

# Identification of HIV-1 Entry Inhibitors through Ligand-Based Pharmacophore and Molecular Docking Analysis

Angadi Sathish Kumar and Estari Mamidala\*

*Infectious Diseases Research Lab, Department of Zoology, Kakatiya University, Warangal-506009 (Telangana), India*

\*(e-mail: drestari@kakatiya.ac.in; Mobile: 91107 41338)

(Received: June 25, 2024; Accepted: November 20, 2024)

---

## ABSTRACT

**Background:** In the global fight against HIV, the development of novel antiviral drugs targeting critical viral entry mechanisms remains a pressing need. This study aimed to identify potential CCR5 receptor inhibitors as promising candidates for anti-HIV drug discovery. From an initial pool of 122,276,899 compounds obtained from the ZINC database, Lipinski's rule of five was applied to filter for favorable pharmacokinetic properties, resulting in 52,272,894 ligands. A pharmacophore model was then generated using the standard drug Maraviroc. The generated pharmacophore model was used to screen the 52,272,894 ligands, yielding 38,402,967 compounds for further evaluation. Molecular docking simulations were performed using AutoDock 4.0 against the CCR5 receptor protein (PDB: 4MBS). The top 20 ligands were selected based on RMSD values and analyzed in detail. The results revealed that two compounds, ZINC000128130021 and ZINC000257322186, exhibited superior binding energies of -8.27 kcal/mol, surpassing the standard drug Maraviroc (-6.75 kcal/mol). These top compounds demonstrated extensive hydrogen bonding and hydrophobic interactions with key active site residues, as well as remarkably low inhibition constants of 871.63 nM and 862.99 nM, respectively, compared to Maraviroc (11.37  $\mu$ M). The comprehensive screening and selection process, combined with the promising in silico results, suggest that ZINC000128130021 and ZINC000257322186 warrant further in vitro and in vivo evaluation as potential CCR5 receptor inhibitors with therapeutic potential for the treatment of HIV-1 infection.

**Key words:** CCR5, autodock, RMSD, Zinc database, molecular docking

## INTRODUCTION

The HIV-1 virus has spread globally, with around 38.4 million people living with HIV worldwide in 2022 (UNAIDS, 2023). In India, the estimated number of people living with HIV/AIDS in 2022 was approximately 2.3 million (AIDSDATAHUB, 2023). Antiretroviral therapy (ART) is the recommended treatment for individuals with HIV, as it helps them maintain a similar life expectancy to those without the virus by reducing the viral load (Tu W et al., 2017). However, the successful implementation of antiviral drug therapy has been hindered by high costs, toxicity, and poor patient adherence (Smith et al., 2009). Consequently, the development of novel potential antiviral drugs is crucial for the treatment of HIV-1 infection and AIDS.

Current anti-HIV drugs primarily target the viral enzymes Reverse Transcriptase, Integrase, and Protease, as well as the CCR5 entry protein. The chemokine receptor type 5 (CCR5) serves as a physiological receptor for leukocyte chemoattractants and is also an important cell entry co-receptor for HIV-1 (Dsouza and Harden, 1996). The interaction between the HIV-1 glycoprotein gp120 and

the cellular receptor CD4 is a critical initial step in viral entry, leading to a conformational change in gp120 and further interaction with the CCR5 or CXCR4 co-receptors (Shepherd et al., 2011). CCR5 belongs to the large family of chemokine receptors, which are expressed on the surface of lymphocytes and other cell types, and are involved in signalling and the coordination of immune responses (Huttenrauch et al., 2005).

The concept of a pharmacophore was first introduced by Ehrlich in 1909, and the modern definition was established by Lemot Kier in 1967, appearing in a publication in 1971 (Güner and Bowen, 2014; Wolber et al., 2008). According to the International Union of Pure and Applied Chemistry (IUPAC), a pharmacophore model is "an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response" (Wermuth et al., 1998). 3D Pharmacophore models capture the nature and three-dimensional arrangement of chemical functionalities in ligands that are relevant for molecular interactions with the macromolecular target, such as hydrophobic areas, aromatic ring systems, hydrogen bond

acceptors, hydrogen bond donors, negatively ionizable groups, and positively ionizable groups (Van Drie 2013). This concept is fundamental in drug design and discovery, as it helps in understanding how different molecules might interact with biological targets and influence their activity. Therefore the objective of this study was to utilize molecular docking and pharmacophore-based screening to identify potential CCR5 receptor inhibitors as novel anti-HIV drug candidates.

## **MATERIALS AND METHODS**

### **Pharmacophore Generation**

This study utilized Maraviroc, a standard drug, to generate a pharmacophore model. Maraviroc was obtained from the ZINC database (<https://zinc.docking.org/>) with the ZINC number ZINC000100003902 and was available in the SDF (Structured Data Format) format. The SDF format of Maraviroc was then converted into a PDB (Protein Data Bank) format using the OpenBabel software. The converted Maraviroc structure in PDB format was uploaded into the Pharmit web server (<https://pharmit.csb.pitt.edu/>) to generate the pharmacophore features. The pharmacophore model generated from Maraviroc consisted of 1 hydrogen bond acceptor and 3 hydrophobic features. The use of Maraviroc as the standard drug to generate the pharmacophore model, followed by the screening of the ligand pool and molecular docking, highlights the systematic approach taken in this study to identify promising lead compounds for the treatment of HIV-related diseases targeting the CCR5 receptor.

### **Protein Preparation**

The molecular docking process aims to simulate the optimal conformation based on the complementarity and pre-organization between the ligand and the receptor, which can predict and obtain the binding affinity and interactive mode (Morris and Lim-Wilby, 2008; Berman et al., 2000). In this study, the CCR5 protein structure was obtained from the Protein Data Bank (PDB) (<https://www.rcsb.org/>) with the PDB ID: 4MBS and saved in the SDF (Structured Data Format) format. The SDF file of the CCR5 protein was converted into a PDB (Protein Data Bank) format using the OpenBabel software. Discovery Studio is a software package for molecular modeling and simulation, developed by Dassault Systèmes BIOVIA. It includes a range of tools for molecular docking, virtual screening, protein

modeling, and analysis of molecular dynamics simulations. The molecular docking component is used to predict the binding mode of a ligand to a target protein and to estimate the strength of the interaction between them (Agu et al., 2023).

From the protein file (4MBS) obtained from the PDB database, the water molecules and the bound ligand were deleted using Biovia Discovery Studio. Prior to performing the docking, the bond orders of the ligands were assigned, hydrogen atoms were added, and the water molecules that did not involve in the interactions were removed. In AutoDock 4.2, Gasteiger charges were calculated for each atom of the macromolecule, instead of using the Kollman charges employed in the previous versions of the program (Panigrahi et al., 2020).

### **Ligand Preparation**

From the pharmacophore generation process, 20 ligands were selected based on their RMSD (Root-Mean-Square Deviation) values. These ligands were obtained from the ZINC database (<https://zinc.docking.org/>) in the SDF (Structured Data Format) format. To prepare the ligands for molecular docking, all 20 ligands were converted from the SDF format into the PDB (Protein Data Bank) format using the OpenBabel software. This conversion ensures that the ligand structures are in a compatible format for the subsequent molecular docking simulations against the CCR5 receptor protein.

### **Grid Generation**

The grid box for the molecular docking simulations was generated using the receptor grid generation feature. This step is essential as docking of the ligands cannot be performed without the prior generation of the grid. The receptor grid generation requires a prepared protein structure with appropriate bond orders and formal charges. The grid generation process involves four tabs: receptor site constraints and rotatable groups (Sahayarayan et al., 2021). The grid points along the X, Y, and Z axes were set to encompass the whole receptor, with the grid center positioned at the geometric center of the target protein, which in this case is the CCR5 receptor (PDB: 4MBS). Both the protein and the ligands were converted to the .pdbqt format, which is the required file format for the AutoDock docking software used in this study.

### **Molecular Docking Methodology**

The molecular docking simulations were performed using the AutoDock 4.0 program,

which implements an empirical free energy function and the Lamarckian Genetic Algorithm (LGA) for ligand conformational searching (Morris et al., 1998). The LGA is a hybrid of a genetic algorithm and a local search algorithm. This algorithm first builds a population of individual conformations (genes), each representing a different random conformation of the docked molecule. The Lamarckian aspect of the algorithm allows the new generation of individuals to inherit the local search adaptations of their parents.

An extended PDB format, termed the PDBQT file, was used for the coordinate files, which includes the atomic partial charges. AutoDock Tools was utilized to create the PDBQT files from the traditional PDB files (Makhouri et al., 2018; Khodade et al., 2007). Rapid energy evaluation was achieved by pre-calculating the atomic affinity potentials for each atom in the ligand molecule. In the AutoGrid procedure, the target enzyme (CCR5 receptor, PDB: 4MBS) was embedded on a three-dimensional grid, and the energy of interaction of each atom in the ligand was calculated.

The key Lamarckian Genetic Algorithm docking parameters were set as follows: a population size of 150 members, a maximum of 2.5 million energy evaluations, a maximum of 27,000 generations, and a single top individual automatically surviving to the next generation. The mutation rate was set at 0.02, and the crossover rate was 0.8. Ten docking runs were performed, with the initial positions and conformations randomized. The probability of executing a local search on a member of the population was set to 0.06. Both the unbound target (4MBS) and the unbound ligands were treated as rigid in the docking setup.

The docking results were analyzed using AutoDock Tools, which provides various methods to analyze the docking simulations, such as conformational similarity, visualization of the binding site and its energy, and other parameters like intermolecular energy and inhibition constant (Park et al., 2006). For each ligand, the ten best poses were generated and scored using the AutoDock 4.2 scoring functions.

## RESULTS AND DISCUSSION

### Pharmacophore Evaluation

This study started with a pool of 122,276,899 compounds obtained from the ZINC database (accessed on 3rd April 2024). To reduce the

ligand count and focus on compounds with favourable pharmacokinetic properties, Lipinski's rule of five was applied. This rule considers parameters such as molecular weight (<500), hydrogen bond acceptors (<10), hydrogen bond donors (<5), octanol-water partition coefficient ( $\log P < 5$ ), rotatable bonds (<10), polar surface area (PSA < 140), and the number of aromatic rings (<10). After applying these criteria, the ligand count was reduced to 52,272,894 compounds.

The next step involved the generation of a pharmacophore model using the Maraviroc ligand and the Pharmit server (Figure 1). This pharmacophore model was then used to screen the 52,272,894 ligands, resulting in the selection of 38,402,967 ligands that fit the pharmacophore features. Based on the RMSD (Root-Mean-Square Deviation) values the top 7 compounds were selected which are showing lowest RMSD value among all. These selected ligands were then subjected to molecular docking against the CCR5 receptor protein (PDB: 4MBS) to assess their binding affinity and interactions. The top 7 ligands were then saved in SDF (Structured Data Format) for further analysis and comparison.

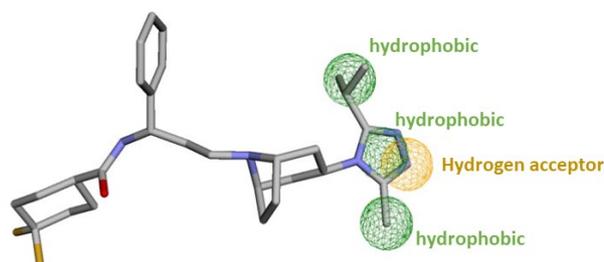


Fig. 1. Pharmacophore models based on the Crystal Structure of the CCR5 Chemokine Receptor with Maraviroc (PDB entry 4MBS) and generated with Pharmit server.

### Molecular Docking Analysis

Based on the analysis of the molecular docking results, the compounds ZINC000128130021 and ZINC000257322186 have the best binding energy of -8.27 kcal/mol, indicating the strongest binding affinity to the CCR5 receptor protein (PDB: 4MBS) among the tested compounds. The third-best binding energy is -7.5 kcal/mol for ZINC000049171945, followed by -7.33 kcal/mol for ZINC000019237289, -7.28 kcal/mol for ZINC000034055701, -7.11 kcal/mol for ZINC000257289465, -6.75 kcal/mol for ZINC000257332661, and -6.75 kcal/mol for the standard drug MARAVIROC (ZINC00010003902) (Table 1).



Fig. 3. 3D and 2D visualization of molecular docking results of ZINC000257322186 with Crystal Structure of the CCR5 Chemokine Receptor (4MBS)

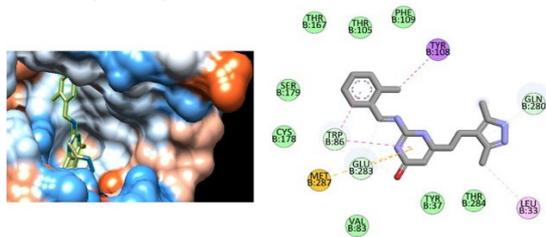


Fig. 4. 3D and 2D visualization of molecular docking results of ZINC000019237289 with Crystal Structure of the CCR5 Chemokine Receptor (4MBS)

When comparing the top two compounds, ZINC000128130021 and ZINC000257322186, with the standard drug Maraviroc (ZINC000100003902), several key differences can be observed. Both ZINC000128130021 and ZINC000257322186 have a superior binding energy of  $-8.27$  kcal/mol, which is better than the  $-6.75$  kcal/mol binding energy of Maraviroc. In terms of hydrogen bonds, ZINC000128130021 forms 2 hydrogen bonds, while Maraviroc does not form any hydrogen bonds. Both top compounds also exhibit more extensive hydrophobic interactions compared to Maraviroc (Figure 5), involving a larger number of residues such as Tyr89, Cys178, Ser180, Thr167, Thr105, Thr284, Gln280, Glu283, Trp86, Phe109, and Met287. Regarding the inhibition constant, the top two compounds have significantly lower inhibition constants of  $871.63$  nM and  $862.99$  nM, respectively, compared to the  $11.37$   $\mu$ M inhibition constant of Maraviroc. These results suggest that ZINC000128130021 and ZINC000257322186 have superior binding affinity, hydrogen bonding, and hydrophobic interactions, as well as more potent inhibition of the CCR5 receptor protein compared to the standard drug Maraviroc.

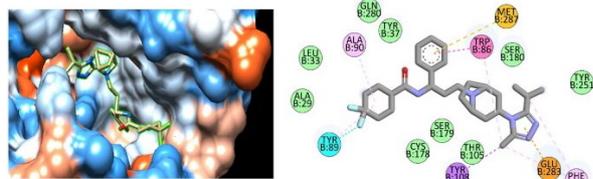


Fig. 5. 3D and 2D visualization of molecular docking results of ZINC000100003902 (Maraviroc) with Crystal Structure of the CCR5 Chemokine Receptor (4MBS).

Both ZINC000128130021 and

ZINC000257322186 show interactions with several of the key active site residues, including Tyr37, Trp86, Tyr89, Glu283, Tyr108, Phe109, and Ile198. ZINC000128130021 specifically interacts with Tyr37 and Tyr108 through hydrogen bonds, while ZINC000257322186 does not form any hydrogen bonds with the active site residues. In contrast, the standard drug Maraviroc also interacts with several active site residues, such as Tyr37, Glu283, Tyr89, and Tyr251, but it does not form any hydrogen bonds with these residues. These findings indicate that ZINC000128130021 have the potential to be more potent inhibitors of the CCR5 receptor protein, targeting key active site residues through a combination of hydrogen bonding and hydrophobic interactions.

The results of this molecular docking study highlight the potential of two compounds, ZINC000128130021 and ZINC000257322186, as potent inhibitors of the CCR5 receptor protein (PDB: 4MBS). These compounds exhibited the strongest binding affinity, as indicated by their superior binding energies of  $-8.27$  kcal/mol, which are significantly better than the standard drug Maraviroc ( $-6.75$  kcal/mol) (Tsamis et al., 2003).

The interactions of the top compounds with the known active site residues of the CCR5 receptor, such as Tyr37, Trp86, Tyr89, Glu283, Tyr108, Phe109, and Ile198, suggest that they may be able to effectively target and inhibit the receptor's function. Notably, ZINC000128130021 forms two hydrogen bonds with the key active site residues Tyr37 and Tyr108, which could contribute to its enhanced binding affinity (Kondru et al., 2008). Furthermore, the extensive hydrophobic interactions exhibited by both ZINC000128130021 and ZINC000257322186, involving a larger number of residues compared to Maraviroc (Figure 5), may also play a crucial role in their potent inhibition of the CCR5 receptor. This combination of hydrogen bonding and hydrophobic interactions has been reported to be a common feature of potent CCR5 inhibitors. The significantly lower inhibition constants of the top compounds,  $871.63$  nM and  $862.99$  nM, respectively, compared to Maraviroc ( $11.37$   $\mu$ M), further support their potential as more effective CCR5 inhibitors. This finding aligns with previous studies that have emphasized the importance of low inhibition constants in the development of potent CCR5 antagonists (Tsamis et al., 2003; Kondru et al., 2008).

The results of this molecular docking study suggest that ZINC000128130021 and ZINC000257322186 are promising lead compounds for the development of new CCR5

receptor inhibitors. Their superior binding affinity, interactions with key active site residues, and potent inhibition constants warrant further investigation, such as in vitro and in vivo studies, to validate their therapeutic potential in targeting the CCR5 receptor.

## CONCLUSION

The molecular docking analysis conducted in this study has identified two promising compounds, ZINC000128130021 and ZINC000257322186, as potent inhibitors of the CCR5 receptor protein (PDB: 4MBS). These compounds exhibited the strongest binding affinity, with binding energies of -8.27 kcal/mol, which is significantly better than the standard drug Maraviroc (-6.75 kcal/mol). The top compounds demonstrated extensive interactions with key active site residues, including hydrogen bonding and hydrophobic interactions, suggesting their ability to effectively target and inhibit the CCR5 receptor. Furthermore, the remarkably low inhibition constants of 871.63 nM and 862.99 nM for ZINC000128130021 and ZINC000257322186, respectively, compared to the 11.37  $\mu$ M inhibition constant of Maraviroc, further underscores their potential as more effective CCR5 antagonists. These findings indicate that ZINC000128130021 and ZINC000257322186 warrant further investigation, such as in vitro and in vivo studies, to validate their therapeutic potential for targeting the CCR5 receptor and their possible development as novel anti-HIV drug candidates.

## REFERENCES

- AIDSDATAHUB, 2023 Indian AIDS Update. India Country Slides 2024. Available online: <https://www.aidsdatahub.org/resource/india-country-slides> (accessed on 9 June 2024).
- Agu, P. C., Afiukwa, C. A. and Orji, O. U. (2023). Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management. *Sci. Rep.* **13**: 13398.
- Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissing, H.; Shindyalov, I. N. and Bourne, P. E. (2000). The protein data bank. *Nucleic Acids Res.* **28**: 235-242.
- Dsouza M. P. and Harden, V. A. (1996). Chemokines and HIV-1 second receptors—confluence of two fields generates optimism in AIDS research. *Nat. Med.* **2**: 1293-1300.
- Güner, O. F. and Bowen, J. P. (2014). Setting the record straight: the origin of the pharmacophore concept. *J. Chem. Inf. Model.* **54**: 1269-83.
- Huttenrauch, F. Pollok-Kopp, B. and Oppermann M. G. (2005). protein-coupled receptor kinases promote phosphorylation and beta-arrestin-mediated internalization of CCR5 homo- and hetero-oligomers. *J. Biol. Chem.* **280**: 37503-37515.
- Khodade, P., Prabhu, R., Chandra, N. (2007). Parallel implementation of Autodock. *J. App. Crystal.* **40**: 598-599.
- Kondru, R., Zhang, J., Ji, C., Mirzadegan, T., Rotstein, D., Sankuratri, S. and Dioszegi, M. (2008). Molecular interactions of CCR5 with major classes of small-molecule anti-HIV CCR5 antagonists. *Mol. Pharmacol.* **73**: 789-800.
- Makhouri, F. R. and Ghasemi, J. B. (2018). In Silico Studies in Drug Research Against Neurodegenerative Diseases. *Curr Neuropharmacol.* **16**: 664-725.
- Morris, G. M. and Lim-Wilby, M. (2008). Molecular docking. *Methods Mol. Biol.* **443**, 365–382.
- Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A.J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19**: 1639-1662.
- Panigrahi, D., Mishra, A. and Sahu, S. K. (2020). Rational in silico drug design of HIV-RT inhibitors through G-QSAR and molecular docking study of 4-arylthio and 4-aryloxy-3-iodopyridine-2(1-H)-one derivative. *J. Basic Appl. Sci.* **48**: 1-18.
- Park, H., Lee, J. and Lee, S. (2006). Critical assessment of the automated AutoDock as a new docking tool for virtual screening. *Proteins* **65**: 549-554.
- Sahayarayan, J. J., Rajan, K. S., Vidhyavathi, R., Nachiappan, M., Prabhu, D., Alfarraj, S., Arokiyaraj, S. and Daniel, A. N. (2021). In-silico protein-ligand docking studies against the estrogen protein of breast cancer using pharmacophore based virtual screening approaches. *Saudi J. Biol. Sci.* **28**: 400-407.
- Smith, J. L.; Bu, W.; Burdick, R. C.; Pathak, V. K. (2009). Multiple ways of targeting APOBEC3- virion infectivity factor interactions for anti-HIV-1 drug development. *Trends Pharmacol. Sci.* **30**: 638-646.
- Shepherd, A. J., Loo, L. and Mohapatra, D. P. (2011). Chemokine co-receptor CCR5/CXCR4-dependent modulation of Kv2.1 channel confers acute neuroprotection to HIV-1 glycoprotein gp120 exposure. *PLoS ONE* **8**: 76698.
- Tu, W., Nyandiko, W. M. and Liu, H. (2017). Pharmacokinetics-based adherence measures for antiretroviral therapy in HIV-infected Kenyan children. *J. Int. AIDS Soc.* **20**: 21157.
- Tsamis, F, Gavrillov, S., Kajumo, F., Seibert, C., Kuhmann, S., Ketas, T. and Palani, A. (2003). 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**: 5201-5208.
- UNAIDS, 2023 Global AIDS Update: The Path That Ends AIDS; July 2023. UNAIDS, AIDS info website; accessed June 2024, available at: <http://aidsinfo.unaids.org/>. UNAIDS, 2023 Core epidemiology slides; July 2023.
- Van Drie J. H. (2013). Generation of three-dimensional pharmacophore models. *Wiley Interdis Rev.* **3**: 449-464.
- Wolber, G., Seidel, T., Bendix, F. and Langer, T. (2008). Molecule-pharmacophore superpositioning and pattern matching in computational drug design. *Drug Discov. Today* **13**: 23-29.
- Wermuth, C. G., Robin Ganellin, C, Lindberg, P. and Mitscher, L. A. (1998). Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1997). *Annu. Rep. Med. Chem.* **33**: 385-395.