August 2013
Ph.D. Dissertation

Molecular Modeling Study of Human Chemokine CCR5 Receptor

Graduate School of Chosun University

Department of Bio-New Drug Development

Gadhe Changdev Gorakshnath

Molecular Modeling Study of Human Chemokine CCR5 Receptor

인간 케모카인 CCR5 수용체의 분자 모델링 연구

23rd August 2013

Graduate School of Chosun University

Department of Bio-New Drug Development

Gadhe Changdev Gorakshnath

Molecular Modeling Study of Human Chemokine CCR5 Receptor

Advisor: Prof. Seung Joo Cho

This dissertation is submitted to the Graduate School of Chosun University in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Science

April 2013

Graduate School of Chosun University

Department of Bio-New Drug Development

Gadhe Changdev Gorakshnath

This is to certify that the Ph.D. dissertation of Gadhe Changdev Gorakshnath has successfully met the dissertation requirements of Chosun University.

Chairman: **Chosun Univ, Prof. Song Yub Shin** ⋯⋯⋯⋯⋯

Member: **Chosun Univ, Prof. Eun Ae Kim** ⋯⋯⋯⋯⋯

Member: **Chosun Univ, Prof. Ryu Seol** ⋯⋯⋯⋯⋯

Member: **Chonnam Univ, Prof. Chang Hyun Đam** ⋯⋯⋯

Member: **Chosun Univ, Prof. Seung Joo Cho** ⋯⋯⋯⋯⋯

June 2013

Graduate School of Chosun University

Contents

List of Tables

List of Figures

Abstract

Molecular Modeling Study of Human Chemokine CCR5 Receptor

Gadhe Changdev Gorakshnath
Advisor: Prof. Seung Joo Cho, Ph.D.
Department of Bio-New Drug Development
Graduate School of Chosun University

Human Chemokine CCR5 receptor is an integral membrane protein which aid entry of human immunodeficiency virus type 1 (HIV-1) into host cell. HIV-1 causes acquired immunodeficiency syndrome (AIDS), which has infected millions of people worldwide. CCR5 is a membrane protein and a member of G-protein coupled receptor (GPCR) superfamily. Unavailability of CCR5 X-ray crystal structure hampers in-depth research of it. So, it is a challenging task to design inhibitors for this integral membrane protein as well as to find out exact binding mode of action for any known inhibitor. To address this problem, we developed a homology model for the human CCR5, based on the recently reported CXCR4 (PDB code: 3ODU) X-ray crystal structure. This was subsequently used to map the binding mode of several CCR5 inhibitors. The piperidine-4-carboxamide derivatives were used for receptor based 3D-QSAR studies, whereas oxamino-piperidino-piperidine amide (OPPA) derivatives were used for MD simulations. 3D-QSAR methods such as, comparative molecular field analysis

-xi-

(CoMFA) and comparative molecular similarity indices analysis (CoMSIA) were performed using receptor based alignments for all the inhibitors. The CoMFA model was obtained with the good statistics (q^2=0.722, NOC=4, Q^2=0.619 and r^2=0.884), while in CoMSIA, SHDA model (q^2=0.712, NOC=4, Q^2=0.522 and r^2=0.825) gave satisfactory results. The robustness of the developed CoMFA and CoMSIA models was checked with the test set molecules, which were excluded during the model development (r^2_{pred} for CoMFA, 0.680 and CoMSIA, 0.705). The CoMFA and CoMSIA contour maps were plotted back onto CCR5 and structural variations with biological activity were explained in the presence of CCR5. Molecular dynamics (MD) simulation for the three OPPA derivatives were performed by embedding complexes (ligands-CCR5) into lipid bilayers. 20 nanosecond MD simulations were performed independently for each complex. MD simulation identifies salt bridge contact between tertiary nitrogen of inhibitors and carboxyl of Glu283 is important. Residues from TM1-TM3 and TM5-TM7 participated in the close interactions. Our MD simulation results are in agreement with previous experimental studies. Combined usage of 3D-QSAR and MD simulation could be efficient to guide the design and development of novel inhibitors.

Keywords: CCR5, Homology modeling, Docking, QSAR, MD simulation

초 록

인간 케모카인 CCR5 수용체의 분자 모델링 연구

가데 창데브 고라그스나트
지도교수: 조 승 주
바이오신약개발학과
조선대학교 대학원

인간 케모카인 CCR5 수용체는 후천성면역결핍 바이러스 1 형이 숙주세포에 들어가는 것은 돕는 통합된 막단백질이다. HIV-1 은 후천성 면역결핍증을 유발하여 전세계적으로 수백만명의 사람들이 감염되었다. CCR5 는 막단백질이면서, G 단백질에 결합된 수용체군의 하나이다. CCR5 의 x 선구조가 밝혀지지 않았기 때문에, 이 단백질의 깊은 이해가 어렵다. 따라서 이 단백질의 억제제가 어떻게 결합하는지와, 나아가서는 어떠한 기작을 가지고 있는지를 연구하는 것이 중요하다. 이 문제를 해결하기 위하여, 본 연구에서는 최근에 보고된 비슷한 CXCR4 의 구조에 근거하여, 인간 CCR5 에 대한 호몰로지 모델을 만들었다. 이것은 다양한 CCR5 억제제의 결합모드를 결정하는데 사용되었다. piperidine-4-carboxamide 유도체들은 수용체기반의 3D-QSAR 연구에 사용되었고, oxamino-piperidino-piperidine amide (OPPA) 유도체들은 분자동력학적인

연구를 하였다. 3D-QSAR 모델로는 수용체기반의 배열을 활용하여 비교분자장 분석(CoMFA)과 비교분자 유사성인덱스 분석(CoMSIA)과 같은 방법을 사용하였다. CoMFA (q2=0.722, NOC=4, Q2=0.619 and r2=0.884)와 CoMSIA (q2=0.712, NOC=4, Q2=0.522 and r2=0.825) 의 결과로 얻은 통계량은 만족할 만 하였다. 이 모형의 안정성은 검증집합 분자들로서 다시 검증하였다. (r2pred CoMFA, 0.680; CoMSIA, 0.705) CoMFA 와 CoMSIA 의 컨투어 맵은 CCR5 속에 그렸고, CCR5 와의 구조활성관계를 얻을 수 있었다. 3 개의 OPPA 유도체들의 분자동력학적 시뮬레이션은 리간드와 CCR5 의 컴플렉스모형은 리피드 2 층으로 넣음으로서 행할 수 있었다. 20 나노초동안의 분자동력학적인 시뮬레이션이 각각의 컴플렉스에 대하여 행하였졌다. 이 결과로 억제제들에 있는 터셔리 아민과 Glu283 의 카복실기가 중요하다는 것을 알았다. TM1-TM3 의 잔기들과 TM5-TM7 의 잔기들이 이들의 상호작용에 중요하다는 것을 알았다. 본 연구의 결과는 기존의 연구결과와 같은 결론을 었었다. 3D-QSAR 과 MD 시뮬레이션을 합하면, 새로운 억제제를 설계하고 개발하는데 효율적일 수 있다.

키워드 : CCR5, 상동 모델링, 도킹, QSAR, MD 시뮬레이션

1. Introduction

G-protein coupled receptor (GPCR) is a largest pool of membrane proteins structures in human genome, which represents the richest source of drug targets for the pharmaceutical industry (1). Previous report revealed that the human genome contains 800 unique GPCRs, of which 460 were predicted to be olfactory receptors (2). On the basis of sequence similarity within seven transmembrane (TM), these receptors can be clustered into five families, namely: rhodopsin family (701 receptors), adhesion family (24 receptors), frizzled/taste family (24 receptors), secretin family (15 receptors) and glutamate family (15 receptors) (2). Chemokines are classified into the subgroups based on the conserved N-terminal cysteine such as CXC, CC, C and CX3C (3). The chemokines regulate development, activation, and recruitment of leukocyte through binding and activation of seven transmembrane (TM) GPCRs (4).

Reports on joint United Nations program of acquired immunodeficiency syndrome (UNAIDS) shows that by the end of 2011, 34 million people were living with the HIV/AIDS. 2.5 millions new HIV/AIDS infections were occurred in 2011 and 1.7 millions AIDS related death occurred (5). AIDS remain world's major health challenge and majority of infections are from sub-Saharan Africa. For the treatment of acquired immunodeficiency syndrome (AIDS) along with reverse transcriptase (RT) and protease (PR) inhibitors, integrase (IN), CCR5 was found to be an important target for AIDS treatment. Today's most useful drug treatment is highly active antiretroviral therapy (HAART) (6), which is targeted against RT and PR enzyme. There are more

than 20 drugs available for the treatment of HIV/AIDS such as nucleoside reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs), protease inhibitors (PIs), integrase inhibitors, entry inhibitors (including fusion inhibitors), maturation inhibitors and cellular inhibitors. However, toxicities and emergence of drug resistant virus necessitate the examination of complimentary target for therapeutic development (7, 8).

HIV primarily infects vital cells in the human immune system such as helper T-cells (specifically CD4+ T-cells), macrophages and dendritic cells (9, 10). It enters into these cells by the adsorption of glycoprotein (11) on its surface to receptors on the target cell followed by fusion of the viral envelope with the cell membrane and the release of the HIV capsid into the cell (12). HIV infection leads to low levels of CD4+ T-cells through three main mechanisms: Firstly, direct viral killing of infected cells; secondly, increased rates of apoptosis in infected cells; and thirdly, killing of infected CD4+ T-cells by CD8 cytotoxic lymphocytes that recognize infected cells. CCR5 receptor is a member of GPCR superfamily (13, 14). It has been identified as a primary co-receptor on CD4+ cells for entry of macrophage-tropic (M-tropic or R5) HIV-1 isolates. Individuals homozygous for a 32-base pair deletion in the gene for CCR5 do not express functional receptor on their cell surfaces and have been identified as being highly resistant to HIV-1 infection whereas heterozygous individuals, who possess only one intact CCR5 allele advance more slowly to AIDS compared with patients having no deletion (15-17).

However, no structural information regarding the precise CCR5 binding site of any known ligand is available, and we considered that a detailed understanding of the binding modes of existing inhibitors would help design more potent antagonists. A number of studies have been undertaken on CCR5 modeling using computational methods, in which approaches, such as quantitative structure-activity relationship (QSAR) analysis, pharmacophore modeling, virtual screening, molecular docking, molecular dynamics simulation, and membrane simulations, were used to deduce the interaction mechanisms of novel inhibitors. Structure-based methods are expected to provide better results than ligand-based methodologies. Virtual screening and QSAR studies based on 1-(3,3-diphenylpropyl)-piperidinylamide derivatives have shown that for high affinity binding, chemical and structural requirements can be identified using one dimensional (1D) physicochemical properties, 2D topological descriptors, and 3D properties, such as steric, electrostatic, hydrophobic, and hydrogen bond acceptor/donor fields around a family of aligned molecules (18, 19). Furthermore, combined molecular modeling techniques, such as comparative receptor modeling, 3D-QSAR, docking, and virtual screening have been used to identify novel entry inhibitors (20), and a docking-based 3D-QSAR approach has been utilized (21) with protein modeling, MD simulation, automated docking, and 3D-QSAR analysis to investigate the detailed interaction of CCR5 with its antagonists. Previous report (22) on dual antagonists docking into the CCR2 and CCR5 suggested the conserved interaction patterns among chemokine receptors. Recently, the design of CCR5 allosteric antagonist's mutagenesis

and modeling approach was successfully used, which shows N-terminal region and ECL2 are important in the interactions (23). Homology modeling, docking and MD simulation study was also performed by Shahlaei et al. (24).

To develop novel antagonistic agents with high potency and selectivity, detailed information is required about structural features influencing biological activity towards CCR5 receptor. To date, crystal structure of CCR5 or its complex with any ligand is not available and interactions between the CCR5 receptor and its antagonists are not known. In the absence of crystal structure of target protein, homology modeling could be a better approach to generate 3D model of target protein. The 3D molecular structures of ligand-protein complexes are considered foundations of structure-based drug design. Frequently, 3D data are available for the protein and drug separately, but infrequently for the protein/drug complex. Molecular docking is a process that enables a ligand and its receptor protein to be arranged in their near native 3D conformation. MD simulations can provide intimate details of individual particles motions as a function of time, and thus, can be used to tackle specific questions about ligand–receptor interactions without resorting to laboratory-based experimentation. Although experiments play a crucial role in the validation of simulation methodologies, as comparisons of simulation and experimental data serve to confirm the validities of calculated results and provide criteria for improving simulation methods.

Here, we performed 3D modeling of CCR5, docking study of known inhibitors, QSAR and MD simulation in explicit lipid bilayers, to know the structural requirement for potent inhibitors and important interacting residues.

2. Materials and Methods

2.1. Sequence alignment

Primary sequence of human CCR5 was retrieved from UniProt database (Accession number: P51681) (http://www.uniprot.org/). The protein BLAST (25) of CCR5 sequence was performed to identify the known homologous three dimensional (3D) structure, which can be further used as a template to model CCR5. Prediction of the TM helices is a difficult task and plays a major role in the investigation of membrane proteins. In protein data bank (PDB) relatively small numbers of crystal structures are reported for the membrane proteins, because of the difficulty in crystallization process (26). The recently reported crystal structure of CXCR4 (PDB code: 3ODU) (27) was used as a template structure to model the human CCR5. Sequence alignment was performed using CLUSTALW 2 program (http://www.ebi.ac.uk/Tools/msa/clustalw2) (28).

2.2. Homology modeling

Actual homology modeling was performed using MODELLER 9v4 program (29). Alignment obtained from CLUSTALW program was used as input and 100 models were developed. MODELLER derived models composed of non-hydrogen atoms. Final model for further analysis was chosen based on the lowest molecular probability density function (MolPdf) score and lowest backbone root mean square

deviation (RMSD) with the template. Stereochemical property of the selected model was assessed using PROCHECK (30), ERRAT (31), ProSA (32), evaluation servers.

2.3. Molecular docking

Docking study for the known inhibitors (TAK779 and TAK220, Figure 1) of CCR5 was performed. However, for receptor based 3D-QSAR, docking study of potent piperidine-4-carboxamide (Table 1-3) (33-36) derivative (compound **85**) was performed in CCR5. For MD simulation, oxamino-piperidino-piperidine amide (OPPA) (37, 38) derivatives were used and docking study of 14-OPPA, 25-OPPA and 37-OPPA (Figure 1) was performed.
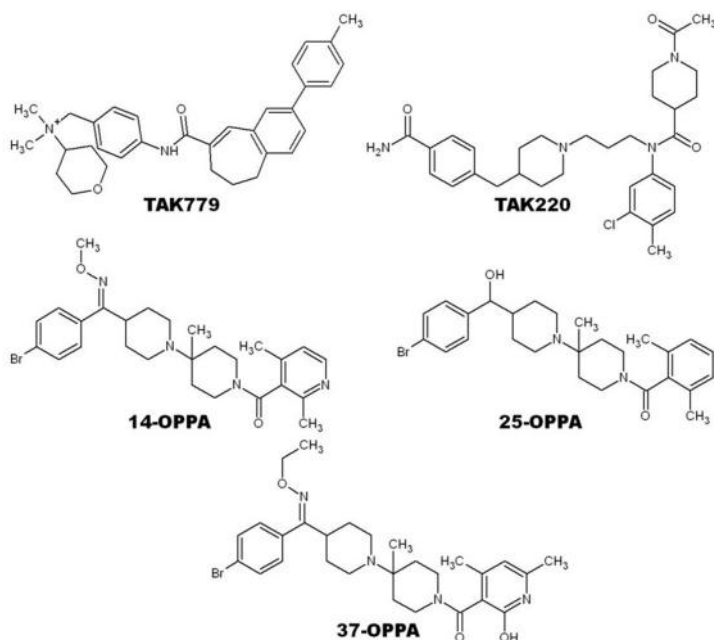
Figure 1: Chemical structure of CCR5 inhibitors.

Binding site for docking experiment was defined based on the previous mutational reports (39-42), as well as knowledge of important acidic amino acid residue (Glu283) in TM7. All the molecules were drawn using SYBYL 8.1 (43) modeling package and minimized using conjugate gradient minimizer with Tripos force field, finally partial atomic charges were assigned using Gasteiger-Hückel method. Autogrid was used to assign grid points, center of grid was assigned onto the OE2 of the Glu283, and box was expanded in such a way so that it covers all the important binding site residues. AUTODOCK (44) program with default settings was used for the docking calculations and Lamarckian genetic algorithm was employed. Final docked conformations of ligands were selected from the lowest energy cluster for further computational study.

2.4. Data set

Data set for 3D-QSAR was selected from previous report (33-36) (Table 1-3). The docked conformation of compound **85** was used to modify other compounds of dataset. The common substructure of template molecule was used to draw the molecules, and variable part was minimized inside CCR5 binding site. Gasteiger-Hückel partial charges were applied to each molecule. All original IC_{50} values of each inhibitor were converted into pIC_{50} ($-logIC_{50}$) in order to use the data as the dependent variable in both CoMFA and CoMSIA models. Dataset consists of 122 molecules with wide range of activity (pIC_{50} = 5.28-9.53), which was divided into training (97

compounds) and test (25 compounds) set by maximum dissimilarity algorithm. QSAR models were generated using training set and validated by test set.

2.5. CoMFA and CoMSIA

CoMFA was designed by the Cramer *et al.* (45) in 1988, which can be used on aligned data set to develop 3D-QSAR model. CoMFA calculates steric and electrostatic fields at each lattice intersection points and grid box dimensions were expanded in each direction by 4 Å from the coordinates of each molecule. 0.5 Å grid spacing was used. Steric field in CoMFA was calculated by Lennard Jones potential, whereas electrostatic field was calculated by Columbic potential. Generated fields were scaled by CoMFA_STD setting with the default energy cutoff +30 kcal/mol. The distance dependent dielectric constant 1.00 was used. A sp^3 hybridized carbon atom having +1 charge served as a probe for calculating steric and electrostatic fields. Any singularities were avoided at the atomic position in the CoMSIA (46) field, because a Gaussian type distance dependence of physicochemical properties was adopted. Hence no arbitrary cutoff was required. Similarity indices were calculated using probe atom with charge +1, radius 1 Å, hydrophobicity +1, H-bond donor and acceptor property +1, and 0.3 attenuation factor (a) for Gaussian type distance.

Table 1 Structures of CCR5 inhibitors their actual activities and predicted activities**.**

Sr. No.	R^1	R^2	X	R^3	Actual activity pIC$_{50}$	Predicted activity	
						CoMFA	CoMSIA
01*		-	CH	CO(4-F-Ph)	5.72	6.26	6.15
02		-	CH	Ph	5.44	5.77	6.13
03*		-	COH	4-Cl-Ph	5.28	6.02	5.12
04*		-			5.68	6.20	5.58
05		-	CH	CH$_2$Ph	6.32	6.12	6.23
06		-	CH	(CH$_2$)$_2$Ph	5.96	6.11	5.49
07*		-	CH	OPh	5.80	6.31	6.19
08		-	CH	OCH$_2$Ph	5.82	5.97	5.34
09		-	CH	CH$_2$(4-F-Ph)	6.51	6.43	6.26

Table 2 Structures of CCR5 inhibitors their actual activities and predicted activities.
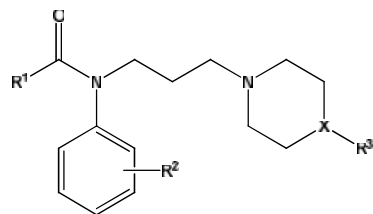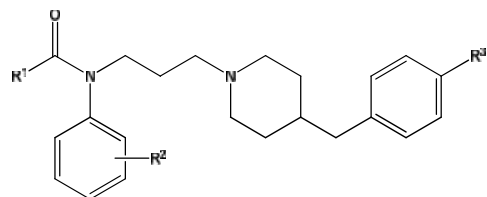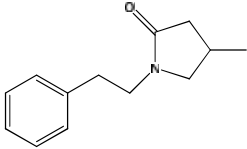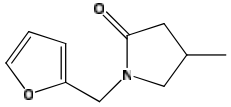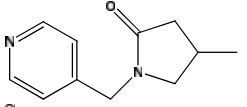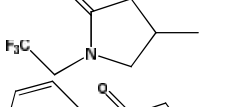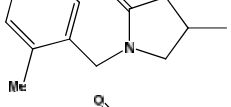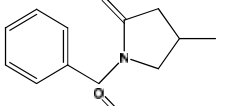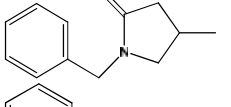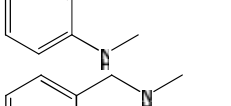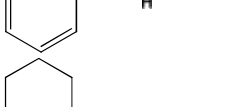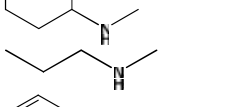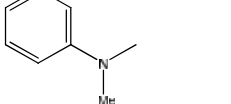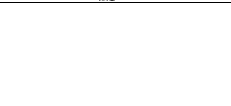
Sr. No.	R¹	R²	R³	Actual Activity	Predicted Activity	
				pIC$_{50}$	CoMFA	CoMSIA
10		2-Me	H	5.33	5.90	6.23
11		3-Me	H	6.80	6.31	6.46
12		4-Me	H	7.05	6.63	6.49
13		4-*t*-Bu	H	6.80	6.39	6.73
14		3,4-(CH$_2$)$_3$-	H	6.74	6.54	6.57
15*		4-MeO	H	5.82	6.23	6.25
16		3-Cl	H	6.92	6.28	6.64
17		4-Cl	H	5.92	6.43	6.46
18		3,4-diCl	H	7.24	6.58	6.86
19		3-Cl, 4-F	H	6.22	6.24	6.71
20		3,4-diF	H	5.68	6.14	6.43
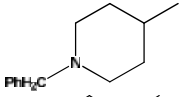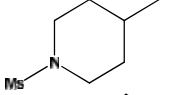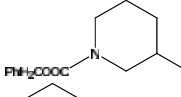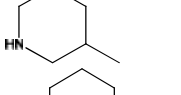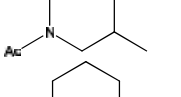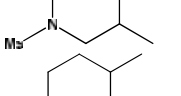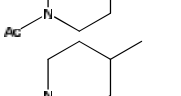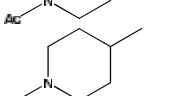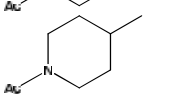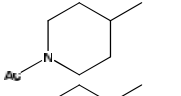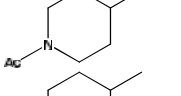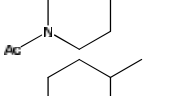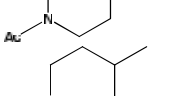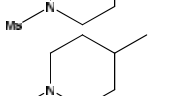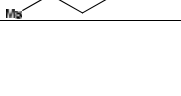
21*		3-CF$_3$	H	6.29	6.07	6.55
22		4-CF$_3$	H	5.46	6.20	6.60
23		3-CN	H	5.39	5.93	6.09
24		3,4-diCl	F	7.30	6.81	6.90
25		H	H	6.89	6.78	7.03
26		H	H	6.96	7.26	7.03
27		H	H	6.64	7.03	7.18
28		H	H	6.51	6.49	6.61
29		H	H	7.42	7.27	7.09
30		H	H	7.48	7.50	7.11
31		H	H	7.07	7.17	7.17
32		H	H	6.66	6.67	7.15
33		H	H	6.48	7.15	7.10
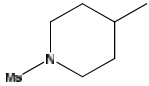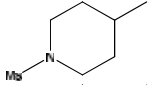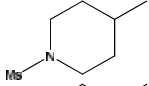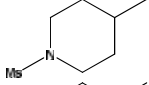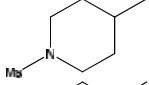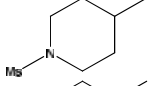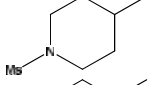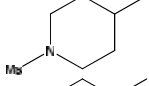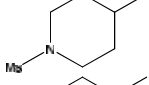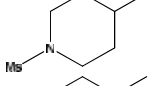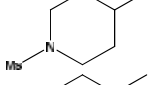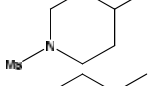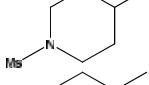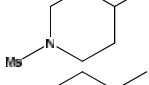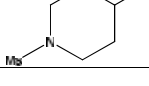
34		H	H	6.44	6.30	6.28
35*		H	H	7.09	7.31	6.74
36*		H	H	6.62	7.24	6.51
37		H	H	6.24	6.10	5.63
38		H	H	7.12	6.43	6.44
39		H	H	7.47	7.26	7.40
40		3-Cl	H	7.36	7.44	7.51
41		3,4-diCl	H	7.37	7.77	7.98
42		H	H	7.74	7.24	7.33
43		H	H	6.82	6.61	7.47
44*		H	H	7.62	7.02	7.59
45		H	H	7.21	7.07	7.16
46		H	H	5.18	5.03	5.69

47		H	H	5.64	5.75	5.83
48*		H	H	6.35	6.08	6.06
49		H	H	6.04	6.19	6.09
50*		H	H	5.70	6.26	6.24
51		H	H	8.23	7.82	7.75
52		H	F	8.11	8.03	7.79
53		H	F	7.89	7.56	7.43
54		H	F	7.74	8.03	7.89
55		H	F	8.18	7.76	7.71
56		H	F	7.12	7.57	7.76
57*		H	F	7.85	7.71	7.79
58		H	F	7.82	8.06	7.88
59*		H	F	6.55	7.86	7.93
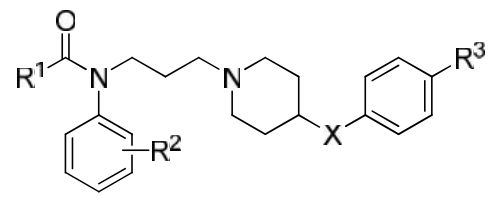
No.	Structure	R1	R2			
60	HOOC-phenyl-NH	H	F	6.46	7.02	7.53
61	H₂NOC-phenyl-NH	H	F	7.54	7.29	6.92
62	MeO-phenyl-NH	H	F	7.52	7.58	7.84
63	MeS-phenyl-NH	H	F	7.82	8.00	8.06
64*	MeO₂S-phenyl-NH	H	F	7.59	7.69	8.48
65	Cl-phenyl-NH	4-Me	H	7.96	8.33	8.00
66	Cl-phenyl-NH	3-Cl	H	7.72	7.56	7.97
67*	Cl-phenyl-NH	3,4-diCl	H	7.21	8.26	8.39
68	Cl-phenyl-NH	H	SO₂Me	8.92	8.51	8.60
69*	Cl-phenyl-NH	H	SO₂(morpholino)	9.00	8.48	8.59
70	PhH₂CCOC-piperidine	H	H	7.32	7.21	7.14
71	HN-piperidine	H	H	5.37	5.95	5.41
72	H₂COC-piperidine	H	H	7.80	8.05	7.83
73	isobutyryl-piperidine	H	H	7.41	6.98	6.79
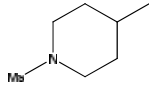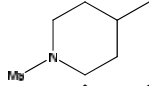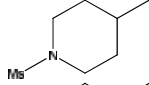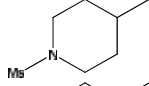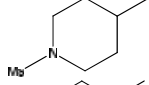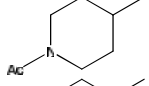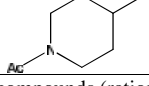
74		H	H	7.48	7.33	7.13
75		H	H	7.40	7.52	7.63
76		H	H	5.92	5.96	6.01
77		H	H	6.11	5.96	5.85
78		H	H	6.17	5.85	6.25
79*		H	H	6.74	5.74	5.63
80		3-Cl	F	8.52	8.46	8.27
81		4-Me	F	8.55	8.81	8.11
82		3,4-diCl	H	8.72	8.51	8.46
83		3,4-diCl	F	8.92	8.72	8.50
84		3-Cl, 4-Me	H	9.21	8.74	8.48
85		3-Cl, 4-Me	F	9.54	8.85	8.41
86		3-Cl, 4-*i*-Pr	F	7.74	8.80	8.70
87*		3-Cl, 4-MeO	F	7.52	8.60	8.28
88		3,4-diCl	F	8.48	8.16	8.27
89		3-Cl, 4-Me	F	9.30	8.42	8.31

90		3,4-diCl	CF_3	8.06	7.91	8.40
91		3,4-diCl	NO_2	8.62	8.32	8.47
92		3,4-diCl	NH_2	8.04	8.22	7.91
93		3,4-diCl	NHAc	8.23	8.38	7.92
94		3,4-diCl	NHMs	8.66	8.37	8.55
95		3,4-diCl	Morpholino	8.55	8.68	8.85
96		3,4-diCl	OMe	8.10	8.30	8.30
97		3,4-diCl	SMe	8.51	8.53	8.46
98		3,4-diCl	SO_2Me	8.66	8.52	9.03
99		3,4-diCl	SO_2Et	8.72	8.81	9.15
100*		3,4-diCl	$SO_2i\text{-}Pr$	8.82	8.55	8.58
101		3,4-diCl	SO_2NH_2	8.47	8.62	8.56
102		3,4-diCl	SO_2NHMe	8.82	8.92	8.76
103		3,4-diCl	SO_2NMe_2	8.92	8.92	8.78
104		3,4-diCl	$SO_2(morpholino)$	8.89	9.22	8.89

Table 3 Structures of CCR5 inhibitors their actual activities and predicted activities**.**

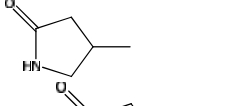| Sr. No.	R^1	R^2	X	R^3	Actual Activity pIC$_{50}$	Predicted Activity CoMFA	Predicted Activity CoMSIA
105		3,4-diCl	CH$_2$	4-Ms	8.66	8.65	8.74
106*		3,4-diCl	S	4-F	8.77	7.76	7.83
107		3,4-diCl	SO	4-F	8.49	8.98	8.59
108*		3,4-diCl	SO$_2$	4-F	8.54	8.06	8.23
109*		3,4-diCl	NH	4-F	7.70	7.73	7.33
110*		3,4-diCl	NHSO$_2$	4-F	8.34	7.73	6.84
111		3,4-diCl	NHCO	4-F	6.14	8.06	8.21
112		3,4-diCl	CH$_2$	4-CN	8.77	7.90	7.94
113		3,4-diCl	CH$_2$	4-CO$_2$Me	8.34	8.73	8.39
114		3,4-diCl	CH$_2$	4-COOH	7.18	6.82	5.88
115		3,4-diCl	CH$_2$	4-CONH$_2$	8.05	8.40	8.06

116		3,4-diCl	CH_2	3-$CONH_2$	8.55	8.24	8.22
117*		3,4-diCl	CH_2	2-$CONH_2$	7.92	7.96	7.90
118		3,4-diCl	CH_2	4-CONHMe	8.29	7.90	8.14
119		3,4-diCl	CH_2	4-CONHt-Bu	7.96	8.06	8.05
120		3,4-diCl	CH_2	4-$CONMe_2$	7.72	7.94	8.00
121		3,4-diCl	CH_2	4-$CONH_2$	8.42	8.44	8.64
122*		3-Cl, 4-Me	CH_2	4-$CONH_2$	8.46	8.66	8.66

Test set compounds (rationally designed) indicated by *

2.6. Partial Least Square (PLS)

In CoMFA and CoMSIA regression analysis was carried out using PLS method (47). To derive 3D-QSAR models, CoMFA and CoMSIA descriptors were used as an independent variable and pIC_{50} value as dependent variable. The cross-validation analysis was performed using leave-one-out (LOO) method, wherein one compound was removed from the dataset and its activity was predicted using the model derived from rest of the molecules. To speed up analysis and to reduce the noise, column with values (r) below 2.0 kcal/mol was filtered off. Non-cross-validated analysis was performed with the optimum number of components (ONC), which were obtained during LOO cross-validation. The q^2 value was selected from the analysis

which had lowest standard error of prediction (SDEP) and ONC. q^2 was calculated from the following formula.

$$q^2 = 1 - \frac{\sum_{x}(x_{pred} - x_{actual})^2}{\sum_{x}(x_{actual} - x_{mean})^2}$$

where, $_{pred}$, $_{actual}$ and $_{mean}$ are predicted, actual, and mean values of the target property (pIC$_{50}$) respectively.

Progressive scrambling (48) was performed to validate the model and to check the dependency of chance correlation. Fifty independent scrambling were carried out with the critical point value of 0.85 (49). The non-parametric redundancy corrected internal prediction parameter (Q^2) was also calculated.

2.7. Predictive correlation coefficient (r^2_{pred})

Predictive ability of the developed CoMFA and CoMSIA models were assessed using twenty five test set compounds. Geometry optimization and energy minimization of test set compounds were same as the training set compounds. The predictive correlation coefficient (r^2_{pred}), based on the test set molecules was computed by using following formula.

$$r^2_{pred} = \frac{(SD - PRESS)}{SD}$$

where, SD is sum of squared deviation between the biological activity of test set molecules and mean activity of the training set molecules. PRESS is the predictive

residual sum of square and is calculated by taking difference in predicted and actual activity of test set molecules. For all conventional analysis (non cross validation) the 'minimum sigma' standard deviation threshold was set to 2.0 kcal/mol.

2.8. MD simulation

MD simulations were carried out using GROMACS 4.5.1 simulation package (50) with 53A6 GROMOS96 force field (51). Ligand topology and coordinates was generated using PRODRG (http://davapc1.bioch.dundee.ac.uk/prodrg/) server (52). Complexes to be simulated (14-OPPA-CCR5, 25-OPPA-CCR5 and 37-OPPA-CCR5) were packed in explicit bilayers of 128 dipalmitoylphosphatidylcholine (DPPC). The topology and coordinates of DPPC were taken from the previously described source of Tieleman et al. (http://moose.bio.ucalgary.ca) (53). Complex structure was inserted into the center of DPPC bilayers, by keeping -helices parallel to Z-axis using INFLATEGRO script (54). Overlapping lipid molecules (2 from upper and 4 from lower layer) were deleted, and alternate compression (26 times) and energy minimization was performed to achieve the desired results. Compressed complex-DPPC systems were hydrated (2486 water molecules) using pre equilibrated simple point charge (SPC) water model (55) with periodic boundary conditions. The total charge of each system was neutralized with the 14 Cl atoms by replacing water molecules. Final system had dimensions 67.4   67.6   61.2 Å.

Constant volume, constant temperature (NVT) ensemble was adopted at 0.1 ps coupling constant and 323 K temperature with the time duration of 500 ps.

Berendsen coupling algorithm was used to heat system (56), with separate coupling of solutes (protein, ligand and DPPC) and solvents (Water and Cl). The SETTLE algorithm (57) was used to keep water molecule bond lengths at their equilibrium values, and the LINCS algorithm (58) was employed to keep all remaining bonds constrained. The time step for integrating the equations of motion was 0.002 ps. The particle mesh ewald (PME) method with a cutoff value of 12 Å was used to calculate long-range electrostatics (59). Followed by NVT ensemble, constant pressure, constant temperature (NPT) simulation was invoked. Pressure was coupled semi-isotropically (coupling constant 1.0 and compressibility 4.5 X 10^{-5} bar^{-1}), which resulted in independent coupling of lateral and perpendicular pressure. In this simulation, 5.0 ps coupling constant, 1 bar constant pressure was used for 1000 picoseconds (ps). After stabilization of both temperature and pressure, production run for 20000 ps was performed using a Nose´-Hoover thermostat [60] at 323 K and a Parrinello-Rahman barostat at 1 bar (61).

3. Results

3.1. Analysis of sequence alignment

Blast search shows that 3ODU as top template with 35 % sequence identity, which is higher than the traditionally used bovine rhodopsin structure and $_2$ adrenergic receptors. In blast search lower E value ($1e^{-33}$) is considered to be good, and it reflects alignment is real and not occurred by chance. 3ODU was chosen as a template structure to model CCR5. CLUSTALW program was employed to generate alignment between target (CCR5) and template (3ODU) structure, and resultant alignment is shown in Figure 2. Obtained alignment was checked for the correct transformation of TM regions.


```
                                                       TM1
CCR5  --MDYQVSSPIYDINYYTS----------------EPCQKINVKQIAARLLPPLYSLVFI 42
3ODU  DYKDDDDAGAPEGISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLPTIYSIIFL 60
        *  :  ::..   .*  ***          ***  :  :  ::   :**.:**.:*:
                              TM2
CCR5  FGFVGNMLVILILINCKRLKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMCQ 102
3ODU  TGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPFWAVDAVANWYFGNFLCK 120
      *.*** ****::    *:*.**** * *.*.::**.*:.:*.**** *.*:* *** .*:
             TM3                                        TM4
CCR5  LLTGLYFIGFFSGIFFIILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLP 162
3ODU  AVHVIYTVNLYSSVWILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLLTIP 180
      :  :*.:::*.:::: ::::*****.***. : .*. *.  *.*: *.: ::*
           TM5
CCR5  GIIFTRSQKEGLHYTCSSHFPYSQYQFWKNFQTLKIVILGLVLPLLVMICYSGILKTLL 222
3ODU  DFIFANVSEADDRYICDRFYPN---DLWVVFQFQHIMVGLILPGIVILSCYCIIISKLS 237
      .:**:. .: . :* *. .:*     ::*  :: ::*.**.:** :*:: **. *:..*
                                TM6
CCR5  RCRNEKKRHRAVRLIFTIMIVYFLFWAPYNIVLLLNTFQEFFGLNN-CSSSNRLDQAMQV 281
3ODU  HSKGHQKRK-ALKTTVILILAFFACWLPYYIGISIDSFILLEIIKQGCEFENTVHKWISI 296
      :.:..:**: *::   . :::.:* * ** * : ::*  :   ::   .* :.: .:
           TM7
CCR5  TETLGMTHCCINPIIYAFVGEKFRNYLLVFFQKHIAKRFCKCCSIFQQEAPERASSVYTR 341
3ODU  TEALAFFHCCLNPILYAFLGAKFK------------------TSAQHALTSGRPLEVLFQ 338
      **:*.: ***:***.***:* **:                  *  : :. *. .* :

CCR5  STGEQEISVGL 352
3ODU  -----------
```

Figure 2: Sequence alignments between CCR5 and 3ODU. TM helices positions were indicated at the top of alignment by grey cylinder.

3.2. Homology model analysis

Sequence alignment generated by CLUSTALW was used as an input to develop 3D-model of CCR5 using MODELLER. 100 models were generated for CCR5. Best model was chosen based on the lower molecular probability density function (MolPdf) score, and lesser backbone root mean square deviations (RMSD) (0.179 Å). Cross-checking was done to ensure the correct TM prediction according to the template structure. The conserved disulfide bonds (Cys20-Cys269 and Cys101-Cys178) were generated in model. Selected model for further study was checked for its stereochemical quality (Figure 3), non-bonded interaction pattern (Figure 4) and knowledge based energy profile (Figure 5).
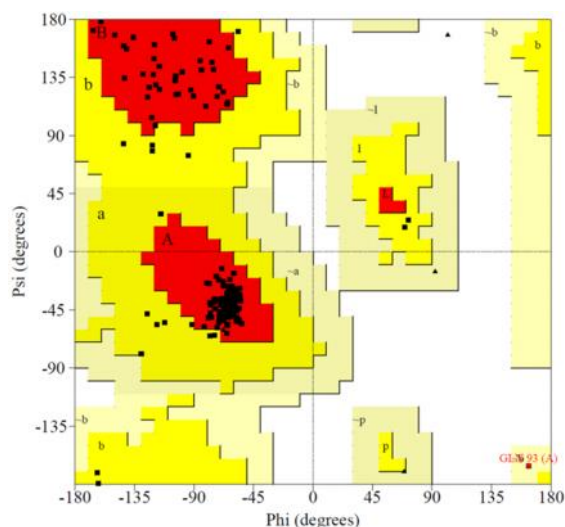
Figure 3: PROCHECK plot of the derived 3D-CCR5 model. The red color indicates residues in the most favored region, yellow indicates residues in an additionally

allowed region, faint yellow indicates residues in a generously allowed region, and white indicates the disallowed region.
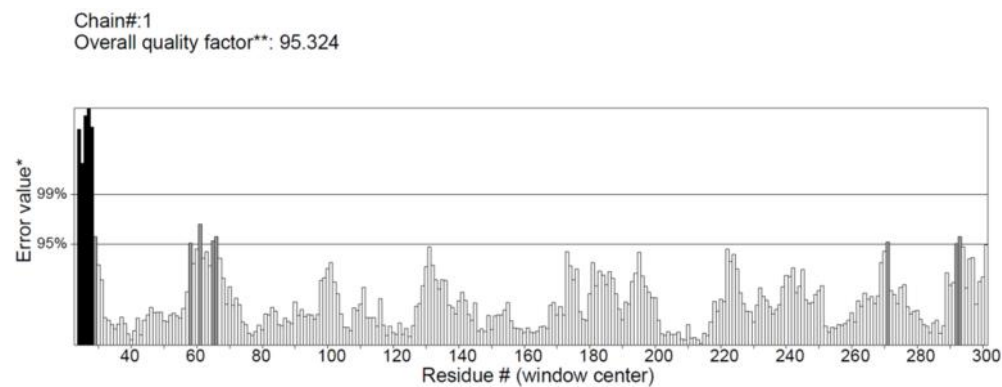
Figure 4: ERRAT plot for the CCR5 model.
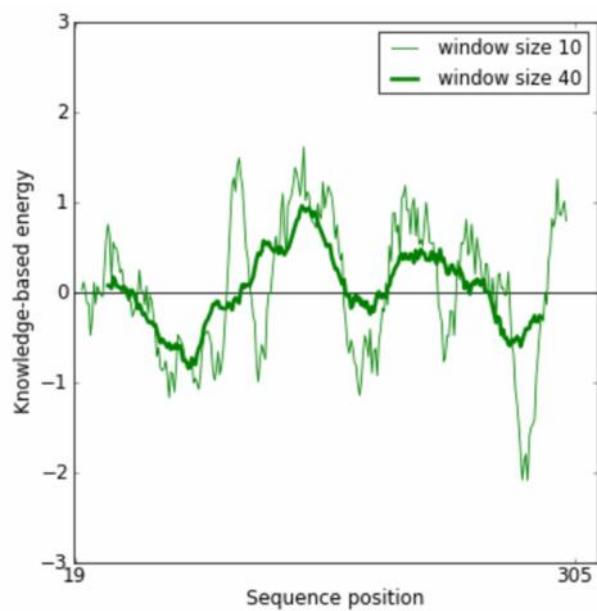
Figure 5: ProSA energy plot (knowledge based) for the CCR5 model (Z-score = -3.03).

Ramachandran plot shows that 99.6 % residues were in most favored and additionally allowed regions, whereas 0.4 % residues were in generously allowed region and no residue was in a disallowed region. ERRAT plot analysis showed an overall model quality of 95.32 %, indicating good model according to template structure plot. Another structure validation parameter, namely, goodness factor (G-factor) was explored. For a good structure, the overall value of G-factor should be above -0.5 and/or close to zero. Our developed CCR5 model had a G-factor value of zero, indicating that the developed model is of good quality. Developed model was validated using ProSA knowledge based energy, which obtained overall Z-score of -3.03. From the above analyses we concluded that the developed model was good for further computational analysis.

3.3. Docking analysis

Molecular docking is a most powerful mean to analyze interaction between two molecules. From the mutational reports on CCR5, it is apparent that the binding site was located in TM1 (Tyr37), TM2 (Trp86, Tyr89), TM3 (Tyr108, Phe109, Phe113), TM5 (Ile198), TM6 (Tyr251) and TM7 (Glu283) with the partial involvement of extracellular loop 2 (ECL2).

3.3.1. Binding mode of TAK779

Binding mode of TAK779 inside CCR5 active site was shown in Figure 6. Binding mode was selected from the topmost cluster with more negative binding

energy. It was observed that quaternary 'N' of TAK779 make salt bridge contact (4 Å) with Glu283.
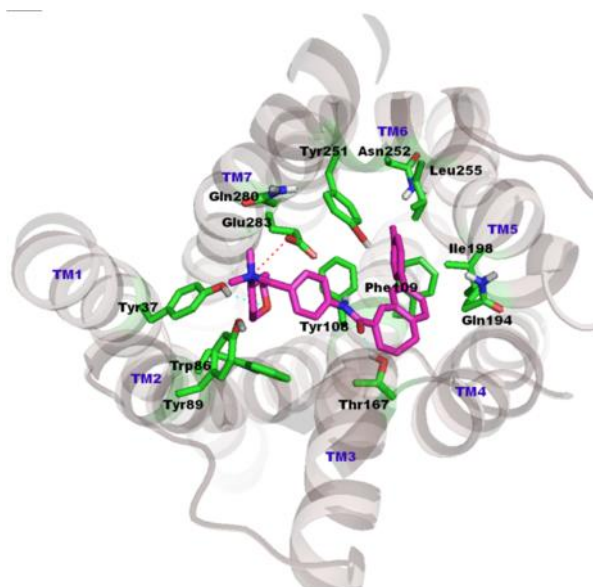
Figure 6: Top view of putative binding pocket for CCR5 along with the docked pose of TAK779. TM helices are shown in transparent light brown, whereas, constructed binding pocket residues were shown in green sticks. All the TM regions are labeled by violet color on the top of helices. The hydrogen bonds between TAK779 and CCR5 residues are indicated by dotted cyan color, whereas salt bridge contact between quaternary 'N' and Glu283 is denoted by dotted magenta line. Figure generated using Pymol program (http://www.pymol.org).

Two hydrogen bonds were observed between tetrahydro-2H-pyran and Tyr37 and central amide 'O' and Thr167. Other part of ligand (4-methyl-phenyl) was docked

into aromatic cluster formed by the Tyr108, Phe109, Ile198, Tyr251, Asn252 and Leu255. Residues from TM4 were not involved into interactions.

3.3.2. Binding mode of TAK220
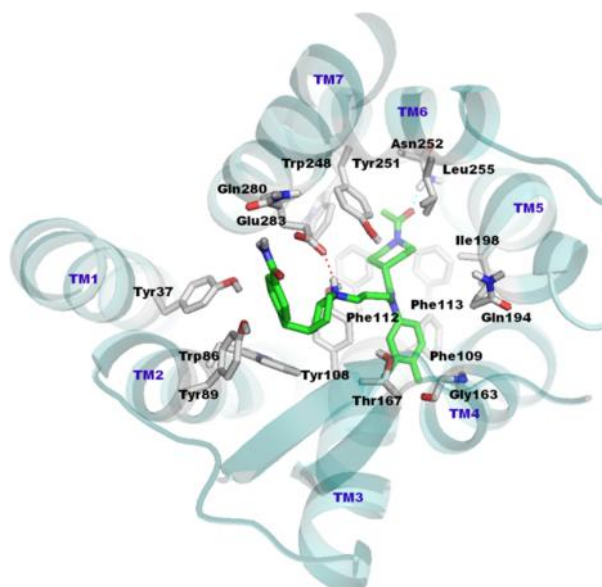
Figure 7: Docked mode of TAK220 inside CCR5. TM helices are shown in faint cyan, whereas residues were shown in white stick. TAK220 displayed in green stick. Hydrogen bond between TAK220 and Asn252 shown by dotted cyan line, while hydrogen bond between basic 'N' of TAK220 and Glu283 was shown by dotted magenta line. TM numbers were displayed on each TM helix.

Docked mode of TAK220 (Figure 7) revealed that the interactions are mainly hydrophobic and hydrophilic in nature. Protonated 'N' of TAK220 was hydrogen bonded (2.1 Å) with the Glu283 and acetyl 'O' bonded (1.8 Å) with Asn252. The central phenyl ring was accommodated into a hydrophobic pocket lined by the residues

Tyr108, Phe109, Gly163, Thr167 and Gln194, whereas N-acetyl-piperidine ring was situated into Phe112, Phe113, Ile198, Trp248, Tyr251, Asn252, Leu255 and Glu283. At other end, 4-benzamide was docked into Tyr37, Trp86, Tyr89, Tyr108, Gln280 and Glu283.

3.3.3. Binding mode of compound 85
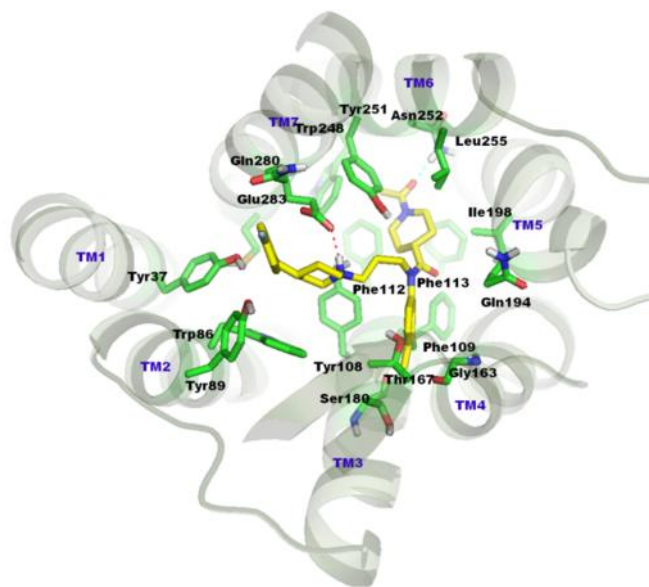
Figure 8: Docked mode of compound **85** inside CCR5. TM helices are shown in faint smudge color, whereas residues are shown in green stick. TAK220 was displayed in yellow stick. Hydrogen bond between compound **85** and Asn252 shown by dotted cyan line, while hydrogen bond between protonated 'N' of compound **85** and Glu283 was shown by dotted magenta line. TM numbers were displayed on each TM helix.

Compound 85 was docked in a similar orientation as TAK220 in CCR5 (Figure 8). Central 3-Cl, 4-Me phenyl was docked into Tyr108, Phe109, Gly163,

Thr167, Ser189 and Gln194, whereas N-acetyl piperidine was situated deep into a pocket formed by the residues Ile198, Trp248, Tyr251, Asn252 and Leu255. Hydrogen bond was established between carbonyl 'O' and Asn252. Another hydrogen bond was found between protonated 'N' of compound **85** and Glu283. 4-F-phenyl was placed into the residues Tyr37, Trp86, Tyr89, Tyr108, Gln280, Glu283 and Met287.

3.3.4. Binding mode of 14-OPPA

Figure 9 represents docked mode of the 14-OPPA inside CCR5. It was scrutinized that the nature of interaction between 14-OPPA and CCR5 was hydrophobic. Tertiary 'N' of 14-OPPA makes salt bridge contact (3.8 Å) with acidic Glu283. 2,4-Dimethyl-3-pyridinyl ring interacts with the hydrophobic residuesTrp86, Tyr89, Trp94, Leu104, Thr105, Thr167, Cys178 and Ser179. It was also determined that this ring and Trp86 interacts through the - stacking interactions. 4-Bromophenyl ring at other side docked into a pocket formed by the Tyr108, Phe109, Phe112, Phe113, Ile198, Val199, Tyr251 and Leu255, whereas ethoxime interacted with the Thr105, Thr167, Gln194 and Thr195. Central polar part of ligand was situated near Met279, Gln280 and Glu283.
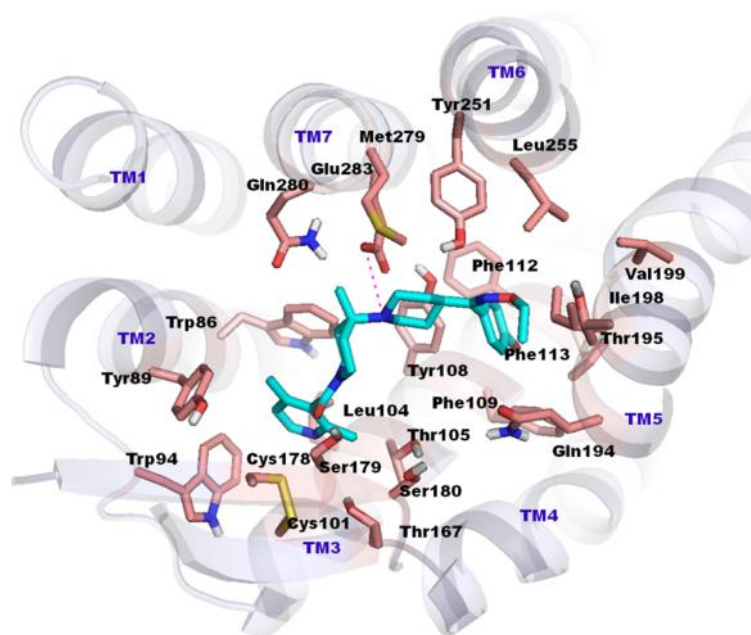
Figure 9: Docked mode of the 14-OPPA inside CCR5. Helices are shown in transparent sky blue color and residues are shown by faint brown color. Ligand is shown in cyan color and salt bridge contact (4.2 Å) shown by dotted magenta color. TM helix numbers are shown on top of each helix.

3.3.5. Binding mode of 37-OPPA

Figure 10 represents binding mode of 37-OPPA inside CCR5. We have obtained similar kind of binding mode for the 37-OPPA as 14-OPPA, because of similar scaffold structure. Orientation of 4-bromophenyl ring was differed than in 14-OPPA docked mode. 4-Bromophenyl moiety of ligand was docked into a pocket formed by TM3 (Tyr108, Phe112), TM5 (Ile198, Val199 and Leu201) and TM6

(Tyr251, Asn252 and Leu255). Ethoxime moiety was docked into the TM3 (Thr105, Tyr108, Phe109), TM4 (Gly145) and TM5 (Gln194, Thr195 and Ile198). Salt bridge contact between central basic nitrogen and acidic Glu283 was detected at a distance of 3.8 Å.
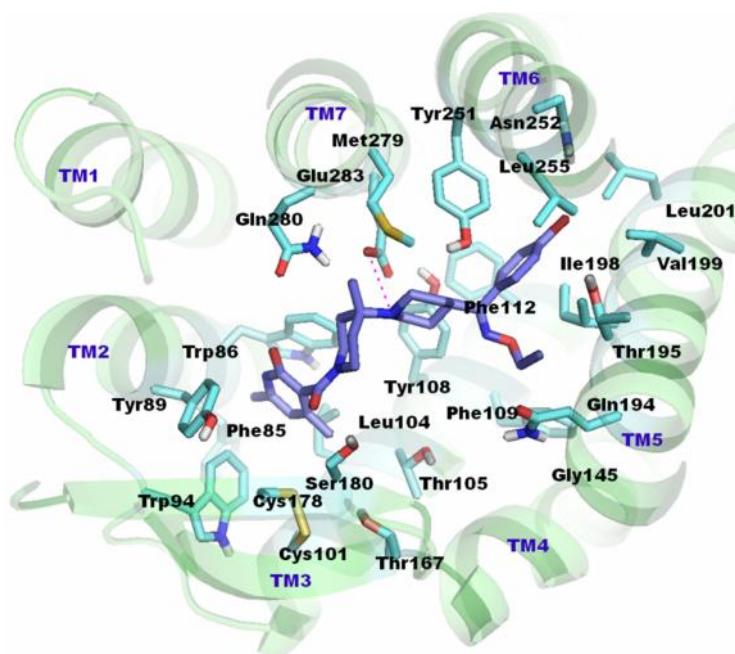
Figure 10: Docked mode of the 37-OPPA inside CCR5. Helices are shown in transparent green color and residues are shown by cyan color. Ligand is shown in violet color and salt bridge contact by dotted magenta color. TM helix numbers are shown on top of each helix.

3.3.6. Binding mode of 25-OPPA

Figure 11 represents the docked mode of lowest active compounds from OPPA series inside CCR5. It was observed that TM1 do not participate in the interaction with ligand. Salt bridge contact (4 Å) was detected between the tertiary 'N' of ligand and acidic Glu283 of CCR5. 2,6-Dimethylphenyl moiety of ligand and Trp86 interacted through the - stacking interactions. This moiety docked into the Trp86, Tyr89, Trp94, Leu104 and Thr105. However, residues from ECL2 such as Thr167, Cys178, Ser179 and Ser180 were also interacted with ligand. At other side, 4-bromophenyl moiety was docked into a hydrophobic pocket lined by the residues Tyr108, Phe109, Phe112, Phe113 and Ile198.
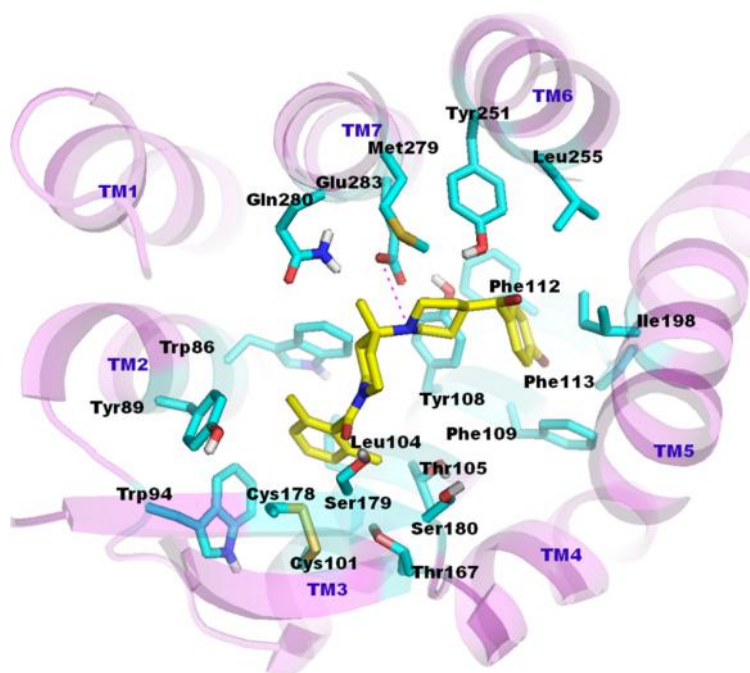
Figure 11: Docked mode of the 25-OPPA inside CCR5. Helices are shown in transparent magenta color and residues were shown by cyan color. Ligand is shown in yellow color and salt bridge contact by dotted magenta color. TM helix numbers are shown on top of each helix.

3.4. 3D-QSAR Analysis

3.4.1. CoMFA model

Atom by atom matching was performed to align entire molecules of dataset over template (compound **85**), which is shown superimposed in Figure 12.
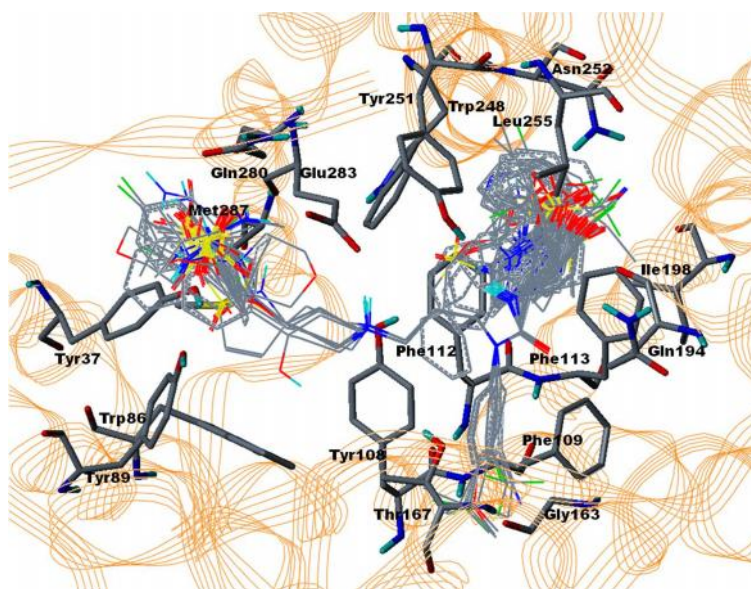
Figure 12: Molecular alignment of the dataset compounds over template (compound 85) inside CCR5 cavity. Aligned dataset molecules were represented by lines and active site residues were shown by stick model. CCR5 is shown by orange helices.

We developed good predictive CoMFA models for the training set (97 CCR5 inhibitors). Rationally designed model showed $q^2 = 0.722$, $r^2 = 0.884$ with SEE = 0.375 and 4 ONC. Further robustness of the developed model was assessed by 1,000 runs of bootstrapping and it provided BS-$r^2 = 0.907$ with SD = 0.013. Progressive scrambling was done to correct redundancy of q^2, and it showed a value of 0.619. The predicted and actual activities of training set compounds are given in Tables 1-3 and plotted in Figure 13A. Statistical summary of the final CoMFA model is given in Table 4.

3.4.2. CoMSIA model

We developed different CoMSIA models (Table 5) using Gasteiger-Hückel charge, 0.5 Å grid spacing and steric (S), electrostatic (E), hydrophobic (H), hydrogen bond acceptor (A) and hydrogen bond donor (D) parameters. A final CoMSIA model was obtained with SHDA parameters. The developed model displayed $q^2 = 0.712$, $r^2 = 0.825$ with ONC = 4. The SEE (0.461) was lower and ANOVA test (F = 108.6) showed good prediction. Statistical summary of the final CoMSIA model was given in Table 4. The q^2 was corrected for redundancy by progressive scrambling and a good value (0.522) was obtained. Further robustness of the developed CoMSIA model was checked with 1,000 runs of bootstrapping analyses, producing a BS-$r^2 = 0.851$ with BS-SD = 0.020. Higher bootstrapping value indicated that model was sufficiently predictive for further analysis. The corresponding field contribution values for SHDA parameters in the model development were 13.4, 27.8, 32.1 and 26.7 %. The predicted

and actual activities of training set compounds were given in Tables 1-3 and plotted in

Figure 13B. Histogram of different CoMSIA models was depicted in Figure 14.
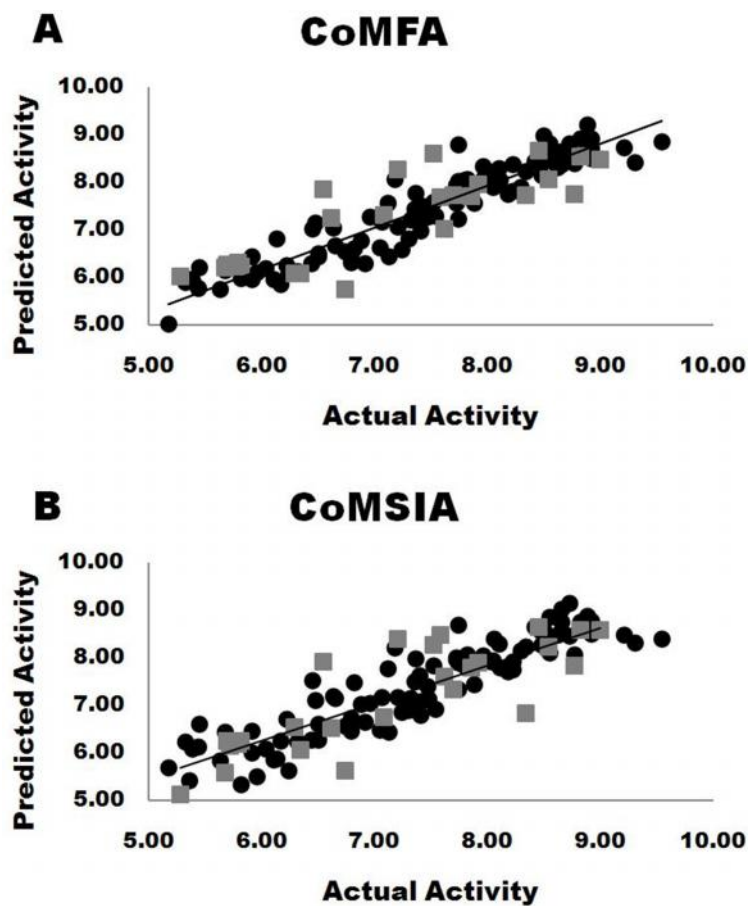
Figure 13: Actual *versus* predicted activity of training set (solid black circle) and test

set (grey square) compounds in (A) CoMFA and (B) CoMSIA models.
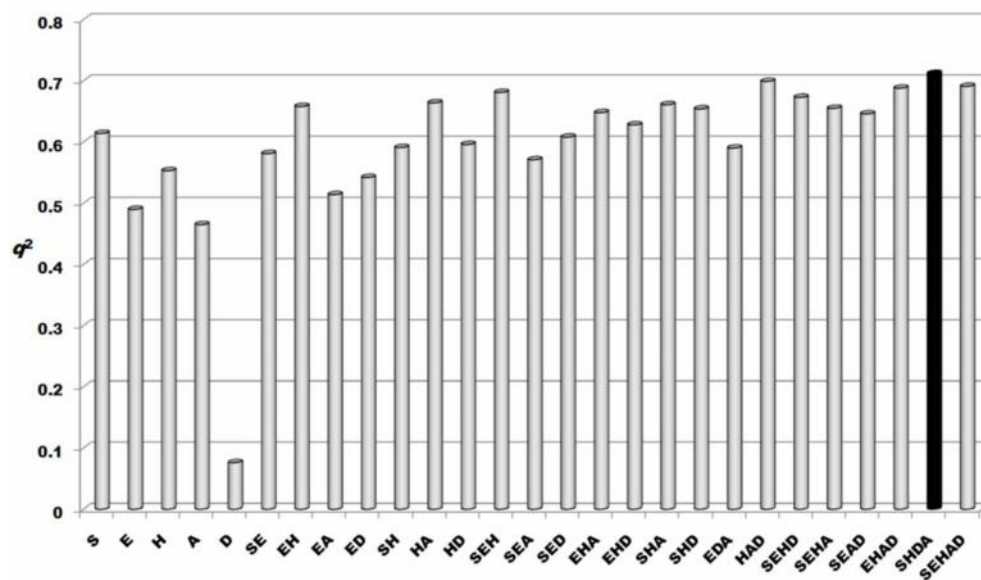
Figure 14: Histogram of twenty seven CoMSIA models and q^2 values. Model (SHDA) selected for further analysis is shown in solid black cylinder.

3.4.3. Predictive r^2_{pred}

CoMFA and CoMSIA models were developed for training set (97 compounds) and validated using test set (25 compounds). During the model development test set molecules were excluded. The r^2_{pred} for selected CoMFA model was 0.680 and for CoMSIA model it was 0.705, which indicates the reliability of developed models. The graph plot of actual and predicted activities of test set compounds were shown in Figure 13.

Table 4. PLS summary results of the CoMFA and CoMSIA models.

Parameter	CoMFA (Rational)	CoMSIA (SHDA)
[a]q^2	0.722	0.712
[b]NOC	4	4
[c]SDEP	0.581	0.591
[d]Q^2	0.619	0.522
[e]r^2	0.884	0.825
[f]SEE	0.375	0.461
[g]F-value	175.7	108.6
[h]BS-r^2	0.907	0.851
[i]BS-SD	0.013	0.020
[j]r^2_{pred}	0.680	0.705
Field Contributions %		
S	60.1	13.4
E	39.9	-
H	-	27.8
A	-	32.1
D	-	26.7

[a]cross validated correlation coefficient after leave one out (LOO). [b]Number of component. [c]Standard error of prediction after LOO. [d]Scrambled correlation coefficient after 50 individual scrambling. [e]Non cross validated correlation coefficient. [f]Standard error of estimate after non cross validation. [g]ANOVA test value. [h]Correlation coefficient after 1000 runs of bootstrapping analysis. [i]Standard deviation after 1000 runs of bootstrapping analysis. [j]Predictive correlation coefficient.

Table 5. Statistical summary of CoMSIA models.

	[a]NOC	[b]q^2	[c]Q^2	[d]r^2	[e]SEE	[f]Fvalue	[g]BS-r^2	[h]BS-SD	Percentage Contribution (%)				
									S	E	H	A	D
CoMSIA													
S	6	0.614	0.513	0.782	0.520	53.93	0.825	0.024	100	-	-	-	-
E	2	0.490	0.434	0.589	0.699	67.27	0.650	0.040	-	100	-	-	-
H	3	0.553	0.490	0.685	0.615	67.43	0.727	0.027	-	-	100	-	-
A	4	0.465	0.522	0.597	0.700	34.04	0.646	0.051	-	-	-	100	-

D	3	0.076	0.489	0.177	0.994	6.685	0.176	0.084	-	-	-	-	100
SE	3	0.581	0.490	0.720	0.580	79.58	0.767	0.046	31.4	68.6	-	-	-
EH	4	0.658	0.522	0.815	0.474	101.1	0.859	0.040	-	51.2	48.8	-	-
EA	4	0.514	0.521	0.691	0.613	51.37	0.756	0.049	-	60.8	-	39.2	-
ED	4	0.542	0.521	0.728	0.575	61.41	0.784	0.030	-	64.9	-	-	35.1
SH	2	0.591	0.434	0.688	0.609	103.5	0.720	0.028	38.4	-	61.6	-	-
HA	4	0.664	0.522	0.790	0.505	86.45	0.821	0.041	-	-	51.4	48.6	-
HD	4	0.596	0.521	0.762	0.530	75.53	0.789	0.031	-	-	59.8	-	40.2
SEH	6	0.681	0.513	0.893	0.364	125.3	0.922	0.017	18.8	43.5	37.7	-	-
SEA	4	0.571	0.489	0.704	0.596	73.75	0.764	0.038	25.0	46.5	-	28.5	-
SED	4	0.608	0.521	0.780	0.516	81.77	0.817	0.026	25.4	44.9	-	-	29.7
EHA	4	0.648	0.522	0.808	0.483	96.57	0.855	0.010	-	33.8	41.1	25.1	-
EHD	3	0.628	0.490	0.772	0.523	105.2	0.809	0.020	-	45.4	26.5	-	28.1
SHA	4	0.661	0.522	0.802	0.490	93.41	0.861	0.021	17.4	-	42.4	40.3	-
SHD	3	0.654	0.489	0.759	0.538	97.46	0.772	0.032	24.9	-	39.3	-	35.8
EDA	4	0.590	0.521	0.771	0.527	77.64	0.796	0.027	-	37.5	-	29.2	33.3
HAD	4	0.699	0.522	0.817	0.472	102.5	0.850	0.017	-	-	34.1	38.1	27.8
SEHD	4	0.673	0.520	0.837	0.446	117.7	0.870	0.023	15.7	34.4	28.6	-	21.2
SEHA	6	0.655	0.513	0.892	0.367	123.4	0.920	0.013	13.6	32.0	33.3	21.1	-
SEAD	4	0.646	0.522	0.799	0.494	91.33	0.837	0.030	17.5	28.7	-	26.5	27.2
EHAD	4	0.688	0.521	0.837	0.445	117.8	0.857	0.021	-	25.8	29.3	23.7	21.3
SHDA	**4**	**0.712**	**0.522**	**0.825**	**0.461**	**108.6**	**0.851**	**0.020**	**13.4**	**-**	**27.8**	**32.1**	**26.7**
SEHAD	4	0.691	0.520	0.836	0.447	117.0	0.863	0.024	11.5	23.6	23.0	20.4	21.5

[a]Optimum number of components as determined by the PLS leave one out cross validation study. [b]Cross validated correlation coefficient by LOO. [c]Scrambled correlation coefficient. [d]Non cross validated correlation coefficient. [e]Standard error of estimate. [f]F-test value. [g]Correlation coefficient after 1000 runs of bootstrapping. [h]Standard deviation for 1000 runs of bootstrapping. S = steric, E = electrostatics, H =

hydrophobic, A = hydrogen bond acceptor, D = hydrogen bond donor. Model used for CoMSIA analysis is marked in **bold face**.

3.5. Contour map Analysis

3.5.1. CoMFA contour map

CoMFA contour map represents the steric and electrostatic requirements for improving the potency. CoMFA steric maps are represented by the green and yellow color, where green represents the steric bulk is favorable for enhancing affinity, which is correspond to increased Lennard-Jones potential. While, yellow denotes steric bulk is unfavorable for enhancing activity and reduced Lennard-Jones potential could be expected at that region. Electrostatic contour maps are represented by the blue and red color. Blue color represents the points where increase in Columbic potential by electropositive group enhances the affinity, whereas red color designates the region where electronegative group are favorable for the affinity. CoMFA steric stdev contour map (Figure 15) along with the most potent molecule (compound **85**) was shown in CCR5 active site. Beneath the central phenyl ring at 3, 4 positions, green contour maps were observed which indicates that the smaller substituent at this region could be beneficial for inhibitory activity. Binding pocket at this region was surrounded by the Tyr108, Phe109, Gy163 and Thr167. Para position of the phenyl ring was substituted by the *t*-butyl, methoxy and *i*-propyl groups in the compound **13** (pIC_{50} = 6.80), **15** (pIC_{50} = 5.82) and **86** (pIC_{50} = 7.74), which showed reduced inhibitory activity. It

indicates that the binding pocket at this region is sterically restricted and larger substituents are unfavorable for inhibitory activity.
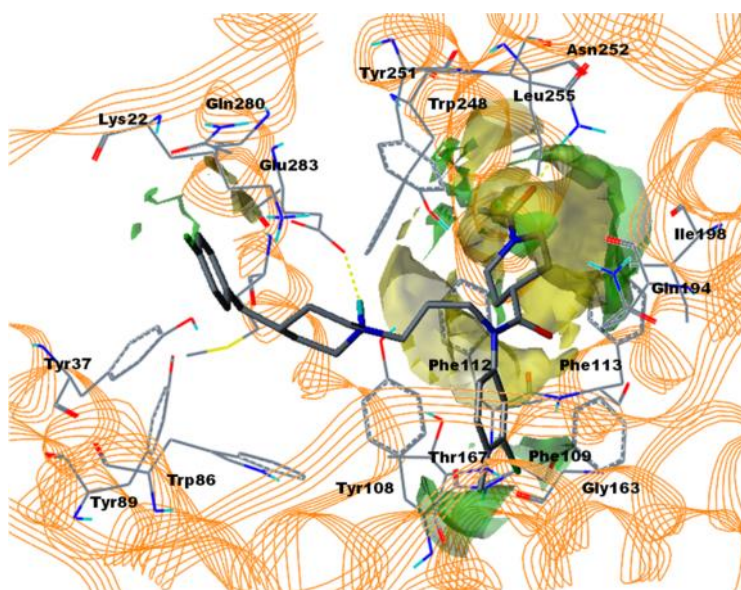
Figure 15: Superposed CoMFA steric stdev coefficient contour maps in CCR5 active site. Transparent green represents sterically favorable region, whereas yellow color represents sterically unfavorable region. Compound **85** showed in stick model and CCR5 in orange helices.

CoMFA steric stdev electrostatic contour map is shown in Figure 16. In vicinity of R^3 substituent red and blue contours were observed, which indicates that the para position with electronegative substituent is highly desirable for the improved inhibitory activity. This is the reason why these compounds **68** (pIC_{50} = 8.92), **69** (pIC_{50} = 9.00), **84** (pIC_{50} = 9.21) and **89** (pIC_{50} = 9.30) are highly active. In contrast, at para position electropositive substitutions like amino, acetamino and methoxy

containing compounds **92** (pIC_{50} = 8.04), **93** (pIC_{50} = 8.23) and **96** (pIC_{50} = 8.10) have reduced binding affinity. Blue contour underneath central phenyl ring indicates the favorability of small electropositive substituents. The R^3 position is surrounded by the Tyr37, Trp86, Tyr89 and Gln280.
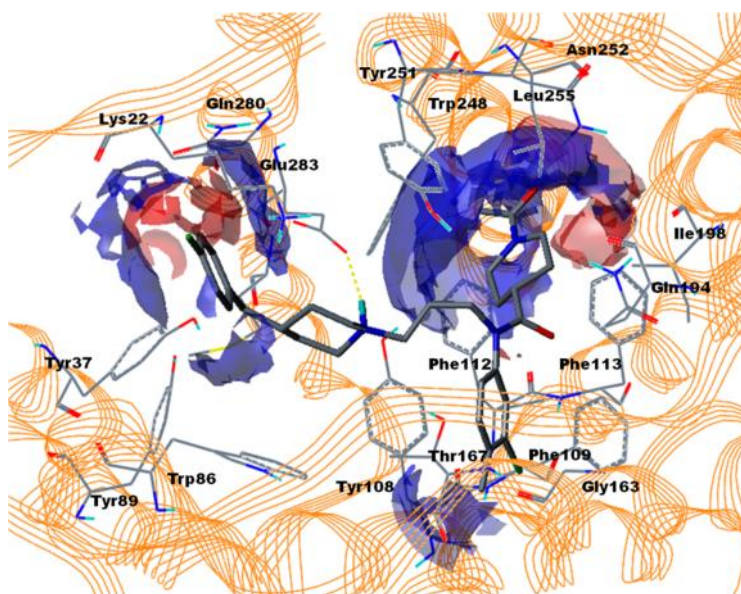
Figure 16: CoMFA electrostatic stdev coefficient contour maps for the most active compound (**85**) in CCR5 receptor. Contour maps are shown in transparent blue and red color. Blue contour indicates electropositive favorable and red contour represents electronegative favorable for activity. CCR5 is shown in orange helices.

3.5.2. CoMSIA contour map

CoMSIA steric stdev coefficient contour map plotted onto the CCR5 interface is shown in Figure 17. A small green contour was observed near R^1 substituent indicating that sterically bulky substituents are favorable. Binding pocket at this region

is lined by the Trp248, Tyr251, Asn252 and Leu255. Beside green contour two yellow contours were observed, which indicates that substituent penetrating to that site are unfavorable for the inhibitory activity. For higher inhibitory effect, at R^1 position polar moieties need to be in proper position and orientation. For e.g. 1-acetylpiperidin-3-yl in compound **78** ($pIC_{50} = 6.17$) *versus* 1-acetylpiperidin-4-yl group afforded compound **72** ($pIC_{50} = 7.80$). 3-substituted piperidine at R^1 position did not offered any advantage in CCR5 inhibition. At R^3 position medium yellow contour map was observed, which indicates the substituent penetrating that direction are unfavorable for inhibitory activity, because the pocket is sterically restricted.
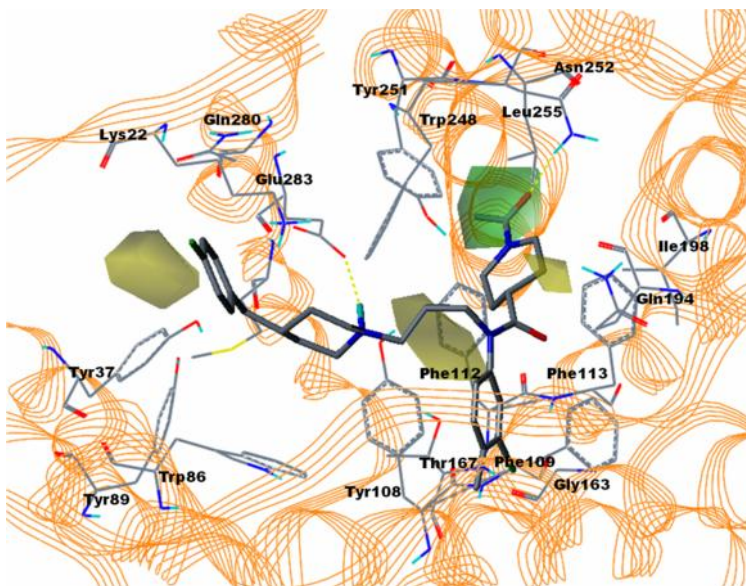
Figure 17: Superposed CoMSIA steric stdev coefficient contour maps in CCR5 active site. Transparent green represents sterically favorable region, whereas yellow color

represents sterically unfavorable region. Compound **85** showed in stick model and CCR5 in orange helices. Residues are shown by the lines.

CoMSIA hydrophobic stdev coefficient contour map was superposed on CCR5 active site and shown in Figure 18. CoMSIA contour map is denoted by the yellow and white color, where yellow represents hydrophobically favorable region and white indicates unfavorable for hydrophobic groups. Yellow contour map was observed in the vicinity of 1-acetylpiperidin-4-yl group, indicating that the place is suitable for the hydrophobic groups to enhance potency. The pocket at this position lined by the residues Ile198, Trp248, Tyr251, Asn255 and Leu255.
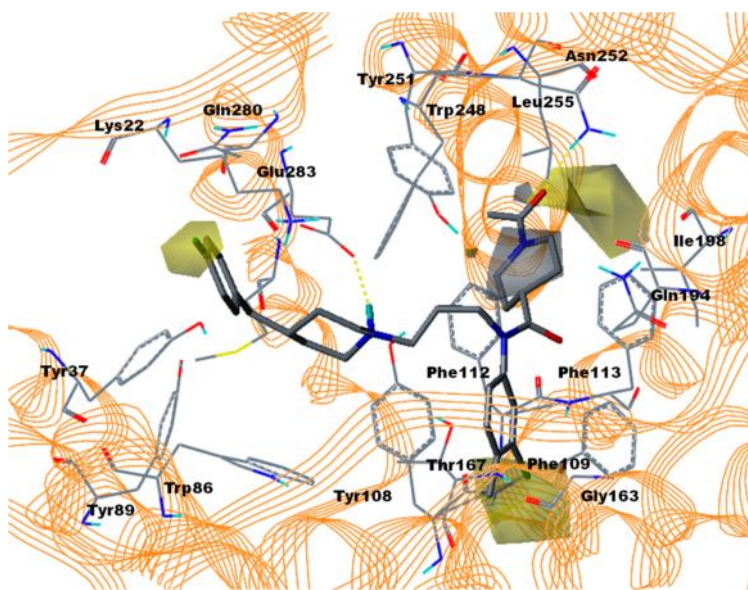
Figure 18: CoMSIA hydrophobic stdev coefficient contour map is superposed on CCR5 active site. Yellow represents hydrophobically favorable region and white

indicates unfavorable for hydrophobic substituents. Compound **85** represented by the stick model, active site residues by the line model and CCR5 by orange helices.

However, medium sized white contour nearby yellow contour was observed, indicating that unfavorable position for hydrophobic groups. A medium sized yellow contour map was observed in the vicinity of central phenyl ring, which indicates that the small hydrophobic substitutions are most welcomed for improved inhibitory potency.

CoMSIA hydrogen bond acceptor stdev coefficient contour map is shown in Figure 19. Contour map was superposed on CCR5 active site.
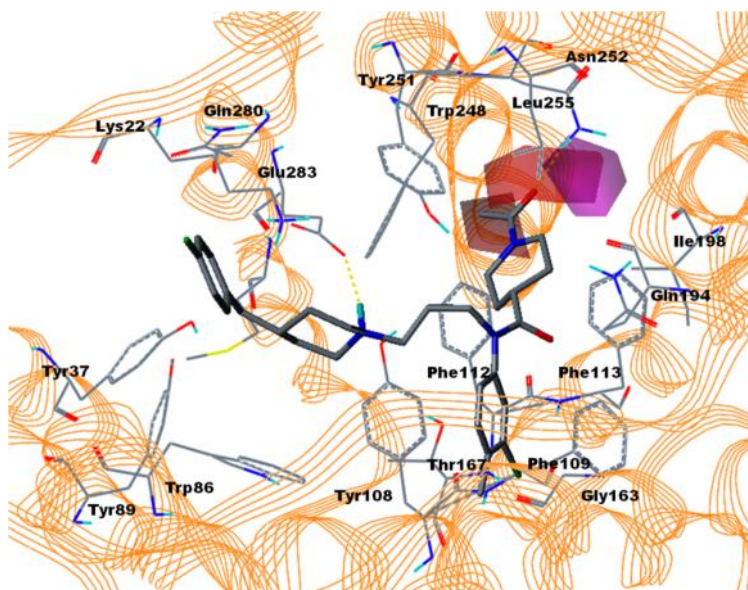
Figure 19: Hydrogen bond acceptor stdev coefficient contour maps was shown superposed on the CCR5 active site. Transparent magenta denotes favorable for hydrogen bond acceptor, whereas, red indicates unfavorable for hydrogen bond

acceptor. Active site residues were shown by lines and ligand by stick model. CCR5 was shown by orange helices.

Hydrogen bond acceptor contour map was denoted by the magenta and red. Magenta denotes hydrogen bond acceptor favorable and red indicates hydrogen bond acceptor unfavorable. Our docked model shows that acetyl 'O' forms hydrogen bond with Asn252. In case of compound **71** (pIC_{50} = 5.37), where R^1 position is de-acetylated shows reduced potency, indicating that the importance of hydrogen bond for the higher inhibitory potency. Lesser the hydrogen bond distance stronger the hydrogen bond and vice versa. Similarly in compound **75** (pIC_{50} = 7.40), the R^1 group was substituted by methylsulfonyl group which showed slightly lesser activity than the compound **72** (pIC_{50} = 7.80), which had acetyl group at respective position.
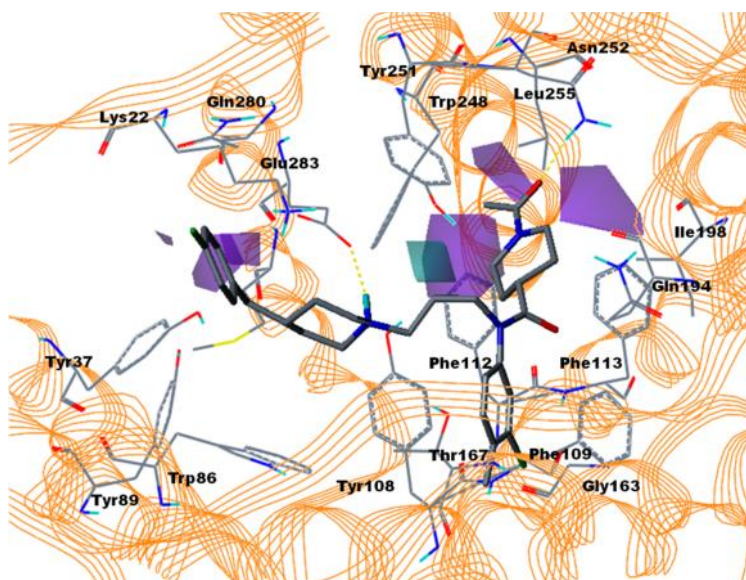
Figure 20: Hydrogen bond donor stdev coefficient contour maps was shown superposed on the CCR5 active site. Transparent cyan denotes favorable for hydrogen bond donor, whereas, violet indicates unfavorable for hydrogen bond donor. Active site residues were shown by lines and ligand by stick model. CCR5 was shown by orange helices.

Hydrogen bond donor contour map was shown in Figure 20. Transparent violet contours near Phe112, Ile198 and Val255 shows that these positions are unfavorable for hydrogen bond donor group because these residues do not have hydrogen bond donor atom. However, small cyan contour was observed near Tyr251, which indicates favorable position for hydrogen bond donor. Hydroxyl group of Tyr251 can donate the hydrogen bond to a ligand which has corresponding hydrogen bond acceptor atom.

3.6. MD simulation analysis

MD simulation analyses of ligand-protein (14-OPPA-CCR5, 25-OPPA-CCR5 and 37-OPPA-CCR5) complexes were performed for the 20000 ps using GROMACS simulation package. During system preparation longer axis of box was kept parallel to CCR5 long axis. Inflategro script was used to pack the lipid molecules around CCR5. Total six lipid (DPPC) molecules were deleted from the system. Systems were compressed and minimized alternately to remove the steric clashes between CCR5 and lipids. Figure 21 represents the simulated system containing solvent (water, Cl⁻) and solute molecules (ligand-CCR5-DPPC). NVT ensemble was simulated at 323 K and NPT ensemble was simulated at 1 bar pressure.
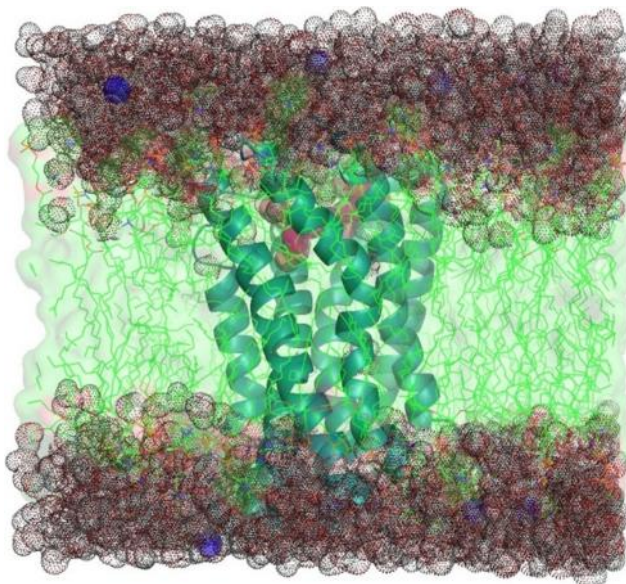
Figure 21: CCR5 embedded in DPPC bilayers. Ligand represented by magenta sphere. CCR5 is shown by the cyan helices, DPPC by green lines, chloride atoms by solid blue sphere and water molecules by dotted sphere.

3.6.1. Backbone RMSD and potential energy analyses

Backbone atoms RMSD of CCR5 were computed from three simulated systems and plotted as a function of time in Figure 22A. It indicated that the proteins RMSD was raised upto ~2000 ps and then maintained throughout simulation time in all system. CCR5 structures in all the simulated systems were affected by the lipid environment, which was evident from Figure 22A. For 14-OPPA-CCR5 system it was observed that after equilibration RMSD was fluctuated between 0.35 nm to 0.41 nm with an average of 0.38 ± 0.03 nm. However, in case of 25-OPPA-CCR5 system, the

backbone RMSD increased constantly up to first 5000 ps and leveled off for rest of the simulation period. After 5000 ps of equilibration, RMSD constantly fluctuated in 0.33 and 0.39 nm, with an average of 0.36 ± 0.03 nm. It was detected that for first 5000 ps CCR5 was equilibrated and afterward fluctuated in 0.35 - 0.41 nm frequency, with an average of 0.38 ± 0.03 nm.
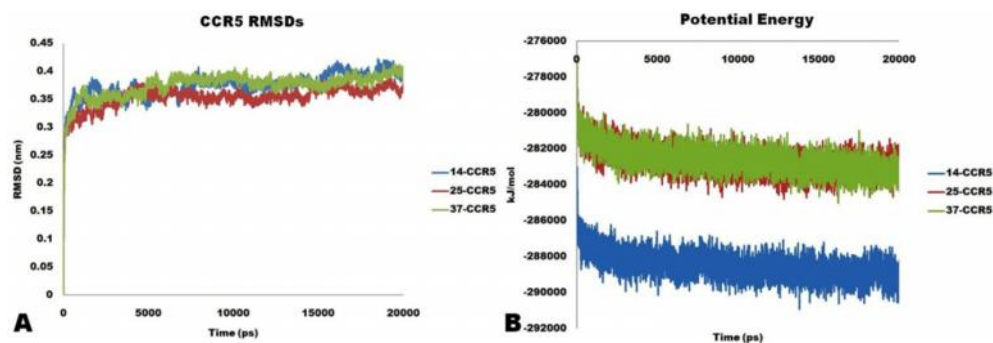
Figure 22: (A) Backbone atoms RMSD of CCR5 in three simulated systems as a function of time. (B) Potential energy of the three systems plotted as function of time. Color markers for each system were given at the right of each figure.

Protein structure stability was assessed by calculating potential energy as a function of time and represented in Figure 22B. For 14-OPPA-CCR5 system all the structures were laid within -2.83e+05 to -2.90e+05 kJ/mol range, with an average of -2.89e+05 kJ/mol. Small deviation in potential energy was observed in last 5000 ps simulation time. Potential energy for 25-OPPA-CCR5 (Figure 22B) system was calculated, which showed that the entire structures were positioned within the -2.77e+05 to -2.85e+05 kJ/mol energy range. It was also observed that the potential

energy decreased gradually. The average potential energy of last 5000 ps of 20,000 ps was -2.83e+05 kJ/mol. Similarly, potential energy for 37-OPPA-CCR5 system was calculated which indicated that the potential energy was gradually decreased throughout simulation period. The average energy of last 5000 ps was calculated and it was -2.83e+05 kJ/mol. All the structures had energies between -2.77e+05 to -2.85e+05 kJ/mol.

3.6.2. Temperature and inhibitors RMSD analyses
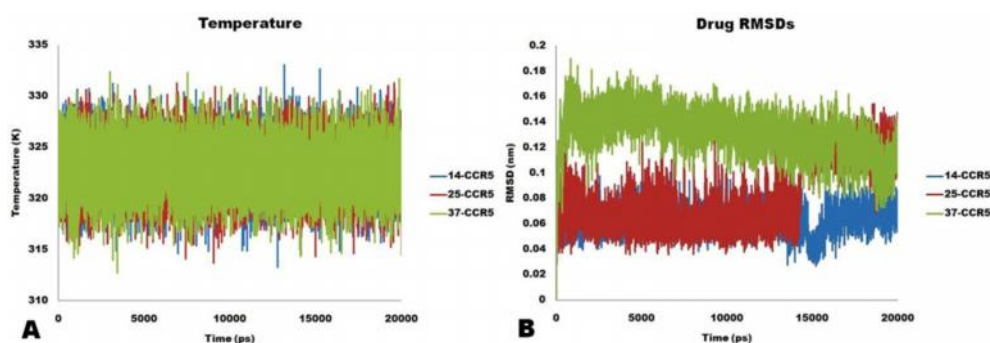
Figure 23: (A) Graph of temperature as a function of time during simulations. Average temperature of each simulated system was 300 K. (B) RMSDs of simulated inhibitors (14-OPPA, 25-OPPA and 37-OPPA) as a time function.

Temperature of each simulated system was monitored throughout simulation period and depicted in Figure 23A. Average temperature was 323 K in each system for 20000 ps simulation, which indicated that the simulations were performed at constant temperature. RMSDs of ligand (14-OPPA) showed that it remained constant (~0.065 nm) up to 14,500 ps, then dropped to ~0.05 nm, and after 16,000 ps rose and leveled off

at ~0.065 nm throughout the remainder of the simulation. The RMSD for 25-OPPA
(Figure 23B) showed that up to 14,192 ps the average RMSD was ~0.065 nm and
afterward it elevated to ~0.116 nm for rest of the simulation period. For 37-OPPA
another trend of RMSD was noted. The RMSD constantly rose up to 5000 ps and
afterward gradually decreased throughout 20,000 ps simulation. The average RMSD for
37 was found to be 0.137 nm. After MD simulations, the complexes (14-OPPA-CCR5,
25-OPPA-CCR5, and 37-OPPA-CCR5) were extracted from trajectories and 100 steps
minimization was performed using the steepest descent method with the aim of
removing bad contacts from the crude complex structures.

3.6.3. RMSF and radius of gyration (Rg) analyses

Atomic fluctuations also called root mean square fluctuations was calculated
for the CCR5 in three simulated systems. RMSF of backbone atoms of CCR5 were
calculated and depicted in Figure 24A. In case of 14-OPPA-CCR5 system greater
fluctuations were observed for the intracellular loop 2 (ICL2) (Lys138, Gln184), ICL3
residues (Lys228, His231) and C-terminal residue (Phe304). For the 25-OPPA-CCR5
system it was observed that the very few fluctuations were exceeded 0.3 nm (Ile23,
Ile27 and Leu174). For the 37-OPPA-CCR5 system, it was scrutinized that most of the
residues were fluctuated less than 0.2 nm and very few exceeded 0.3 nm (Arg60,
Tyr184, Arg225, Arg232, Arg267 and Glu302). From this figure it was also evident
that average fluctuations were below 0.2 nm. We can conclude from the above results
that the N-terminal, C-terminal and intracellular loops were fluctuated considerably,

and -helices and -sheets are stable. During MD simulations it was observed that the

Tyr108, Tyr251 and Glu283 forms hydrogen bond and stabilized the architecture of

TMs.
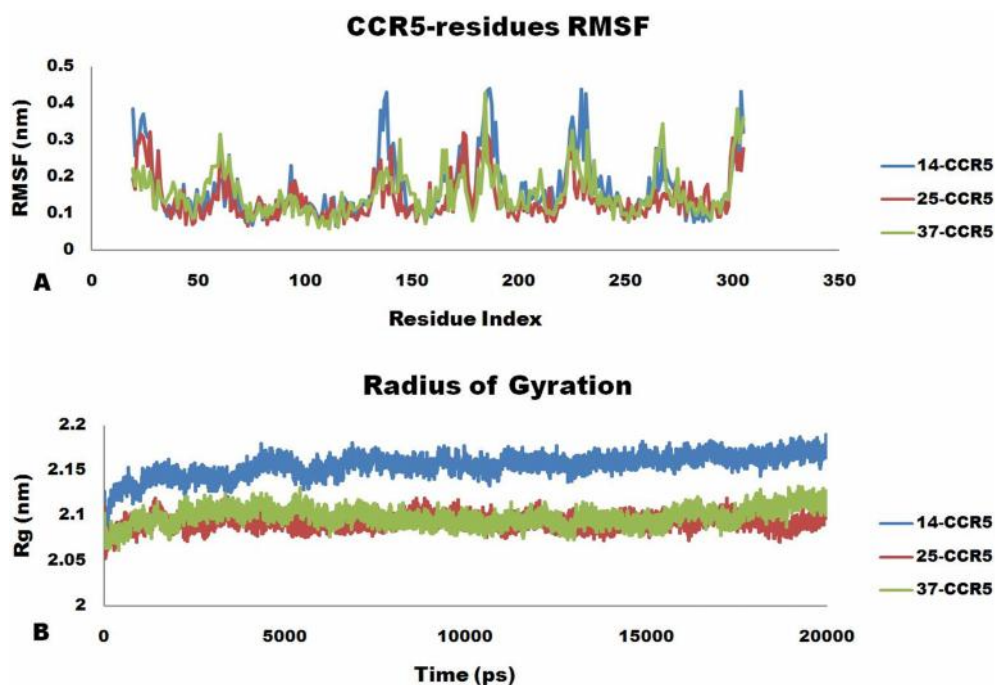
Figure 24: (A) RMSF of CCR5 in the simulated systems. (B) Radius of gyration as a function of time for the simulated systems.

Radius of gyration (Figure 24B) was calculated to measure the compactness of

CCR5 throughout simulation period. For the 14-OPPA-CCR5 system, a slow increase

in Rg from ~2.075 to ~2.185 nm was observed, which indicates CCR5 compactness

was slowly diminished. However, in 25-OPPA-CCR5, the Rg value was found to be

increased from ~2.05 to ~2.12 nm up to ~4000 ps and afterward it maintained between

~2.07 and ~2.12 nm, which indicated that CCR5 structure was compact throughout the simulation. For the 37-OPPA-CCR5 system, Rg was increased up to 5000 ps from ~2.062 to ~2.13 nm, afterward up to 15,000 ps simulation it was maintained and again rose up in last 5000 ps. It indicates that the compactness of CCR5 was maintained between 5000 and 15,000 ps and afterward it slowly diminished the CCR5 packing.

3.6.4. Characterization of loop regions

CCR5 consists of uncharacterized N-terminal and C-terminal region. Also it contains six loops, namely ECL1-3 and ICL1-3. In CCR5 highly conserved two disulfide bridges (Cys20-Cys269, Cys101-Cys178) are present and previously it was reported that they are important in the ligand binding (62). Figure 25 depicts the calculated distance between the S-S of disulfide bridges.
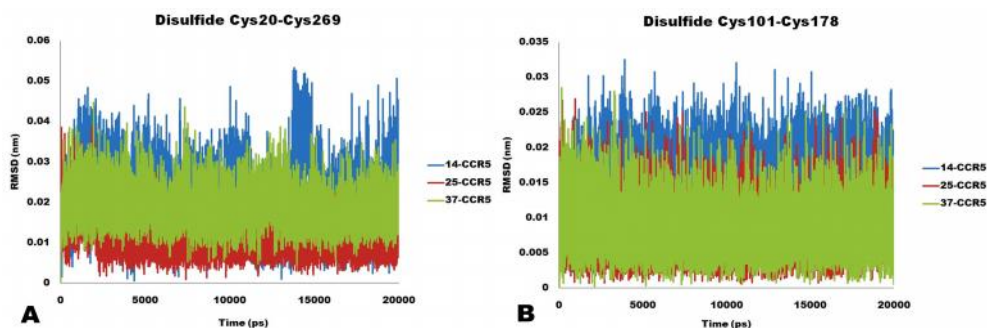
Figure 25: Distance between disulfide bridges as a function of time. (A) Disulfide Cys20-Cys269 and (B) Cys101-Cys178 bridges distance in three simulated system.

The average RMSD of Cys20-Cys269 was ~0.052 nm in 14-OPPA-CCR5, ~0.015 nm in 25-OPPA-CCR5 and ~0.02 nm in 37-OPPA-CCR5 systems. However,

for Cys101-Cys178 it was ~0.015 nm in 14-OPPA-CCR5, ~0.012 nm in 25-OPPA-CCR5 and ~0.011 nm in 37-OPPA-CCR5. From this result we could say that disulfide bridges were intact throughout simulations and helped to maintain the TM architecture of CCR5.

N-terminal (Pro19-Ala30), ECL1 (Ala90-Gln102), ECL2 (Thr167-Ile198) and ECL3 (Gln261-Gln277) were treated as different groups and RMSD was calculated to know the deviation from original position.

Figure 26 depicts the RMSD of (A) N-terminal, (B) ECL1, (C) ECL2 and (D) ECL3 as a function of time. RMSD of N-terminal region was ~0.35 nm for 14-OPPA-CCR5 system. However, it was less than ~0.3 nm for the other two simulated systems. For the 14-OPPA-CCR5 system, it was observed that ECL2 loop was more dynamic and fluctuated with maximum amplitude of ~0.5 nm. The average RMSD for ECL1 was ~0.082 ± 0.002 nm and for ECL3 it was ~0.135 ± .002 nm. However, for 25-OPPA-CCR5 system it was noticed that ECL1 has RMSD of ~0.225 nm, whereas it reached to ~0.4 nm for ECL2 and ~0.21 nm for ECL3. For 37-OPPA-CCR5 system it was scrutinized that ECL1 was stable with the maximum of RMSD reached to ~0.16 nm. ECL2 RMSD was ~0.3 nm and ECL3 was ~0.275 nm with an average of ~0.225 nm. From these results we can conclude that the N-terminal region and ECL2 loop are more dynamic among the simulated systems, which is in line with the previous reports (62-65).
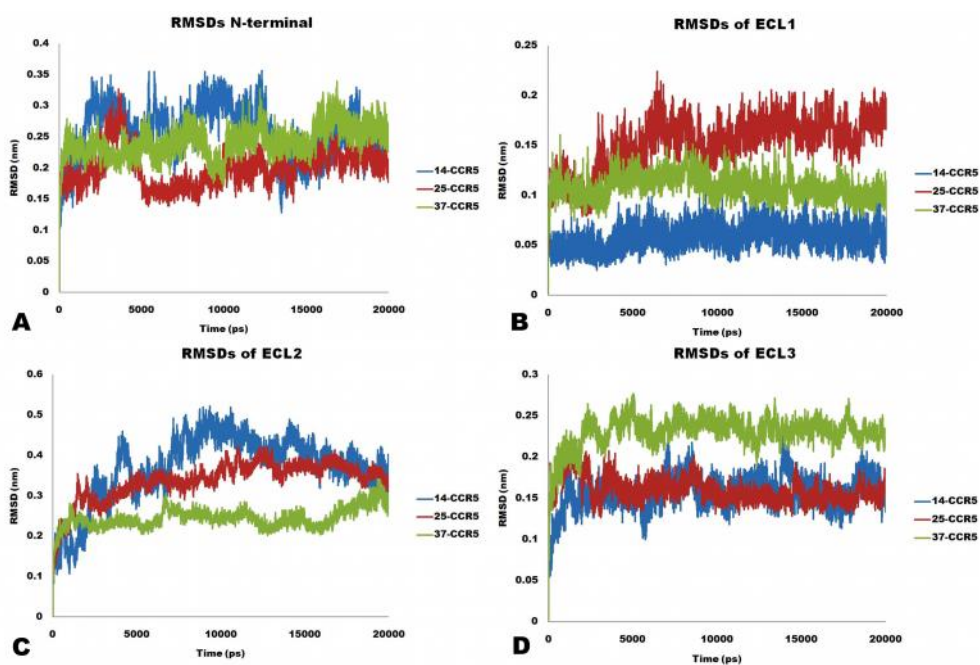
Figure 26: RMSD of (A) N-terminal, (B) ECL1, (C) ECL2 and (D) ECL3 regions as a function of time.

3.7. Binding modes analyses after MD simulation

3.7.1. 14-OPPA-CCR5 interaction

20000 ps MD simulation was performed for the complex structure embedded in DPPC lipid bilayers. Complex was extracted and minimization was performed for 100 steps using steepest descent algorithm in order to remove bad contacts. Figure 27 represents the binding mode of 14-OPPA after MD simulations. It was observed that the some residues were changed after MD simulation. Pyridine, ethoxime and bromophenyl was reoriented into the active site and stabilized in CCR5. 2,4-dimethyl-

3-pyridinyl ring was stabilized into a hydrophobic pocket lined by the residues Tyr37, Trp86, Tyr89 and Ala90.
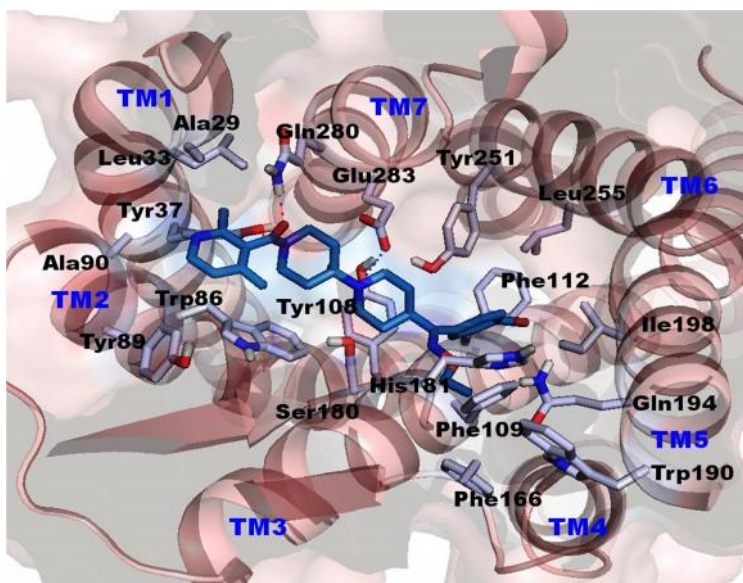
Figure 27: Top view of MD simulated binding mode of 14-OPPA inside CCR5. TM helices are shown in brown, whereas active site residues were shown by white-blue sticks. Salt bridge (3.8 Å) contact was shown by blue dotted line and hydrogen bond contact by red dotted line. Ligand is shown by violet. TM helices were numbered at the top.

The docking predicted - stacking interaction between pyridine and Trp86 was disappeared and edge-to-edge interactions were observed with pyridine and Tyr37 and Tyr89. It was also detected after MD simulation that Tyr37 came closer to pyridine ring, which was not nearer in docked mode. It indicated that Tyr37 is an important residue in ligand-CCR5 interaction. After MD simulation it was observed that ethoxime

was relocated into a pocket Thr105, Phe109, Ser180, His181, Phe166, Trp190 and Gln194. Bromophenyl ring was moved to a pocket lined by the residues Tyr108, Phe109, Phe112, Ile198, Tyr251 and Leu255 from its original position in Phe112, Ile198, Trp248, Tyr251, Asn252 and Leu255. After MD simulation bromophenyl exhibited - stacking interactions with Phe109 and face-to-edge with Tyr251. The salt bridge distance (0.42 nm) identified in docking was reduced after MD simulation (0.38 nm), indicated the importance of it in the ligand-CCR5 interactions. Halogen-interaction (0.38 nm) was detected between Br and Phe112. After MD simulation it was observed that ligand forms one hydrogen bond interaction with Gln280, which was not detected in docked model.

3.7.2. 37-OPPA-CCR5 interaction

Figure 28 depicted the MD simulated mode of 37-OPPA in CCR5 cavity. After 20000 ps simulations complex was extracted and minimized for 100 steps using steepest descent algorithm in order to remove bad contacts. Salt bridge contact was detected between basic 'N' of ligand and carboxyl of Glu283 at a distance of 3.6 Å. Ligand positioned into CCR5 cavity like 14-OPPA. Hydrophobic interactions were detected between bromophenyl and Phe109, Phe112, Ile198, Trp248, Tyr251 and Leu255. On the other hand it is visible that pyridine ring was stabilized into TM2, TM3 and TM7 containing residues Trp86, Tyr89, Trp94, Leu104, Tyr108, Cys178 and Gln280. Internal hydrogen bonding were observed in Tyr108-Glu283-Tyr251, which gave stability to TM architecture. However, it was also noted that the ethoxime moiety

was located into a pocket formed by the residues Thr105, Phe109, Ser160, Trp190 and Gln194.
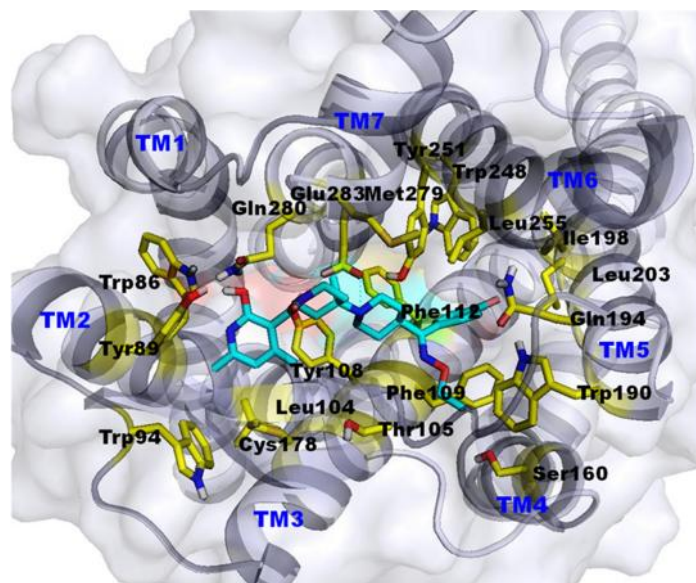
Figure 28: Top view of MD simulated binding mode of 37-OPPA inside CCR5. TM helices are shown in white-blue, whereas active site residues were shown by yellow sticks. Salt bridge (3.6 Å) contact was shown by cyan dotted line. Ligand is shown by cyan color. TM helices were numbered at the top.

3.7.3. 25-OPPA-CCR5 interaction

Figure 29 represents the MD simulated binding mode of 25-OPPA inside CCR5 cavity. Salt bridge contact was detected between tertiary 'N' and acidic Glu283 at a distance of 3.8 Å. Ligand was tightly packed into the TM helices and internal hydrogen bonding was detected between Tyr108-Glu283-Tyr251, which maintained the structural organization of CCR5. Strong   -   stacking interaction was detected between

2,6-dimethylphenyl and Trp86 and hold ligand tightly near it. Face-to-edge interactions were detected between 2,6-dimethylphenyl and Tyr89 and Trp94. Para bromophenyl was stabilized into a hydrophobic cavity lined by the residues Tyr108, Phe112, Ile198, Val199, Leu203, Tyr251 and Leu255. Edge-to-edge interactions were observed between para-bromophenyl and Phe112 and Tyr251. In MD simulation it was observed that the central polar part of ligand was stabilized into the Tyr108, Gln280 and Glu283 residues. Polar hydroxyl group interacted with the Thr195 and Ile198.
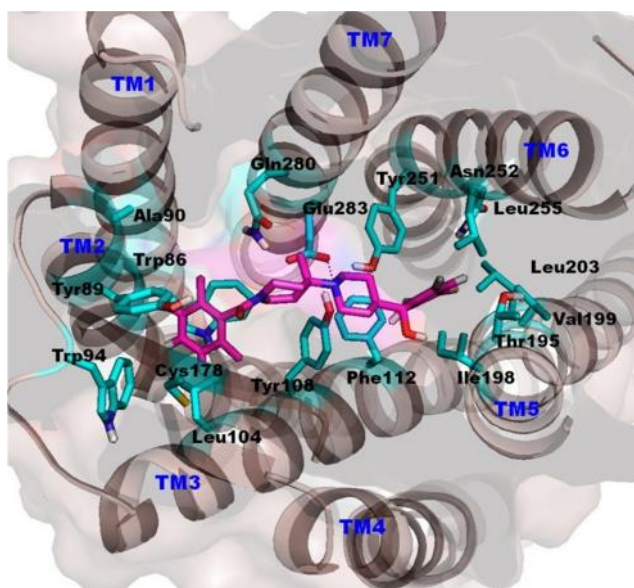
Figure 29: Top view of MD simulated binding mode of 25-OPPA inside CCR5. TM helices are shown in brown, whereas active site residues were shown by cyan sticks. Salt bridge (3.8 Å) contact was shown by blue dotted line. Ligand is shown by magenta stick. TM helices were numbered at the top.

2D-schematic plot interactions of inhibitors and CCR5 were shown in Figures 30 and 31.
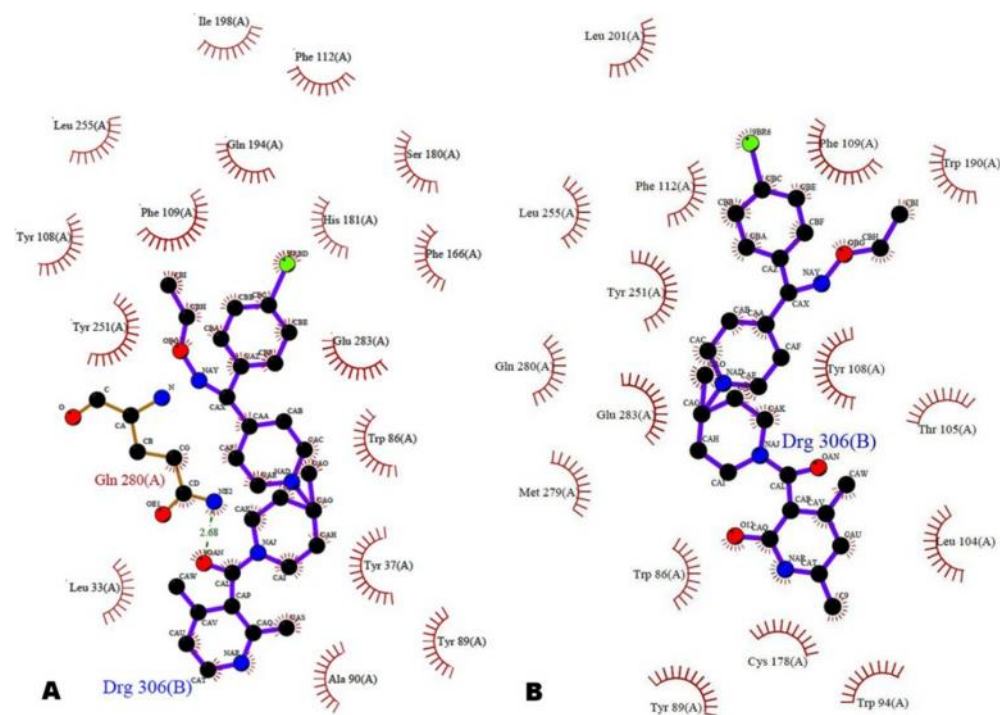
Figure 30: 2D-schematic plot interactions between (A) 14-OPPA-CCR5 and (B) 37-OPPA-CCR5 were generated using Ligplot server. Hydrogen bond was shown by the green dashed line. Interacting residues were shown by the half moon containing spoke. Ligands are shown by the violet color stick.
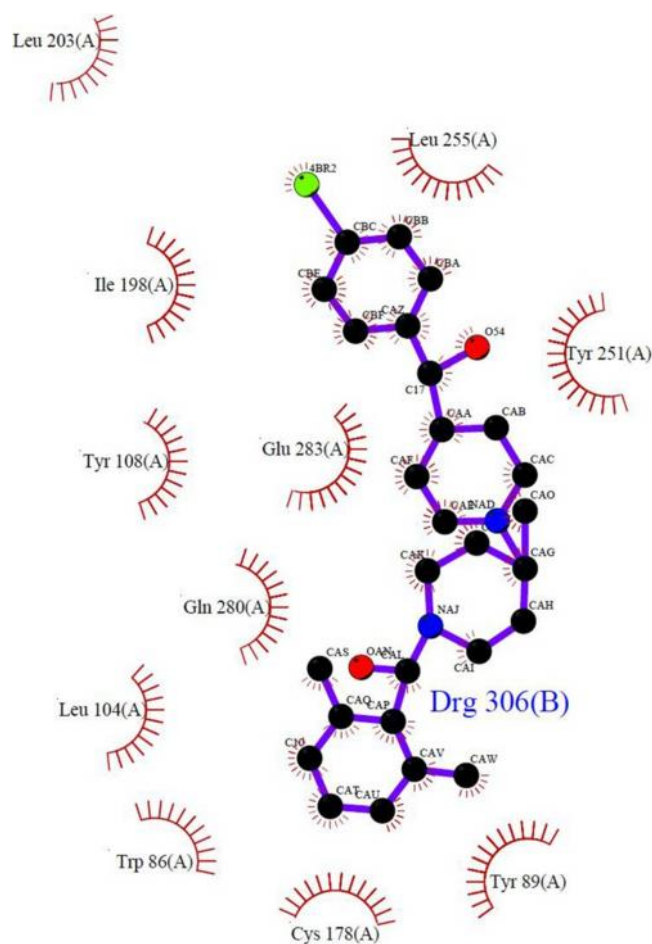
Figure 31: 2D-schematic plot interactions between 25-OPPA-CCR5 were generated using Ligplot server. Interacting residues were shown by the half moon containing spoke. Ligand is shown by the violet color stick.

4. Discussion

In this study we performed comparative modeling, docking, 3D-QSAR and MD simulation of CCR5 inhibitors. 3D-model of CCR5 was developed based on the recently reported CXCR4 (PDB code: 3ODU) (27) structure. We chose this structure over traditional bovine rhodopsin and $_2$-adrenergic receptors because higher sequence identity and larger binding pocket volume. CCR5 shares 35.4 % sequence identity with CXCR4, which is higher than bovine rhodopsin (23.3 %) and $_2$-adrenergic (25.4 %) receptors. Most of the previous modeling reports are based on the bovine rhodopsin and $_2$-adrenergic receptors (21, 24, 66-70). However previous report stated the importance of larger binding pocket volume in GPCR, which was developed based on the bovine rhodopsin structure (71). Quality of developed CCR5 model was assessed using various validation servers (Figures 2-4) (30-32), which indicated the reliability of model. A putative binding pocket for the docking study was defined according to the previous mutagenesis reports and knowledge of important acidic amino acid residue in TM7 (Glu283) (39-42).

Molecular docking of the TAK779, TAK220, compound 85, 14-OPPA, 25-OPPA and 37-OPPA were done to know the interaction at atomic level. Autodock program was used for the docking study. Our docking results are in line with the previous reports, which states that the inhibitors dock into TM1, TM2, TM3, TM5, TM6 and TM7 (39-42, 72). Our results are in line with the previous study reported by Li et al., (68), which states that Glu283 (TM7) is an important residue in the positively

charged inhibitor binding to CCR5. Also, this report confirmed that Tyr108 (TM3), Ile198 (TM5), Tyr251 (TM6) and Glu283 (TM7) are the crucial residues in all the tested compounds (Anibamine, Aplaviroc, Maraviroc and Vicriviroc) and in agreement with our results for TAK779, TAK220 and compound **85** binding. Our results are in line with previously published mutagenesis studies, which concluded that mutation of Glu283Ala reduces the binding affinity of TAK779 (40, 69). As with the CCR5 inhibitors, positively charged or basic nitrogen atoms, as well as aromatic groups, are the key pharmacophore elements of antagonists.

Docked pose of TAK779 (Figure 5) indicated that the hydrogen bond interaction of the tetrahydropyran rings oxygen atom with the hydroxyl group of Tyr37 (TM1) is essential for the inhibitory effect, which is consistent with previously published results (69). Furthermore, the docked model places only the N,N-dimethyl-N-benzyl-tetrahydro-2H-pyran-4-aminium moiety of TAK779 in the pocket formed by TM1, TM2 and TM7. However, the larger 4-methylphenyl-6,7-dihydro-5H-benzocycloheptenyl moiety, on the other hand, is predicted to bind outside this pocket and to interact with residues in TM5 and TM6, which is consistent with a previous report (73). TAK220 finds out similar binding site as of TAK779 and interacted with CCR5 residues. Our result is in accordance with the previous report, which indicated that Gly163 and Ile198 interact with TAK220 (41, 73). It was also observed that the Asn252 and Leu255 interacts with TAK220, which is in line with previous Alanine scanning mutagenesis reports indicating that mutation of these two residues are

resistant to antiviral effect of TAK220 (41). Similar kind of binding mode and interaction were observed for the compound **85**. It has been previously demonstrated that for inhibitor effectiveness basic nitrogen and hydrophobic aromatic groups are highly desirable (38, 74, 75).

Docking based 3D-QSAR was done on piperidine-4-carboxamide derivatives. All the inhibitors of dataset were sketched based on the docked conformation of template (compound **85**) and minimized inside CCR5 by keeping constraint on core part. 3D-QSAR was used to identify the requirement of high affinity binder to CCR5. CoMFA and CoMSIA were utilized to develop 3D-QSAR models. In this study, rational selection was done based on the 2D fingerprints as molecular topological descriptors contain information about molecular fragments. Here, designed subset was used as internal validation (training) set for 3D-QSAR. Model designed from this set was utilized into the prediction of remaining compounds (test set). We developed robust and predictive CoMFA and CoMSIA models in terms of q^2, r^2, Q^2 and r^2_{pred}. Previous study stated that the $q^2 > 0.5$ is not always a sufficient measure to assess reliability of QSAR model. But, external validation is the only way to establish a reliable QSAR model (76). Thus, we validated developed models by test set (25 compounds), which gave satisfactory r^2_{pred} for both CoMFA (0.680) and CoMSIA (0.705). To strengthen the developed QSAR models we performed progressive scrambling. Developed CoMFA and CoMSIA models were validated for redundancy correction of LOO q^2. Previously it is reported that the Q^2 0.350 signifies that the

originally unperturbed LOO QSAR model is robust and predictive. Our CoMFA and CoMSIA models were validated using progressive scrambling, which has obtained the satisfactory Q^2 values for CoMFA (0.619) and CoMSIA (0.522).

Developed CoMFA and CoMSIA models were plotted back onto CCR5 to know the requirements for the improved inhibitory potency. Contour maps revealed that, at the R^1 position, filling the cavity with the bulky group containing a polar substituent is highly desirable for inhibitory activity, because the ligand interacts through a hydrogen bond with the receptor. Our 3D-QSAR models with the receptor information demonstrated the importance of electrostatic, hydrogen bond donor and acceptor parameters for the CCR5 antagonism. Moreover, molecular view of ligand-protein interactions demonstrated how the ligand is interacting with the receptor, and can give correct guidelines to optimize ligands with enhanced inhibitory potency.

Our group is actively involved in chemokine receptors (CCR2 and CCR5) modeling (77, 78). In particular, previously we published ligand based HQSAR, CoMFA and CoMSIA study for 38 CCR5 inhibitors (77). Current study differed from the previous study in the sense that we used receptor based conformation for the most potent ligand (compound **85**), and all the other ligands were modeled inside the receptor. In addition, two well known inhibitors (TAK779 and TAK220) were docked to map the binding site in CCR5, and to locate the crucial interacting residues. Very recently, we performed the modeling study of dual inhibitor (CCR2 and CCR5) and identified the residues important for dual antagonists in CCR2 and CCR5 cavities

through an in silico modeling and MD simulation (78). Our interaction model of ligand-CCR5 revealed the importance of Tyr89 and Gly167.

Docking and MD simulation was performed for the oxamino-piperidino-piperidine amide (OPPA) derivatives (14-OPPA, 25-OPPA and 37-OPPA). MD simulations can provide intimate details of individual particles motions as a function of time, and thus, can be used to tackle specific questions about ligand-receptor interactions without resorting to laboratory-based experimentation. Although experiments play a crucial role in the validation of simulation methodologies, as comparisons of simulation and experimental data serve to confirm the validities of calculated results and provide criteria for improving simulation methods. We used three derivatives of OPPA for docking and MD simulations. 20000 ps MD simulations were performed for the complexes (ligand-CCR5-DPPC-Sol-Cl). All the simulations were performed under constant temperature (300 K) and pressure (1 bar) conditions.

Potential energy of the simulated systems were calculated, which indicated the energy decreased gradually and systems were stabilized. Our results showed that the disulfide bridges (Cys20-Cys269 and Cys101-Cys178) were intact after MD simulations, indicating its importance to keep structural organization of TM, which is in accordance with the previous reports (62, 63). Our MD simulation results indicated that the N-terminal region and ECL2 (Figure 25) are the most dynamic loops, which is supported by the previous studies (62-65) indicating that dynamic nature of loop regions are necessary for viral entry. MD simulation results indicated that the simulated

inhibitors interacted through the salt bridge and hydrophobic interactions. MD simulation of the most potent compound (14-OPPA) indicated that it was oriented into the pocket and interacted with new residues such as Ala29, Ala90, Ser180, His181, Phe166, and Trp190. Docking identified salt bridge distance was decreased after MD simulation for all the simulated inhibitors, indicating the importance of Glu283, which is in line with previous reports (40, 69). MD simulation result suggests that the Ala29 (TM1) and Ala90 (TM2) are important in the interaction with inhibitors, which is supported by the previous report indicating that mutation by aliphatic amino acid residues altered binding of TAK779, AD-101 and SCH-C (69). All the simulated inhibitors (14-OPPA, 25-OPPA and 37-OPPA) showed hydrophobic interaction with the Trp86 (TM2), which is in accordance with previous report (67) stated that mutation of Trp86Ala showed significant reduction in binding affinities of TAK779, TAK220, Maraviroc and Vicriviroc. MD simulation results of 25-OPPA shows the importance of Thr195, which is supported by the previous report (67) indicating mutation of Thr195Ala reduced the binding affinity for Aplaviroc, Maraviroc, Vicriviroc and TAK779 in CCR5.

The entire MD simulated models found Ile198 as an important residue in the interactions. Our results in line with prior mutational studies indicating Ile198 is a vital residue in the binding pocket (39, 41, 73, 78). Nishikawa *et al*. (41) explained the importance of Leu255 (TM6) for interaction with TAK779 and TAK220, which is in accordance with our MD simulation results. From MD simulation results we can say

that the most potent inhibitor (14-OPPA) interacts with all seven TMs, whereas medium potent inhibitor (37-OPPA) interact with all the TMs except TM1, and least potent inhibitor devoid of any interaction from TM1 and TM4. After MD simulation it was observed that the salt bridge distance was reduced, which indicates the importance of this interaction. In fact it seems that Glu283 anchors the ligand and pulled towards it (39, 42). Nishikawa *et al.* also demonstrated that the mutation of Glu283 abolished the activity of TAK779 and TAK220 (41).

Our MD simulation results are in line with the most of previous mutagenesis reports (39-42, 67, 69, 73, 79), but in contrast with the Shahlaei *et al.* report (24). However, inhibitors (14-OPPA, 25-OPPA and 37-OPPA) in the current MD study identified essential residues (Trp86, Tyr89, Tyr108, Phe112, Ile198, Tyr251, Leu255 and Gln280) and interacted closely, which is in line with previous reports (39-42, 69). Partial involvement of ECL2 was observed in the all MD simulated binding modes.

Here, we proposed general binding modes of OPPA derivatives based on MD simulation results of the most potent (14-OPPA), medium potent (37-OPPA) and least potent (25-OPPA) inhibitors. We find some trend for the simulated inhibitors, such as each inhibitor shows essential salt bridge contact with Glu283, which seems to be necessary for inhibitory activity. Also central polar region (TM7) flanked by the two aromatic pockets (pocket 1, TM1-3 and pocket 2, TM3-6) in CCR5 and shared by simulated inhibitors. All the simulated inhibitors showed important interactions with

the Trp86, Tyr89, Tyr108, Phe112, Ile198, Tyr251, Leu255 and Gln280 residues. Our

MD simulation results could be exploited to design and develop novel CCR5 inhibitors.

5. Conclusion

This work addresses the homology modeling, docking, 3D-QSAR and MD simulation study of CCR5. Computational 3D-model of CCR5 was developed using recently reported CXCR4 X-ray crystal structure as a template. Developed model was validated using RAMACHANDRAN plot, ERRAT plot and ProSA energy. Molecular docking study was performed for the potent CCR5 inhibitors inside 3D-model of CCR5. Docked conformations of TAK779, TAK220 and compound **85** were selected according to the previous reports on important residues. Docked conformation of compound **85** was considered as a bioactive conformation and rest of the molecules were developed based on the bioactive conformations. Predictive 3D-QSAR models in terms of q^2, r^2 and Q^2 were developed using CoMFA and CoMSIA. Developed QSAR models were validated using test set compounds, which gave satisfactory prediction (r^2_{pred}). CoMFA/CoMSIA models were plotted as 3D coefficient contour maps onto CCR5 active site to explain in detail SAR. Our 3D-QSAR models allow us to focus on those regions, where steric, electrostatic, hydrophobic and hydrogen bond donor/acceptor parameters plays important role in the interactions of inhibitors with CCR5. Contour map analyses revealed that at R^1 position polar substituents with proper orientation is highly desirable for maximal inhibitory activity. Also our model indicates the importance of hydrogen bond acceptor/donor interaction for CCR5 activity. Para position of central phenyl ring is well-tolerated for steric and hydrophobic groups, which displayed favorable green and yellow contour

To know the important residues of CCR5 and detailed interaction mechanism of ligand-CCR5, we performed MD simulation of three CCR5 inhibitors (14-OPPA, 25-OPPA and 37-OPPA) in explicit lipid bilayers. 20000 ps MD simulations were performed for the three complexes (ligand-CCR5-DPPC-Sol-Cl). It was observed that the substituents reoriented into the binding site and stabilized the conformations. Potential energy and RMSDs indicated that both ligands and CCR5 were stable in the simulations. Our simulation indicated importance of salt bridge to hold ligand tightly inside CCR5. Simulation results also suggested Trp86, Tyr89, Tyr108, Phe112, Ile198, Tyr251, Leu255, and Gln280 are important in the binding locus, which is in accordance with previous mutational studies.

Obtained results could be exploited to design and develop novel CCR5 inhibitors with the improved potency.

References

[1]   Tyndall, J.D., and Sandilya, R., GPCR agonists and antagonists in the clinic, Med. Chem., 1 (2005) 405-421.

[2]   Fredriksson, R., Lagerstrom, M.C., Lundin, L.G., and Schioth, H.B., The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints, Mol. Pharmacol., 63 (2003) 1256-1272.

[3]   Murphy, P.M., Baggiolini, M., Charo, I.F., Hebert, C.A., Horuk, R., Matsushima, K., Miller, L.H., Oppenheim, J.J., and Power, C.A., International Union of Pharmacology. XXII. Nomenclature for chemokine receptors, Pharmacol. Rev., 52 (2000) 145-176.

[4]   Gerard, C., and Rollins, B.J., Chemokines and disease, Nature Immunol., 2 (2001) 108-115.

[5]   http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology /2012/gr2012/20121120_UNAIDS_Global_Report_2012_en.pdf Accessed on April 15, 2013.

[6]   Caliendo, A.M, and Hirsch M.S., Combination therapy for infection due to human immunodeficiency virus type 1, Clin. Infect. Dis., 18 (1994) 516-524.

[7]   Appelt, K., Crystal structures of HIV-1 protease-inhibitor complexes, Perspect. Drug Discovery Des., 1 (1993) 23-48.

[8]   Condra, J.H., Schleif, W.A., Blahy, O.M., Gabryelski, L.J., Graham, D.J., Quintero, J.C., Rhodes, A., Robbins, H.L., Roth, E., and Shivaprakash, M., In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors, Nature, 374 (1995) 569-571.

[9]   Furtado, M.R., Callaway, D.S., Phair, J.P., Kunstman, K.J., Stanton, J.L., Macken, C.A., Perelson, A.S., and Wolinsky, S.M., Persistence of HIV-1 transcription in peripheral-blood mononuclear cells in patients receiving potent antiretroviral therapy, N. Eng. J. Med., 340 (1999) 1614-1622.

[1 0]   Koot, M., van't Wout, A.B., Kootstra, N.A., de Goede, R.E., Tersmette, M., and Schuitemaker, H., Relation between changes in cellular load, evolution of viral phenotype, and the clonal composition of virus populations in the course of human immunodeficiency virus type 1 infection, J. Infect. Dis., 173 (1996) 349-354.

[1 1]   Chan, D.C., and Kim, P.S., Endocytic entry of HIV-1, Curr. Biol., 10 (2000) 1005-1008.

[1 2]   Piot, P., Bartos, M., Ghys, P.D., Walker, N., and Schwartländer, B., The global impact of HIV/AIDS, Nature, 410 (2001) 968-973.

[1 3]   Strader, C.D., Fong, T.M., Tota, M.R., Underwood, D., and Dixon, R.A.F., Structure and function of G protein-coupled receptors, Ann. Rev. Biochem., 63 (1994) 101-132.

[1 4]   Cocchi, F., DeVico, A.L., Garzino-Demo, A., Arya, S.K., Gallo, R.C., and Lusso, P., Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the major HIV-suppressive factors produced by CD8+ T cells, Science, 270 (1995) 1811-1815.

[1 5]   Liu, R., Paxton, W.A., Choe, S., Ceradini, D., Martin, S.R., Horuk, R., MacDonald, M.E,. Stuhlmann, H., Koup R.A., and Landau, N.R., Cell, 86 (1996) 367-377.

[1 6]   Marmor, M., Sheppard, H.W., Donnell, D., Bozeman, S., Celum, C., Buchbinder, S., Koblin B., and SeageIii, G.R., Homozygous and heterozygous CCR5-Delta32 genotypes are associated with resistance to HIV infection, J. Acquir. Immune. Defic. Syndr., 27 (2001) 472-481.

[1 7]   Samson, M., Libert, F., Doranz, B.J., Rucker, J., Liesnard, C., Farber, C.M. , Saragosti, S., Lapoumeroulie, C., Cognaux J., and Forceille, C., Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene, Nature, 382 (1996) 722-725.

[1 8]   Afantitis, A., Melagraki, G., Sarimveis, H., Koutentis, P.A., Markopoulos, J., and Igglessi-Markopoulou, O., Investigation of substituent effect of 1-(3,3-diphenylpropyl)- piperidinyl phenylacetamides on CCR5 binding affinity using QSAR and virtual screening techniques, J. Comput.-Aided Mol. Des., 20 (2006) 83-95.

[19] Aher, Y.D., Agrawal, A., Bharatam, P.V., and Garg, P., 3D-QSAR studies of substituted 1-(3,3-diphenylpropyl)- piperidinyl amides and ureas as CCR5 receptor antagonists, J. Mol. Model., 13 (2007) 519-529.

[20] Carrieri, A., Pérez Nueno, V.I., Fano, A., Pistone, C., Ritchie, D.W., and Teixidó, J., Biological profiling of anti- HIV agents and insight into CCR5 antagonist binding using in silico techniques, ChemMedChem, 4 (2009) 1153-1163.

[21] Xu, Y., Liu, H., Niu, C., Luo, C., Luo, X., Shen, J., and Jiang, H., 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, Bioorg. Med. Chem., 12 (2004) 6193-6208.

[22] Kothandan, G., Gadhe, C.G., and Cho, S.J., Structural insights from binding poses of CCR2 and CCR5 with clinically important antagonists: A combined in silico study, PLoS ONE, 7 (2012) e32864.

[23] Metz, M., Bourque, E., Labrecque, J., Danthi, S.J., Langille, J., Harwig, C., Yang, W., Darkes, M.C., Lau, G., Santucci, Z., Bridger, G.J., Schols, D., Fricker, S.P., and Skerlj, R.T., Prospective CCR5 small molecule antagonist compound design using a combined mutagenesis/modeling approach, J. Am. Chem. Soc., 133 (2012) 16477-16485.

[24] Shahlaei, M., Madadkar-Sobhani, A., Mahnam, K., Fassihi, A., Saghaie, L., and Mansourian, M., Homology modeling of human CCR5 and analysis of its

binding properties through molecular docking and molecular dynamics simulation, Biochimica et Biophysica Acta (BBA)-Biomembranes, 1808, (2011) 802-817.

[２５]   Altschul, S., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res., 25 (1997) 3389-3402.

[２６]   Tusnady, G., Dosztanyi, Z., and Simon, I., Transmembrane proteins in the Protein Data Bank: identification and classification. Bioinformatics, 20 (2004) 2964-2972.

[２７]   Wu, B., Chien, E.Y.T., Mol, C.D., Fenalti, G., Liu, W., Katritch, V., Abagyan, R., Brooun, A., Wells, P., Bi, F.C., Hamel, D.J., Kuhn, P., Handel, T.M., Cherezov, V., and Stevens, R.C., Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists, Science, 330 (2010) 1066-1071.

[２８]   Thompson, J.D., Higgins, D.G., and Gibson, T.J., CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, Nucleic Acids Res., 22 (1994) 4673-4680.

[２９]   Eswar, N., John, B., Mirkovic, N., Fiser, A., Ilyin, V.A., Pieper, U., Stuart, A.C., Marti-Renom, M.A., Madhusudhan, M.S., Yerkovich, B., and Sali, A.,

Tools for comparative protein structure modeling and analysis, Nucleic Acids Res., 31 (2003) 3375-3380.

[30] Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M., PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr., 26 (1993) 283-291.

[31] Colovos, C., and Yeates, T.O., Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci., 9 (1993) 1511-1519.

[32] Wiederstein, M., & Sippl, M.J., ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res., 35 (2007) W407-W410.

[33] Imamura, S., Ishihara, Y., Hattori, T., Kurasawa, O., Matsushita, Y., Sugihara, Y., Kanzaki, N., Iizawa, Y., Baba, M., and Hashiguchi, S., CCR5 antagonists as anti-HIV-1 agents. 1. Synthesis and biological evaluation of 5-oxopyrrolidine-3-carboxamide derivatives, Chem. Pharm. Bull., 52 (2004) 63-73.

[34] Imamura, S., Kurasawa, O., Nara, Y., Ichikawa, T., Nishikawa, Y., Iida, T., Hashiguchi, S., Kanzaki, N., Iizawa, Y., Baba, M., and Sugihara, Y., CCR5 antagonists as anti-HIV-1 agents. Part 2: Synthesis and biological evaluation of N-[3-(4-benzylpiperidin- 1-yl) propyl]-N,N0-diphenylureas, Bioorg. Med. Chem., 12 (2004) 2295-2306.

[3 5]   Imamura, S., Nishikawa, Y., Ichikawa, T., Hattori, T., Matsushita, Y., Hashiguchi, S., Kanzaki, N., Iizawa, Y., Baba, M., and Sugihara. Y., CCR5 antagonists as anti-HIV-1 agents. Part 3: Synthesis and biological evaluation of piperidine-4-carboxamide derivatives, Bioorg, Med. Chem., 13 (2005) 397-416.

[3 6]   Imamura, S., Ichikawa, T., Nishikawa, Y., Kanzaki, N., Takashima, K., Niwa, S., Iizawa, Y., Baba, M, and Sugihara. Y., Discovery of a piperidine-4-carboxamide CCR5 antagonist (TAK-220) with highly potent Anti-HIV-1 activity, J. Med. Chem., 49 (2006) 2784-2793.

[3 7]   Palani, A., Shapiro, S., Clader, J.W., Greenlee, W. J., Cox, K., Strizki, J., and Baroudy, B.M., Discovery of 4-[(Z)-(4-bromophenyl)-(ethoxyimino) methyl]-1'-[(2, 4-dimethyl-3-pyridinyl) carbonyl]-4'-methyl-1, 4'-bipiperidine N-oxide (SCH 351125): an orally bioavailable human CCR5 antagonist for the treatment of HIV infection, J. Med. Chem., 44 (2001) 3339-3342.

[3 8]   Palani, A., Shapiro, S., Josien, H., Bara, T., Clader, J.W., Greenlee, W.J., Cox, K., Strizki, J.M., and Baroudy, B.M., Synthesis, SAR, and biological evaluation of oximino-piperidinopiperidine amides. 1. Orally bioavailable CCR5 receptor antagonists with potent anti-HIV activity. J. Med. Chem., 45 (2002) 3143-3160.

[3 9]   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., and Moore, J.P.,

A binding pocket for a small molecule inhibitor of HIV-1 entry within the transmembrane helices of CCR5, Proc. Natl Acad. Sci. U.S.A., 97 (2000) 5639-5644.

[4 0] 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., Binding of 2-aryl-4-(piperidin-1-yl) butanamines and 1, 3, 4-trisubstituted pyrrolidines to human CCR5: a molecular modeling- guided mutagenesis study of the binding pocket, Biochemistry 42 (2003) 1544-1550.

[4 1] Nishikawa, M., Takashima, K., Nishi, T., Furuta, R.A., Kanzaki, N., Yamamoto, Y., and Fujisawa. J., Analysis of binding sites for the new small-molecule CCR5 antagonist TAK-220 on human CCR5, Antimicrob. Agents Chemother., 49 (2005) 4708-4715.

[4 2] 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., and Dragic. T., Analysis of the mechanism by which the small-molecule CCR5 antagonists SCH-351125 and SCH-350581 inhibit human immunodeficiency virus type 1 entry, J. Virol., 77 (2003) 5201-5208.

[4 3] SYBYL 8.1 Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA (2008).

[4 4]   Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., and Olson. A.J., Automated docking using a Lamarckian genetic algorithmand an empirical binding free energy function, J. Comput. Chem., 19 (1998) 1639-1662.

[4 5]   Cramer III, R.D., Patterson, D.E., and Bunce. J.D., Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins, J. Am. Chem. Soc., 110 (1988) 5959-5967.

[4 6]   Klebe, G., Abraham, U., and Mietzner, T., Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity, J. Med. Chem., 37 (1994) 4130-4146.

[4 7]   Dunn III, W.J., Wold, S., Edlund, U., Hellberg, S., and Gasteiger. J., Multivariate structure activity relationships between data from a battery of biological tests and an ensemble of structure descriptors: The PLS method, Quant. Struct.-Act. Relat., 3 (1984) 131-137.

[4 8]   Clark, R.D., and Sprous, D.G., Validating models based on large datasets. In Rational Approaches to Drug Design; Holtje, H.D., and Sippl, W., (Eds.). Prous Science SA, Barcelona, (2001), 475-485.

[4 9]   Clark, R.D., and Fox, P.C., Statistical variation in progressive scrambling. J. Comput. Aided. Mol. Des., 18 (2004) 563-576.

[5 0]   Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., and Berendsen, H.J.C., GROMACS: Fast, flexible, and free, J. Comput. Chem., 26 (2005) 1701-1718.

[5 1]   Oostenbrink, C., Villa, A., Mark, A.E., and Van Gunsteren, W.F., A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS forcefield parameter sets 53A5 and 53A6, J. Comput. Chem., 25 (2004) 1656-1676.

[5 2]   Aalten, D.M.F., Bywater, R., Findlay, J.B.C., Hendlich, M., Hooft, R.W. W., and Vriend, G., PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules, J. Comput.-Aided Mol. Des., 10 (1996) 255-262.

[5 3]   Tieleman, P., MacCallum, J., Ash, W., Kandt, C., Xu, Z., and Monticelli, L., Membrane protein simulations with a united-atom lipid and all-atom protein model: lipid_protein interactions, side chain transfer free energies and model proteins, J. Phys. Condens. Matter, 18 (2006), S1221-S1234.

[5 4]   Kandt, C., Ash, W.L., and Peter Tieleman, D., Setting up and running molecular dynamics simulations of membrane proteins, Methods, 41 (2007) 475-488.

[5 5]   Hermans, J., Berendsen, H.J.C., Van Gunsteren, W.F., and Postma, J.P.M., A consistent empirical potential for water-protein interactions, Biopolymers, 23 (1984) 1513-1518.

[5 6]   Berendsen, H.J.C., Postma, J.P.M., Van Gunsteren, W.F., DiNola, A., and Haak, J.R., Molecular dynamics with coupling to an external bath, J. Chem. Phys., 81 (1984) 3684-3690.

[5 7]   Miyamoto, S., and Kollman, P.A., SETTLE: An analytical version of the SHAKE and RATTLE algorithm for rigid water models, J. Comput. Chem., 13 (1992) 952-962.

[5 8]   Hess, B., P-LINCS: A parallel linear constraint solver for molecular simulation, J. Chem. Theory Comput., 4 (2008) 116-122.

[5 9]   Darden, T., York, D., and Pedersen, L., Particle mesh Ewald: An N.log (N) method for Ewald sums in large systems, J. Chem. Phys., 98 (1993) 10089-10092.

[6 0]   Hoover, W.G., Canonical dynamics: Equilibrium phase space distributions, Phys. Rev. A, 31 (1985) 1695-1697.

[6 1]   Parrinello, M., and Rahman, A., Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys., 52 (1981) 7182-7190.

[6 2]   Genoud, S., Kajumo, F., Guo, Y., Thompson, D., and Dragic, T., CCR5-mediated human immunodeficiency virus entry depends on an amino-terminal gp120-binding site and on the conformational integrity of all four extracellular domains, J. Virol., 73 (1999) 1645-1648.

[6 3]   Blanpain, C., Lee, B., Vakili, J., Doranz, B. J., Govaerts, C., Migeotte, I., Sharron, M., Dupriez, V., Vassart, G., Doms, R. W., and Parmentier, M.,

Extracellular cysteines of CCR5 are required for chemokine binding, but dispensable for HIV-1 coreceptor activity, J. Biol. Chem., 274 (1999) 18902-18908.

[6 4] Horuk, R., Molecular properties of the chemokine receptor family, Trends Pharmacol. Sci., 15 (1994) 159-165.

[6 5] Liu, S., Shi, X., Liu, C., and Sun, Z., Characterize dynamic conformational space of human CCR5 extracellular domain by molecular modeling and molecular dynamics simulation, J. Mol. Struct.: Theochem, 673 (2004) 133-143.

[6 6] Fano, A., Ritchie, D. W., and Carrieri. A., Modeling the structural basis of human CCR5 chemokine receptor function: from homology model building and molecular dynamics validation to agonist and antagonist docking, J. Chem. Inf. Model., 46 (2006) 1223-1235.

[6 7] Kondru, R., Zhang, J., Ji, C., Mirzadegan, T., Rotstein, D., Sankuratri, S., and Dioszegi. M., Molecular interactions of CCR5 with major classes of small-molecule anti-HIV CCR5 antagonists, Mol. Pharmacol., 73 (2008) 789-800.

[6 8] Li, G., Haney, K.M., Kellogg, G.E., and Zhang. Y., Comparative docking study of anibamine as the first natural product CCR5 antagonist in CCR5 homology models, J. Chem. Inf. Model., 49 (2009) 120-132.

[6 9] Seibert, C., Ying, W., Gavrilov, S., Tsamis, F., Kuhmann, S. E., Palani, A., Tagat, J.R., Clader, J.W., McCombie, S.W., Baroudy, B.M., Smith, S.O.,

Dragic, T., Moore, J.P., and Sakmar, T.P., Interaction of small molecule inhibitors of HIV-1 entry with CCR5, Virology, 349 (2006) 41-54.

[7 0]   Wang, T., & Duan, Y., Binding modes of CCR5-targetting HIV entry inhibitors: Partial and full antagonists, J. Mol. Graph. Model., 26 (2008) 1287-1295.

[7 1]   Kimura, S.R., Tebben, A.J., and Langley, D.R., Expanding GPCR homology model binding sites via a balloon potential: A molecular dynamics refinement approach, Proteins, 71 (2008) 1919-1929.

[7 2]   Govaerts, C., Bondue, A., Springael, J. Y., Olivella, M., Deupi, X., Le Poul, E., Wodak, S.J., Parmentier, M., Pardo, L., and Blanpain, C., Activation of CCR5 by chemokines involves an aromatic cluster between transmembrane helices 2 and 3, J. Biol. Chem., 278 (2003) 1892-1903.

[7 3]   Paterlini, M.G., Structure modeling of the chemokine receptor CCR5: implications for ligand binding and selectivity, Biophys. J., 83 (2002) 3012-3031.

[7 4]   Shiraishi, M., Aramaki, Y., Seto, M., Imoto, H., Nishikawa, Y., Kanzaki, N., Kanzaki, N., Okamoto, M., Sawada, H., Nishimura, O., Baba, M., and Fujino, M., Discovery of novel, potent, and selective small-molecule CCR5 antagonists as anti-HIV-1 agents: synthesis and biological evaluation of anilide derivatives with a quaternary ammonium moiety, J. Med. Chem., 43 (2000) 2049-2063.

[7 5]   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., and Strizki, J.M., Piperazine-based CCR5 antagonists as HIV-1 inhibitors. II. Discovery of 1-[(2,4-dimethyl-3-pyridinyl)carbonyl]-4-methyl-4-[3(S)-methyl-4-[1(S)-[4-(trifluoromethyl)phenyl]ethyl]-1-piperazinyl]-piperidine N1-oxide (Sch-350634), an orally bioavailable, potent CCR5 antagonist, J. Med. Chem., 44 (2001) 3343-3346.

[7 6]   Golbraikh, A., and Tropsha, A., Beware of q2! J. Mol. Graph. Model., 20 (2002) 269-276.

[7 7]   Gadhe, C.G., Lee, S.H., Madhavan, T., Kothandan, G., Choi, D., and Cho, S.J., Ligand based CoMFA, CoMSIA and HQSAR analysis of CCR5 antagonists, Bull. Korean Chem. Soc., 31 (2010) 2761-2779.

[7 8]   Kothandan, G., Gadhe, C.G., and Cho, S.J., Structural insight from binding poses of CCR2 and CCR5 with clinically important antagonists: a combined in silico study, PLoS ONE 7 (2012) e32864.

[7 9]   Billick, E., Seibert, C., Pugach, P., Ketas, T., Trkola, A., Endres, M. J., Murgolo, N.J., Coates, E., Reyes, G.R., Baroudy, B.M., Sakmar, T.P., Moore, J.P., and Kuhmann, S.E., The differential sensitivity of human and rhesus macaque CCR5 to smallmolecule inhibitors of human immunodeficiency virus type 1 entry is explained by a single amino acid difference and suggests a mechanism of action for these inhibitors, J. Virol., 78 (2004) 4134-4144.

Appendix

Appendix A

Acknowledgments

The completion of my doctoral studies would not have been possible without the help of many people. I may not name all of them here, but I remain grateful to them all.

First of all, I would like to express my sincere gratitude to my advisor respected **Prof. Seung Joo Cho**, for his love, guidance, patience, support and friendship both professionally and personally during the entire course of this dissertation. I would like to express my earnest appreciation for his informative and insightful guidance, his encouragement and patience, as well as his communicable passion towards work and life. I have learned not only scientific techniques from him, but also the scientific thinking and attitude. This dissertation could not have been accomplished without his wisdom and guidance. I am really glad that he gave me an opportunity to work with him. He could not even realize how much I have learned from him.

I would like to extend my warm regards to **Dr. Kishorebabu Bandamravuri, Dr. Gopal, Dr. Ankita De, Dr. Pawan Shahi**, **Dr. Bijoy Swaminathan, Dr. Murlidhar Chaurashiya, Dr. Thirumurthy Madhavan, and Dr. Kathir** for their wonderful support and motivation throughout Chosun University life.

I would also like to thank to my colleague that I have worked with**, Mr. Gugan,** for his continuous support in laboratory from the very first day of Chosun University.

Specially, I would like to thanks **Abhijeet Takawale, Gaurav Bedse, Kailas Karkhile and Pradeep Chaudhari** for being great friends and their motivations.

The chain of my gratitude would be definitely incomplete if I would forget to thanks my friends **Merlin, Vijay Sankar, Hari, Sudhan, Inbarasan, Binu, Rajesh, Palani, Muthu, Arvind, Anand, Pavithra, Piyush, Mrs. Yun Hwa Ko, Miss. Han Na Choi and Miss. Hyun** Chosun university friends, Chonnam University friends and all my Korean friends, Words just cannot describe the help they have done to me. I have been fortunate to be surrounded by countless good-hearted, friendly, and helpful people who have in one way or another helped me through the last four years in Gwangju.

My deepest gratitude goes to the most important people of my life, **Gorakshnath Gadhe** (father), **Nandabai Gadhe** (mother), **Dnyaneshwar Gadhe** (brother) and **Savita** (sister). My parents raised me wisely and well, and any good qualities anyone may have observed in me should be attributed to them. Their blessings have helped me to become what I am today. Last, but not the least, I offer my humble salutations to GOD for everything.

This thesis is dedicated to all of them.

Appendix B

List of Publications

1. Kothandan G., **Gadhe C.G**., Cho S.J., Theoretical characterization of Galanin receptor type 3 (Gal3) and its interaction with agonist (GALANIN) and antagonists (SNAP 37889 and SNAP 398299): An in silico analysis. *Chem. Biol. Drug Des*., 81 (2013) 757-774.

2. Madhavan T., **Gadhe C.G**., Kothandan G., Cho S.J., Enhancement of P-gylcoprotein modulators of arylmethylaminephenyl derivatives: an integrative modeling approach. *Med. Chem. Res*., 22 (2013) 2511-2523.

3. **Gadhe C.G**., Kothandan G., Cho S.J., Binding site exploration of CCR5 using in silico methodologies: a 3D-QSAR approach. *Arch. Pharm. Res*., 1 (2013) 6-31.

4. **Gadhe C.G**., Kothandan G., Cho S.J., Computational modeling of human coreceptor CCR5 antagonist as a HIV-1 entry inhibitor: using an integrated homology modeling, docking, and membrane molecular dynamics simulation analysis approach. *J. Biomol. Struct. Dyn*., (2012) DOI:10.1080/07391102.2012.732342.

5. Kothandan G., Madhavan T., **Gadhe C.G**., Cho S.J., A combined 3D QSAR and pharmacophore-based virtual screening for the identification of potent p38 MAP kinase inhibitors: an in silico approach. *Med. Chem. Res*., 22 (2013) 1773-1787.

6. **Gadhe C.G**., Kothandan G., Cho S.J. Large variation in electrostatic contours upon addition of steric parameters and the effect of charge calculation schemes in CoMFA on mutagenicity of MX analogues. *Mol. Simulat*., 38 (2012) 861-871.

7. Kothandan G[#]., **Gadhe C.G**[#]., Cho S.J., Structural insights from binding poses of CCR2 and CCR5 with clinically important antagonists: a combined in silico study. *PLoS ONE*, 7 (2012) e32864.

8. Madhavan T., Chung J.Y., Kothandan G., **Gadhe C.G**., Cho S.J., 3D-QSAR studies of JNK1 inhibitors utilizing various alignment methods. *Chem. Biol. Drug Des*., (2012) 79 53-67.

9. Kothandan G., **Gadhe C.G**., Madhavan T., Choi C.H., Cho S.J., Docking and 3D-QSAR (quantitative structure activity relationship) studies of flavones, the potent inhibitors of p-glycoprotein targeting the nucleotide binding domain. *Eur. J. Med. Chem*., 46 (2011) 4078-4088.

10. **Gadhe C.G**., Kothandan G., Madhavan T., Cho S.J., Molecular modeling study of HIV-1 gp120 attachment inhibitors. *Med. Chem. Res*., 21 (2011) 1892-1904.

11. Madhavan T., **Gadhe C.G**., Kothandan G., Lee K., Cho S.J., Various Atomic Charge Calculation Schemes of CoMFA on HIF-1 Inhibitors of Moracin Analogs. *Int. J. Quant. Chem*., 112 (2011) 995-1005.

12. Madhavan T., Kothandan G., **Gadhe C.G**., Cho S.J., QSAR analysis on PfPK7 inhibitors using HQSAR, CoMFA, and CoMSIA. *Med. Chem. Res*., 21 (2011) 681-693.

13. Kothandan G., **Gadhe C.G**., Madhavan T., Cho S.J., Binding site analysis of CCR2 through in silico methodologies: docking, CoMFA, and CoMSIA. *Chem. Biol. Drug Des*., 78 (2011) 161-174.

14. **Gadhe C.G**., Madhavan T., Kothandan G., Cho S.J., In silico quantitative structure-activity relationship studies on P-gp modulators of tetrahydroisoquinoline-ethyl-phenylamine series. *BMC Struct. Biol*., 11 (2011) 5.

15. **Gadhe C.G**., Lee S.H., Madhavan T., Kothandan G., Choi D., Cho S.J., Ligand Based CoMFA, CoMSIA and HQSAR Analysis of CCR5 Antagonists. *Bull. Korean Chem. Soc*., 31 (2010) 2761-2770.

16. **Gadhe C.G**., Madhavan T., Kothandan G., Lee T.B., Lee K., Cho S.J. Various Partial Charge Schemes on 3D-QSAR Models for P-gp Inhibiting Adamantyl Derivatives. *Bull. Korean Chem. Soc*., 32 (2011) 1604-1612.

 # indicates equally contributed