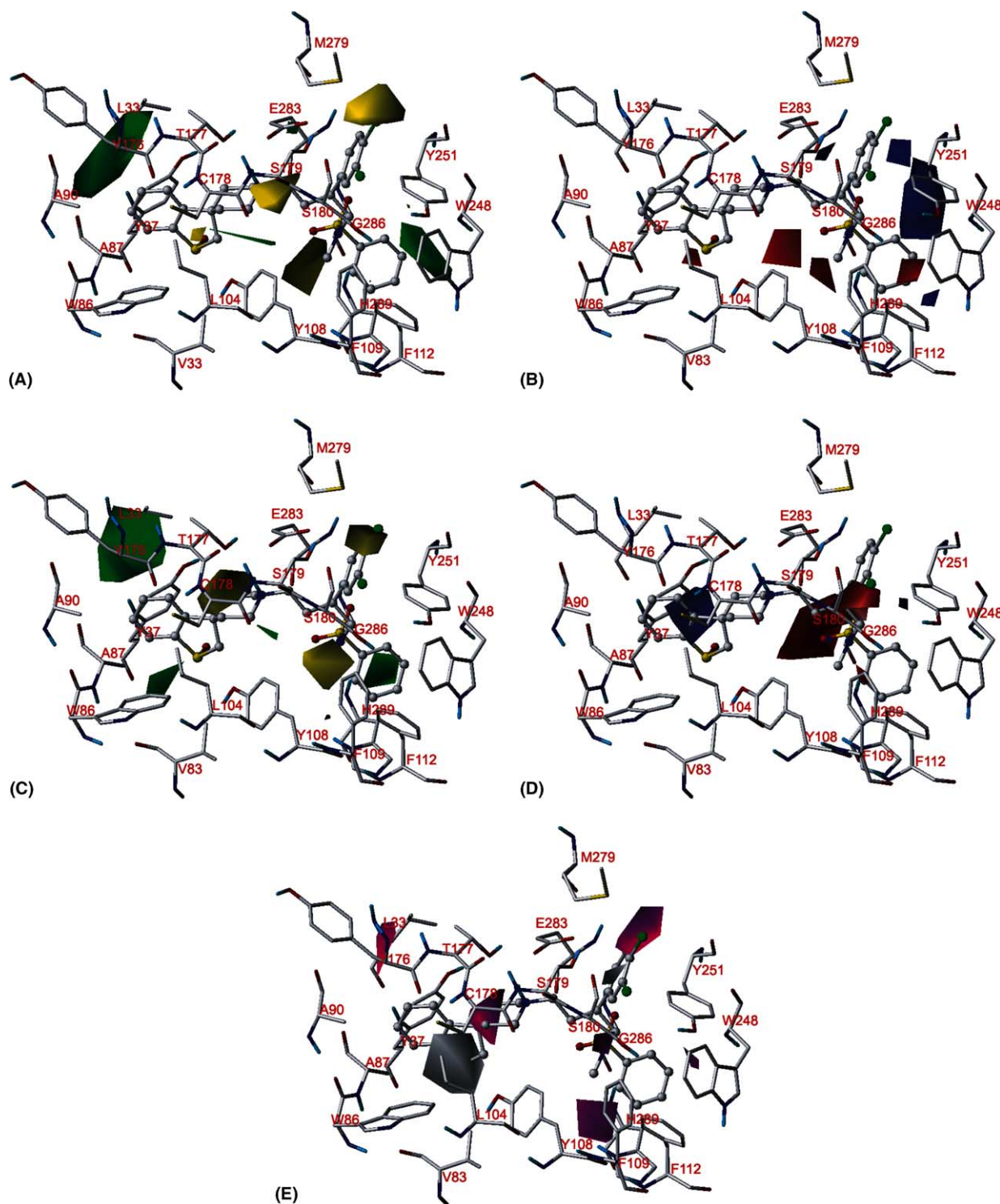


Figure 7. Contour maps as compared with the topology of compound 2–CCR5 complex. Only the residues relevant to the discussion are shown for clarity. Steric (A) and electrostatic (B) field distributions of CoMFA and steric (C), electrostatic (D), and hydrophobic (E) field distribution of CoMSIA were shown, respectively. The residues are represented as sticks, and the ligand is shown in ball-and-stick. Sterically favored areas are in green; sterically unfavored areas are in yellow. Positive potential favored areas are in blue; positive potential unfavored areas are in red. Hydrophobic favored areas are in magenta; hydrophilic favored areas are in white.

ligand design plus a predictive model for scoring novel synthetic candidates. The predictive activity from 3D QSAR and their consistency to CCR5 antagonist activities indicate the validity of those models.

Acknowledgements

We thank Professor Arthur J. Olson for his kindness in offering us the AutoDock 3.0.3 program. We gratefully

acknowledge financial support from National Natural Science Foundation of China (Grants 20102007, 29725203, and 20072042), the State Key Program of Basic Research of China (Grant 2002CB512802), and 863 Hi-Tech Program (Grants 2002AA233061, 2001AA235051, and 2001AA235041).

References and notes

- UNAIDS. *Report on the Global HIV/AIDS Epidemic, 2000*. XIII International AIDS Conference, Durban, South Africa, 2000.
- Pomerantz, R. J. *J. Am. Med. Assoc.* **1999**, *282*, 1177–1179.
- Chun, T. W.; Fauci, A. S. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10958–10961.
- Furtado, M. R.; Callaway, D. S.; Phair, J. P.; Kunstman, K. J.; Stanton, J. L.; Macken, C. A.; Perelson, A. S.; Wolinsky, S. M. *N. Engl. J. Med.* **1999**, *340*, 1614–1622.
- Finzi, D.; Blankson, J.; Siliciano, J. D.; Margolick, J. B.; Chadwick, K.; Pierson, T.; Smith, K.; Lisiewicz, J.; Lori, F.; Flexner, C.; Quinn, T. C.; Chaisson, R. E.; Rosenberg, E.; Walker, B.; Gange, S.; Gallant, J.; Siliciano, R. F. *Nat. Med.* **1999**, *5*, 512–517.
- Berger, E. A.; Doms, R. W.; Fenyo, E. M.; Korber, B. T.; Littman, D. R.; Moore, J. P.; Sattentau, Q. J.; Schuitemaker, H.; Sodroski, J.; Weiss, R. A. *Nature* **1998**, *391*, 240.
- Strader, C. D.; Fong, T. M.; Tota, M. R.; Underwood, D.; Dixon, R. A. *Annu. Rev. Biochem.* **1994**, *63*, 101–132.
- Cocchi, F.; DeVico, A. L.; Garzino-Demo, A.; Arya, S. K.; Gallo, R. C.; Lusso, P. *Science* **1995**, *270*, 1811–1815.
- Liu, R.; Paxton, W. A.; Choe, S.; Ceradini, D.; Martin, S. R.; Horuk, R.; MacDonald, M. E.; Stuhlmann, H.; Koup, R. A.; Landau, N. R. *Cell* **1996**, *86*, 367–377.
- Michael, N. L.; Chang, G.; Louie, L. G.; Mascola, J. R.; Dondero, D.; Bix, D. L.; Sheppard, H. W. *Nat. Med.* **1997**, *3*, 338–340.
- Willoughby, C. A.; Rosauer, K. G.; Hale, J. J.; Budhu, R. J.; Mills, S. G.; Chapman, K. T.; MacCoss, M.; Malkowitz, L.; Springer, M. S.; Gould, S. L.; DeMartino, J. A.; Siciliano, S. J.; Cascieri, M. A.; Carella, A.; Carver, G.; Holmes, K.; Schleif, W. A.; Danzeisen, R.; Hazuda, D.; Kessler, J.; Lineberger, J.; Miller, M.; Emini, E. A. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 427–431.
- Hale, J. J.; Budhu, R. J.; Mills, S. G.; MacCoss, M.; Malkowitz, L.; Siciliano, S.; Gould, S. L.; DeMartino, J. A.; Springer, M. S. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1437–1440.
- Hale, J. J.; Budhu, R. J.; Holson, E. B.; Finke, P. E.; Oates, B.; Mills, S. G.; MacCoss, M.; Gould, S. L.; DeMartino, J. A.; Springer, M. S.; Siciliano, S.; Malkowitz, L.; Schleif, W. A.; Hazuda, D.; Miller, M.; Kessler, J.; Danzeisen, R.; Holmes, K.; Lineberger, J.; Carella, A.; Carver, G.; Emini, E. A. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2741–2745.
- Hale, J. J.; Budhu, R. J.; Mills, S. G.; MacCoss, M.; Gould, S. L.; DeMartino, J. A.; Springer, M. S.; Siciliano, S. J.; Malkowitz, L.; Schleif, W. A.; Hazuda, D.; Miller, M.; Kessler, J.; Danzeisen, R.; Holmes, K.; Lineberger, J.; Carella, A.; Carver, G.; Emini, E. A. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2997–3000.
- Lynch, C. L.; Hale, J. J.; Budhu, R. J.; Gentry, A. L.; Mills, S. G.; Chapman, K. T.; MacCoss, M.; Malkowitz, L.; Springer, M. S.; Gould, S. L.; DeMartino, J. A.; Siciliano, S. J.; Cascieri, M. A.; Carella, A.; Carver, G.; Holmes, K.; Schleif, W. A.; Danzeisen, R.; Hazuda, D.; Kessler, J.; Lineberger, J.; Miller, M.; Emini, E. A. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3001–3004.
- Lynch, C. L.; Willoughby, C. A.; Hale, J. J.; Holson, E. J.; Budhu, R. J.; Gentry, A. L.; Rosauer, K. G.; Caldwell, C. G.; Chen, P.; Mills, S. G.; MacCoss, M.; Berk, S.; Chen, L.; Chapman, K. T.; Malkowitz, L.; Springer, M. S.; Gould, S. L.; DeMartino, J. A.; Siciliano, S. J.; Cascieri, M. A.; Carella, A.; Carver, G.; Holmes, K.; Schleif, W. A.; Danzeisen, R.; Hazuda, D.; Kessler, J.; Lineberger, J.; Miller, M.; Emini, E. A. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 119–123.
- Dorn, C. P.; Finke, P. E.; Oates, B.; Budhu, R. J.; Mills, S. G.; MacCoss, M.; Malkowitz, L.; Springer, M. S.; Daugherty, B. L.; Gould, S. L.; DeMartino, J. A.; Siciliano, S. J.; Carella, A.; Carver, G.; Holmes, K.; Danzeisen, R.; Hazuda, D.; Kessler, J.; Lineberger, J.; Miller, M.; Schleif, W. A.; Emini, E. A. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 259–264.
- Finke, P. E.; Meurer, L. C.; Oates, B.; Mills, S. G.; MacCoss, M.; Malkowitz, L.; Springer, M. S.; Daugherty, B. L.; Gould, S. L.; DeMartino, J. A.; Siciliano, S. J.; Carella, A.; Carver, G.; Holmes, K.; Danzeisen, R.; Hazuda, D.; Kessler, J.; Lineberger, J.; Miller, M.; Schleif, W. A.; Emini, E. A. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 265–270.
- Finke, P. E.; Meurer, L. C.; Oates, B.; Shah, S. K.; Loebach, J. L.; Mills, S. G.; MacCoss, M.; Castonguay, L.; Malkowitz, L.; Springer, M. S.; Gould, S. L.; DeMartino, J. A. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2469–2473.
- Finke, P. E.; Oates, B.; Mills, S. G.; MacCoss, M.; Malkowitz, L.; Springer, M. S.; Gould, S. L.; DeMartino, J. A.; Carella, A.; Carver, G.; Holmes, K.; Danzeisen, R.; Hazuda, D.; Kessler, J.; Lineberger, J.; Miller, M.; Schleif, W. A.; Emini, E. A. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2475–2479.
- Tagat, J. R.; McCombie, S. W.; Steensma, R. W.; Lin, S.; Nazareno, D. V.; Baroudy, B.; Vantuno, N.; Xu, S.; Liu, J. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2143–2146.
- Tagat, J. R.; Steensma, R. W.; McCombie, S. W.; Nazareno, D. V.; Lin, S. I.; Neustadt, B. R.; Cox, K.; Xu, S.; Wojcik, L.; Murray, M. G.; Vantuno, N.; Baroudy, B. M.; Strizki, J. M. *J. Med. Chem.* **2001**, *44*, 3343–3346.
- McCombie, S. W.; Tagat, J. R.; Vice, S. F.; Lin, S. I.; Steensma, R.; Palani, A.; Neustadt, B. R.; Baroudy, B. M.; Strizki, J. M.; Endres, M.; Cox, K.; Dan, N.; Chou, C. C. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 567–571.
- 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*, 5698–5703.
- Dragic, T.; Trkola, A.; Thompson, D. A.; Cormier, E. G.; Kajumo, F. A.; Maxwell, E.; Lin, S. W.; Ying, W.; Smith, S. O.; Sakmar, T. P.; Moore, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5639–5644.
- Finke, P. E.; Meurer, L. C.; Oates, B.; Mills, S. G.; MacCoss, M.; Malkowitz, L.; Springer, M. S.; Daugherty, B. L.; Gould, S. L.; DeMartino, J. A.; Siciliano, S. J.; Carella, A.; Carver, G.; Holmes, K.; Danzeisen, R.; Hazuda, D.; Kessler, J.; Lineberger, J.; Miller, M.; Schleif, W. A.; Emini, E. A. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 265–270.
- Dorn, C. P.; Finke, P. E.; Oates, B.; Budhu, R. J.; Mills, S. G.; MacCoss, M.; Malkowitz, L.; Springer, M. S.; Daugherty, B. L.; Gould, S. L.; DeMartino, J. A.; Siciliano, S. J.; Carella, A.; Carver, G.; Holmes, K.; Danzeisen, R.; Hazuda, D.; Kessler, J.; Lineberger, J.; Miller, M.; Schleif, W. A.; Emini, E. A. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 259–264.

28. Song, M.; Breneman, C. M.; Sukumar, N. *Bioorg. Med. Chem.* **2004**, *12*, 489–499.
29. Castonguay, L. A.; Weng, Y.; Adolfsen, W.; Di Salvo, J.; Kilburn, R.; Caldwell, C. G.; Daugherty, B. L.; Finke, P. E.; Hale, J. J.; Lynch, C. L.; Mills, S. G.; MacCoss, M.; Springer, M. S.; DeMartino, J. A. *Biochemistry* **2003**, *42*, 1544–1550.
30. 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*, 739–745.
31. Huang, X.; Xu, L.; Luo, X.; Fan, K.; Ji, R.; Pei, G.; Chen, K.; Jiang, H. *J. Med. Chem.* **2002**, *45*, 333–343.
32. Broughton, H. B. *J. Mol. Graph. Model.* **2000**, *18*, 247–257.
33. Cramer, M.; Cramer, R. D., III; Jones, M. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.
34. Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, *37*, 4130–4146.
35. Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219–3228.
36. Sybyl version 6.8; Tripos Associates; St. Louis, MO, 2000.
37. Insight II. 2000. User Guide, MSI Inc., San Diego, USA, 2000.
38. Thompson, J. D.; Higgins, D. G.; Gibson, T. J. *Nucleic Acids Res.* **1994**, *22*, 4673–4680.
39. Henikoff, S.; Henikoff, J. G. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 10915–10919.
40. Pearson, W. R. *Methods Enzymol.* **1990**, *183*, 63–98.
41. Huang, X. Q.; Jiang, H. L.; Luo, X. M.; Chen, K. X.; Ji, R. Y.; Cao, Y.; Pei, G. *Acta Pharmacol. Sin.* **2000**, *21*, 521–528.
42. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.
43. Castonguay, L. A.; Weng, Y.; Adolfsen, W.; Di Salvo, J.; Kilburn, R.; Caldwell, C. G.; Daugherty, B. L.; Finke, P. E.; Hale, J. J.; Lynch, C. L.; Mills, S. G.; MacCoss, M.; Springer, M. S.; DeMartino, J. A. *Biochemistry* **2003**, *42*, 1544–1550.
44. Tsamis, F.; Gavrilov, S.; Kajumo, F.; Seibert, C.; Kuhmann, S.; Ketas, T.; Trkola, A.; Palani, A.; Clader, J. W.; Tagat, J. R.; McCombie, S.; Baroudy, B.; Moore, J. P.; Sakmar, T. P.; Dragic, T. *J. Virol.* **2003**, *77*, 5201–5208.
45. Govaerts, C.; Bondue, A.; Springael, J. Y.; Olivella, M.; Deupi, X.; Le Poul, E.; Wodak, S. J.; Parmentier, M.; Pardo, L.; Blanpain, C. *J. Biol. Chem.* **2003**, *278*, 1892–1903.
46. Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Wang, J. M.; Ross, W. S.; Simmerling, C.; Darden, T.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 7. Users' Manual*. University of California, 2002.
47. Morris, G. M.; Goodsell, D. S.; Huey, R.; Hart, W. E.; Halliday, R. S.; Belew, R. K.; Olson, A. J. Autodock Version 3.0.3. The Scripps Research Institute, Molecular Graphics Laboratory, Department of Molecular Biology, 1999.
48. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1999**, *19*, 1639–1662.
49. Solis, F. J.; Wets, R. J. B. *Math. Oper. Res.* **1981**, *6*, 19–30.
50. Morris, G. M.; Goodsell, D. S.; Huey, R.; Olson, A. J. *J. Comput. Aided Mol. Des.* **1996**, *10*, 293–304.
51. Purcel, W. P.; Singer, J. A. *J. Chem. Eng. Data* **1967**, *12*, 235–246.
52. Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. *J. Comb. Chem.* **1999**, *1*, 55–68.
53. Perera, L.; Foley, C.; Darden, T. A.; Stafford, D.; Mather, T.; Esmon, C. T.; Pedersen, L. G. *Biophys. J.* **2000**, *79*, 2925–2943.
54. Orry, A. J. W.; Wallace, B. A. *Biophys. J.* **2000**, *79*, 3083–3094.
55. Huang, X.; Shen, J.; Cui, M.; Shen, L.; Luo, X.; Ling, K.; Pei, G.; Jiang, H.; Chen, K. *Biophys. J.* **2003**, *84*, 171–184.
56. Bissantz, C.; Bernard, P.; Hibert, M.; Rognan, D. *Proteins* **2003**, *50*, 5–25.
57. POV-Ray-Team. POV-Ray version 3, 1999 (www.povray.org).
58. Adamian, L.; Liang, J. *J. Mol. Biol.* **2001**, *311*, 891–907.
59. Ulmschneider, M. B.; Sansom, M. S. *Biochim. Biophys. Acta* **2001**, *1512*, 1–14.
60. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. *Protein Eng.* **1995**, *8*, 127–134.
61. Huang, X. Q.; Jiang, H. L.; Luo, X. M.; Chen, K. X.; Ji, R. Y.; Cao, Y.; Pei, G. *Acta Pharmacol. Sin.* **2000**, *21*, 521–528.
62. Chong, L. T.; Duan, Y.; Wang, L.; Massova, I.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14330–14335.
63. Wang, L.; Duan, Y.; Stouten, P.; De Lucca, G. V.; Klabe, R. M.; Kollman, P. A. *J. Comput. Aided Mol. Des.* **2001**, *15*, 145–156.