Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbamem

Homology modeling of human CCR5 and analysis of its binding properties through molecular docking and molecular dynamics simulation

Mohsen Shahlaei [a,b], Armin Madadkar-Sobhani [c,d,*], Karim Mahnam [e], Afshin Fassihi [b,*],
Lotfollah Saghaie [b], Mahboubeh Mansourian [b]

[a] Department of Medicinal Chemistry, School of Pharmacy, Kermanshah University of Medical Sciences, Kermanshah, Iran
[b] Department of Medicinal Chemistry, School of Pharmacy and Isfahan Pharmaceutical Sciences Research Center, Isfahan University of Medical Sciences, Isfahan, Iran
[c] Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran
[d] Department of Life Sciences, Barcelona Supercomputing Center, C\ Jordi Girona 31, Edificio Nexus II, 08028 Barcelona, Spain
[e] Biology Department, Faculty of Science, Shahrekord University, Shahrekord, Iran

A B S T R A C T

In this study, homology modeling, molecular docking and molecular dynamics simulation were performed to explore structural features and binding mechanism of some inhibitors of chemokine receptor type 5 (CCR5), and to construct a model for designing new CCR5 inhibitors for preventing HIV attachment to the host cell. A homology modeling procedure was employed to construct a 3D model of CCR5. For this procedure, the X-ray crystal structure of bovine rhodopsin (1F88A) at 2.80 Å resolution was used as template. After inserting the constructed model into a hydrated lipid bilayer, a 20 ns molecular dynamics (MD) simulation was performed on the whole system. After reaching the equilibrium, twenty-four CCR5 inhibitors were docked in the active site of the obtained model. The binding models of the investigated antagonists indicate the mechanism of binding of the studied compounds to the CCR5 obviously. Moreover, 3D pictures of inhibitor–protein complex provided precious data regarding the binding orientation of each antagonist into the active site of this protein. One additional 20 ns MD simulation was performed on the initial structure of the CCR5–ligand 21 complex, resulted from the previous docking calculations, embedded in a hydrated POPE bilayer to explore the effects of the presence of lipid bilayer in the vicinity of CCR5–ligand complex. This article is part of a Special Issue entitled Protein translocation across or insertion into membranes

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Chemokines are a large family of chemotactic proteins that have a key role in the immune and inflammatory response of various diseases and disorders including asthma and allergic diseases as well as autoimmune diseases such as rheumatoid arthritis (RA) [1]. Chemokines control leukocyte activation and immigration to sites of inflammation via interaction with a family of G-protein coupled receptors (GPCRs). The chemokine receptor type 5 (CCR5) functions physiologically as a receptor for the leukocyte chemoattractants RANTES, MIP-1a, and MIP-1b, and it has also recently been indicated to function physiologically as one of the important cell entry co-receptors for HIV-1 [2].

In lesions of multiple sclerosis (MS), CCR5 has been discovered on activated myeloid microglial cells and infiltrating T cells. CCR5 antagonists therefore might be helpful in suppressing the chronic inflammatory symptoms of this disorder.

Since the discovery of CCR5 as a co-receptor for HIV-1 cell entry, there has been an increased attempt in the pharmaceutical industry to develop CCR5 antagonists [3]. Undeniably, drug development process is often faced with the circumstances that a potent ligand should be designed for a given protein for which no experimentally determined structure is yet obtained. The most well-known examples are almost certainly the GPCRs, which have a significant role in many biochemical processes.

At present, 50% of all recently developed drugs are targeted against G-protein coupled receptors [4]. There are strong evidences that all GPCRs have a common fold [5]. Their general structure is consisting of a counterclockwise array of seven transmembrane (TM) α-helices of ~25 to 35 residues long (TMs 1 to 7), which extend across the cellular membrane linked by three extracellular (EC1 to 3) and three intracellular loops (IC1 to 3) [6–8]. The N-terminal region of GPCRs, which vary in length and function, is located on the extracellular part of the lipid membrane, while the C-terminal region is on the intracellular part.

In order to be suitable for a structure-based drug design process, the structure of the target protein must be available with a certain

* Corresponding authors. A. Madadkar-Sobhani is to be contacted at: Department of Bioinformatics, Institute of Biochemistry and Biophysics (IBB), University of Tehran, Tehran, Iran. Tel.: +98 21 66969261; fax: +98 21 66969261. A. Fassihi, Isfahan Pharmaceutical Sciences Research Center, 81746-73461, Isfahan, Iran. Tel.: +98 311 7922562; fax: +98 311 6680011.
E-mail addresses: armin@ibb.ut.ac.ir (A. Madadkar-Sobhani),
fassihi@pharm.mui.ac.ir (A. Fassihi).

level of accuracy. This is typically achieved by X-ray crystallography procedure. For membrane proteins whose crystallization does not take place as simply as their aqueous counterparts, obtaining a high-resolution structure is a particularly complicated job. The lipid environment in the membrane can have a serious impact on the conformation of the membrane protein [9], and it is not easy to obtain this final conformation via crystallography.
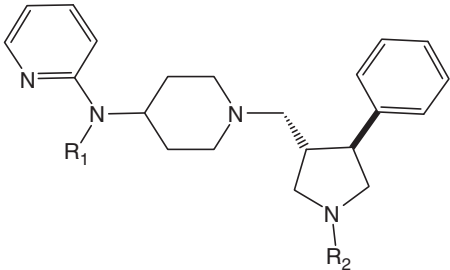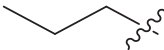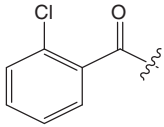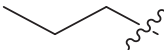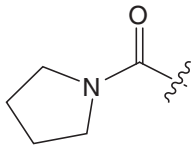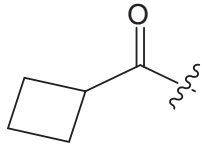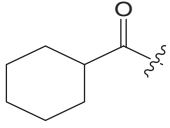
In the lack of an experimentally established crystal structure of a given protein, homology modeling is the best alternative to construct a reasonable three-dimensional (3D) model of the target. At present, homology modeling is the most accurate technique for 3D structure prediction of proteins [10]. The obtained models will be appropriate

for a wide variety of applications, such as structure-based molecular design and study of mechanism of action of their agonists or antagonists [11–13]. This method is capable to offer a rational structural model for a given membrane protein sequence with related templates sharing more than 25% sequence identity [14,15].

Several studies have been reported on molecular modeling of CCR5. Huang et al. have constructed CCR5 models associated with MIP-1b and RANTES to explore the residues crucial for the interaction between CCR5 and its natural ligands [16]. In another study, Paterlini et al. constructed a 3D model of CCR5 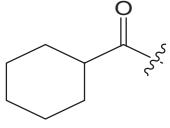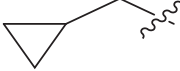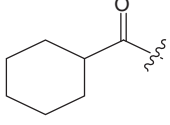to clarify the binding site and selectivity of the antagonist TAK779 [17]. Arseniev et al. also explored the conformational characteristics of CCR5 extracellular domain by structural modeling of

Table 1
General structures and structural details of compounds used in the docking procedure.

Compound	R_1	R_2	pIC_{50}	ΔG binding (kcal/mol)	Docking energy (kcal/mol)
1			7.85	−5.26	−8.49
2			8.05	−4.58	−7.45
3			8.00	−4.62	−8.14
4			7.92	−5.35	−8.22
5			8.22	−5.11	−8.34
6			8.05	−5.3	−6.34

Table 1 (*continued*)
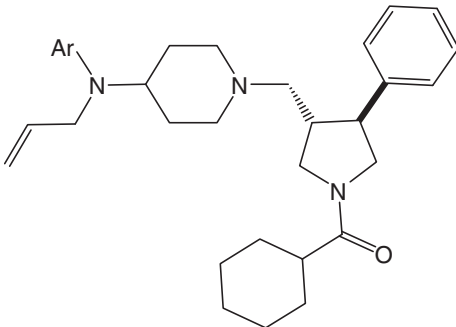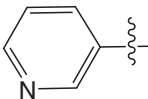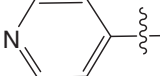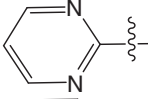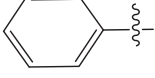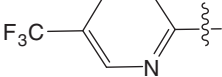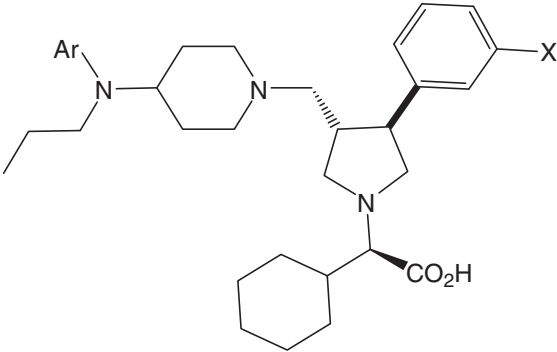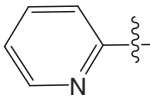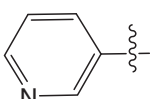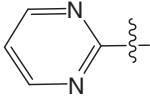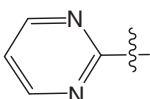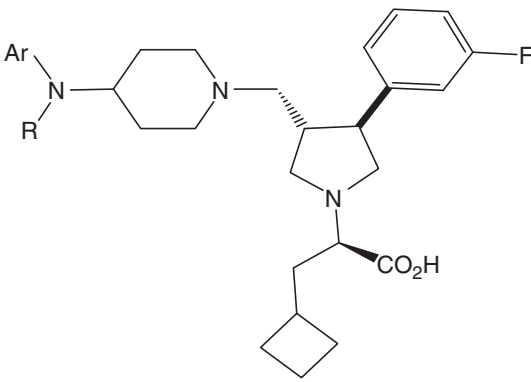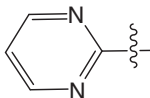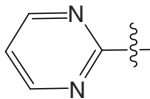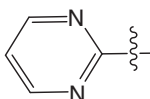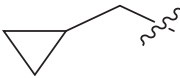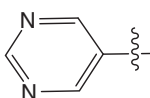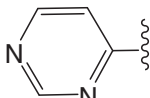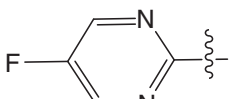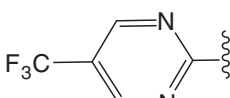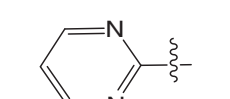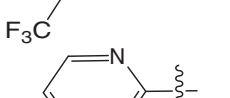
Compound	Ar	pIC$_{50}$	ΔG binding (kcal/mol)	Docking energy (kcal/mol)
7		7.64	−5.59	−9.1
8		7.18	−5.42	−7.2
9		7.92	−5.25	−7.63
10		7.38	−6	−9.06
11		7.68	−5.68	−8.64

Compound	Ar	X	pIC$_{50}$	ΔG binding (kcal/mol)	Docking energy (kcal/mol)
12		H	8.41	−5.68	−9.91
13		H	8.30	−5.29	−9.43
14		H	8.18	−5.06	−7.98
15		F	8.80	−4.82	−8.68

Table 1 (*continued*)

Compound	Ar	R	pIC$_{50}$	ΔG binding (kcal/mol)	Docking energy (kcal/mol)
16	pyrimidin-2-yl	n-C$_3$H$_7$-	8.74	−4.8	−7.89
17	pyrimidin-2-yl	n-C$_2$H$_5$-	8.82	−4.06	−8.11
18	pyrimidin-2-yl	cyclopropylmethyl	8.96	−4.46	−7.87
19	pyrimidin-5-yl	n-C$_3$H$_7$-	8.77	−4.29	−8.58
20	pyrimidin-4-yl	n-C$_3$H$_7$-	8.92	−4.76	−8.49
21	5-F-pyrimidin-2-yl	n-C$_3$H$_7$-	9.40	−4.79	−8.46
22	5-F$_3$C-pyrimidin-2-yl	n-C$_3$H$_7$-	9.30	−4.59	−7.98
23	4-F$_3$C-pyrimidin-2-yl	n-C$_3$H$_7$-	8.92	−4.53	−8.82
24	4-MeO-pyrimidin-2-yl	n-C$_3$H$_7$-	9.10	−5.26	−7.27

CCR5 [18]. In another report, Liu and coworkers elucidated some three-dimensional models of CCR5 using homology modeling approach followed by 1 ns molecular dynamics (MD) simulation [19]. A common shortcoming shared by the above-mentioned studies is that the authors explored CCR5 structure/function without considering influence of lipid bilayer on its folding. In other words, MD simulation was carried out without inserting protein in lipid bilayer. Here in this study, thanks to modern software and hardware, our primary objective is to investigate the presence of lipid bilayer in molecular dynamics simulation and its effects on dynamic behavior of the developed CCR5 model.

Previous experimental studies demonstrate that different binding interactions of ligands to a receptor lead to different receptor conformations [20]. Thus, we faced two challenges: (1) how to construct the 3D models of the CCR5 receptor; and (2) how to identify the binding conformations of the receptor in complex with the ligand. To circumvent these difficulties, several modeling and simulation approaches were utilized in this study. In brief, the computational steps were as follows:

1. Several 3D models of the CCR5 receptor were built by means of homology modeling method based on the various alignments and the X-ray crystal structure of bovine rhodopsin.
2. An initial 20 ns MD simulation was performed on the constructed CCR5 receptor model inserted in a hydrated POPE (palmitoyloleoyl phosphatidyl ethanolamine) bilayer.
3. To find the possible ligand binding conformation of CCR5, twenty-four ligands were docked into the final structure extracted from the MD trajectories of CCR5 by using the molecular docking procedure, including prediction of the free energy of binding of ligands to the receptor. The conformation of the most biologically active ligand was selected as the initial conformation for further simulations. Among the various conformations of this ligand obtained from the docking procedure, the conformation with the lowest binding free energy to CCR5 receptor was chosen.
4. An additional 20 ns MD simulation was performed on the initial structure of CCR5–ligand complex resulted from the above docking calculations, embedded in a hydrated POPE bilayer. Then the resulting ligand–receptor complex was investigated by analyzing all of the modeling, docking and simulation results.

It is accepted that a homology model is not the same as an experimentally determined structure of a given protein. However, it is still very helpful in assisting us to comprehend the binding modes of CCR5 and its antagonists. This keeps researchers away from clear drawbacks in further design procedure by using the combination of homology modeling and other computational methods. It should be noted that in the chronological era of development and application of computational drug design, utilization of available experimentally determined protein structures is a recently applied tool [21]. Several studies point out improved performance of a protein structure-based approach in computational drug design and pharmacophore modeling [22,23]. We hope results of the current study shed some light to the binding mode of CCR5 antagonists for further computational studies.

2. Experimental

2.1. Preparation of data set

All the biological data used here were derived from literature [24]. General chemical structures and the structural details of these compounds are reported in Table 1. Data are reported as pIC_{50} (Table 1) for binding versus MIP-1a [25]. pIC_{50} is the dependent variable that characterizes the biological parameter for the developed docking model. The structures of molecules were drawn and optimized using HyperChem (HyperCube Inc., Gainesville, FL). Semi-empirical AM1 method with Polak–Ribiere algorithm until the root mean square gradient of 0.01 kcal mol^{-1} was used as the optimization method.

2.2. Template searching, transmembrane helix prediction and sequence alignment

The primary sequence of human CCR5 was obtained from SWISS-PROT database (Accession number P51681) [26]. BLAST (Basic Local Alignment Search Tool) at NCBI was used to find the homologous

proteins with known structures to be employed as the template in the process of CCR5 homology modeling [27].

The crystal structure of bovine rhodopsin (1F88A) at 2.80 Å resolution was obtained as the modeling template of CCR5 from the protein data bank [28]. To help define the transmembrane (TM) regions, TM helices of bovine rhodopsin structure were assigned using DSSP [29]. Several helix prediction methods were applied in a direct manner using their websites to allocate putative TM helix segments of CCR5 including: MEMSAT 1.5 [30], PORTER [31], SABLE [32], split. pmfst [33], DOS [34], HMMTOP [35], Mobyle Pasteur [36], OCTOPUS [37], TMHMM [38], TMpred [39], TOPCONS [40]. After predicting helices by these servers, a so-called consensus approach was employed. In this method for each residue, a consensus prediction was calculated by counting the number of servers that predicted the residue as being in a helix. For example, consensus 100% was allocated to a residue that was predicted as helix by all of the 11 servers. The same procedure was used to predict TM helices of bovine rhodopsin for comparing with the results of DSSP.

2.3. Homology modeling

In homology modeling step, we would like to look for an experimentally determined structure of high sequence identity with the CCR5 receptor. To align the sequence of human CCR5 receptor with that of bovine rhodopsin, the CLUSTALW program was employed directly from its website at http://www2.ebi.ac.uk/CLUSTALW [41]. The alignment was adjusted manually based on the results of TM helix prediction procedure and conserved key residues of GPCRs suggested by Baldwin et al [6].

MODELLER [42] version 9.2 was used to build homology models of CCR5 from bovine rhodopsin crystallographic structure. From the alignments, 3D models containing all non-hydrogen atoms were obtained automatically using the methods implemented in MODELLER. From the 1000 model generated with MODELLER for each alignment, the one corresponding to the lowest value of the probability density function (pdf) and fewest restraints violations was selected for further analysis. An *ab initio* method implemented in the MODELLER that has been demonstrated to predict the conformations of loop regions was used to refine some of the loops of the selected model.

The root mean square deviations (RMSDs) of the models relative to the template (1F88) were calculated using MODELLER. The RMSD differences from template geometry for bond lengths and bond angles were also calculated using MODELLER. The overall stereochemical quality of the final developed model for each CCR5 model was assessed by the program PROCHECK [43]. G-factor was calculated for the developed model using PROCHECK. Environment profile of final developed model was checked using Verify-3D (Structure Evaluation Server) [44].

2.4. Docking protocol

Docking was performed by AutoDock 4.0 program, using the implemented empirical free energy function and the Lamarckian Genetic Algorithm (LGA) [45]. The grid maps were calculated using AutoGrid. In all dockings, a grid map with $80 \times 80 \times 80$ points and a grid-point spacing of 0.375 Å was applied. Because the location of the ligand in the complex was known, the maps were centered on the ligand's binding site on Cα of Phe112. For all docking parameters, standard values were used as described before [45], except the amount of independent docking runs performed for each docking simulation, which was set to 200. Cluster analysis was performed on the docked results using a root mean square (RMS) tolerance of 0.5 Å, and the initial coordinates of the ligand were used as the reference structure.

2.5. Molecular dynamics simulations

Molecular dynamics simulation was performed in two phases. In the first phase, the minimum energy conformation of CCR5 obtained from protein homology modeling process was used as the starting structure for MD simulation. The components of the simulation system were protein, lipid and water. In the second phase, the obtained conformation of CCR5 from the first phase of MD as well as the docked ligand was employed in the MD simulation process.

All MD simulations were carried out by the GROMACS 4 package [46] using the ffgmx force field (Gromos87). Solvent (i.e. water and ions), lipid, protein and ligand were coupled separately to a temperature bath. The lipid bilayer was described using a previously developed topology file (Tieleman, see http://moose.bio.ucalgary.ca) [47]. The best model from homology modeling step was inserted at the center of the POPE bilayer with its long axis normal to the membrane–water interface. The α-helical domain of the receptor was placed at the same level as the lipid bilayer. Overlapping lipid and water molecules were discarded to remove overlaps between atoms of protein and lipids. Water molecules were represented using a simple point charge (SPC216) model [48]. Thirty Cl^- counter-ions were added by replacing water molecules to ensure the overall charge neutrality of the simulated system that comprised of CCR5, ligand, 226 POPEs, and 9305 water molecules. Final POPE–CCR5 system was placed in a box with the dimensions of $96 \times 95 \times 100$ (all in Å) with 43412 atoms in total. Periodic boundary conditions were applied in all three directions of space. The force field parameters of ligand were obtained from PRODRG web server [49]. At first, an energy minimization process was carried out. After energy minimization process, position restraint procedure was performed in association with NVT and NPT ensembles.

An NVT ensemble was adopted at constant temperature of 323 K and with a coupling constant of 0.1 ps with time duration of 100 ps. After stabilization of temperature an NPT ensemble was performed. In this phase a constant pressure of 1 bar was employed with a coupling constant of 5.0 ps with time duration of 1 ns [50]. NPT ensemble was finished after pressure stabilization. The coupling scheme of Berendsen was employed in both of NVT and NPT ensembles. The particle mesh Ewald (PME) method interaction was used [51]. A 12 Å cutoff for long-range and the Lincs algorithm for covalent bond constraints were applied [52].

2.6. Molecular images and evaluation of molecular dynamics

All the molecular images and animations were produced using VMD and rendered by Tachyon ray tracer [53]. Schematic two-dimensional representations of the docking results were produced using LIGPLOT [54]. The trajectories were analyzed using the standard tools included in the GROMACS distribution. Time interval for RMSD calculation was 5 ps.

3. Results and discussion

3.1. Sequence analysis, helix prediction, and homology modeling

In the homology modeling phase, we look for an experimentally determined structure of high "sequence identity" with the CCR5. The building of protein models by homology modeling normally proceeds along a series of well-defined and commonly accepted steps: (1) sequence alignment between the target and the template; (2) building an initial model; (3) refining the model; and (4) evaluating the quality of the model [12,15].

The crystal structure of A-chain of bovine rhodopsin was the structurally homologous protein that was found by BLAST analysis and hence chosen as template for developing the CCR5 model. BLAST results of CCR5 and bovine rhodopsin showed score of 78.2 and E value of 5×10^{-15}.

Achieving a high accuracy 3D model of CCR5 using the homology modeling approach is a very challenging procedure. The reason is the fact that none of the members of the chemokine receptor class has been crystallized and no 3D structures of chemokine receptors was determined experimentally yet. In spite of its low identity with CCR5 (the identity is 20.6%), bovine rhodopsin can be employed as a template for CCR5. The reason is it belongs to the GPCR family and its crystal structure exhibits a clear conformational feature of a seven transmembrane helix bundle, which is a common characteristic in all GPCRs proteins. In addition, the sequence identity in the transmembrane segments between CCR5 and bovine rhodopsin is 30%, which makes it possible to generate a model. According to a previous study [55], if two proteins have 30% "segment identity," then their Cα atoms are within 3.5 Å of their correct position.

The sequence homology between CCR5 and bovine rhodopsin is not strong enough for building a model flawlessly. Therefore, to facilitate describing the TM regions of CCR5, the secondary structure was predicted by several TM prediction methods. Since a small number of high-resolution structures are available for such proteins (0.5% of the PDB entries), the prediction of transmembrane helices has a significant role in the study of membrane proteins. Here, the prediction methods were utilized against bovine rhodopsin with a known 3D structure and CCR5. The distributions of helix lengths in the two proteins were examined (Table 2) and consensus 8%, 16%, 25%, 50%, 75%, and 100% of methods, which predicted α-helices for each sequence were assigned (Fig. 1). All of these methods work with different algorithms and predict different length and size of helices. As shown in Table 2, all methods predicted seven TM helices except TMHMM, and TOPCONS, which predicted 6 TM helices for CCR5. In addition, these methods were used to predict TM helices of bovine rhodopsin that were somehow different from DSSP. The 50% consensus for the rhodopsin was almost in good agreement with DSSP assignment (data was not shown) and hence was selected for assessment of helices predicted in the constructed model.

Table 2
Locations of the helical TM segments of CCR5 predicted by different commonly used methods.

Helix	Split.pmfst	DOS	HMMTOP	MEMSAT	Mobyle Pasteur	OCTOPUS	PORTER	SABLE	TMHMM	TMpred	TOPCONS	Bovine rhodopsin Crystal structure (DSSP)
I	30–57	34–57	36–56	36–56	36–56	33–53	25–57	28–57	34–57	39–56	36–56	34–64
II	65–87	69–86	67–86	67–87	72–92	68–88	64–91	65–91	68–90	68–87	67–87	71–100
III	102–131	103–136	103–124	103–124	104–124	98–128	99–130	99–136	105–124	106–124	104–124	106–139
IV	140–165	143–164	143–167	142–166	141–161	145–165	142–165	145–165	144–166	145–165	145–165	150–172
V	196–223	196–218	198–217	198–218	198–218	196–216	191–223	187–223	199–221	195–215	195–215	196–223
VI	233–263	235–257	235–260	236–257	236–256	236–256	228–260	234–263	234–256	236–256	235–256	243–277
VII	286–301	294–297	291–312	285–301	281–301	278–298	276–315	271–315		281–300		285–321

```
                  1...|...10....|...20....|...30....|...40....|...50....|...60....|...70....|...80....|...90....|...100....|..110....|..120....|
Target sequence   MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVLILILINCKRLKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFIILLTID
Cons.  |   8      -----------------------HHHHHHHHHHHHHHHHHHHHHHHHHHH-----HHHHHHHHHHHHHHHHHHHHHHHH-----HHHHHHHHHHHHHHHHHHHHHHHHH
Cons.  |  16      -----------------------HHHHHHHHHHHHHHHHHHHHHHHHHHHHH------HHHHHHHHHHHHHHHHHHHHHHHH------HHHHHHHHHHHHHHHHHHHHHHHH
Cons.  |  25      -----------------------HHHHHHHHHHHHHHHHHHHHHHHHHHHHH------HHHHHHHHHHHHHHHHHHHHHHHH------HHHHHHHHHHHHHHHHHHHHHHHH
Cons.  |  50      ----------------------------HHHHHHHHHHHHHHHHHHHHH----------HHHHHHHHHHHHHHHHHHHHHH--------------HHHHHHHHHHHHHHHH-
Cons.  |  75      --------------------------------HHHHHHHHHHHHHH----------HHHHHHHHHHHHHHHH-------HHH-----------------HHHHHHHHHHHHHHHH-
Cons.  | 100      --------------------------------HHHHHHHHHHHHH----------HHHHHHHHH--HHH------------HHHHHHHHHHHHH-
H-Score|          ----------------------11122333466AAABBBBBBBBBBBBBBBBBAAA5------13369AAABBBBBBBBBBBAABBB954431-----1333479ABBBBBBBBBBBBBBBBBBBBB5
                  ..130....|..140....|..150....|..160....|..170....|..180....|..190....|..200....|..210....|..220....|..230....|..240....|..250....|
Target sequence   RYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTCSSHFPYSQYQFWKNFQTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHRAVRLIFTIMIVYFLFWAP
Cons.  |   8      HHHHHHHHHH---HHHHHHHHHHHHHHHHHHHHHH-----------------HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH----HHHHHHHHHHHHHHHHHHH
Cons.  |  16      HHHHHHHHHH----HHHHHHHHHHHHHHHHHHHHH------------------HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH---------HHHHHHHHHHHHHHHHHH
Cons.  |  25      HHHHHH----------HHHHHHHHHHHHHHHHHHH------------------HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH-----------HHHHHHHHHHHHHHHHHHH
Cons.  |  50      ----------------HHHHHHHHHHHHHHHHHH-------------------HHHHHHHHHHHHHHHHHHHHHHHH-------------HHHHHHHHHHHHHHHHHHHH
Cons.  |  75      ----------------HHHHHHHHHHHHHHHH-----------------------HHHHHHHHHHHHHHHHHHHHH----------------HHHHHHHHHHHHHHHHH
Cons.  | 100      ------------------HHHHHHHHHHHHH--------------------------HHHH-HHHHHHHHHH------------------HHHHHHHHHHHHHHHH
H-Score|          55544322222---12467ABBBBBBBBBBBBBBBBBAAAA941------------------11112222466ABBBBBABBBBBBBBBBBBA9844433----11111246BBBBBBBBBBBBBBBB
                  ....|..260....|..270....|..280....|..290....|..300....|..310....|..320....|..330....|..340....|..350....|
Target sequence   YNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVTETLGMTHCCINPIIYAFVGEKFRNYLLVFFQKHIAKRFCKCCSIFQQEAPERASSVYTRSTGEQEISVGL
Cons.  |   8      HHHHHHHHHHHH------HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH--------------------------------
Cons.  |  16      HHHHHHHHHHH-------------HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH---------------------------------
Cons.  |  25      HHHHHHHHHH-----------HHHHHHHHHHHHHHHHHHHHHH---------------------------------------------
Cons.  |  50      HHHHHHH--------------------HHHHHHHHHHHHHHHH----------------------------------------------
Cons.  |  75      HHHHHH-----------------------------HHH-----------------------------------------------
Cons.  | 100      HHHHHH-----------------------------------------------------------------------------
H-Score|          BBBBBB6444222-------1111122333555567777767789998764333333333333222----------------------------------
```

Fig. 1. Consensus of 8%, 16%, 25%, 50%, 75% and 100% of each residue in CCR5 predicted as helix by 11 different helix prediction methods. Each residue has a score from 1 to 12 based on how many methods predicted it as helix. A and B in scoring scale are equal to 10, 11.

The sequence alignment between bovine rhodopsin and human CCR5 receptor was produced using ClustalW program. Then it was manually edited according to the known key conserved residues of the GPCRs family of proteins suggested by Baldwin et al. [6] and TM helix prediction methods (Fig. 2).

Based on the sequence alignment as an input, the MODELLER software generates a large number of spatial restraints from the template structure and constructs a molecular model of the CCR5. The restraints are usually attained by assuming that the corresponding distances and angles between aligned residues in the template and the target proteins are comparable. The output of this process is a homology model of CCR5.

More than 90% of the members of the class A of GPCRs (rhodopsin like) have a disulfide bridge between the transmembrane helix 3 (TM3) and the extracellular loop 2 (EC2) [56]. However, there are evidences of a disulfide bond between residues Cys101 at the beginning of TM3 and Cys178 in the middle of extracellular loop 2 (EL2) in CCR5. This disulfide bond was made and kept as a constraint in the homology modeling process.

The possible applications of generated protein models depend mainly on the quality of the obtained models. The quality of the generated model of CCR5 was checked by several methods. To study the orientation of the helices of CCR5, we overlaid the transmembrane regions of the generated model with the template. From the obtained RMSD value (1.87 Å) for CCR5 and rhodopsin, we can deduce that there is a good agreement between the helices.

Further checked by PROCHECK, the final CCR5 model indicates that more than 97% of residue φ–ψ angles are in the favored or additional allowed regions of Ramachandran plot (Fig. 3A). Again, it indicates that the final obtained 3D model of CCR5 is satisfactory. With respect to Ramachandran plot, it is observed only three residues are in disallowed region. Residues located in the unfavorable regions are far from the substrate-binding domain, indicating that these residues may not affect the ligand–protein binding simulations.

A parameter produced by PROCHECK indicating the quality of covalent and bond angle distance is the G-factor. The G–factors were −0.01°, 0.17° and −0.1° for dihedrals, covalent and in overall, respectively. The overall main-chain and side-chain parameters, as

```
sp|P51681|CCR5_HUMAN   MDYQVSSPIYDINYYTS-------EPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVIL 53
sp|P02699|OPSD_BOVIN   MNGTEGPNFYVPFSNKTGVVRSPFEAPQYYLAEPWQFSMLAAYMFLLIMLGFPINFLTLY 60
                       *:   ..:*    .:       * . *    :*.. *::::** *:*.:

sp|P51681|CCR5_HUMAN   ILINCKRLKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQ--WDFGNTMCQLLTGLYFIG 111
sp|P02699|OPSD_BOVIN   VTVQHKKLRTPLNYILLNLAVADLFMVFGGFTTTLYTSLHGYFVFGPTGCNLEGFFATLG 120
                       : :: *:*::  :  **:***::**:***:::     : *:: :  ** * *:*  : :*

sp|P51681|CCR5_HUMAN   FFSGIFFIILLTIDRYLAVHAVFALKARTVTFGVVTSVITWVVAVFASLPGII-FTRSQ 170
sp|P02699|OPSD_BOVIN   GEIALWSLVVLAIERYVVVCKPMSNFRFG-ENHAIMGVAFTWVMALACAAPPLVGWSRYI 179
                       .:: :::*:*.***:.*  :.: ::   ....:: .:**:*:*: .: :: ::*

sp|P51681|CCR5_HUMAN   KEGLHYTCSSHFPYSQYQFWKNFQTLKIVILGLVLPLLVMVICYSGILKTLLRCRNE--- 227
sp|P02699|OPSD_BOVIN   PEGMQCSCGIDYYTPHEETNNESFVIYMFVVHFIIPLIVIFFCYGQLVFTVKEAAAQQQE 239
                       **:: :*. .: .: :  :: ::  .|.:**:*:..:**. :: *: .. :

sp|P51681|CCR5_HUMAN   -----KKRHRAVRLIFTIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVT 282
sp|P02699|OPSD_BOVIN   SATTQKAEKEVTRMVIIMVIAFLICWLPYAGVAFY--------IFTHQGSDFGPIFMTIP 291
                       *  .:...*::: ::*.:.: * :** * :     : . ..*:    * :.

Sp|P51681|CCR5_HUMAN   ETLGMTHCCINPIIYAFVGEKFRNYLLVFFQKHIAKRFCKCCSIFQQEAPERASSVYTRS 342
sp|P02699|OPSD_BOVIN   AFFAKTSAVYNPVIYIMMNKQFRNCMVTTL----------CCGKNPLGDDEASTTVSKTE 341
                       :. *. .  **:** ::::::** **: :       **.      * :::* . .

sp|P51681|CCR5_HUMAN   TGEQEISVGL 352
sp|P02699|OPSD_BOVIN   TSQVAPA--- 348
                       *.:    :
```

Fig. 2. CLUSTALW pairwise sequence alignment between the target (CCR5) and template (1F88). The known conserved residues for the GPCRs suggested by Baldwin et al. [6] are marked using red boxes.

evaluated by PROCHECK, are all very favorable. No clash between residues of the model has also been identified in the viewer. The final step of testing was the packing quality of each residue as evaluated by the Verify-3D method, which represents the profile achieved with respect to the residues. The compatibility score above zero in the Verify-3D graph is corresponding to acceptable side-chain environments (Fig. 4). This suggests that the model has overall self-consistency in terms of sequence–structure compatibility.

3.2. MD simulation: Phase I

Clarification of ligand binding mechanisms is the essential step to achieve more selective and potent lead compounds for a given target. To accomplish this goal, we need to construct a model of protein and then let it feel its natural environment. Molecular dynamic simulation is one of the best methods for such a refinement. After constructing a 3D structure of CCR5, in order to investigate the conformational variations of the protein within the hydrated lipid environment, the MD simulation for this model was performed in an explicit lipid bilayer environment.

The developed model of CCR5 was inserted in a pre-equilibrated lipid bilayer consisting of 340 molecules of POPE. Placing of the alpha helices of CCR5 into the lipid bilayer core was carried out in such a way that alpha helices were perpendicular to the membrane plane and protein-overlapping lipids were removed.

In order to impair the incorrect interactions imposed to the 7 TMs during homology modeling, successive relaxation method was applied to the POPE–CCR5 system. As a result, constraints between TM helices were reduced in sequential steps to avoid unwanted structural drifts during production phase of the MD. The POPE–CCR5 system remained firm after the relaxation as very slight drift in energy, temperature, or lipids density was monitored during the MD (data not shown).

Fig. 5A shows the time history of RMSD for protein structure immersed in lipid bilayer relative to the starting structure (the output of the homology modeling process). As it is evident, backbone RMSD was about 4.7 Å after 20 ns of simulation, and was not increased significantly after 15 ns of simulations. The RMSD value implies that this protein structure has been affected by its environment dramatically. As it is indicated in Fig. 5A, the protein folding process shows
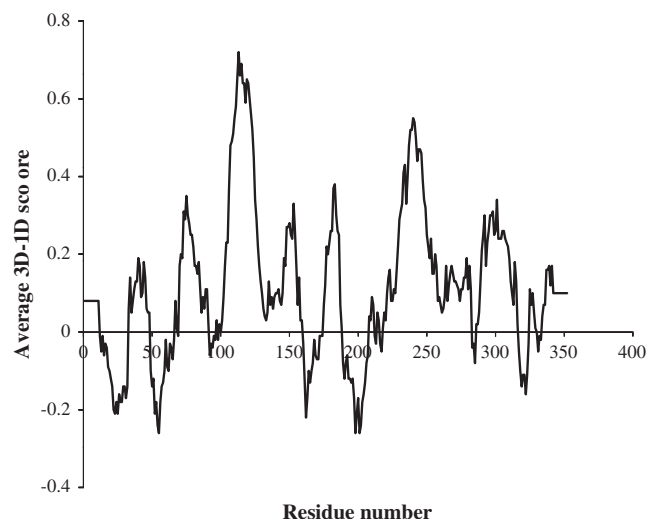

Fig. 4. The 3D profiles of human CCR5 model that was generated using Verify-3D server. Overall compatibility score above zero indicates residues are reasonably folded.

good reliability since the structures of the developed model are stable during the MD simulation. Therefore, the MD simulation was essential to specify geometry of CCR5 in the vicinity of lipid bilayer. In Fig. 5B variation in total energy versus time in 20 ns of MD is illustrated, which shows small fluctuation in the last 5 ns of MD.

Furthermore, the average temperature in 20 ns of MD simulation at 323 K was equal to 323 ± 1.4 K for the investigated system (Fig. 5C). Therefore, the extracted equilibrium structure at 323 K belonging to the CCR5 was obtained under stable temperature conditions. These facts show that energy conservation was satisfied in MD simulation.

After structural refinement of CCR5 model by MD simulation in phase I, the geometric quality of the backbone conformation, namely all the tests performed in homology modeling step, was carried out again and the quality of the model was confirmed (data not shown). In brief, the quality of the backbone geometry of the developed model, the residue interactions, the residue contacts, and the dynamic stability of
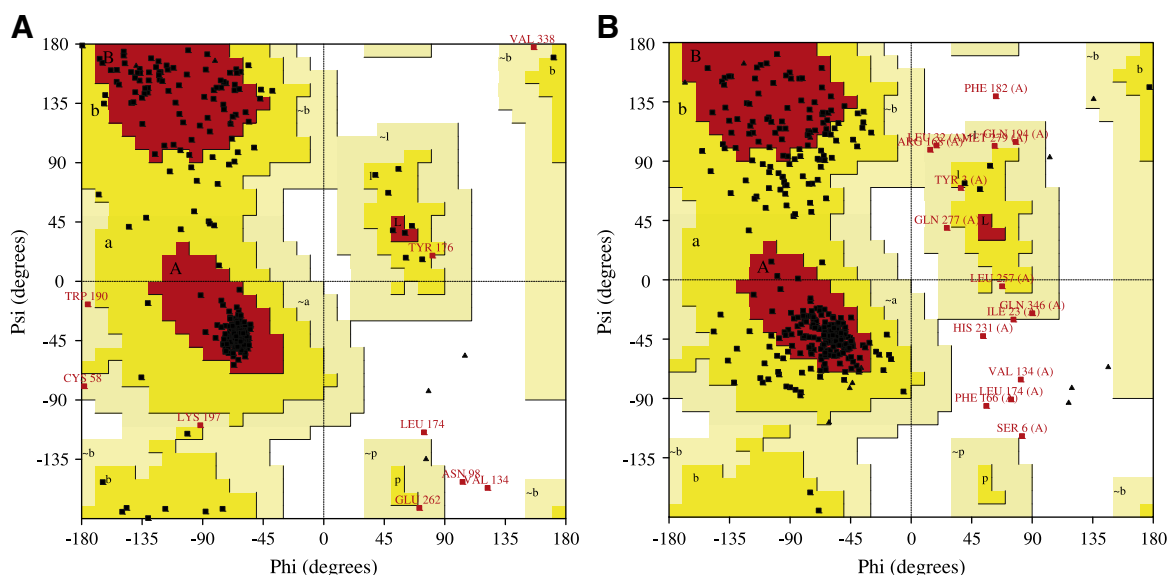
Fig. 3. Ramachandran plot of the φ–ψ distribution of CCR5 produced by PROCHECK (A) after homology modeling and, (B) after MD simulation Phase II. [A, B, L] most favored regions; [a, b, l, p] additional allowed regions; [~a, ~b, ~l, ~p] generously allowed regions; white areas are disallowed regions.
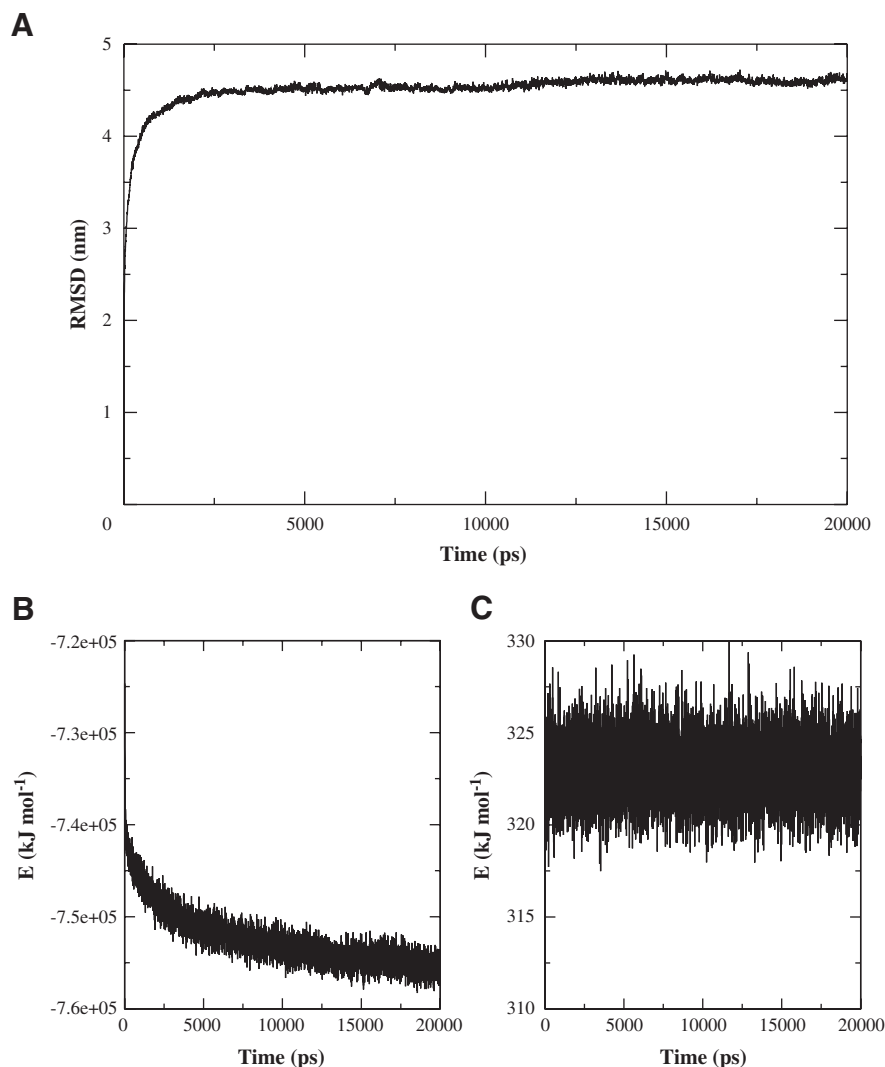
Fig. 5. (A) Time dependence of the RMSDs (Å) from homology model of CCR5 for the backbone atoms in the 20 ns MD simulation of Phase I. (B) The variation of the total energy in investigated system during first phase of MD simulation. (C) Time dependence of the temperature during the first phase of MD simulation.

the structure are within the limits established for the reliable structures. Passing all the quality assurance tests by the developed model, proposes that this model is a good estimation of CCR5 structure, and can be used to describe various protein–ligand interactions and to study the relationship between the structure and function.

3.3. Molecular docking

Molecular docking is the most widespread method for the calculation of protein–ligand interactions. Molecular docking is an efficient technique to predict the potential ligand binding site(s) on the whole protein target. Dynamical studies on binding modes are essential to elucidate key structural characteristics and interactions and they provide helpful data for designing effective CCR5 antagonists. Hence, in order to explore the predictability as well as the characteristics of the binding pocket of the developed model and to make the rational design of novel and more selective antagonists of CCR5 possible, molecular docking was carried out on CCR5 binding pocket using a set of CCR5 antagonists shown in Table 1. The 200 docking conformations for each investigated antagonist were separated into clusters according to RMSD tolerance of 0.5 Å. As well as RMSD cluster analysis, AutoDock also uses binding free energy assessment to assign the best binding conformation. Energies estimated by AutoDock are described by intermolecular energy

(including van der Waals, hydrogen bonding, desolvation, and electrostatic energies), internal energy, and torsional free energy. Among these calculated energies by AutoDock, the first two provide the docking energy while the sum of the first and the third items account for the binding energy. Among all interactions occurring in the active site, the electrostatic interaction between the ligand and the receptor is the most significant, because in most cases it can assign the strength of binding and the exact position of the ligand in the active site [57,58].

Experimental binding affinities (as pIC_{50} values) and predicted values (as binding ΔG) by LGA dockings of the 24 compounds of 1,3,4-trisubstituted pyrrolidine are summarized in Table 1. Investigated compounds selected for molecular docking have some common structural features, a pyrrolidine ring in the center of the molecule and the three side chains (arms) in positions 1, 2 and 4. The CCR5 receptor is supposed to distinguish these structures and stabilize the ligand binding. A potential binding site of CCR5 was verified according to the previous studies of site-directed mutagenesis [59–62]. The pocket is located in the extracellular side of the trans-membrane domain and partly covered by the second extracellular loop (EL2).

The docking results show that all of the studied compounds occupy an almost similar space in the binding site. The best possible binding mode of ligand 21 as the most biologically active compound in the CCR5 active site is illustrated in Fig. 6.
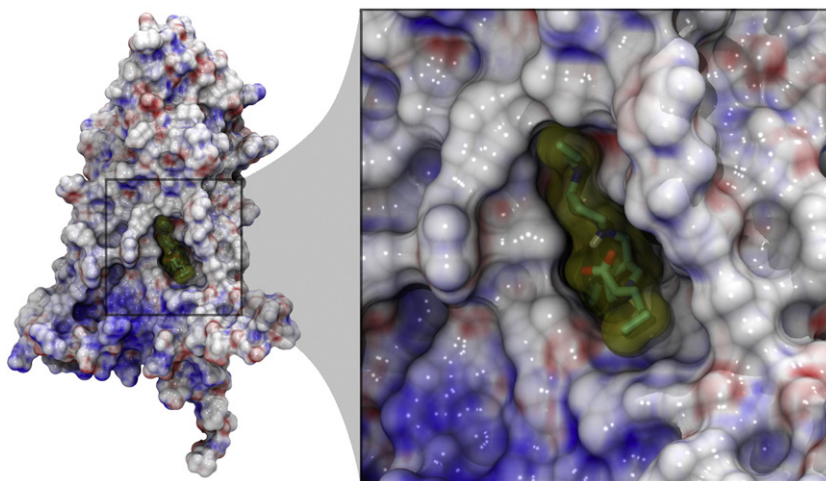
Fig. 6. The structure of compound 21 as the most potent compound in the binding cavity of CCR5. The molecule 21 represented by a stick model inside a solvent excluded surface (SES) in yellow, and colored by elements.

With respect to the docking results, Tyr108 and Phe112 in TM3 and Phe85 in TM2 and Val46 in TM1 assemble one hydrophobic cluster, which accommodates the 5-fluropyrimidine moiety of the ligand. In addition, the His315 forms a cation-aromatic interaction with the carboxylic group of the ligand. It can be seen clearly in the docking results that the flourobenzene ring of ligand 21 is surrounded by hydrophobic residues, Leu70 in TM2 and Gly301 in TM7, mainly through the hydrophobic interaction. In addition, the nitrogen atom of fluropyrimidne ring forms hydrogen bond to Pro294 as a hydrogen-bond donor (Fig. 7).

As can be seen in Table 1, among the first series of compounds (1–6), removal of chlorine atom leads to an increase in the activity. This is probably because of the decrease in steric hindrance and the increase of interaction between amide group and Tyr251, which improves H-bond formation. Simple changes in substitutions of pyrrolidine nitrogen, as with the urea derivative (compound 2), or the cycloalkyl derivatives (compounds 3 and 4), also provides similar activity.

Comparing the binding modes of docked antagonists confirms that one arm of these molecules occupies the region within the Glu302, Gly301 in TM6 and Lue70 and Asn71 in TM2 while the second arm occupies the area neighboring to the Pro294 and Ilu295 in TM6,

Phe112 in TM3 and Ile42 in TM1. The third arm is surrounded by Met49 and Leu50 in TM1 and Glu302 in TM7.

LIGPLOT software was used to investigate the hydrophobic and hydrogen bonding interactions to validate the developed model of the interaction between ligand and receptormore. As shown in 2D schematic interaction model of antagonist 21 with the CCR5 (Fig. 8), there are hydrophobic interactions between the studied ligand and CCR5. These interactions exist between Pro70 and Val46 and the carbon atoms of fluropyrimidine moiety of inhibitor 21, also same hydrophobic contact exists between carbon atoms of flurobenzene ring and Gly30 and Glu301, suggesting that more hydrophobic interactions around this area should improve the inhibitory activities. There is also a hydrogen bond between antagonist 21 and the Pro294 proposing that more hydrophilic substituents on inhibitors in this area should increase the activity.

The docking results of other investigated ligands were in accordance with the obtained results of compound 21. In all docking results three hydrophilic residues, Arg305, Glu302 and His315 act as anchors for the carbonyl moiety of ligands and almost in all of them a hydrogen bond has formed between Pro294 and hydrogen acceptor atom of ligand.
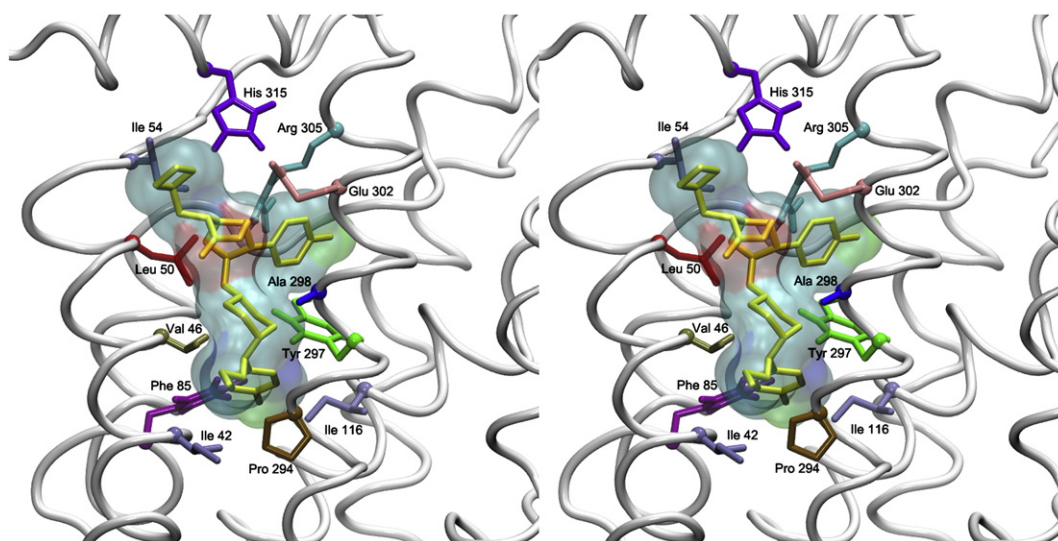
Fig 7. Stereoview of the predicted binding site of CCR5 by initial docking. Compound 21 represented as yellow stick model inside a solvent excluded surface (SES) colored by elements. Important residues in the pocket are indicated by different colors.
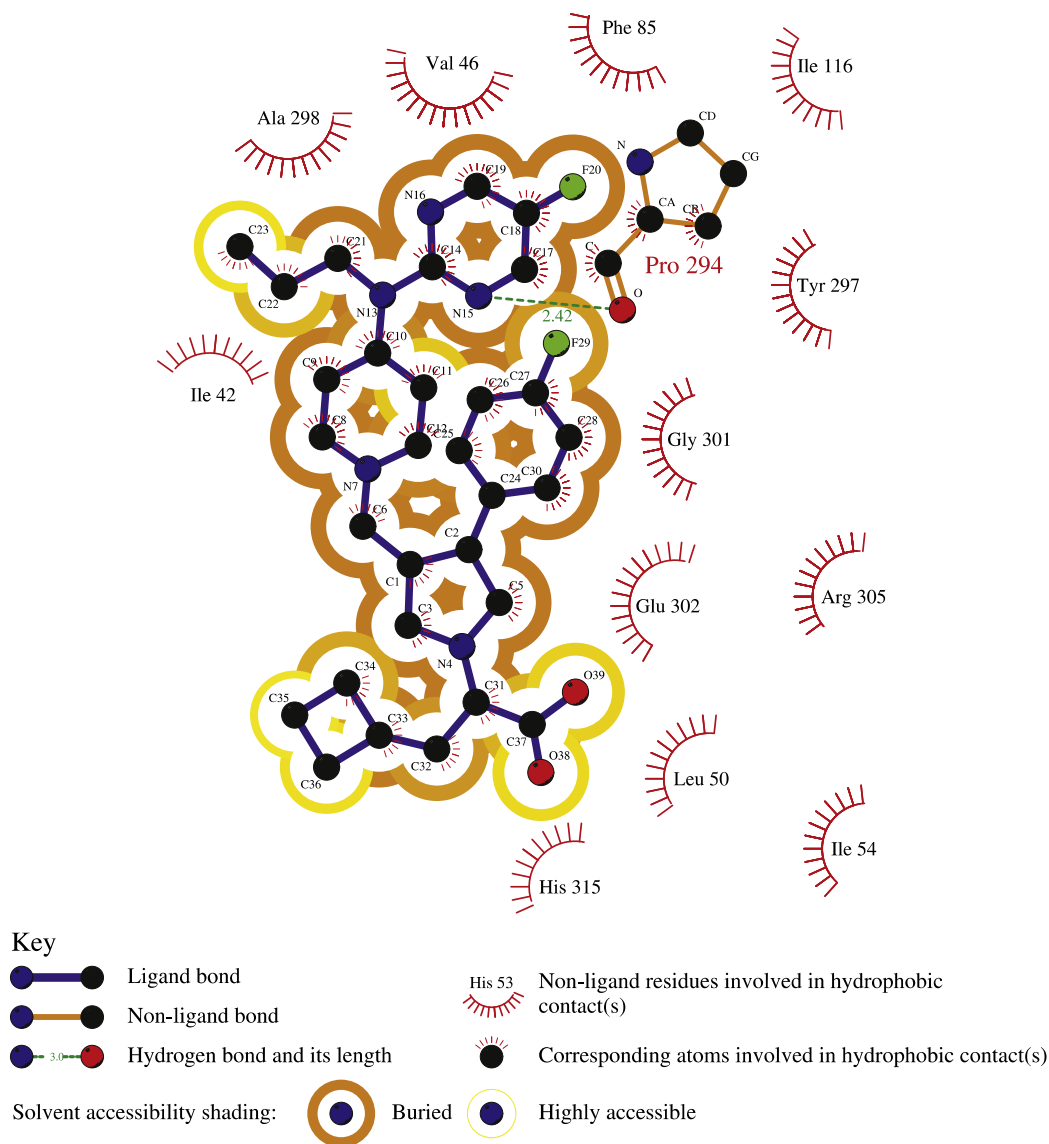
Fig. 8. Two-dimensional scheme of interactions between the compound 21 and CCR5 generated by LIGPLOT. Only the more important residues for binding are shown after initial docking. Brown stick models present the important residues in the active site, thick purple stick models present the inhibitor, green dotted lines are hydrogen bonds and dashed half-moons present hydrophobic interactions with the corresponding amino acid residues of CCR5.

3.4. MD simulation: Phase II

The obtained docking results of CCR5 antagonists allowed us to propose a general binding mode of these ligands and to determine residues involved in the ligand recognition. Nevertheless, we decided to perform a second MD simulation on the ligand–protein complex for further investigation of binding modes of ligands and to explain the effects of ligand binding on the conformation of protein.

Hence, ligand 21–CCR5 complex was selected as a representative for MD simulation. This complex was used as the starting conformation for an additional MD run. The aim of the second MD simulation was to get more precise ligand–receptor models in a state close to the natural conditions and to explore the binding modes of the ligands further. Although molecular docking offers reasonable binding structures for investigated ligands, the MD simulation can account for even the smallest variances.

One drawback of the MD simulation of proteins in lipid bilayer is that its success depends on various methodological concepts such as type of force field, types of applied constraints, and the accuracy of found solution for the equations of motion [63,64]. Especially, among

these issues the treatment of electrostatic interactions among the investigated components of the system is worthy to give special attention [65]. Since lipid biomembrane systems are extremely charged, lipid molecules are either polar or charged and they interact with each other, the polar water environment, counterions, and proteins. Correct treatment of electrostatic interactions in MD simulations of proteins in the presence of lipid biomembrane is consequently one of the most significant matters in this subject. Various algorithms are employed to treat electrostatic interactions of MD simulation of proteins in biomembrane such as the Ewald summation method, its variants [66], and the fast multiple method [67]. Especially, the particle-mesh Ewald (PME) technique has been used increasingly often in biomembrane simulations [68]. Thus, we employed this algorithm for treating electrostatic interactions among CCR5, ligand, Cl^- as counterion and biomembrane.

The conformational stability of the CCR5–antagonist complex in the simulation procedure was assessed by carrying out a 20 ns molecular dynamics simulation of model in a lipid (POPE) bilayer. The final conformation of the CCR5–ligand complex in POPE and water is illustrated in Fig. 9.
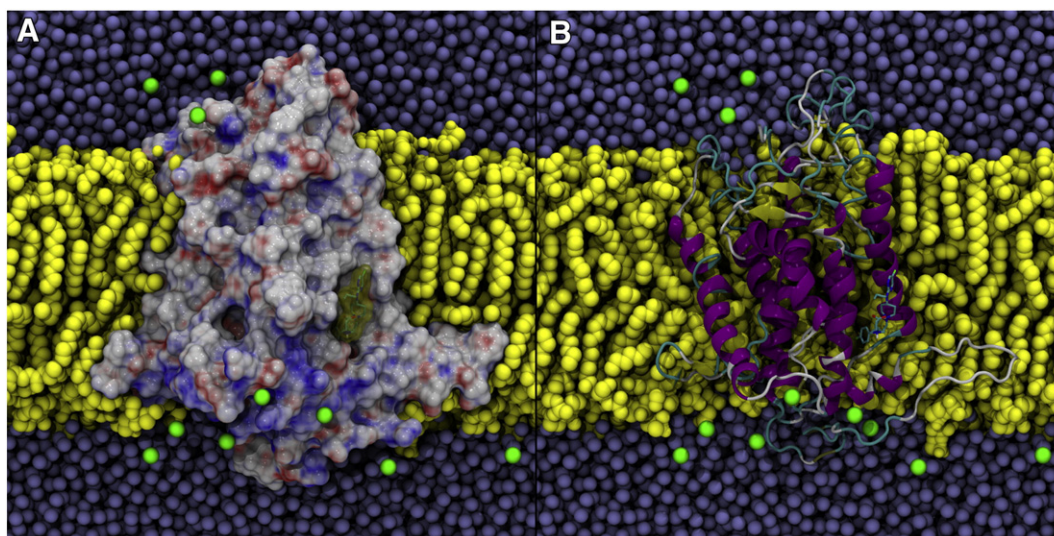
Fig. 9. CCR5–antagonist complex in hydrated POPE lipid bilayer after 20 ns MD simulation in Phase II, lipids in yellow and waters in blue. The front half of the lipid bilayer and two thirds of waters were not shown for the sake of clarity. (A) CCR5 receptor appeared as solvent excluded surface (SES), colored by electrostatic potential. (B) CCR5 receptor represented as ribbon, colored by secondary structure elements.

The trajectories were stable during the whole production part of the 20 ns MD simulation run. The trajectory stability was checked and was corroborated by the analysis of RMSD (Fig. 10A), total energy (Fig. 10B) and temperature (Fig 10C) as functions of time for the CCR5. Fig. 10A reveals that the RMSD values for the CCR5 has a rising in the first 3000 ps and then stay stable in the rest of the simulation time. The average RMSD for the CCR5 model when measured from 20 ns was found to be 0.56 ± 0.03 nm.

The total root mean square fluctuation or RMSF of the peptidic backbone plus side chains (in solid black), and the RMSF of the peptidic backbone (N, Cα and C) (in red dashed) were depicted in Fig. 11 for the developed geometry of CCR5 in lipid bilayer over the last 5 ns of the simulation. Fig. 11 shows that the residues at N-terminal and C-terminal regions have more or less equal RMSF values. It was discovered that during the dynamics simulations very few fluctuations gone beyond 1 Å and even less fluctuations exceeded 1.5 Å for the total protein. The residues 130–145 with fluctuations close to 1.8 Å observed in the dynamics plots were located close to the end of TM3, IL2 and the beginning of TM4, which indicates that these regions of protein are more unstable than other regions of protein during MD simulation. It can be seen that the protein–ligand complex immersed among lipid bilayer and water molecules stays in equilibrium throughout the entire MD. Then, we deduce that the MD simulation has constructed an improved and more relaxed structure of protein and ligand, which can be analyzed for further studies.

Importantly, simulation of CCR5 both in the presence and the absence of ligand led to the same final peptide conformations and orientations. There are only slight differences in the backbone dihedral angle values, and these differences slightly affect the position of some residues in the Ramachandran plot. As it can be seen in Fig. 3B, presence of the ligand in the active site of protein during MD simulation leads to displacement of some residues into the disallowed regions.

The RMSD of the ligand was calculated based on the MD simulation of the system to obtain information on position fluctuations and movements of ligand's atoms. Fig. 12 shows that the RMSD of the compound 21 from the initial conformation, increased to 0.23 Å after 1 ns and then leveled off to nearly 14 ns. At this point of simulation, RMSD of ligand dropped suddenly and after about 3 ns rose up again to 0.23 Å and leveled off. This shows that, after 17 ns simulation and preliminary fluctuations in the magnitude of RMSD of ligand's atoms, the ligand obtained an equilibrium state characterized by the RMSD profile.

At the end of MD simulation position and orientation of ligand in the introduced binding site were changed (Fig. 13) and this important observation indicates useful application of MD simulations after docking of ligands in the binding site. Explorative runs of molecular dynamics simulation on the complex between the receptor and the investigated antagonist revealed that, except for Ile54, His315, Glu302 and Ala298, the rest of residues of active site determined by docking were changed and some new residues such as Ile56, Phe299 and Met49 are positioned in proximity of ligands and could participate in the interaction.

Docking results of the compound 21 to CCR5 model showed that at the end of MD simulation a new hydrogen bonding was found to exist between docked molecule and Glu302 of CCR5 model and previous hydrogen bond between ligand and Pro294 was vanished. The new hydrogen bond was formed between the nitrogen atom of pyrrolydine ring of ligand as acceptor and carboxylic group of Glu302 as donor of hydrogen binding. The compound 21 is also stabilized by hydrophobic interactions on the two farthest ends of the molecule. On one end, the propyl group is making stacking interaction with Phe299 as is shown in the 2D LIGPLOT analyses (Fig. 14). Thus, Phe299 may play a key role in orienting the ligand at the active site by hydrophobic interactions. This residue does not exist in the binding site of CCR5 recognized by docking process. Similarly, on the other end, the florobenzene group is surrounded by three hydrophobic residues, His315, Ile54 and Phe56. At the end of MD simulation, the hydrophobic pocket of CCR5, which appears at the end of docking including Gly30 and Glu301, was vanished and hydrophobic interactions between flourobenzene ring of ligand and these residues could not be recognized. The orientation of the flourobenzene and flouropyrimidine rings of the compound 21 was not similar in binding modes after molecular docking and MD simulation. These results reveal that MD simulation obligate ligand to optimize its orientation and distance to binding site for maximum interaction with receptor. On docking of this ligand with CCR5, the lowest energy conformation did not show any interaction with the His315 and Phe299 side chains due to the absence of appropriate orientation and distance as was observed at the end of MD simulation with CCR5.
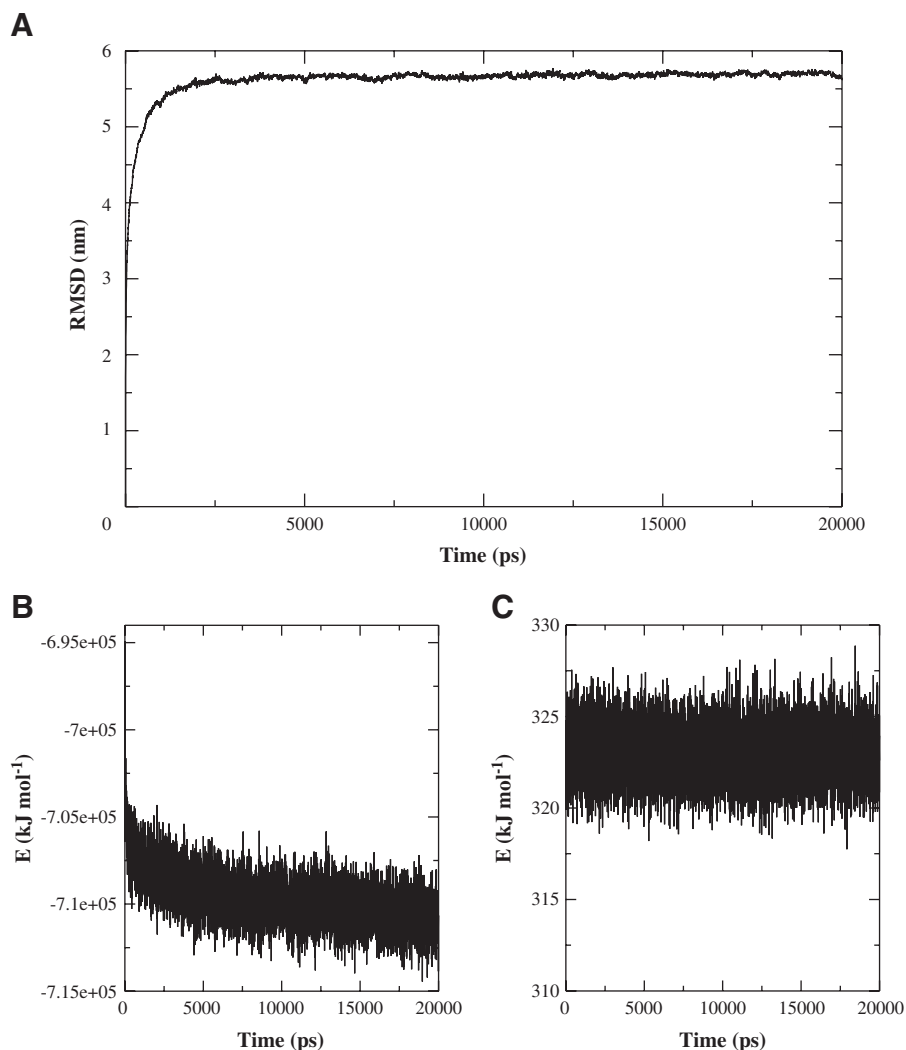
Fig. 10. (A) all atoms RMSD in the 20 ns MD simulation of Phase II. (B) The variation in total energy of system versus time during second phase of MD simulation. (C) The variation in temperature of system versus time during second phase of MD simulation.

4. Conclusion

In this work, homology modeling, molecular docking and molecular dynamics simulation were performed to explore structural features and binding mechanism of some inhibitors of CCR5, and to construct a model for designing new CCR5 inhibitors effective in preventing HIV attachment to the host cell.
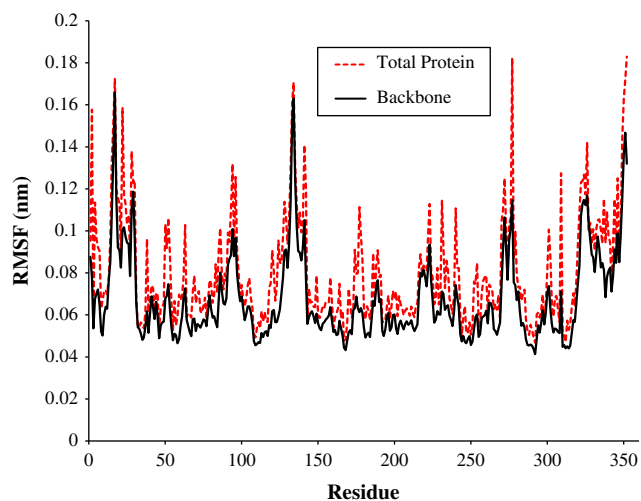
Fig. 11. Fluctuations for the total protein (black lines) and backbone (red dashed line) of CCR5 during MD simulation Phase II.
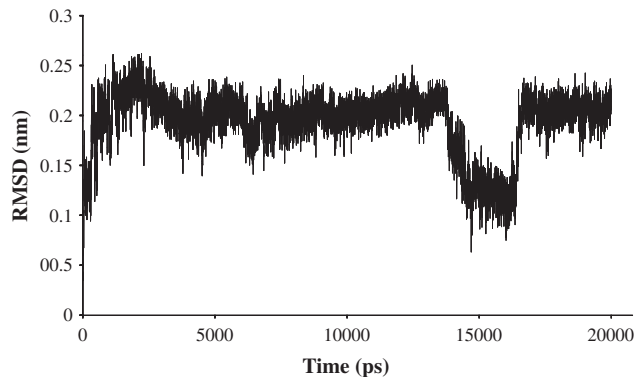
Fig. 12. Graphical representation of RMSD of ligand's atoms from the starting structure of docked ligand as a function of time during MD simulation Phase II.
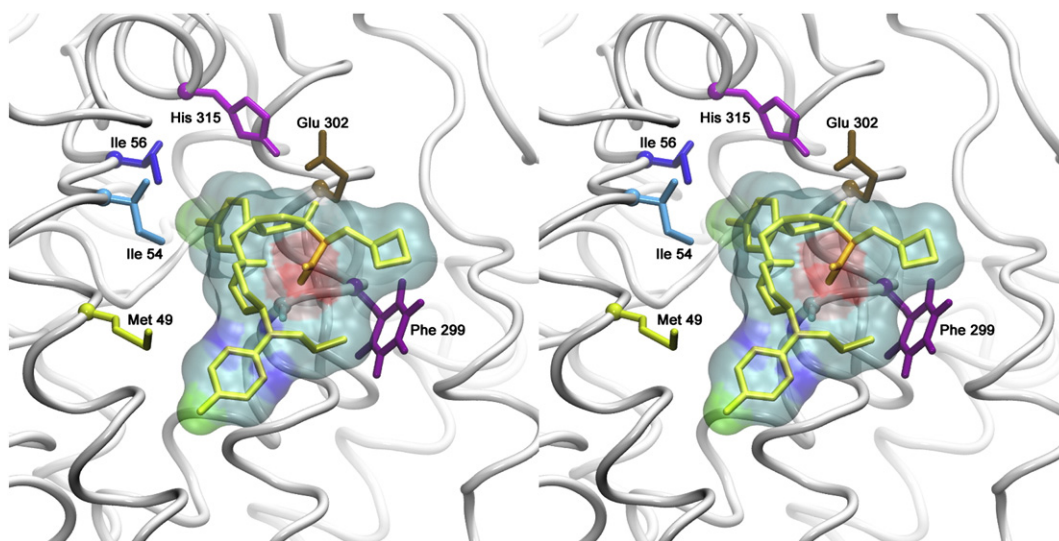
Fig 13. Stereoview of the binding site of CCR5 after MD simulation Phase II. Compound 21 represented as yellow stick model inside a solvent excluded surface (SES) colored by elements. Important residues in the pocket are indicated by different colors.
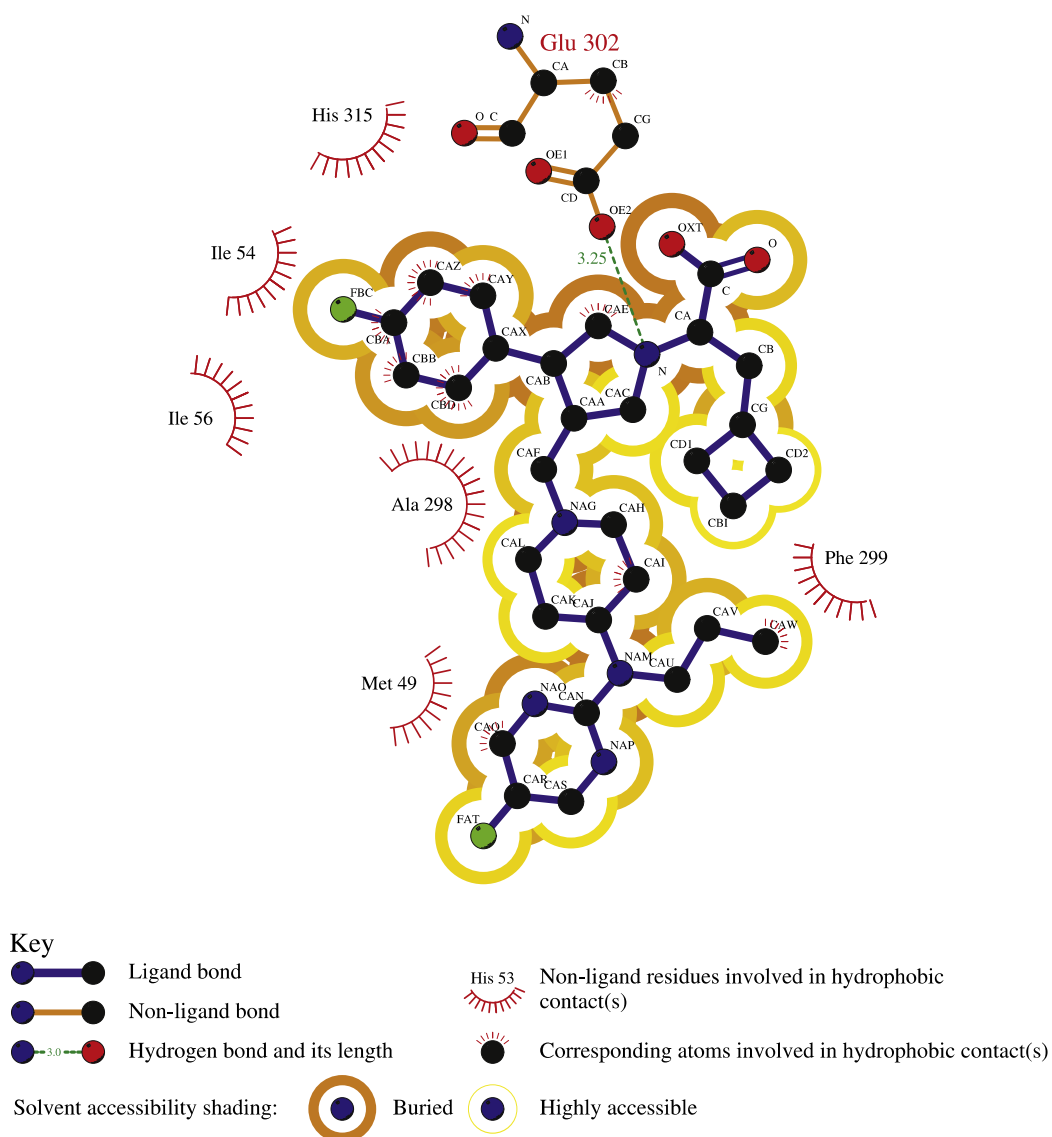
Fig. 14. Intermolecular interactions of compound 21 with CCR5 in 2D flattened space after MD simulation Phase II.

The obtained docking results of CCR5 antagonists allowed us to propose a general binding mode of these ligands to this receptor and to determine residues involved in the ligand recognition. For further investigation of binding modes of ligands and to explain the effects of ligand binding on protein conformation, we decided to perform a second molecular dynamics simulation on the ligand–protein complex.

At the end of the MD simulation, a change in the position and orientation of the ligand in binding site was observed. This important observation indicated that the application of MD simulation after docking of ligands was useful. Explorative runs of molecular dynamics simulation on the receptor–ligand complex revealed that except for Ile54, His315, Glu302 and Ala298, the rest of residues in the active site determined by docking were changed and some new residues such as Ile56, Phe299 and Met49 were positioned in proximity of the ligands and could participate in the interaction.

Acknowledgments

References

[1] R.G. Wei, D.O. Arnaiz, Y.L. Chou, D. Davey, L. Dunning, W. Lee, S.F. Lu, J. Onuffer, B. Ye, G. Phillips, CCR5 receptor antagonists: discovery and SAR study of guanylhydrazone derivatives, Bioorg. Med. Chem. Lett. 17 (2007) 231–234.
[2] M.P. Dsouza, V.A. Harden, Chemokines and HIV-1 second receptors—confluence of two fields generates optimism in AIDS research, Nat. Med. 2 (1996) 1293–1300.
[3] A. Palani, S. Shapiro, J.W. Clader, W.J. Greenlee, S. Vice, S. McCombie, K. Cox, J. Strizki, B.M. Baroudy, Oximino-piperidino-piperidine-based CCR5 antagonists. Part 2: synthesis, SAR and biological evaluation of symmetrical heteroaryl carboxamides, Bioorg. Med. Chem. Lett. 13 (2003) 709–712.
[4] T. Klabunde, G. Hessler, Drug design strategies for targeting G-protein-coupled receptors, Chembiochem 3 (2002) 929–944.
[5] U. Gether, Uncovering molecular mechanisms involved in activation of G protein-coupled receptors, Endocr. Rev. 21 (2000) 90–113.
[6] J.M. Baldwin, G.F.X. Schertler, V.M. Unger, An alpha-carbon template for the transmembrane helices in the rhodopsin family of G-protein-coupled receptors, J. Mol. Biol. 272 (1997) 144–164.
[7] K. Palczewski, G protein-coupled receptor rhodopsin, Annu. Rev. Biochem. 75 (2006) 743–767.
[8] V. Cherezov, D.M. Rosenbaum, M.A. Hanson, S.G.F. Rasmussen, F.S. Thian, T.S. Kobilka, H.J. Choi, P. Kuhn, W.I. Weis, B.K. Kobilka, R.C. Stevens, High-resolution crystal structure of an engineered human beta(2)-adrenergic G protein-coupled receptor, Science 318 (2007) 1258–1265.
[9] A. Ivetac, M.S.P. Sansom, Molecular dynamics simulations and membrane protein structure quality, Eur. Biophys. J. Biophys. Lett. 37 (2008) 403–409.
[10] L. Bordoli, F. Kiefer, K. Arnold, P. Benkert, J. Battey, T. Schwede, Protein structure homology modeling using SWISS-MODEL workspace, Nat. Protoc. 4 (2009) 1–13.
[11] T. Yoneda, S. Yoneda, N. Takayama, M. Kitazawa, H. Umeyama, A homology modeling method of an icosahedral viral capsid: inclusion of surrounding protein structures, Journal of Molecular Graphics & Modelling 17 (1999) 114.
[12] K. Ginalski, Comparative modeling for protein structure prediction, Curr. Opin. Struct. Biol. 16 (2006) 172–177.
[13] M.Y. Li, B.H. Wang, Homology modeling and examination of the effect of the D92E mutation on the H5N1 nonstructural protein NS1 effector domain, J. Mol. Model. 13 (2007) 1237–1244.
[14] A. Tramontano, Homology modeling with low sequence identity, Meth. Companion Meth. Enzymol. 14 (1998) 293–300.
[15] L.R. Forrest, C.L. Tang, B. Honig, On the accuracy of homology modeling and sequence alignment methods applied to membrane proteins, Biophys. J. 91 (2006) 508–517.
[16] N.M. Zhou, Z.W. Luo, J.W. Hall, J.S. Luo, X.B. Han, Z.W. Huang, Molecular modeling and site-directed mutagenesis of CCR5 reveal residues critical for chemokine binding and signal transduction, Eur. J. Immunol. 30 (2000) 164–173.
[17] M.G. Paterlini, Structure modeling of the chemokine receptor CCR5: implications for ligand binding and selectivity, Biophys. J. 83 (2002) 3012–3031.
[18] R.G. Efremov, F. Legret, G. Vergoten, A. Capron, G.M. Bahr, A.S. Arseniev, Molecular modeling of HIV-1 coreceptor CCR5 and exploring of conformational space of its extracellular domain in molecular dynamics simulation, J. Biomol. Struct. Dyn. 16 (1998) 77–90.
[19] S.Q. Liu, X.F. Shi, C.Q. Liu, Z.R. Sun, Characterize dynamic conformational space of human CCR5 extracellular domain by molecular modeling and molecular dynamics simulation, J. Mol. Struct. THEOCHEM 673 (2004) 133–143.
[20] J.J. Kulagowski, H.B. Broughton, N.R. Curtis, I.M. Mawer, M.P. Ridgill, R. Baker, F. Emms, S.B. Freedman, R. Marwood, S. Patel, C.I. Ragan, P.D. Leeson, 3–4-(4-chlorophenyl)piperazin-1-yl methyl-1H-pyrrolo 2, 3-b pyridine: an antagonist

[21] O.F. Guner, History and evolution of the pharmacophore concept in computer-aided drug design, Curr. Top. Med. Chem. 2 (2002) 1321–1332.
[22] T. Steindl, T. Langer, Influenza virus neuraminidase inhibitors: generation and comparison of structure-based and common feature pharmacophore hypotheses and their application in virtual screening, J. Chem. Inf. Comput. Sci. 44 (2004) 1849–1856.
[23] M. Rella, C.A. Rushworth, J.L. Guy, A.J. Turner, T. Langer, R.M. Jackson, Structure-based pharmacophore design and virtual screening for novel angiotensin converting enzyme 2 inhibitors, J. Chem. Inf. Model. 46 (2006) 708–716.
[24] C.A. Willoughby, K.G. Rosauer, J.J. Hale, R.J. Budhu, S.G. Mills, K.T. Chapman, M. MacCoss, L. Malkowitz, M.S. Springer, S.L. Gould, J.A. DeMartino, S.J. Siciliano, M.A. Cascieri, A. Carella, G. Carver, K. Holmes, W.A. Schleif, R. Danzeisen, D. Hazuda, J. Kessler, J. Lineberger, M. Miller, E.A. Emini, 1, 3, 4 trisubstituted pyrrolidine CCR5 receptor antagonists bearing 4-aminoheterocycle substituted piperidine side chains, Bioorg. Med. Chem. Lett. 13 (2003) 427–431.
[25] D.J. Hazuda, P. Felock, M. Witmer, A. Wolfe, K. Stillmock, J.A. Grobler, A. Espeseth, L. Gabryelski, W. Schleif, C. Blau, M.D. Miller, Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells, Science 287 (2000) 646–650.
[26] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R.D. Appel, A. Bairoch, ExPASy: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31 (2003) 3784–3788.
[27] S.F. Altschul, T.L. Madden, A.A. Schaffer, J.H. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[28] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, Nucleic Acids Res. 28 (2000) 235–242.
[29] W. Kabsch, C. Sander, Dictionary of protein secondary structure—pattern-recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (1983) 2577–2637.
[30] D.T. Jones, W.R. Taylor, J.M. Thorton, A model recognition approach to the prediction of all-helical membrane–protein structure and topology, Biochemistry 33 (1994) 3038–3049.
[31] G. Pollastri, A. McLysaght, Porter: a new, accurate server for protein secondary structure prediction, Bioinformatics 21 (2005) 1719–1720.
[32] R. Adamczak, A. Porollo, J. Meller, Combining prediction of secondary structure and solvent accessibility in proteins, Proteins: Struct., Funct., Bioinf. 59 (2005) 467–475.
[33] D. Juretic, A. Jeroncic, D. Zucic, Sequence analysis of membrane proteins with the Web server SPLIT, Croat. Chem. Acta 72 (1999) 975–997.
[34] M. Cserzo, E. Wallin, I. Simon, G. vonHeijne, A. Elofsson, Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method, Protein Eng. 10 (1997) 673–676.
[35] G.E. Tusnady, I. Simon, The HMMTOP transmembrane topology prediction server, Bioinformatics 17 (2001) 849–850.
[36] B. Neron, H. Menager, C. Maufrais, N. Joly, J. Maupetit, S. Letort, S. Carrere, P. Tuffery, C. Letondal, Mobyle: a new full web bioinformatics framework, Bioinformatics 25 (2009) 3005–3011.
[37] H. Viklund, A. Elofsson, OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar, Bioinformatics 24 (2008) 1662–1668.
[38] K. Melen, A. Krogh, G. von Heijne, Reliability measures for membrane protein topology prediction algorithms, J. Mol. Biol. 327 (2003) 735–744.
[39] K. Hofmann, W. Stoffel, TMbase: a database of membrane spanning proteins, Biol. Chem. Hoppe Seyler 347 (1993) 166.
[40] A. Bernsel, H. Viklund, A. Hennerdal, A. Elofsson, TOPCONS: consensus prediction of membrane protein topology, Nucleic Acids Res. 37 (2009) W465–W468.
[41] J.D. Thompson, D.G. Higgins, T.J. Gibson, 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.
[42] A. Sali, T.L. Blundell, Comparative protein modeling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[43] R.A. Laskowski, M.W. Macarthur, D.S. Moss, J.M. Thornton, Procheck—a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[44] R. Luthy, J.U. Bowie, D. Eisenberg, Assessment of protein models with 3-dimensional profiles, Nature 356 (1992) 83–85.
[45] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, J. Comput. Chem. 19 (1998) 1639–1662.
[46] D. Van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: fast, flexible, and free, J. Comput. Chem. 26 (2005) 1701–1718.
[47] D.P. Tieleman, J.L. MacCallum, W.L. Ash, C. Kandt, Z.T. Xu, L. Monticelli, 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.
[48] L. Rivail, C. Chipot, B. Maigret, I. Bestel, S. Sicsic, M. Tarek, Large-scale molecular dynamics of a G protein-coupled receptor, the human 5-HT4 serotonin receptor, in a lipid bilayer, J. Mol. Struct. Theochem. 817 (2007) 19–26.
[49] D.M.F. van Aalten, R. Bywater, J.B.C. Findlay, M. Hendlich, R.W.W. Hooft, G. Vriend, PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules, Journal of Computer-Aided Molecular Design 10 (1996) 255–262.

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

[51] T. Darden, D. York, L. Pedersen, Particle mesh ewald—an n.log(n) method for ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.

[52] B. Hess, H. Bekker, H.J.C. Berendsen, J. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.

[53] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, Journal of Molecular Graphics 14 (1996) 33.

[54] A.C. Wallace, R.A. Laskowski, J.M. Thornton, Ligplot—a program to generate schematic diagrams of protein ligand interactions, Protein Eng. 8 (1995) 127–134.

[55] R. Sanchez, A. Sali, Comparative protein structure modeling in genomics, J. Comput. Phys. 151 (1999) 388–401.

[56] A. Gonzalez, L.S. Duran, R. Araya-Secchi, J.A. Garate, C.D. Pessoa-Mahana, C.F. Lagos, T. Perez-Acle, Computational modeling study of functional microdomains in cannabinoid receptor type 1, Bioorg. Med. Chem. 16 (2008) 4378–4389.

[57] D.M. Garrido, D.F. Corbett, K.A. Dwornik, A.S. Goetz, T.R. Littleton, S.C. McKeown, W.Y. Mills, T.L. Smalley, C.P. Briscoe, A.J. Peat, Synthesis and activity of small molecule GPR40 agonists, Bioorg. Med. Chem. Lett. 16 (2006) 1840–1845.

[58] S.C. McKeown, D.F. Corbett, A.S. Goetz, T.R. Littleton, E. Bigham, C.P. Briscoe, A.J. Peat, S.P. Watson, D.M.B. Hickey, Solid phase synthesis and SAR of small molecule agonists for the GPR40 receptor, Bioorg. Med. Chem. Lett. 17 (2007) 1584–1589.

[59] T. Dragic, A. Trkola, D.A.D. Thompson, E.G. Cormier, F.A. Kajumo, E. Maxwell, S.W. Lin, W.W. Ying, S.O. Smith, T.P. Sakmar, J.P. Moore, A binding pocket for a small molecule inhibitor of HIV-1 entry within the transmembrane helices of CCR5, Proc. Natl Acad. Sci. USA 97 (2000) 5639–5644.

[60] L.A. Castonguay, Y.M. Weng, W. Adolfsen, J. Di Salvo, R. Kilburn, C.G. Caldwell, B.L. Daugherty, P.E. Finke, J.J. Hale, C.L. Lynch, S.G. Mills, M. MacCoss, M.S. Springer, J.A. DeMartino, Binding of 2-aryl-4-(piperidin-1-yl)butanamines and 1, 3, 4-trisubstituted pyrrolidines to human CCR5: a molecular modeling-guided mutagenesis study the binding pocket, Biochemistry 42 (2003) 1544–1550.

[61] F. Tsamis, S. Gavrilov, F. Kajumo, C. Seibert, S. Kuhmann, T. Ketas, A. Trkola, A. Palani, J.W. Clader, J.R. Tagat, S. McCombie, B. Baroudy, J.P. Moore, T.P. Sakmar, T. Dragic, 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.

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

[63] D.P. Tieleman, S.J. Marrink, H.J.C. Berendsen, A computer perspective of membranes: molecular dynamics studies of lipid bilayer systems, Biochim. Biophys. Acta, Rev. Biomembr. 1331 (1997) 235–270.

[64] W.F. van Gunsteren, A.E. Mark, Validation of molecular dynamics simulation, J. Chem. Phys. 108 (1998) 6109–6116.

[65] J. Norberg, L. Nilsson, On the truncation of long-range electrostatic interactions in DNA, Biophys. J. 79 (2000) 1537–1553.

[66] C. Sagui, T.A. Darden, Molecular dynamics simulations of biomolecules: long-range electrostatic effects, Annu. Rev. Biophys. Biomol. Struct. 28 (1999) 155–179.

[67] L. Greengard, V. Rokhlin, A fast algorithm for particle simulations (Reprinted from the Journal of Computational Physics, vol 73, pg 325-348, 1987), J. Comput. Phys. 135 (1997) 280–292.

[68] R.M. Venable, B.R. Brooks, R.W. Pastor, Molecular dynamics simulations of gel (L-beta I) phase lipid bilayers in constant pressure and constant surface area ensembles, J. Chem. Phys. 112 (2000) 4822–4832.