UNIVERSITY OF KWAZULU-NATAL ™

INYUVESI YAKWAZULU-NATALI

Mechanistic Insights and *in silico* Studies on Selected G Protein-Coupled Receptors Implicated in HIV and Neurological Disorders

Patrick Appiah-Kubi

Bsc. (*Chem.*), MMedSc. (*Pharm. Chem.*)

2021

A thesis submitted to the College of Health Sciences at the University of KwaZulu-Natal for the award of the degree of Doctor of Philosophy in Pharmaceutical Chemistry.

Mechanistic Insights and *in silico* Studies on Selected G Protein-Coupled Receptors Implicated in HIV and Neurological Disorders

Patrick Appiah-Kubi

Bsc. (*Chem.*), MMedSc. (*Pharm. Chem.*)

214584533

2021

A thesis submitted to the College of Health Sciences at the University of KwaZulu-Natal for the award of the degree of Doctor of Philosophy in Pharmaceutical Chemistry.

The chapters of this thesis are written as a set of discrete research manuscripts, beginning with a general introduction and final summary. These chapters may typically have been published, accepted in, or submitted to internationally recognised peer-reviewed journals.

This is to certify that the contents of this thesis are the original research work of Mr Patrick Appiah-Kubi.

As the candidate's supervisor, I have approved this thesis for submission.

Supervisor:

Signed: [redacted] Name: Prof Mahmoud E. S. Soliman

Date: 1st June 2021

This thesis is divided into eight chapters, including this one:

Chapter 1: This is an introductory chapter that addresses the background, rationale and relevance of the study and the aim and objectives. The general outline and structure of the thesis conclude this chapter.

Chapter 2: This chapter provides a comprehensive literature review on the two GPCR family studied in this thesis.

Chapter 3: This chapter provides the theoretical underpinnings of biomolecular simulations and computational drug design approaches applied in the thesis. The chapter introduces the general applications of computational approaches in pharmaceutical research. The chapter further discusses biomolecular simulations and the molecular mechanics force fields that govern the simulations of membrane-bound proteins.

Chapter 4: Published work- this chapter is presented in the journal's required format and is the final version of the published manuscript. (***Appendix I: Proof of Published Article***)

Chapter 5: Published work - this chapter is presented in the journal's required format and is the final version of the published manuscript. (***Appendix II: Proof of Published Article***)

Chapter 6: Published work - this chapter is presented in the journal's required format and is the final version of the published manuscript. (***Appendix III: Proof of Published Article***)

Chapter 7: Submitted work - this chapter is presented in the journal's required format and is the final version of the submitted manuscript. (***Appendix IV: Proof of Submitted Manuscript***)

Chapter 8: This final chapter presents a general discussion, concluding remarks, limitations and future research recommendations.

ABSTRACT

G protein-coupled receptors (GPCRs) are the largest membrane protein receptor superfamily involved in a wide range of physiological processes. GPCRs form the major class of drug targets for a diverse array of pathophysiological conditions. Consequently, GPCRs are recognised as drug targets for the treatment of various diseases, including neurological disorders, cardiovascular conditions, oncology, diabetes, and HIV. The recent advancement in GPCR structure resolutions has provided novel avenues to understand their molecular basis of signal transduction, ligand recognition and ligand-receptor interactions. These advances provide a framework for the structure-based discovery of new drugs in targeting GPCRs implicated in the pathogenesis of various human diseases.

In this thesis, the interactions of inhibitors at two dopamine receptor subtypes and C-C chemokine receptor 5 (CCR5) of the Class A GPCR family were investigated. Dopamine receptors and CCR5 are validated GPCR targets implicated in neurological disorders and HIV disease, respectively. The lack of structural information on these receptors limited our comprehension of their antagonists' structural dynamics and binding mechanisms. The recently solved crystal structures for these receptors have necessitated further investigations in their ligand-receptor interactions to obtain novel insights that may assist drug discovery towards these receptors.

This thesis comprehensively investigated the binding profiles of atypical antipsychotics (class I and class II) at the first crystal structure of the D_2 dopamine receptor (D2DR). The class I antipsychotics exhibited binding poses and dynamics different from the class II antipsychotics with disparate interaction mechanistic at D2DR active site. The class II antipsychotics were remarkably observed to establish a recurrent and vital interaction with Asp114 via strong hydrogen bond interactions. Furthermore, compared to class I antipsychotics, the class II antipsychotics were found to engage favourably with the deep hydrophobic pocket of D2DR.

In addition, the structural basis and atomistic binding mechanistic of the preferential selective inhibition at D3DR over D2DR were explored. This study investigated two small molecules (R-VK4-40 and Y-QA31) with substantial selectivity (> 180-fold) for D3DR over D2DR. The selective antagonists adopted shallow binding modes at D3DR while demonstrating a deep hydrophobic pocket binding at D2DR. Also, the vital roles and contribution of critical residues to the selective binding of R-VK4-40 and Y-QA31were identified in D3DR. Structural and

binding free energy analyses further discovered distinct stabilising effects of the selective antagonists on the secondary architecture and binding profiles of D3DR relative to D2DR.

Furthermore, the atomistic molecular interaction mechanism of how slight structural modification between novel derivatives of 1-heteroaryl-1,3-propanediamine (Compd-21 and -34) and Maraviroc significantly affects their binding profiles toward CCR5 were elucidated. This study utilised explicit lipid bilayer molecular dynamics (MD) simulations and advanced analyses to explore these inhibitory disparities. The thiophene moiety substitution common to Compd-21 and -34 was found to enhance their CCR5-inhibitory activities due to complementary high-affinity interactions with residues critical for the gp120 V3 loop binding. The study further highlights the structural modifications that may improve inhibitor competitiveness with the gp120 V3 loop.

Finally, structure-based virtual screening of antiviral chemical database was performed to identify potential compounds as HIV-1 entry inhibitors targeting CCR5. The identified compounds made pertinent interactions with CCR5 residues critical for the HIV-1 gp120-V3 loop binding. Their predicted *in silico* physicochemical and pharmacokinetic descriptors were within the acceptable range for drug-likeness. Further structural optimisations and biochemical testing of the proposed compounds may assist in the discovery of novel HIV-1 therapy.

The studies presented in this thesis provide novel mechanistic and *in silico* perspective on the ligand-receptor interactions of GPCRs. The findings highlighted in this thesis may assist in further research towards the identification of novel drug molecules towards CCR5 and D_2-like dopamine receptor subtypes.

DECLARATION I – PLAGIARISM

I, Patrick Appiah-Kubi, declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other person's data, pictures, graphs or other information unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain another person's writing unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
    a. their words have been re-written, but the general information attributed to them has been referenced.
    b. where their exact words have been used, then their writing has been placed in italics and inside quotation marks and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the internet unless specifically acknowledged, and the source being detailed in the thesis and in the reference sections.

A detailed contribution to publications that form part or/and include research presented in this thesis is stated (include manuscripts submitted, accepted, in the press and published).

Signed: **Patrick Appiah-Kubi**

DECLARATION II – LIST OF THESIS PUBLICATIONS

1. **Patrick Appiah-Kubi**, Fisayo Andrew Olotu, and Mahmoud ES Soliman. "Probing Binding Landscapes and Molecular Recognition Mechanisms of Atypical Antipsychotic Drugs towards the Selective Targeting of D2 Dopamine Receptor." *Molecular Informatics* 38, no. 11-12 (2019): 1900044 https://doi.org/10.1002/minf.201900044. (*Published Article*) (**IF= 2.741**)

Contribution:

P. Appiah-Kubi designed, performed the experiment, data analysis, interpretation of data and manuscript preparation.

FA. Olotu: reviewed and edited the drafted manuscript

M.E.S. Soliman: Supervised the study.

2. **Patrick Appiah-Kubi**, Fisayo Andrew Olotu, and Mahmoud E. S. Soliman. "Elucidating the Disparate Inhibitory Mechanisms of Novel 1-Heteroaryl-1,3-Propanediamine Derivatives and Maraviroc towards C-C Chemokine Receptor 5: Insights for Structural Modifications in HIV-1 Drug Discovery." https://doi.org/10.2174/1573406417666201208122110 *Medicinal Chemistry* (*Published Article*) (**IF= 2.577**)

Contribution:

P. Appiah-Kubi: designed, performed the experiment, data analysis, interpretation of data and manuscript preparation.

FA. Olotu: reviewed and edited the drafted manuscript

M.E.S. Soliman: Supervised the study.

3. **Patrick Appiah-Kubi**, Fisayo Andrew Olotu, and Mahmoud E. S. Soliman. "Exploring the structural basis and atomistic binding mechanistic of the selective antagonist blockade at D_3 dopamine receptor over D_2 dopamine receptor." *Journal of Molecular Recognition*. https://doi.org/10.1002/jmr.2885 (*Published Article*) **(IF= 2.214)**

Contribution:

P.Appiah-Kubi: designed, performed the experiment, data analysis, interpretation of data and manuscript preparation.

FA. Olotu: reviewed and edited the drafted manuscript

M.E.S. Soliman: Supervised the study.

4. **Patrick Appiah-Kubi**, Emmanuel Amarachi Iwuchukwu, and Mahmoud E. S. Soliman. "Structure-based identification of novel scaffolds as potential HIV-1 entry inhibitors towards CCR5 in HIV disease therapy" *Medical Hypothesis*. (*Submitted Manuscript*) **(IF= 1.390)**

Contribution:

P.Appiah-Kubi: contributed to the project by designing, performing all the experimental work, manuscript preparation and writing.

E. A. Iwuchukwu: Manuscript writing and data collation.

M.E.S. Soliman: Supervised the study.

RESEARCH OUTPUT

A- Publications included in this thesis

1. **Patrick Appiah-Kubi**, Fisayo Andrew Olotu, and Mahmoud ES Soliman. "Probing Binding Landscapes and Molecular Recognition Mechanisms of Atypical Antipsychotic Drugs towards the Selective Targeting of D2 Dopamine Receptor." *Molecular Informatics* 38, no. 11-12 (2019): 1900044. https://doi.org/10.1002/minf.201900044. (*Published Article*) (**IF= 2.741**)
2. **Patrick Appiah-Kubi**, Fisayo Andrew Olotu, and Mahmoud ES Soliman. "Elucidating the Disparate Inhibitory Mechanisms of Novel 1-Heteroaryl-1,3-Propanediamine Derivatives and Maraviroc towards C-C Chemokine Receptor 5: Insights for Structural Modifications in HIV-1 Drug Discovery." https://doi.org/10.2174/1573406417666201208122110 *Medicinal Chemistry* (*Published Article*) (**IF= 2.577**)
3. **Patrick Appiah-Kubi**, Fisayo Andrew Olotu, and Mahmoud E. S. Soliman. "Exploring the structural basis and atomistic binding mechanistic of the selective antagonist blockade at D3 dopamine receptor over D2 dopamine receptor." Journal of Molecular Recognition. https://doi.org/10.1002/jmr.2885 (*Published Article*) **(IF= 2.214)**
4. **Patrick Appiah-Kubi**, Emmanuel Amarachi Iwuchukwu, and Mahmoud E. S. Soliman. "Structure-based identification of novel scaffolds as potential HIV-1 entry inhibitors towards CCR5 in HIV disease therapy" *Medical Hypothesis*. (*Manuscript submitted*) **(IF= 1.390)**

B- Publications as co-author not included in this thesis during candidature

1. El Rashedy, Ahmed A., **Patrick Appiah-Kubi**, and Mahmoud ES Soliman. "A Synergistic Combination Against Chronic Myeloid Leukemia: An Intra-molecular Mechanism of Communication in BCR–ABL1 Resistance." *The Protein Journal* 38, no. 2 (2019): 142-150. (**IF = 1.317)**
2. Issahaku, Abdul R., Clement Agoni, Opeyemi S. Soremekun, **Patrick A. Kubi**, Ransford O. Kumi, Fisayo A. Olotu, and Mahmoud ES Soliman. "Same Target, Different Therapeutic Outcomes: The Case of CAY10471 and Fevipiprant on CRTh2 Receptor in Treatment of Allergic Rhinitis and Asthma." *Combinatorial Chemistry & High Throughput Screening* 22, no. 8 (2019): 521-533. (**IF = 1.195)**
3. Aljoundi, Aimen, Ahmed El Rashedy, **Patrick Appiah-Kubi**, and Mahmoud ES Soliman. "Coupling of HSP72 α-Helix Subdomains by the Unexpected Irreversible Targeting of

Lysine-56 over Cysteine-17; Coevolution of Covalent Bonding." *Molecules* 25, no. 18 (2020): 4239. (**IF = 3.267)**

4. Ugbaja, Samuel C., **Patrick Appiah-Kubi**, Monsurat M. Lawal, Nelisiwe S. Gumede, and Hezekiel M. Kumalo. "Unravelling the molecular basis of AM-6494 high potency at BACE1 in Alzheimer's disease: an integrated dynamic interaction investigation." Journal of Biomolecular Structure and Dynamics (2020): 1-13. (**IF= 3.549**)
5. Ugbaja, Samuel C., Zainab K. Sanusi, **Patrick Appiah-Kubi**, Monsurat M. Lawal, and Hezekiel M. Kumalo. "Computational modelling of potent β-secretase (BACE1) inhibitors towards Alzheimer's disease treatment." Biophysical Chemistry 270 (2021): 106536. (**IF= 1.995**)

C- Conference attendance and abstracts

1. **Patrick Appiah-Kubi**, Fisayo Andrew Olotu, and Mahmoud ES Soliman. "Probing Binding Landscapes and Molecular Recognition Mechanisms of Atypical Antipsychotic Drugs towards the Selective Targeting of D_2 Dopamine Receptor". Abstract for poster presentation in the Computational Chemistry and Materials Science SIG session at the Centre for High-Performance Computing (CHPC) 2019 National Conference, 1-5 December 2019 at Birchwood Conference Centre, Boksburg, Johannesburg, South Africa. https://events.chpc.ac.za/event/47/contributions/1114/contribution.pdf

2. **Patrick Appiah-Kubi**, Fisayo Andrew Olotu, and Mahmoud ES Soliman. "Elucidating the Disparate Inhibitory Mechanisms of Novel 1-Heteroaryl-1,3-Propanediamine Derivatives and Maraviroc towards C-C Chemokine Receptor 5: Insights for Structural Modifications in HIV-1 Drug Discovery." Abstract for Mini Oral Talk in the Computational Chemistry and Materials Science SIG session at the Centre for High-Performance Computing (CHPC) 2020 National Conference, 30 November 2020 to 2 December 2020, South Africa.

3. Antibiotic Stewardship and Conservation in Africa Conference held at the University of KwaZulu-Natal, Durban, South Africa, 20-23 October 2019.

DEDICATION

This thesis is dedicated to:

My beloved wife - Mrs Patricia A. Appiah-Kubi

My lovely daughter- Ms Akosua Pokuaa Appiah-Kubi who was born during this candidature.

My late mother- Ms Elizabeth Afia Konadu who died in my absence during this candidature.

All who could not have the opportunity of higher education.

ACKNOWLEDGEMENTS

Ebenezer, thus far, God has brought me. Praises and honour be to His name.

A special appreciation to my parents and families in Ghana, my South African families – (The Nhari Family, the Hlatshwayo family, the Kunene Family, and the Mwaramba Family), and friends who have been my support system. I am forever grateful for the expressed love and care during this period.

I would like to thank my supervisor, Prof Mahmoud E. S. Soliman, for his support and the opportunity to work with him and conduct my research in his research group. God bless you.

To Dr Ahmed A. Elrashedy, thank you for your technical support during the membrane protein simulation in lipid bilayer protocol development.

To my colleagues of the Molecular Bio-computation and Drug Design Research team, I thank you all.

To everyone who has helped me along the way, I say thank you, and God bless you from my heart.

I would acknowledge the University of KwaZulu-Natal College of Health Sciences for funding my doctoral degree.

I acknowledge the Centre for High-Performance Computing (CHPC) (http://www.chpc.ac.za), Cape Town, South Africa, for computational support and Schrodinger Maestro licence.

ABBREVIATIONS, SYMBOLS AND ACRONYMS

3D QSAR	3D Quantitative Structure-Activity Relationships
4E-BP	eukaryotic initiation factor 4E-binding protein 1
AC	Adenylyl cyclase
ADMET	Absorption, Distribution, Metabolism, Excretion and Toxicity
AIDS	Acquired Immune Deficiency Syndrome
AMBER	Assisted Model Building and Energy Refinement
AMPA	α-Amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors
cAMP	3'-5'-cyclic adenosine monophosphate
CCR5	C-C chemokine receptor 5
CDDD	Computational Drug Design and Discovery
CHARMM	Chemistry at HARvard Macromolecular Mechanics
CREB	cAMP response element-binding protein
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
cryo-EM	cryo-Electron Microscopy
D2DR	D_2 dopamine receptor
D3DR	D_3 dopamine receptor
D4DR	D_4 dopamine receptor
DARPP-32	dopamine and cyclic AMP-regulated phosphoprotein, 32 kDa
DPPC	1,2-dipalmitoyl-sn-glycero-3-phosphocholine
ECL	Extracellular loop
FBDD	Fragment-Based Drug Discovery
FEP	Free Energy Perturbation
HIV-1	Human Immunodeficiency Virus 1
$GABA_A$	γ-Aminobutyric acid A
GAFF	General AMBER Force Field
GIRK	G protein-coupled inward rectifier potassium
GPCRdb	G Protein-Coupled Receptor database
GPCRs	G protein-coupled receptors
GROMOS	GROningen Molecular Simulation Package
GSK-3	Glycogen Synthase Kinase-3
ICL	Intracellular loop

IP_3	Inositol triphosphate
LBDD	Ligand-Based Drug Design
LIE	Linear Interaction Energy
MD	Molecular dynamics
MM	Molecular Mechanics
MM/GBSA	Molecular Mechanics/Generalized Born Surface Area
MM/PBSA	Molecular Mechanics Poisson/Boltzmann Surface Area
NMDA	N-methyl-D-aspartate
NMR	Nuclear Magnetic Resonance
mTOR	mammalian Target of Rapamycin
OPLS-AA	Optimised Parameters for Liquid Simulations All-Atom
PI3K	Phosphatidylinositol 3-kinase
PCA	Principal Component Analysis
PDK	Phosphoinositide-dependant kinase
PIP2, PIP3	Phosphatidylinositol 2 and 3
PKA	Protein Kinase A
PKC	Protein kinase C
PLCβ	Phospholipase C isoform β
PME	Particle-Mesh-Ewald
POPC	1,2-dipalmitoyl-sn-glycero-3-phosphocholine
QM/MM	Quantum Mechanics/Molecular Mechanics
RESP	Restricted Electrostatic Potential
RoG or rGyr	Radius of Gyration
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
SASA	Solvent Accessible Surface Area
SBDD	Structure-Based Drug Design
SBVS	Structure-Based Virtual Screening
TI	Thermodynamic Integration
TM	Transmembrane

Amino Acids

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENT

CHAPTER 1

Introduction

1.1 Background and Rationale of the study

G protein-coupled receptors (GPCRs) are physiologically essential membrane proteins involved in the signal transduction of chemicals from the extracellular matrix into the cell. They are implicated in many diseases (Latorraca et al., 2017). The GPCRs family represent one of the significant successful drug targets due to their direct role in major physiological processes regulations. They account for nearly 34% of marketed drugs (Santos et al., 2017). The recent explosion in GPCR structures serves as a reliable starting point for *in silico* drug discovery and biomolecular simulations in accelerating drug design opportunities and mechanistic understanding of these receptors. As a significant prominent family of drug targets known to date, GPCRs continue to be the subject of considerable research efforts directed towards discovering improved therapeutics.

Dopamine receptors and C-C chemokine receptor 5 (CCR5) are validated GPCR drug targets implicated in neurological disorders and HIV disease, respectively (Rangel-Barajas et al., 2015; Tan et al., 2013). Dopamine receptors are an integral member of the family of GPCRs that play vital functions in central nervous system neurotransmitters. Numerous neurological disorders, such as attention deficit hyperactivity disorder, Parkinson's disease, depression, and schizophrenia, have been associated with dysfunction of the dopaminergic system (Klein et al., 2019; Rangel-Barajas et al., 2015; Beaulieu & Gainetdinov, 2011). The existing antipsychotic drugs that target the different dopamine receptors show poor selectivity between the respective dopamine receptor subtypes. Significant adverse effects stem from the lack of selectivity of these antipsychotics. These include metabolic syndrome, cardiovascular hypertension, and neurological side effects, including extrapyramidal reactions and tardive dyskinesia due to the distinctive function of these receptor subtypes, thus reducing the utility of these essential groups of drugs. (Álvarez et al., 2013; Ballon et al., 2014; Kaar et al., 2020). The recent advancement in the structure determination of the D_2-like dopamine receptor provides new research opportunities in understanding binding mechanistic and interactions of antipsychotics at dopamine receptors (Fan et al., 2020; Wang et al., 2018; Chien et al., 2010; Wang et al., 2017).

Another member of the GPCR family involved in the modulation of immune response and as an HIV-1 co-receptor is the C-C Chemokine receptor 5 (CCR5). During entry into CD4+ T-

cells, CCR5 serves as the essential co-receptor of HIV-1. An attractive therapeutic approach to blocking HIV-1 infection and replication is preventing the entry of HIV into host cells (Arimont et al., 2017). The quest to develop inhibitors with the capability of inhibiting HIV entry by targeting CCR5 resulted in the approval of the first and only CCR5 antagonist (Maraviroc) in 2007 (FDA, 2007). Many CCR5 antagonists have been documented, but very few have advanced into clinical use before being discontinued, mainly due to challenging drug-like properties (Qi et al., 2020). The prescription of Maraviroc is limited due to identified factors such as its drug-drug interactions (especially when co-administrated with CYP3A4 inhibitors), CYP450 inhibition, and viral resistance (Garcia-Perez et al., 2015; Peng et al., 2018). Therefore, the rapid global rise in patients diagnosed with HIV necessitates discovering novel therapeutics for HIV treatment with fewer side effects and better efficacy. Recently, 1 Heteroaryl-1,3-propanediamine derivatives have been identified as a series of novel CCR5 antagonists compared to Maraviroc with low cytotoxicity, exceptional *in vitro* anti-HIV-1 profile, and good pharmacokinetic properties (Peng et al., 2018).

In selecting compounds for synthesis or further optimisation, the application of molecular modelling methods in drug design and development plays an essential role. These techniques further provide an atomistic basis in directing further design efforts (Aminpour et al., 2019; Cavasotto et al., 2019). Biomolecular simulations are proven methods for exploring the conformational landscape of GPCRs at the atomic level and inhibitor/drug binding mechanisms for biological targets (Latorraca et al., 2017; Alfonso-Prieto et al., 2019). A molecular level insight into how small molecules interact with their GPCRs and how drugs/small molecules can selectively target subfamily receptors is critical for discovering novel treatments of conditions implicated in the GPCR family.

1.2 Aims and objectives

The purpose of this thesis is to provide novel mechanistic insights into how dopamine receptors and CCR5 receptors interact with their small molecule inhibitors and to apply a structure-based discovery approach to identifying new potential CCR5 inhibitors. The following objectives were followed to achieve the above goal:

1. To provide a comprehensive investigation of the molecular recognition and binding mechanistic of atypical antipsychotics drugs at D_2 dopamine receptor.
   1.1 Predict the binding modes of the selected atypical drugs through molecular docking.
   1.2 Embed the drug-receptor complexes in a lipid bilayer and perform MD simulations.

1.3 Estimate the free energies of the drug-receptor complexes upon binding using MMGB/SA end-point free energy approach.

1.4 Analyse the drug-receptor interactions and characterise residues critical for drug binding.

2. To explore the structural basis and atomistic binding mechanistic of the selective antagonist blockade at D_3 dopamine receptor over D_2 dopamine receptor.

2.1 Predict the binding modes of the D_3 selective antagonists at both D2DR and D3DR binding site.

2.2 Perform MD simulations in a lipid bilayer environment to elucidate the structural and conformational changes associated with the selective binding.

2.3 Identify residues that drive the selectivity and higher binding affinity of the studied compounds.

3. To elucidate the molecular mechanism and structural dynamics of 1-Heteroaryl-1,3-propanediamine derivatives interactions with CCR5 in HIV inhibition.

3.1 Perform all-atom MD simulations of protein-ligand complexes in a lipid bilayer.

3.2 Post-process MD data to estimate the binding free energy.

3.3 Identify residues critical for the higher affinity of the novel derivative relative to Maraviroc.

4. To identify novel potential HIV-1 entry compounds targeting CCR5 using structure-based drug discovery techniques.

4.1 Perform a structure-based virtual high-throughput screening using compounds from the Asinex antiviral database.

4.2 Predict ligand-binding poses and affinities, assess ligand binding mode modes and interactions at receptor target.

4.3 Perform MD simulations on the best hits in a lipid bilayer to evaluate their stability at the CCR5 binding pocket.

4.4 To apply *in silico* techniques in predicting the molecular properties and pharmacokinetics (ADME/Tox) of the top compounds.

1.3 Novelty and Significance of the Study

The crystal structures of the D_3 dopamine receptor (D3DR) bound to eticlopride and D_4 dopamine receptor (D4DR) bound to antipsychotic nemonapride were resolved in 2010 (Chien et al., 2010) and 2017 (Wang et al., 2017), respectively. However, the lack of crystallised ligand-bound D2DR structures before 2018 limited the molecular understanding of ligand recognition and receptor function. The first crystal structure of the D_2 dopamine receptor bound with risperidone solved in 2018 (Wang et al., 2018), followed by haloperidol in 2020 (Fan et al., 2020), provided novel insights. An unanticipated binding mode of risperidone was discovered by the crystal structure of D2DR bound to atypical antipsychotic risperidone (Wang et al., 2018). The observed unexpected binding mode was different from earlier docking studies based on a D2DR homology model that used D3DR, or D4DR solved structures as templates (Duan et al., 2015; Salmas et al., 2017).

Interestingly, the recent crystal structure of the risperidone-bound serotonin 2A receptor (5-HT2AR) also shows the same risperidone binding mode, as observed in D2DR (Kimura et al., 2019). However, the observed distinct risperidone binding mode and interaction in the crystal structure of D2DR and 5-HT2AR could not be replicated by molecular docking approach using a homology modelled structure of D2DR (Wang et al., 2018). The related antipsychotic drugs binding modes at D2DR, therefore, remained uncertain. There is a need to define risperidone and related antipsychotic drugs at the binding pocket D2DR in order to make significant progress in the structure-based design of these drug classes. To further clarify the binding mechanistic of atypical antipsychotics at D2DR, paper I of this thesis addresses this research gap. In this paper, the binding mechanistic and conformational changes associated with six atypical antipsychotics targeting D2DR have been elucidated (Appiah-Kubi et al., 2019).

The current marketed antipsychotic drugs lack selectivity toward a given D_2-like receptor subtype (Li et al., 2016). They are associated with considerable adverse effects, including metabolic syndrome, cardiovascular hypertension, and neurological side effects, including tardive dyskinesia and extrapyramidal reactions (Kaar et al., 2020; Ballon et al., 2014; Álvarez et al., 2013). The selective antagonist inhibition of D3DR over D2DR have been demonstrated to lessen drug-seeking behaviour and associated side effects compared to non-subtype selective antagonists (Andreoli et al., 2003; Higley et al., 2011; Galaj et al., 2015; Manvich et al., 2019). However, the high degree of sequence conservations between D2DR and D3DR, particularly at the ligand-binding pockets and within the transmembrane (TM) domains, remains a challenge to the therapeutic design of subtype-selective antipsychotics.

The discovery of two small molecules, R-VK4-40 and Y-QA31, that substantially inhibited D3DR with >180-fold selectivity over D2DR was reported in recent studies (Shaik et al., 2019; Kumar et al., 2016; Sun et al., 2016). Contrary to the side effect of GSK598,809 and SB277011A, which in the presence of cocaine increase blood pressure (Appel et al., 2015; Appel & Acri, 2018), R-VK4-40, when administered with cocaine, displays no cardiovascular side effects (Jordan et al., 2019). Also, Y-QA31 exhibit antipsychotic effects in cognitive dysfunction, negative and positive symptoms without inducing extrapyramidal side effects in preclinical models of schizophrenia (Sun et al., 2016). However, the structural determinant and atomistic molecular mechanistic by which R-VK4-40 and Y-QA31 achieved their selectivity at D3DR over D2DR have not been elucidated. To address this gap, the paper II of this thesis provides molecular and structural insights into these differential binding mechanistic using meta-analytic computational simulation methods.

Recently, new 1-Heteroaryl-1,3-propanediamine derivatives, Compd-21 and Compd-34 (Peng et al., 2018), were synthesised as CCR5 antagonists. These inhibitors have displayed ~3 times more potency than Maraviroc, with improved pharmacokinetic profiles (Peng et al., 2018). The main structural variation between Maraviroc, Compd-34, and Compd-21 is the substitution of the phenyl group in Maraviroc with thiophen-2-yl and thiophen-3-yl moieties in Compd-21, and Compd-34, respectively. However, atomistic molecular details of the interaction mechanisms of how slight structural variance between these inhibitors (Compd-21, Compd-34, and Maraviroc) significantly affects their binding profiles at the CCR5 receptor was lacking. Paper III of this thesis provided atomistic understanding by identifying molecular properties and receptor interactions, which may be useful in designing more potent HIV-1 entry inhibitors targeting CCR5.

The application of a structure-based virtual screening approach successfully identified novel potential scaffolds from the unexplored Asinex antiviral compound database toward CCR5. The identified compounds provide a basis for further structural optimisation and/or biochemical testing of the identified compounds against CCR5.

Findings from this thesis will expand our understanding of inhibitor binding interactions at the D2-like dopamine and CCR5 receptor as well as assist in identifying novel compounds for further optimisation and characterisation.

References

Alfonso-Prieto, M., Navarini, L. & Carloni, P. 2019. Understanding ligand binding to g-protein coupled receptors using multiscale simulations. *Frontiers in molecular biosciences*, 6: 29.

Álvarez, Y., Pérez-Mañá, C., Torrens, M. & Farré, M. 2013. Antipsychotic drugs in cocaine dependence: a systematic review and meta-analysis. *Journal of substance abuse treatment*, 45(1): 1–10.

Aminpour, M., Montemagno, C. & Tuszynski, J.A. 2019. An overview of molecular modeling for drug discovery with specific illustrative examples of applications. *Molecules*, 24(9): 1693.

Andreoli, M., Tessari, M., Pilla, M., Valerio, E., Hagan, J.J. & Heidbreder, C.A. 2003. Selective antagonism at dopamine D 3 receptors prevents nicotine-triggered relapse to nicotine-seeking behavior. *Neuropsychopharmacology*, 28(7): 1272–1280.

Appel, N.M. & Acri, J.B. 2018. Cardiovascular safety studies on dopamine D3 receptor antagonists and cocaine. *Journal of Pharmacological and Toxicological Methods*, 93: 156.

Appel, N.M., Li, S.-H., Holmes, T.H. & Acri, J.B. 2015. Dopamine D3 receptor antagonist (GSK598809) potentiates the hypertensive effects of cocaine in conscious, freely-moving dogs. *Journal of Pharmacology and Experimental Therapeutics*, 354(3): 484–492.

Appiah-Kubi, P., Olotu, F.A. & Soliman, M.E.S. 2019. Probing Binding Landscapes and Molecular Recognition Mechanisms of Atypical Antipsychotic Drugs towards the Selective Targeting of D2 Dopamine Receptor. *Molecular Informatics*, 38(11–12): 1900044.

Arimont, M., Sun, S.-L., Leurs, R., Smit, M., De Esch, I.J.P. & de Graaf, C. 2017. Structural analysis of chemokine receptor–ligand interactions. *Journal of medicinal chemistry*, 60(12): 4735–4779.

Ballon, J.S., Pajvani, U., Freyberg, Z., Leibel, R.L. & Lieberman, J.A. 2014. Molecular pathophysiology of metabolic effects of antipsychotic medications. *Trends in Endocrinology & Metabolism*, 25(11): 593–600.

Beaulieu, J.-M. & Gainetdinov, R.R. 2011. The physiology, signaling, and pharmacology of dopamine receptors. *Pharmacological reviews*, 63(1): 182–217.

Cavasotto, C.N., Aucar, M.G. & Adler, N.S. 2019. Computational chemistry in drug lead discovery and design. *International Journal of Quantum Chemistry*, 119(2): e25678.

Chien, E.Y.T., Liu, W., Zhao, Q., Katritch, V., Han, G.W., Hanson, M.A., Shi, L., Newman, A.H., Javitch, J.A. & Cherezov, V. 2010. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science*, 330(6007): 1091–1095.

Duan, X., Zhang, M., Zhang, X., Wang, F. & Lei, M. 2015. Molecular modeling and docking study on dopamine D2-like and serotonin 5-HT2A receptors. *Journal of Molecular Graphics and Modelling*, 57: 143–155.

Fan, L., Tan, L., Chen, Z., Qi, J., Nie, F., Luo, Z., Cheng, J. & Wang, S. 2020. Haloperidol bound D 2 dopamine receptor structure inspired the discovery of subtype selective ligands. *Nature communications*, 11(1): 1–11.

FDA, U.S. 2007. FDA notifications. Maraviroc approved as a CCR5 co-receptor antagonist. *AIDS Alert*, 22: 103.

Galaj, E., Manuszak, M., Babic, S., Ananthan, S. & Ranaldi, R. 2015. The selective dopamine D3 receptor antagonist, SR 21502, reduces cue-induced reinstatement of heroin seeking and heroin conditioned place preference in rats. *Drug and alcohol dependence*, 156: 228–233.

Garcia-Perez, J., Staropoli, I., Azoulay, S., Heinrich, J.-T., Cascajero, A., Colin, P., Lortat-Jacob, H., Arenzana-Seisdedos, F., Alcami, J. & Kellenberger, E. 2015. A single-residue change in the HIV-1 V3 loop associated with maraviroc resistance impairs CCR5 binding affinity while increasing replicative capacity. *Retrovirology*, 12(1): 1–20.

Higley, A.E., Kiefer, S.W., Li, X., Gaál, J., Xi, Z.-X. & Gardner, E.L. 2011. Dopamine D3 receptor antagonist SB-277011A inhibits methamphetamine self-administration and methamphetamine-induced reinstatement of drug-seeking in rats. *European journal of pharmacology*, 659(2–3): 187–192.

Kaar, S.J., Natesan, S., McCutcheon, R. & Howes, O.D. 2020. Antipsychotics: mechanisms underlying clinical response and side-effects and novel treatment approaches based on pathophysiology. *Neuropharmacology*, 172: 107704.

Kimura, K.T., Asada, H., Inoue, A., Kadji, F.M.N., Im, D., Mori, C., Arakawa, T., Hirata, K., Nomura, Y. & Nomura, N. 2019. Structures of the 5-HT 2A receptor in complex with the antipsychotics risperidone and zotepine. *Nature structural & molecular biology*, 26(2): 121.

Klein, M.O., Battagello, D.S., Cardoso, A.R., Hauser, D.N., Bittencourt, J.C. & Correa, R.G. 2019. Dopamine: functions, signaling, and association with neurological diseases. *Cellular and molecular neurobiology*, 39(1): 31–59.

Kumar, V., Bonifazi, A., Ellenberger, M.P., Keck, T.M., Pommier, E., Rais, R., Slusher, B.S., Gardner, E., You, Z.-B. & Xi, Z.-X. 2016. Highly selective dopamine D3 receptor (D3R) antagonists and partial agonists based on eticlopride and the D3R crystal structure: new leads for opioid dependence treatment. *Journal of medicinal chemistry*, 59(16): 7634–7650.

Latorraca, N.R., Venkatakrishnan, A.J. & Dror, R.O. 2017. GPCR dynamics: structures in motion. *Chemical reviews*, 117(1): 139–155.

Li, P., L Snyder, G. & E Vanover, K. 2016. Dopamine targeting drugs for the treatment of schizophrenia: past, present and future. *Current topics in medicinal chemistry*, 16(29): 3385–3403.

Manvich, D.F., Petko, A.K., Branco, R.C., Foster, S.L., Porter-Stransky, K.A., Stout, K.A., Newman, A.H., Miller, G.W., Paladini, C.A. & Weinshenker, D. 2019. Selective D 2 and D 3 receptor antagonists oppositely modulate cocaine responses in mice via distinct postsynaptic mechanisms in nucleus accumbens. *Neuropsychopharmacology*, 44(8): 1445–1455.

Peng, P., Chen, H., Zhu, Y., Wang, Z., Li, J., Luo, R.-H., Wang, J., Chen, L., Yang, L.-M. & Jiang, H. 2018. Structure-Based Design of 1-Heteroaryl-1, 3-propanediamine Derivatives as a Novel Series of CC-Chemokine Receptor 5 Antagonists. *Journal of medicinal chemistry*, 61(21): 9621–9636.

Qi, B., Fang, Q., Liu, S., Hou, W., Li, J., Huang, Y. & Shi, J. 2020. Advances of CCR5 antagonists: From small molecules to macromolecules. *European Journal of Medicinal Chemistry*: 112819.

Rangel-Barajas, C., Coronel, I. & Florán, B. 2015. Dopamine receptors and neurodegeneration. *Aging and disease*, 6(5): 349.

Salmas, R.E., Yurtsever, M. & Durdagi, S. 2017. Atomistic molecular dynamics simulations of typical and atypical antipsychotic drugs at the dopamine D2 receptor (D2R) elucidates their inhibition mechanism. *Journal of Biomolecular Structure and Dynamics*, 35(4): 738–754.

Santos, R., Ursu, O., Gaulton, A., Bento, A.P., Donadi, R.S., Bologa, C.G., Karlsson, A., Al-Lazikani, B., Hersey, A. & Oprea, T.I. 2017. A comprehensive map of molecular drug targets. *Nature reviews Drug discovery*, 16(1): 19–34.

Shaik, A.B., Kumar, V., Bonifazi, A., Guerrero, A.M., Cemaj, S.L., Gadiano, A., Lam, J., Xi, Z.-X., Rais, R. & Slusher, B.S. 2019. Investigation of novel primary and secondary pharmacophores and 3-substitution in the linking chain of a series of highly selective and bitopic dopamine D3 receptor antagonists and partial agonists. *Journal of medicinal chemistry*, 62(20): 9061–9077.

Sun, X., Gou, H., Li, F., Lu, G., Song, R., Yang, R., Wu, N., Su, R., Cong, B. & Li, J. 2016. Y-QA31, a novel dopamine D 3 receptor antagonist, exhibits antipsychotic-like properties in preclinical animal models of schizophrenia. *Acta Pharmacologica Sinica*, 37(3): 322–333.

Tan, Q., Zhu, Y., Li, Jian, Chen, Z., Han, G.W., Kufareva, I., Li, T., Ma, L., Fenalti, G., Li, Jing, Zhang, W., Xie, X., Yang, H., Jiang, H., Cherezov, V., Liu, H., Stevens, R.C., Zhao, Q. & Wu, B. 2013. Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science*, 341(6152): 1387–1390.

Wang, S., Che, T., Levit, A., Shoichet, B.K., Wacker, D. & Roth, B.L. 2018. Structure of the D2 dopamine receptor bound to the atypical antipsychotic drug risperidone. *Nature*, 555: 269. http://dx.doi.org/10.1038/nature25758.

Wang, S., Wacker, D., Levit, A., Che, T., Betz, R.M., McCorvy, J.D., Venkatakrishnan, A.J., Huang, X.-P., Dror, R.O. & Shoichet, B.K. 2017. D4 dopamine receptor high-resolution structures enable the discovery of selective agonists. *Science*, 358(6361): 381–386.

CHAPTER 2

2.1 Overview to G protein-coupled receptors

The largest transmembrane receptor superfamily encoded in the human genome are the G protein-coupled receptors (GPCRs), which control several major physiological processes. The human genome alone account for about 800 unique GPCRs and are important therapeutic targets of many potent drugs (Ghosh et al., 2015; Odoemelam et al., 2020). GPCRs can sense the presence of a diverse range of molecules outside the cell and, in response, activate various intracellular signals (Latorraca et al., 2017). The direct role of GPCRs in controlling major physiological processes has made them one of the most effective drug targets and account for almost 35% of all FDA-approved drugs (~700 drugs) (Sriram & Insel, 2018). However, these drugs target only 134 unique GPCRs, primarily the Class A (Rhodopsin) receptors such as histamine receptors, muscarinic receptors, dopamine receptors, adrenoceptors, and serotonin receptors (Garland, 2013; Hauser et al., 2018; Sriram & Insel, 2018).

The ability of GPCRs to convey signals via the cell membrane relies on their potential to undergo conformational changes. Diverse extracellular ligands varying from proteins (such as chemokines) to neurotransmitters to small hormones, bind to the extracellular region and induce structural changes to allow signalling proteins (e.g. arrestins and G proteins) to bind to the intracellular surface of GPCR (Figure 2.1) (Latorraca et al., 2017). The structural basis for GPCR's ligand recognition and the understanding of the process of their dynamic signalling had been a challenge in the past two decades, especially for rational drug design. However, recent progress in GPCR structure biology has provided a novel understanding of GPCRs dynamics, signalling pathways, and ligand binding interactions. Thus, providing novel opportunities for structure-based drug discovery and selective drug development among receptor subtypes.

2.1.1 Classifications of GPCRs

Based on GPCRs amino acid sequence homology and functional similarity, the commonly used classification scheme is the A-F system (Kolakowski Jr, 1994; Attwood & Findlay, 1994). This classification scheme identifies six classes (A-F) of GPCR sequences from humans, animals and plants, with only four categories (A, B, C, and F) identified in humans. Class A, also called the "rhodopsin-like family", comprises hormones, neurotransmitters, and light receptors,

representing about 80 % of GPCRs. Class A is generally distinguished structurally by seven TM helices, an eighth helix, a C-terminal tail, and a short N-terminus.

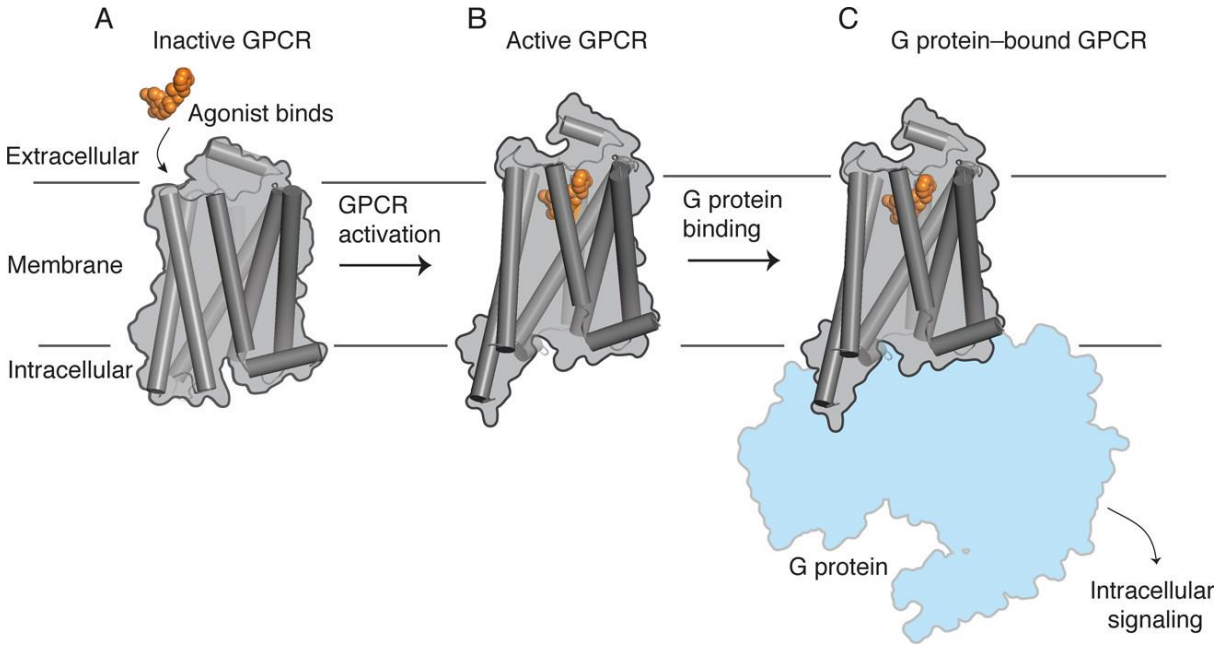


Figure 2.1 GPCR signalling: The binding to an inactive GPCR by an orthosteric agonist ligand (in orange); (B) the active state of the ligand-bound GPCR due to conformational changes in complex formation; and (C) the active state of the GPCR complex binds to a G protein, which then activates the G protein's alpha-subunit (Image adapted from (Latorraca et al., 2017)).

Class B, commonly called the "secretin receptor family", consists of approximately 70 receptors. The structural characteristics of Class B are a seven-transmembrane and an N-terminal of about 120 residues. The metabotropic glutamate family, taste receptors, the GABA receptor and calcium-sensing receptors constitute the Class C GPCRs. Seven TM helices and an N-terminal domain of around 500 residues structurally define these receptors. The fungal pheromone mating receptors, the slime mold cAMP receptors and frizzled/smoothed receptors constitute Class D, Class E and Class F, respectively.

Another GPCR classification system is grouping the human GPCRs into five families known as "GRAFS" (Schiöth & Fredriksson, 2005; Fredriksson et al., 2003). The GRAFS system comprises the families of Class C (Glutamate), Class A (Rhodopsin), Class B2 (Adhesion), Class F (Frizzled/Taste2) and Class B1 (Secretin), based on a phylogenetic analysis of nearly 800 sequences from the human GPCR (Fredriksson et al., 2003) (Figure 2.2). The GRAFS system differs from the A-F system in that it further divides Class B receptors into two families, Secretin (class B1) and Adhesion (class B2). This distinction is based on an initial observation

that the evolutionary background of the families of Secretin and Adhesion differs from one another. Furthermore, it also adds the Taste2 receptors in the Frizzled receptor family.

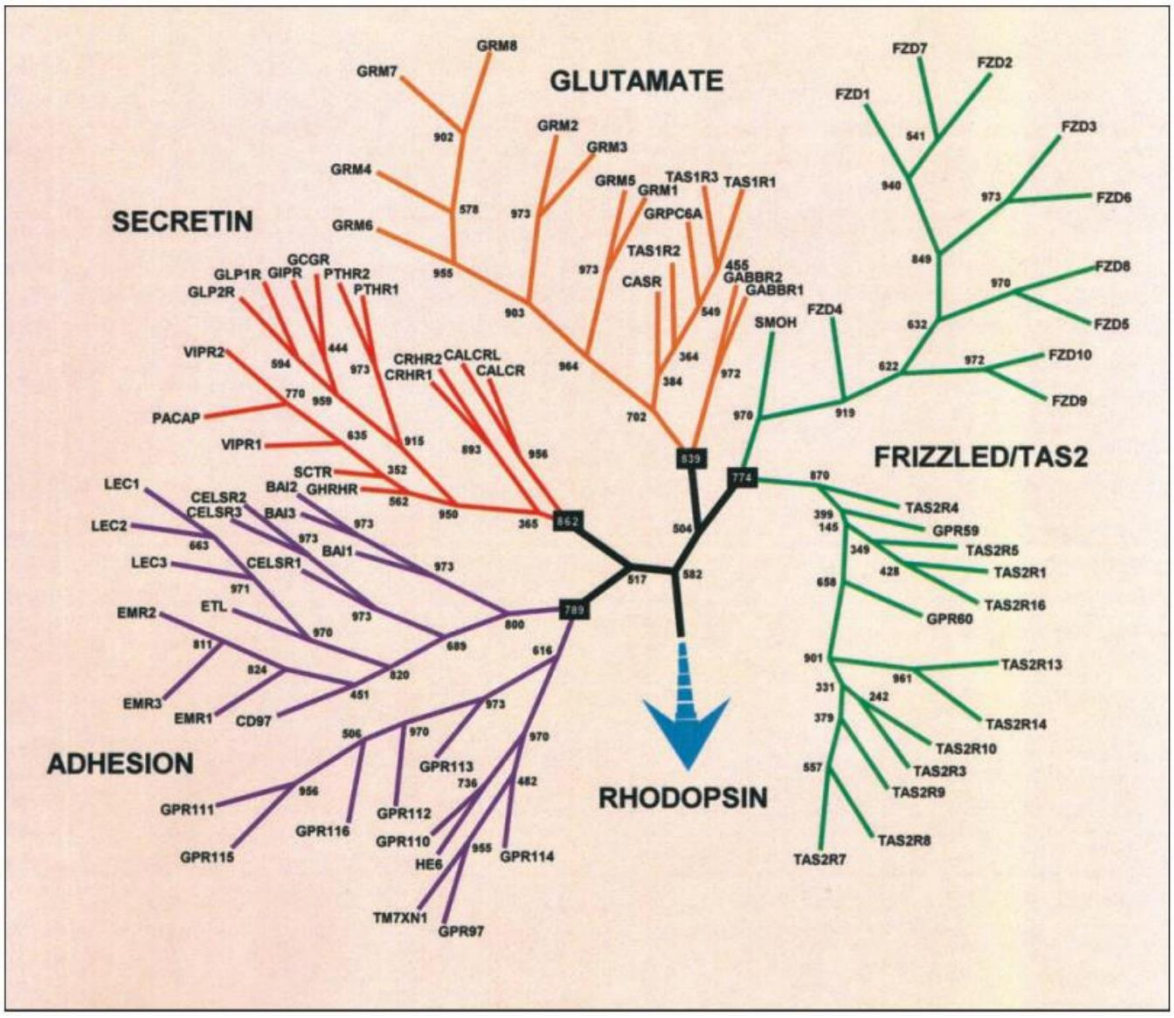


Figure 2.2 Phylogenetic relationship in the human genome of GPCRs (TMI–TMVII). This phylogenetic analysis is based on the "GRAFS" classification system. The different colours of the tree represent the different families (Image adapted from (Fredriksson et al., 2003)).

2.1.2 Structural characteristics of G protein-coupled receptors

Three divisions primarily characterise GPCR structures: (i) an extracellular section consisting of three extracellular loops and the N-terminus forming the opening to the ligand-binding site; (ii) seven alpha-helical transmembrane regions (TM1-TM7); and (iii) an intracellular section consisting of the intracellular amphipathic helix (H8), the three intracellular loops (ICL1-ICL3), and the C-terminus (Figure 2.3) (Heifetz et al., 2020; Latorraca et al., 2017). In general, ligand entry is regulated by the extracellular region; the transmembrane (TM) domain binds ligands, constitutes the structural core and transmits molecular signals through conformational

changes to the intracellular part; and the intracellular region couples with signalling proteins such as G proteins and arrestins (Heifetz et al., 2020).

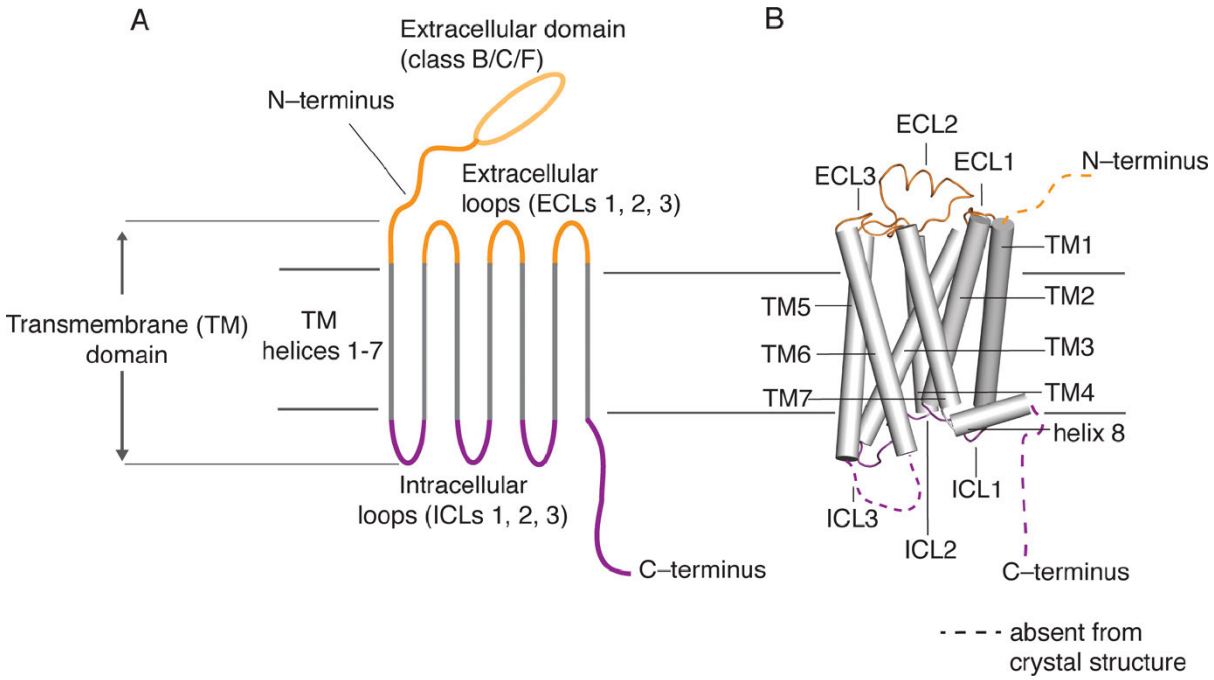


Figure 2.3 GPCRs' general design and structural characteristics. (A) the three extracellular loops (ECLs) and the N-terminus (in orange), the seven transmembrane helices region (in grey), and the three extracellular loops and the C-terminus (in purple). (B) GPCR general architecture displayed in cartoon representation (Image adapted from (Latorraca et al., 2017)).

When triggered by the binding of G protein and an agonist ligand, GPCRs usually change from inactive state to active state conformations. Hydrophobic patterns and different functionally structural signature motifs characterise the strongly conserved transmembrane domains of GPCRs. The NPxxY motif (connecting TM7 and helix 8), the CWxP motif (TM6), and the D(E)RY motif (TM3) compose these motifs. To retain the receptor in its ground state, the ionic interaction (ionic lock) of the TM3 conserved D(E)RY motif has been observed (Rovati et al., 2007).

2.1.2.1 Transmembrane and extracellular residue indexing methods for GPCRs

Superscripts are assigned to residues of the transmembrane regions (TM1-TM7) based on the Ballesteros-Weinstein numbering system (Ballesteros & Weinstein, 1995) to enable comparison of residues among the different GPCRs. The residues of the extracellular and intracellular loop regions are not applied with the Ballesteros-Weinstein numbering approach because of the high variation in sequence and length. The developers of the GPCRdb (Isberg et al., 2016) have used a numbering system for the extracellular and intracellular loop regions

based on the Ballesteros-Weinstein approach for some regions. The Ballesteros-Weinstein nomenclature assigns each transmembrane residue with an index number corresponding to the position of the residue in that transmembrane domain. For example, the conserved GPCRs Asp residue in TM3 (which is $Asp110^{3.32}$ in D3DR and $Asp114^{3.32}$ in D2DR) is denoted as $Asp^{3.32}$ where 3 stands for TM3 and 32 stands for the position of the Asp residue in TM3. This technique enables the different GPCRs to be compared using the most conserved residues as a reference point. The Ballesteros-Weinstein and the GPCRdb numbering schemes are used in this thesis.

2.1.3 GPCR function, activation, and signal transduction pathways

The physiology of most vertebrate relies on GPCR signal transduction. GPCRs can recognise different types of signals, such as small molecules, peptides, photons of light, lipids, hormones, and proteins. GPCRs are crucial nodes of information exchange between the extracellular and intracellular environment of cells. The mechanisms of GPCR activation, signalling and regulation are markedly conserved and exemplify evolutionary convergence (Deupi et al., 2012; Reiter et al., 2012; Reiter & Lefkowitz, 2006) (Figure 2.4**)**.

Generally, in GPCR activation, ligand binding involves the N-terminus, the exposed transmembrane helices, and the extracellular loops. In contrast, G protein-coupling and signal transduction involve the intracellular regions and loops. GPCRs have the classical role of coupling the binding of an extracellular ligand to the binding pocket and activating distinct heterotrimeric G proteins (subunits of Gβ and Gγ, Gβ and Gα), resulting in modulations of downstream effector proteins. The heterotrimeric G proteins subsequently dissociate from the receptor sending signals that produce second messengers, including Ca^{2+}, cAMP, and inositol phosphates (e.g. IP3), which then activates various cellular responses.

The activated GPCR can bind to β-arrestins with high affinity when phosphorylated (Kang et al., 2014; Tian et al., 2014). To desensitise G protein signalling, β-arrestins avert the additional coupling of G proteins to the activated receptor. β-arrestins have been identified as independent signal transducers, controlling signalling processes such as protein synthesis, apoptosis, cell migration and the activation of mitogen-activated protein kinases that regulate the cytoskeleton. GPCRs play a vital role in controlling diverse physiological functions, such as taste, secretion, nervous system control, smell, vision, metabolism, immune response, embryonic growth, cell differentiation, and cell differentiation. As a consequence, GPCR dysfunction is involved in a variety of diseases, including neurological/neurodegenerative

disorders (Guimarães & Thathiah, 2020; Azam et al., 2020; Huang et al., 2017), diabetes (Gupta & Vasudevan, 2019; Sebastiani et al., 2018), cardiovascular disease (J. Wang et al., 2018), obesity (Riddy et al., 2018), cancer (Cerchio & Chen, 2020; Gad & Balenga, 2020), HIV (Brelot & Chakrabarti, 2018), and inflammation (Lin et al., 2017), making GPCRs crucial drug target for pharmaceutical developments.

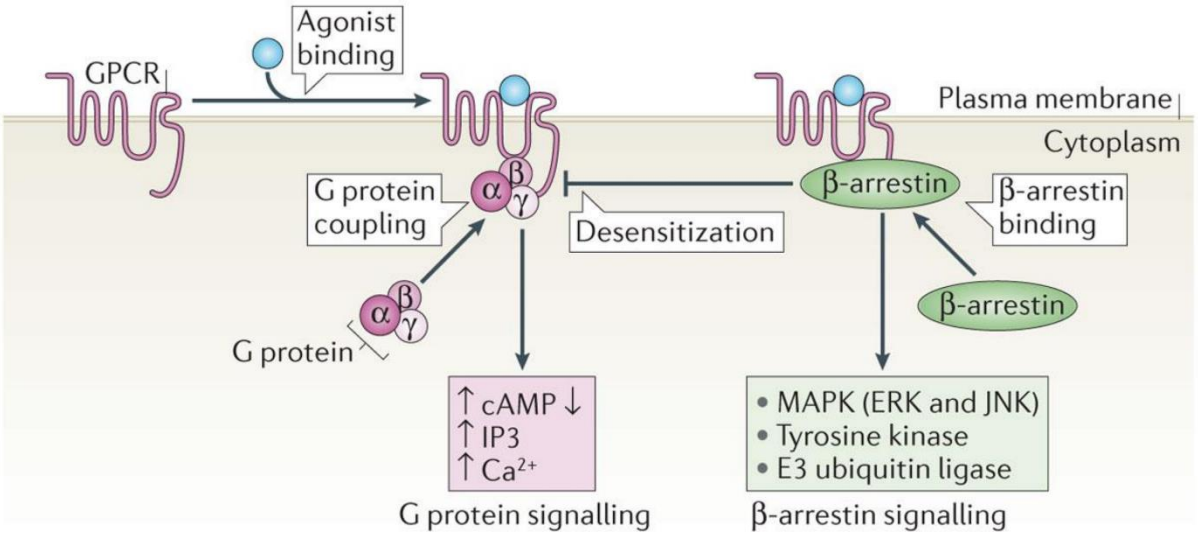


Figure 2.4 A schematic representation of GPCR activation and signalling (Image adapted from (Ghosh et al., 2015)).

2.1.4 Advances in GPCR structure determination

Structural knowledge of GPCRs offers a deeper understanding of their molecular mechanisms for cell signalling and ligand recognition, which assist in the novel development of effective drugs for severe human diseases treatment. The crystallisation and determination of GPCR structures were deemed to be almost impossible before the recent revolution. This was due to the inherent conformational flexibility of GPCRs which posed a challenge in the efforts to obtain their crystal structures. Considerable challenges that hampered the structure determinations and molecular understanding of GPCRs were the inability to get pure, stable, and functional samples.

However, significant progress has been made in GPCR structure determination since the first solved bovine rhodopsin receptor structure in 2000 (Palczewski et al., 2000). This was followed by the first high-resolution crystal structures of the non-rhodopsin-like GPCR of the human β_2-adrenoceptor (β_2AR) bound with the inverse agonist (carazolol) and Fab5 (Rasmussen et al., 2007), and β_2AR in complex carazolol (Cherezov et al., 2007) in 2007. Aside from the detailed interactions demonstrated between β_2AR and carazolol, the structural determination success

also provided novel strategies for optimising and crystallising GPCR structures. The advancements in X-ray crystallographic data collection methods, including X-ray free-electron lasers (XFELs) and micro-focus X-ray sources at synchrotrons, were included in the novel strategies. Also, advances in protein engineering (e.g., thermo-stabilizing mutations (Tate, 2012) and fusion proteins, for instance, T4 lysozyme (Chun et al., 2012)), data acquisition (Liu et al., 2014), and structural biology (e.g., lipid cubic phase (LCP) crystallisation (Caffrey & Cherezov, 2009; Caffrey, 2015) and cryo-electron microscopy (cryo-EM) (Liang et al., 2017)) have increased the number GPCR structures. The application of these novel techniques and guidelines has propelled GPCR structure determination's progress in less than two decades (Table 2.1 and Figure 2.5).

Table 2.1 Statistics of solved GPCR structures obtained from the PDB database, includes multiple receptor complexes (Table extracted from https://gpcrdb.org/structure/statistics accessed on 20 October 2020; Last updated **2020-09-30)**.

Class	**Class A**	**Class B1**	**Class B2**	**Class C**	**Class F**	**Total**
Unique Receptor complexes#	70	10	0	4	3	87
Receptors*****	404	46	0	17	19	486
Receptor-ligand structures*	397	45	0	17	19	478
Active-state structures*	106	31	0	2	6	145
G protein-Receptor structures*	9	30	0	0	0	39

*Orthologues receptors are counted more than once. # A receptor with more than one structure bound to diverse ligands was counted once. The active state is characterised as an intracellular TM bundle that is agonist-bound and open.

Recently, cryo-electron (cryo-EM) microscopy has emerged as a modern technique for determining membrane protein structures (Thal, Vuckovic, et al., 2018; García-Nafría & Tate, 2020), causing a surge in the quality and number of available GPCRs. Several high-resolution solved structures have been solved for chemokine receptors, aminergic receptors, nucleotide receptors, lipid receptors, and peptide receptors. These structures provide insights into receptor function and ligand-recognition of GPCRs and enable the possible comparison of important details between various GPCR subtypes (Qu et al., 2020). Consequently, computational investigations of biomolecules have greatly benefited from the increased crystal structures as they provided a reliable starting structure for computational studies (Hollingsworth & Dror, 2018).

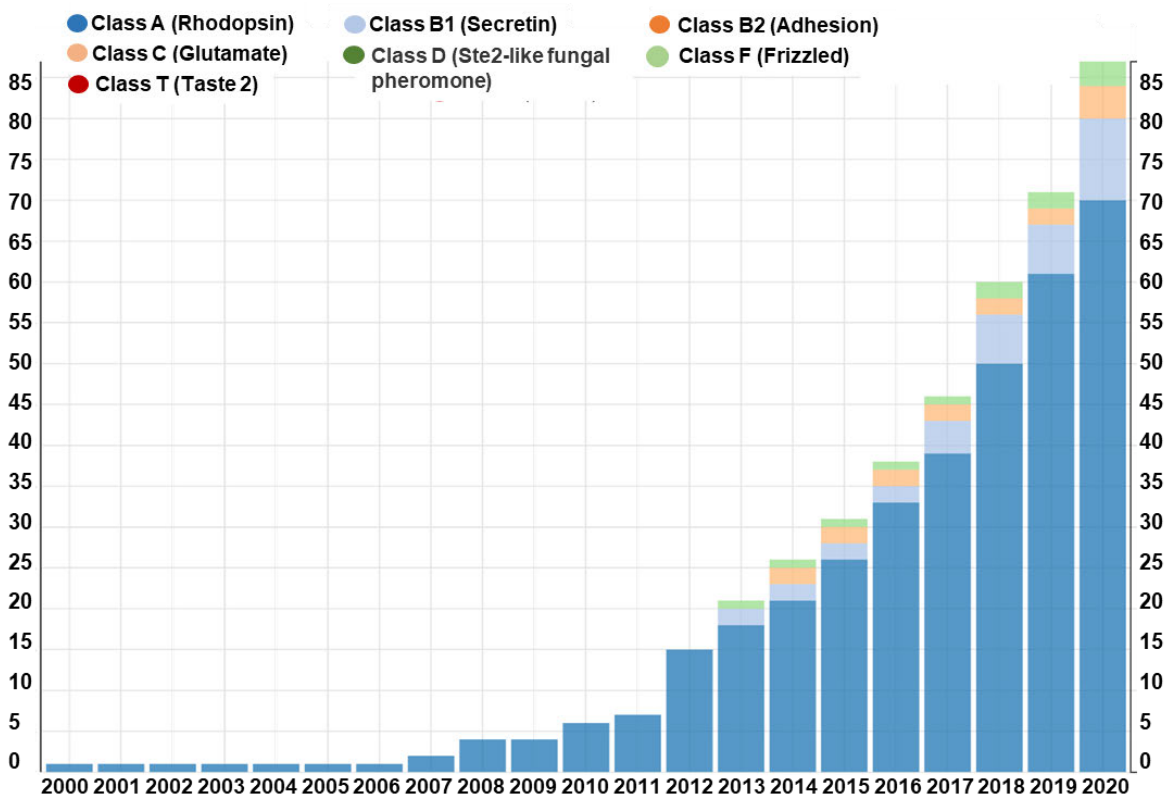


Figure 2.5 The increase in the number of GPCR structure determinations over time for each receptor of the various classes of GPCRs obtained from the GPCRdb database (Image adapted from https://gpcrdb.org/structure/statistics, accessed in October 2020).

2.1.5 G protein-coupled receptors as targets for drugs.

The substantial involvement of GPCRs in various pathophysiological processes makes them an important drug target and constitutes the largest FDA-approved drug family (Allen & Roth, 2011; Rask-Andersen et al., 2014; Santos et al., 2017). The high proportion of drugs targeting the GPCR family can be attributed to various factors, including their ability to bind to drugs with higher affinity, their interactions with several types of chemical entities, and their plasma membrane expression facilitates extracellular molecular interactions (Sriram & Insel, 2018). It is projected that there are nearly 700 validated drugs that target 134 GPCRs, which accounts for about 35% of all FDA approved drugs (Sriram & Insel, 2018). It is estimated that two-thirds of the current drugs targeting GPCRs frequently targets the cyclic adenosine monophosphate (cAMP) signalling pathways (Sriram & Insel, 2018). Drugs targeting GPCRs may act as antagonists or agonists (inverse or neutral agonist), with the majority acting as an antagonist.

The rhodopsin α-subgroup receptors such as the cannabinoid, adrenergic, muscarinic, dopamine, serotonin, and histamine receptors are vital drug targets for anti-histamines, cardiovascular drugs, and antipsychotics (Tyndall & Sandilya, 2005; Jacoby et al., 2006). The

rhodopsin β-subgroup are drug targets pursued in conditions such as hormone-related cancer (Kotake et al., 1999) and pulmonary arterial hypertension (Barst et al., 2004). The chemokine receptors are drug targets because of their function as coreceptors in HIV-1 strain (Onuffer & Horuk, 2002), chronic and acute inflammations (Kraneveld et al., 2010; Elemam et al., 2020) and in cancer immunotherapy (Mollica Poeta et al., 2019). The opioid receptors are vital drug targets for pain, alcoholism and cough treatment (Tyndall & Sandilya, 2005).

2.1.6 GPCRs and Structure-based drug discovery

The current rise in structures of GPCRs has offered valuable knowledge to facilitate structure-based drug design and discovery. These structural insights into the binding of ligands and activation of the receptor, when combined with computational approaches, can significantly accelerate the discovery of new ligands for GPCRs (Qu et al., 2020). Structure-based virtual screening and lead optimisation have proved their immense potential in discovering drugs targeting GPCRs.

Katritch et al. and Carlsson et al. have applied successfully structure-based virtual screening in identifying novel ligands targeting adenosine A_{2A} receptor ($A_{2A}AR$) (Katritch et al., 2010; Carlsson et al., 2010). The screening of 1.4 million compounds by Carlsson et al. using molecular docking against $A_{2A}AR$ resulted in the experimental validation of 20 high-ranking molecules. Finally, seven of the hits were validated as new selective ligands of $A_{2A}AR$ (Carlsson et al., 2010). On over four million compounds, Katritch et al. also used structure-based virtual screening techniques and reported 53 hits against $A_{2A}AR$ (Katritch et al., 2010). In another structure-based study, Rodríguez et al. identified nine experimentally confirmed novel serotonin (5-hydroxytryptamine [5-HT]) subtype-selective ligands with a preference for 5-HT_{1B} over 5-HT_{2B} subtype from 1.3 million compounds (Rodríguez et al., 2014). The nine demonstrated selective compounds exhibited up to 300-fold selectivity towards 5-HT_{1B} with three compounds been agonist of the G protein pathway (Rodríguez et al., 2014).

Other successful structure-based discoveries include the identification of novel ligands at Beta-2 Adrenergic Receptor (β_2AR) (Yakar & Akten, 2014; Kolb et al., 2009); the discovery of the potent opioid analgesic PZM21 with reduced side effects targeting at μ-Opioid-Receptor (μOR) (Manglik et al., 2016); and novel effective HIV entry inhibitors targeting C-C chemokine receptor 5 (Peng et al., 2018). Also, fragment-based screening was performed by Christopher et al. on 3500 compounds at metabotropic glutamate receptor 5 ($mGlu_5$) and identified the negative allosteric modulator pyrimidine 5 (Christopher et al., 2015). Several

candidate drugs developed via structure-based techniques are presently in clinical trials or preclinical development for conditions varying from neurological disorders to cancer (Hauser et al., 2017; Congreve & Marshall, 2010). More recently, the D_2 dopamine receptor crystal structure bound with haloperidol aided in identifying two novel D2DR subtype-selective agonists (O_4SE_6 and O_8LE_6) that exclude agonism at D4DR and D3DR (Fan et al., 2020).

2.2 Overview of Dopamine Receptors

Dopamine is a crucial and dominant catecholamine neurotransmitter in the brain. The dopaminergic receptors mediate the physiological actions of dopamine. Dopamine receptors are essential members of the GPCRs superfamily of Class A membrane receptors (Beaulieu & Gainetdinov, 2011; Thal, Glukhova, et al., 2018). There are five closely related but different dopamine receptor subtypes to date, namely: D_5, D_4, D_3, D_2, and D_1 dopamine receptors. These subtypes are classified further into D_2-like receptors (D_4, D_3 and D_2) and D_1-like receptors (D_5 and D_1) based on structure and pharmacological role (Baik, 2013). The D_1 and D_2 dopamine receptors are primarily and abundantly expressed in the brain (D_1 is highly expressed), and the two are rarely co-expressed in the same cells (Missale et al., 1998; Baik, 2013). Two D_2 dopamine receptor (D2DR) isoforms are available, namely the long (D2L) and the short (D2S) isoforms produced by alternative splicing (Giros et al., 1989; Dal Toso et al., 1989; Shioda, 2017). These isoforms are similar, except that they differ in the insertion of only 29 amino acids in the D2L intracellular loop, which has been posited to assist in determining the specificity of the second messenger (Żuk et al., 2020; Baik, 2013; Giros et al., 1989).

Dopamine receptors expression is mainly abundant in the central nervous system (CNS) and the peripherals such as kidneys, blood vessels, retina, heart, the renin-angiotensin system, and adrenals controlling the release of catecholamine (Beaulieu & Gainetdinov, 2011). Figure 2.6 depicts the dopaminergic pathways, dopamine receptors and dopamine distributions in the central and peripheral systems. Four dopaminergic pathways have been established in the mammalian brain, namely: mesocortical, mesolimbic, nigrostriatal, and tuberoinfundibular (Figure 2.6B). These neurons play an essential function in the CNS, such as cognition, locomotor activity, emotion, appetite, reward, attention, sleep, learning, and working memory. Dopamine can regulate endocrine, heart rate, kidney function, cardiovascular function, gastrointestinal motility and so on in the periphery (Vallone et al., 2000). For D_2-like dopamine receptors, dopamine has a higher affinity ranging from 10 to 100-fold than for D_1-like dopamine receptors. The D_1 dopamine receptor has the lowest dopamine affinity (Beaulieu & Gainetdinov, 2011; Tritsch & Sabatini, 2012).

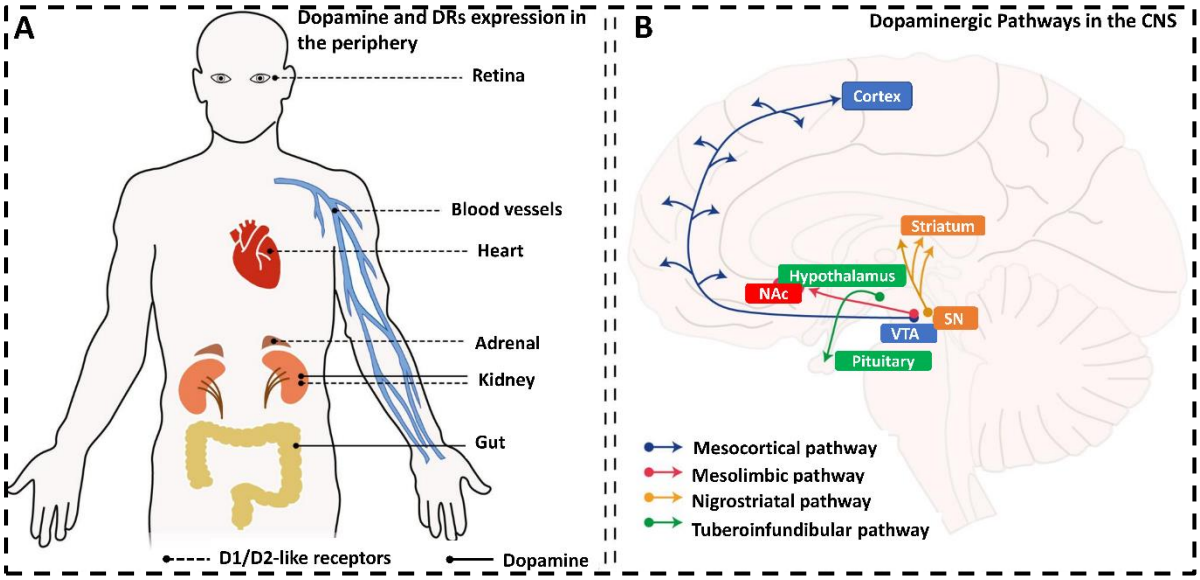


Figure 2.6 The peripheral expression of dopamine receptors and dopamine (A) and the central nervous system distribution of four major dopaminergic pathways (B). The ventral tegmental area (VTA) is the centre of the mesocorticolimbic system: dopaminergic neurons are transmitted by the mesocortical pathway (blue) to the cortex and by the mesolimbic pathway (red) to the nucleus accumbens. The tuberoinfundibular pathway (green) is formed by dopaminergic neurons projecting from the hypothalamic nuclei to the pituitary. In contrast, in the projection of the substantia nigra (SN) to the striatum, dopamine neurons form the nigrostriatal pathway (orange) (Klein et al., 2019)).

2.2.1 Dopamine receptor signalling

The binding of dopamine to a receptor produces intracellular responses depending on the kind of dopamine receptor stimulated. The downstream signalling of dopamine primarily includes G proteins, although G protein-independent signalling pathways can also be involved in dopamine receptor signalling (Luttrell & Lefkowitz, 2002). The D_2-like and D_1-like dopamine receptors are functionally different in the manner they modulate their intracellular signalling pathways.

The adenosine triphosphate (ATP) is converted into 3'-5'-cyclic adenosine monophosphate (cAMP) as a result of adenylyl cyclase (AC) protein activation induced by the coupling of D_1-like receptors to Gαs/olf protein (Sunahara & Taussig, 2002). The cAMP then stimulates the activity of the protein kinase A (PKA) by interacting with the catalytic and regulatory subunits of PKA, which then induce the release of catalytic subunits for the phosphorylation of various substrates (Akimoto et al., 2013) (Figure 2.7). The DARPP-32 phosphorylation at the Thr34 residue causes inhibition of phosphatase-1 (PP1) protein (Nishi, Snyder, et al., 1999). In contrast, the Thr75 of DARPP-32, when phosphorylated by cyclin-dependent kinase 5 (Cdk5),

inhibits PKA activity and thus suppresses the signalling of D_1-like dopamine receptors (Bibb et al., 1999; Undieh, 2010). It has also been documented that activation of the D_1-like dopamine receptor controls the electrochemical gradient through Na+K+-ATPase, pumping potassium in and sodium out of cells. It has also been shown that activation of D_1-like dopamine receptors inhibits Na+K+-ATPase through PKA and PKC signalling pathways in the striatum (Gomes & Soares-da-Silva, 2002; Nishi, Fisone, et al., 1999) (Figure 2.7).

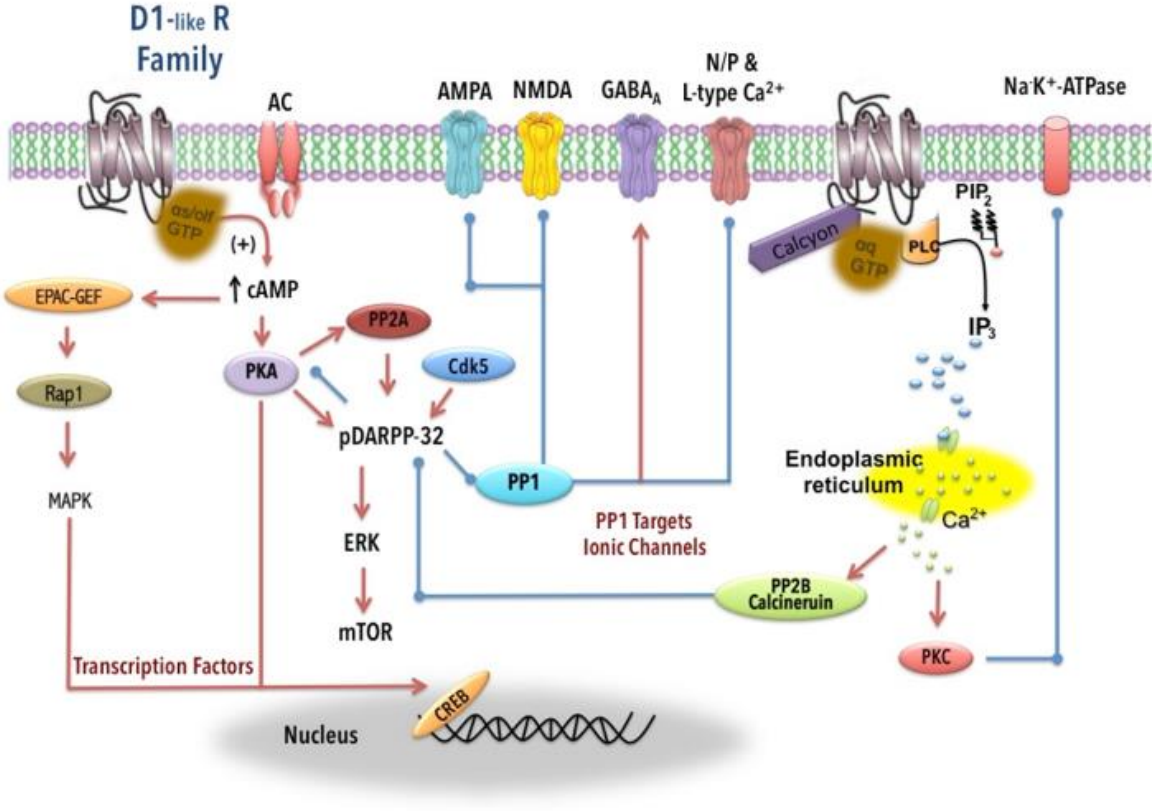


Figure 2.7 The intracellular signalling pathways of D_1-like dopamine receptors, showing dopamine-mediated effects through complex activation of intracellular signals. Red arrows indicate stimulatory effects, blue lines ending with circles for inhibitory effects, and plausible activation indicated by the dashed red arrow (Image adapted from (Rangel-Barajas et al., 2015)).

The $G\alpha_{i/o}$ class of G proteins is the central mediator of D_2-like signalling (i.e., D_4, D_3, and D_2 dopamine receptors). Contrary to D1-like receptors, D2-like receptors coupling to Gαi/o protein inhibits (AC) and decrease the intracellular concentration 3'-5'-cyclic adenosine monophosphate (cAMP), resulting in the blocking of protein kinase A (PKA) activity (Beaulieu & Gainetdinov, 2011; Akimoto et al., 2013) (Figure 2.8). After the stimulation of D_2-like receptors, DARPP-32 phosphorylation at Thr34 is usually decreased because of diminished PKA activation.

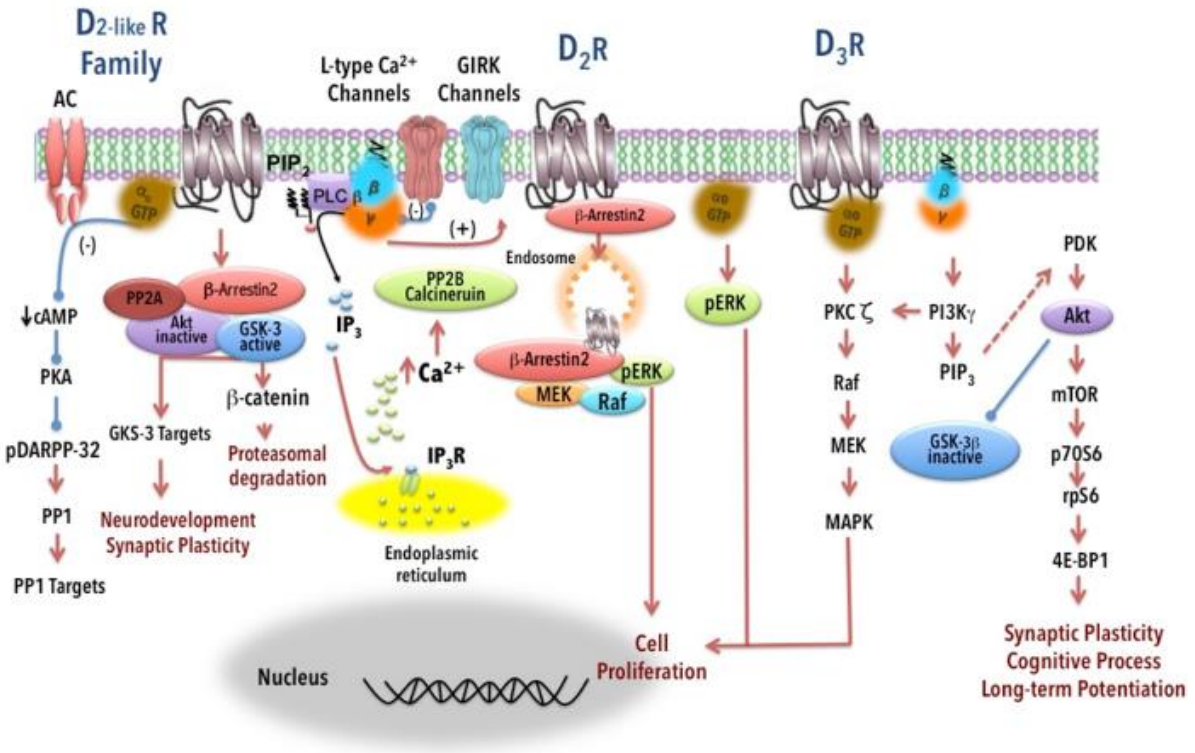


Figure 2.8 The intracellular signalling pathways of D_2-like dopamine receptors, showing dopamine-mediated effects through complex activation of intracellular signals. These signals relate to functions including proteasomal degradation, cell proliferation, neurodevelopment, and cognitive process. Red arrows indicate stimulatory effects, blue lines ending with circles for inhibitory effects, and plausible activation indicated by the dashed red arrow (Image adapted from (Rangel-Barajas et al., 2015)).

2.2.2 Structure determination of dopamine receptors

Advances in GPCR structure determination have led to the crystallisation of all the D2-like dopamine receptors; however, none of the D1-like receptors has crystallised yet (Table 2.2 and Figure 2.9). The first experimental structure of G Protein coupled to D2DR embedded in a lipid membrane was resolved recently by cryo-electron microscopy (cryo-EM) (Yin et al., 2020). These resolved structures have provided novel insights into the recognition of ligands and the dopamine receptor activation process.

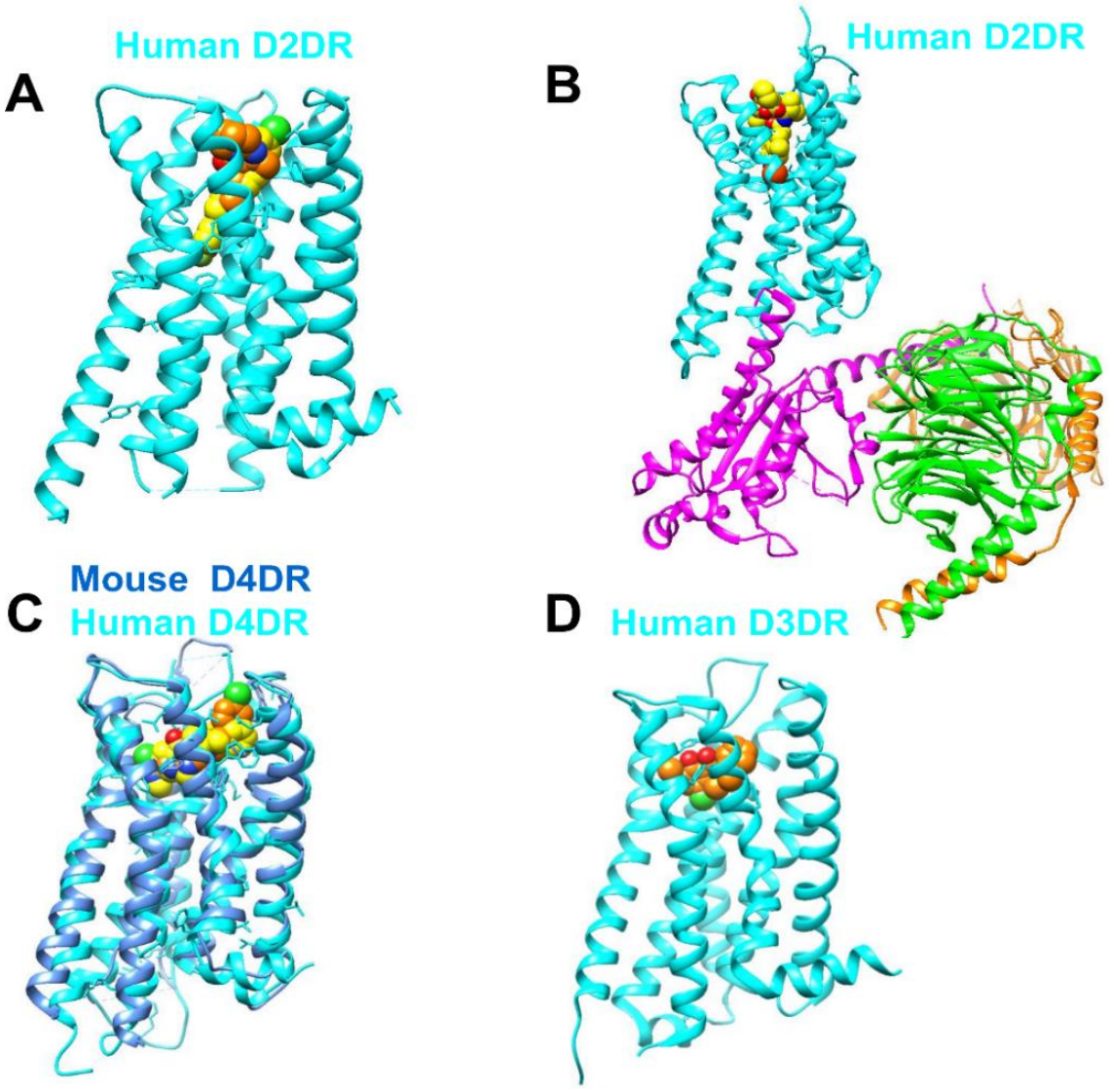


Figure 2.9 Crystal structures: (A) human D_2 dopamine receptor in complex with risperidone antagonist (orange carbon spheres) and haloperidol antagonist (yellow carbon spheres). (B) human D_2 dopamine receptor–G-protein complex with the bromocriptine agonist (yellow carbon sphere). The receptor is shown in cyan, $G\alpha_i$ in magenta, $G\beta$ in green, and $G\gamma$ in orange. (C) Human D_4 dopamine receptor (cyan) in complex with nemonapride (yellow carbon spheres) and the mouse D_4 dopamine receptor bound to the subtype-selective antagonist L745870 (orange carbon sphere) and (D) the human D_3 dopamine receptor bound to the D2/D3 agonist eticlopride (orange carbon spheres) (Image prepared with UCSF Chimera program by author).

Table 2.2 List of solved structures dopamine receptors.

Dopamine receptors								
Receptor	**Species**	**Ligand**	**Ligand Function**	**PDB code**	**Resolution (Å)**	**Year**	**Reference**	**Method**
D2	Human	Bromocriptine G-protein	Agonist Signalling-protein	6VMS	3.80	2020	(Yin et al., 2020)	cryo-EM
		Haloperidol	Antagonist	6LUQ	3.10	2020	(Fan et al., 2020)	x-ray
		Risperidone	Antagonist	6CM4	2.87	2018	(S. Wang et al., 2018)	
D3	Human	Eticlopride	Antagonist	3PBL	2.89	2010	(Chien et al., 2010)	
D4	Human	Nemonapride	Antagonist	5WIU	1.96	2017	(Wang et al., 2017)	
				5WIV	2.14			
	Mouse	L745870	Antagonist	6IQL	3.50	2019	(Zhou et al., 2019)	

2.2.3 Diseases implicated in the dysfunction of the dopaminergic system

The dopaminergic pathway is involved in diverse physiological processes, including motor behaviour, neuroendocrine function, cognitive function, and emotion (Klein et al., 2019). However, the dysfunction of the dopaminergic signalling pathways is involved in several neurological and psychiatric conditions, including schizophrenia, Parkinson's disease, bipolar disorder, substance abuse or addiction, Huntington's disease, attention deficit hyperactivity disorder (ADHD), and depression (Rangel-Barajas et al., 2015; Maggio et al., 2015; Beaulieu & Gainetdinov, 2011; Heidbreder & Newman, 2010; Klein et al., 2019).

2.2.4 Drugs targeting D_2-like dopamine receptors and recent advances in drug discovery

The dopamine receptors are the primary pharmacological targets of all the existing antipsychotic medications. The D2-like receptors (D2/D3) are the essential targets of these antipsychotics within the dopamine receptor subtypes. Agonists targeting D_2-like dopamine receptors are subdivided into ergoline agonists (including cabergoline, pergolide, bromocriptine, and lisuride) and non-ergoline agonists (including pramipexole and ropinirole) (Brooks, 2000). Apomorphine is a D_2-like, and D_1-like dopamine receptor agonist demonstrated to improve the symptoms of Parkinson's disease and is administered via subcutaneous injection (Schwab, 1951; Brooks, 2000). Antipsychotics acting as antagonists have been classified as first-generation or typical and second-generation or atypical. The second-generation antipsychotics demonstrate potent antagonism at the D_2 dopamine receptor and serotonin 2A receptor (5-HT_{2A}) (Figure 2.10 and 2.11).

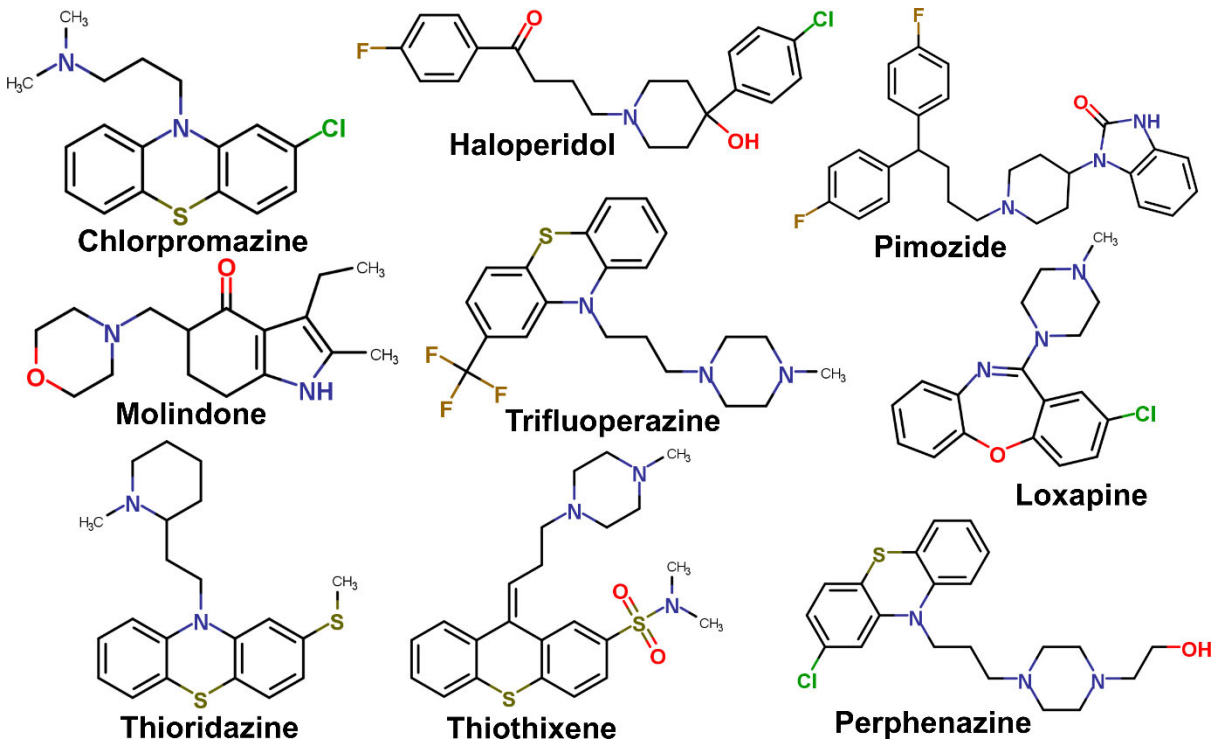


Figure 2.10 2D chemical structures of commonly prescribed first-generation (typical) antipsychotic medications (Image prepared by author).

The existing FDA-approved antipsychotic drugs lack selectivity toward a given D_2-like receptor subtype, and they also interact with other GPCRs, for instance, serotonergic, adrenergic, cholinergic, and histaminergic receptors (Li et al., 2016; Moritz et al., 2018). Consequently, there has been a recent paradigm shift towards selective drug targeting of the diverse dopamine receptor subtypes in the treatment of neurological conditions, such as drug addiction, Parkinson's disease, and schizophrenia. The recent shift is imperative since the approved antipsychotics are accompanied by substantial adverse effects such as metabolic syndrome, cardiovascular hypertension, and neurological side effects, including extrapyramidal reactions and tardive dyskinesia (Kaar et al., 2020; Ballon et al., 2014; Álvarez et al., 2013). These undesirable effects reduce patient compliance with medications and the quality of life (Novick et al., 2010; Lieberman et al., 2005). The selective targeting of each D_2-like receptor subtype has been shown to produce fewer side effects (Li et al., 2016; Holmes et al., 2004). Antagonists with higher selectivity for D_3 dopamine receptor (D3DR) over D_2 dopamine receptor (D2DR) have been shown to show promising results in reducing cocaine and opioid reward and are highly effective in mitigating relapse to drug-seeking behaviour in preclinical models (Andreoli et al., 2003; Higley et al., 2011; Galaj et al., 2015). The availability of crystal structures for all the D_2-like dopamine receptors provides a reference

point for the rational design of novel scaffold molecules toward D_2-like subtype-selective drugs.

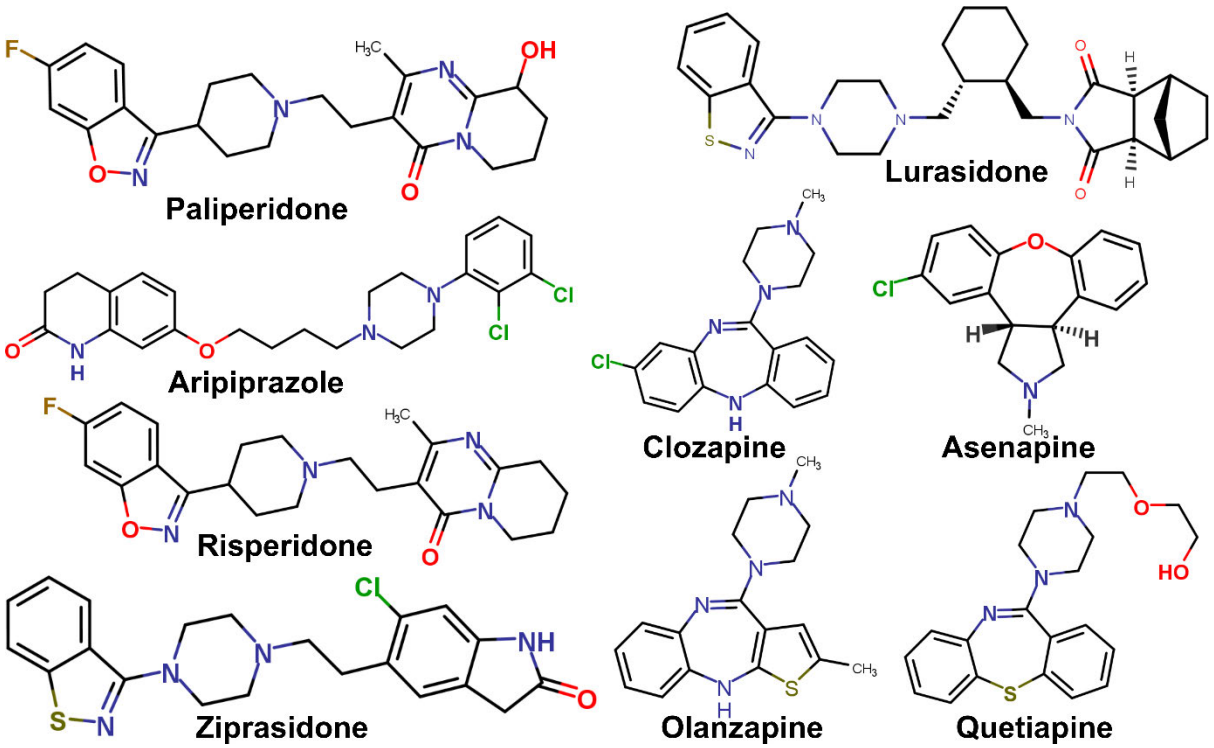


Figure 2.11 2D chemical structures of commonly prescribed second-generation (atypical) antipsychotic medications (Image prepared by author).

2.3 Overview of CC Chemokine receptors

Chemokines are chemotactic cytokines that signal proteins that regulate different functions, such as recruitment of immune cells and immune surveillance. Approximately 50 chemokine endogenous ligands are identified in mice and humans, making chemokines the most common class of cytokines (Griffith et al., 2014). They have been shown to be essential for human immune system development and homeostasis and are necessary for all destructive or protective immune and inflammatory activities (Hughes & Nibbs, 2018). These chemokines are categorised into the CX3C, CXC, CC, and XC subfamilies based on the structure and the number of N-terminal cysteine residues. The chemokines are also designated as "R" to denote the receptor and consist of seven N-terminus and C-terminus transmembrane helices common to the GPCR family (Murphy, 2002).

CCR5 with endogenous antagonists including MIP-1α (CCL3), MIP-1β (CCL4), and RANTES (Regulated upon Activation, Normal T cell Expressed and Presumably Secreted, CCL5) has been recognised as a functional GPCR (Combadiere et al., 1996; Samson et al., 1996). The cotaxin, the monocyte chemotactic proteins (MCP-4, MCP-3, and MCP-1), and several other

CC chemokines were later identified to bind CCR5 with varied affinities and efficiencies during receptor activation (Blanpain et al., 1999). Table 2.3 shows the identified chemokines as endogenous ligands of CCR5. The quest to develop chemokine antagonists started when the chemokine receptors CCR5 and CXR4 were identified as coreceptors of the human HIV-1 virus (Deng et al., 1996). The major surface coreceptors that promote HIV-1 entry are CCR5 and CXCR4, of which CCR5 is dominant (51 %) in cases of HIV-1 infection. The R5-tropic strains of HIV-1 primarily use the coreceptor CCR5 during the development of a new infection. The determinant (X4 tropic or R5 tropic) of the virus tropism depends on the selectivity of the coreceptor (Berger et al., 1998).

Table 2.3 Chemokine receptors and chemokines associated with CCR5

Chemokine receptor	**Chemokines**	
	Common Name	Systematic Name
CCR2, CCR4	MCP-1	CCL2
CCR1, CCR4, CCR5	MIP-1α	CCL3
CCR5, CCR3, CCR1	RANTES	CCL5
CCR5	MIP-1β	CCL4
CCR5, CCR3, CCR2	MCP-4	CCL13
CCBP2, CCR5	LD78β	CCL3L1
CCR5, CCR3, CCR2, CCR1	MCP-3	CCL7
CCR1, CCR2B, CCR3, CCR5	MCP-2	CCL8

2.3.1 Structure determination of Chemokine Receptors

Presently, the crystal structures of seven different chemokine receptors—CCR9, CCR7, CCR5, CCR2, CXCR4, CXCR1 and the homologue US28 of the viral chemokine–receptor CX3CR1 have been determined (Table 2.4). The resolved structures of the chemokine receptors provide valuable insights into the structural basis of ligand recognition and chemokine-receptor activation. These structures reveal various ligand binding sites as they mostly come bound with chemokine ligands, small molecule antagonists, or modified peptides (Figure 2.12).

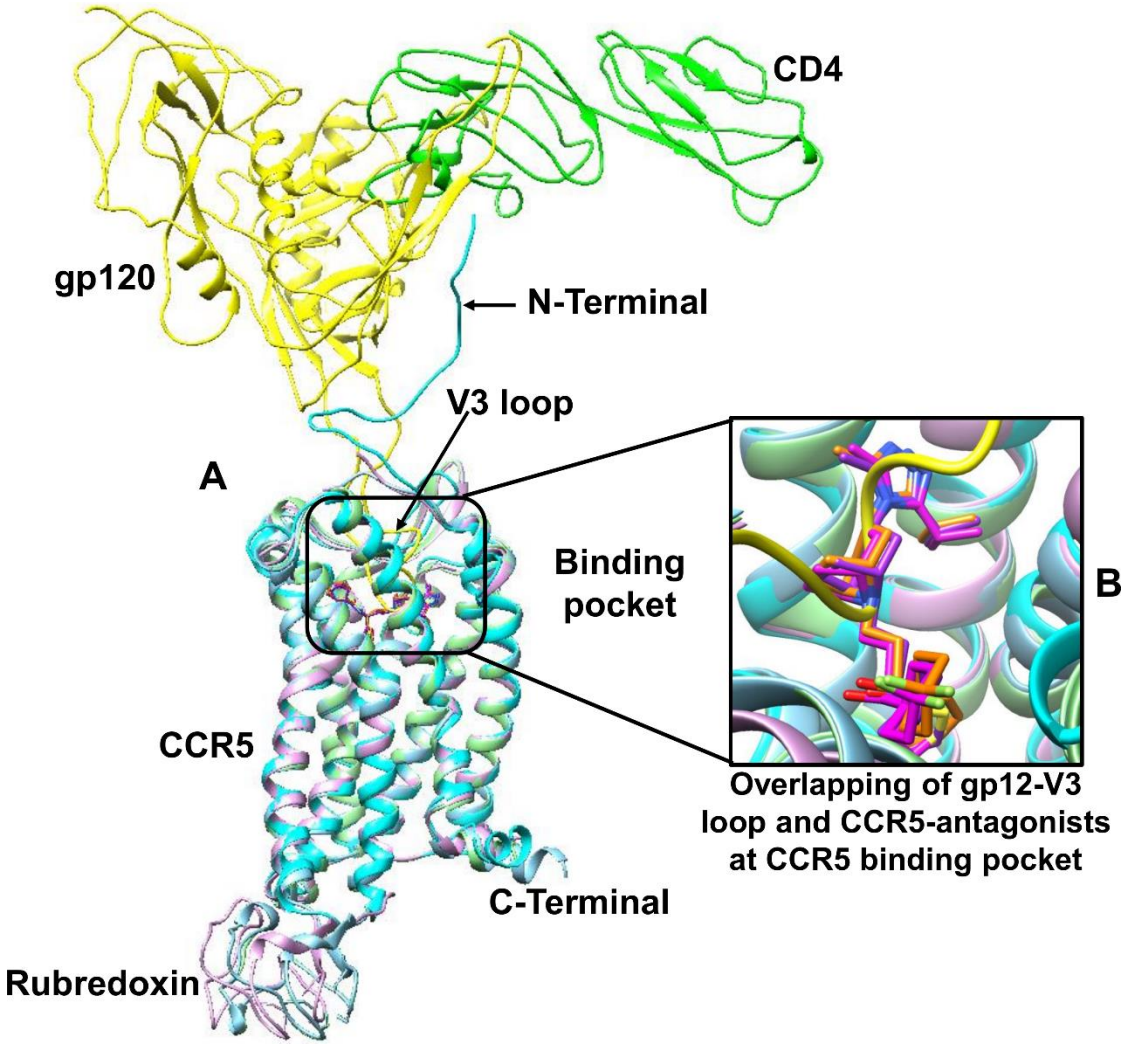


Figure 2.12 Superposition of CCR5 crystal structures in complex with gp120 and antagonists. (**A**) The full length unmodified CCR5 (cyan) in complex with gp120 (yellow) and CD4 (green) (PDB ID: 6MEO); CCR5 (sky blue) in complex with Maraviroc (orange) (PDB ID: 4MBS): CCR5 (light green) in complex with Compound 21 (Magenta) (PDB ID: 6AKX); and CCR5 (plum) in complex with Compound 34 (Purple) (PDB ID: 6AKY). (**B**) gp120 V3 loop overlaps with CCR5-antagonists (Maraviroc, Compound 21 and Compound 32) at CCR5 binding pocket (Image prepared with UCSF Chimera program by author).

The first crystal structure of CCR5 was determined bound with Maraviroc in 2013 (Tan et al., 2013). The solved CCR5 structure served as the basis in a structure-based optimisation and design of novel 1-Heteroaryl-1,3-propanediamine derivatives relying on the functional and binding properties of Maraviroc at CCR5 (Peng et al., 2018). Two of these novel compounds (Compound 21 and Compound 34) demonstrated their potential as candidate drugs as HIV entry inhibitors with improved anti-HIV-1 activity and tolerable pharmacokinetic profile compared with Maraviroc (Peng et al., 2018). Additionally, the Structure of the full-length unmodified Structure of CCR5 in complex with gp120 and CD4 was solved, providing novel insights into CCR5 ligand recognition (Shaik et al., 2019). This structure showed that the V3

loop of gp120 overlaps with Maraviroc at the CCR5 binding site demonstrating a direct competitive inhibition of gp120 binding by Maraviroc contrary to previous views of allosteric inhibition (Shaik et al., 2019).

Table 2.4 Overview of the crystal structures of the solved chemokine receptor and their corresponding ligands.

Chemokine receptors						
Receptor	**Species**	**Ligand**	**PDB code**	**Technique**	**Year**	**Reference**
CCR5	Human	Maraviroc (antagonist)	4MBS	X-ray	2013	(Tan et al., 2013)
		[5P7]CCL5 (antagonist)	5UIW		2017	(Zheng et al., 2017)
		Compound 21 (antagonist)	6AKX		2018	(Peng et al., 2018)
		Compound 34 (antagonist)	6AKY			
		HIV-1 envelope spike, Cd4 complex	6MEO	cryo-EM		(Shaik et al., 2019)
			6MET			
CCR2		BMS-681 (orthosteric) & CCR2-RA-[R] (allosteric) antagonists	5T1A	X-ray	2016	(Zheng et al., 2016)
		MK-0812 (antagonist)	6GPS		2019	(Apel et al., 2019)
		MK-0812 (antagonist)	6GPX			
CCR7		Cmp2105 (allosteric antagonist)	6QZH			(Jaeger et al., 2019)
CCR9		Vercirnon (antagonist)	5LWE		2016	(Oswald et al., 2016)
CXCR4		CVX15 (antagonist)	3OE0		2010	(Wu et al., 2010)
		IT1t (antagonist)	3OE9			
			3OE8			
			3OE6			
			3ODU			
		vMIP-II (antagonist)	4RWS		2015	(Qin et al., 2015)
CXCR1		Ligand free	2LNL	Solid-state NMR	2012	(Park et al., 2012)
US28		CX3CL1 (agonist)	4XT1	X-ray	2015	(Burg et al., 2015)
		CX3CL1 (agonist)	4XT3			
		Ligand free	5WB1		2018	(Miles et al., 2018)
		CX3CL1.35 (agonist)	5WB2			
cryo-EM = cryo-Electron Microscopy; X-ray = X-ray crystallography. Solid-state NMR = Solid-state Nuclear Magnetic Resonance						

2.3.2 CC Chemokine receptors 5 (CCR5) in HIV Infection

CCR5 is a member of the GPCR family of signal transducers and plays a significant function in the initial stages of HIV-1 infection as a functional coreceptor for HIV-1 viral entry (Tan et al., 2013). CCR5 constitutes an extracellular N-terminus, a C-terminus, three extracellular loops (ECLs), three intracellular loops (ICLs), and seven transmembrane helices. The CCR5 second extracellular loop (ECL2) and N-terminus are crucial for HIV interactions during viral entry. Compared to the viral coreceptor CXCR4 and receptor CD4, the CCR5 has an additional advantage as a cellular target since it is relatively redundant for normal immune function (Askew et al., 2016). CXCR4 and CD4 are involved in vital processes in immune function (Berger et al., 1999; Nagasawa et al., 1996), limiting their effectiveness as antiretroviral therapeutic targets. The critical role played by CCR5 in HIV-1 infection was demonstrated when the CCR5 gene (CCR5-Δ32) naturally occurring mutation conferred resistance to HIV-1 infection (Allers et al., 2011). Additionally, individuals possessing heterozygous Δ32 more slowly progressed to AIDS than do homozygous individuals with the wild-type gene (Liu et al., 1996; Paxton et al., 1998). Also, the density levels of CCR5 on $CD4^+$ T cells correlates positively with RNA viral loads (de Roda Husman et al., 1997) and, when untreated, progresses to AIDS (Reynes et al., 2000). CCR5 is essential in stabilising the conformational changes induced by CD4, which are competent for promoting fusion (Shaik et al., 2019).

The infection of HIV into a new host requires entry into susceptible target cells. Viral entry into host cells is facilitated by the HIV envelope glycoproteins gp120 (receptor binding) and gp41 (fusion). The mechanism of HIV entry occurs in several steps (Figure 2.13). The initial stage of viral entry is the attachment of gp120 to the CD4 receptor. The binding to CD4 induces a conformational change exposing the binding pocket of the coreceptor (CCR5 or CXCR4) (Tran et al., 2012). The CD4 binding site and the V1 and V2 regions shift from the trimer centre, causing the central gp41 stalk and the V3 loop to be exposed (Liu et al., 2008). The gp120 bridging sheet and the V3 loop then attach to either CCR5 or CXCR4 coreceptor exposed binding pocket subject to the R5 or R4 virus, respectively. The coreceptor engagement induces a second conformational change that activates gp41 fusion into the host cell membrane (Cormier & Dragic, 2002; Wilen et al., 2012b). In detail, a pre-fusion intermediate is formed by gp41 whereby extended helices are formed by the N-terminal heptad repeat-1 (HR1) (light green) and the C-terminal heptad repeat-2 (HR2) (dark green) of gp41, with the fusion peptide (FP) (yellow) been inserted into the host cell membrane (Figure 2.13)(Chan & Kim, 1998; Falkenhagen & Joshi, 2018). When HR1 is inserted with HR2, the gp41 folds back on itself,

leading to the formation of a six-helix bundle. The viral and cellular membranes are brought together by this conformational change which causes viral membranes and cellular lipids to mix. After the creation of a fusion pore resulting from lipid mixing, the content of virions is released into the cytoplasm (Wilen et al., 2012a).

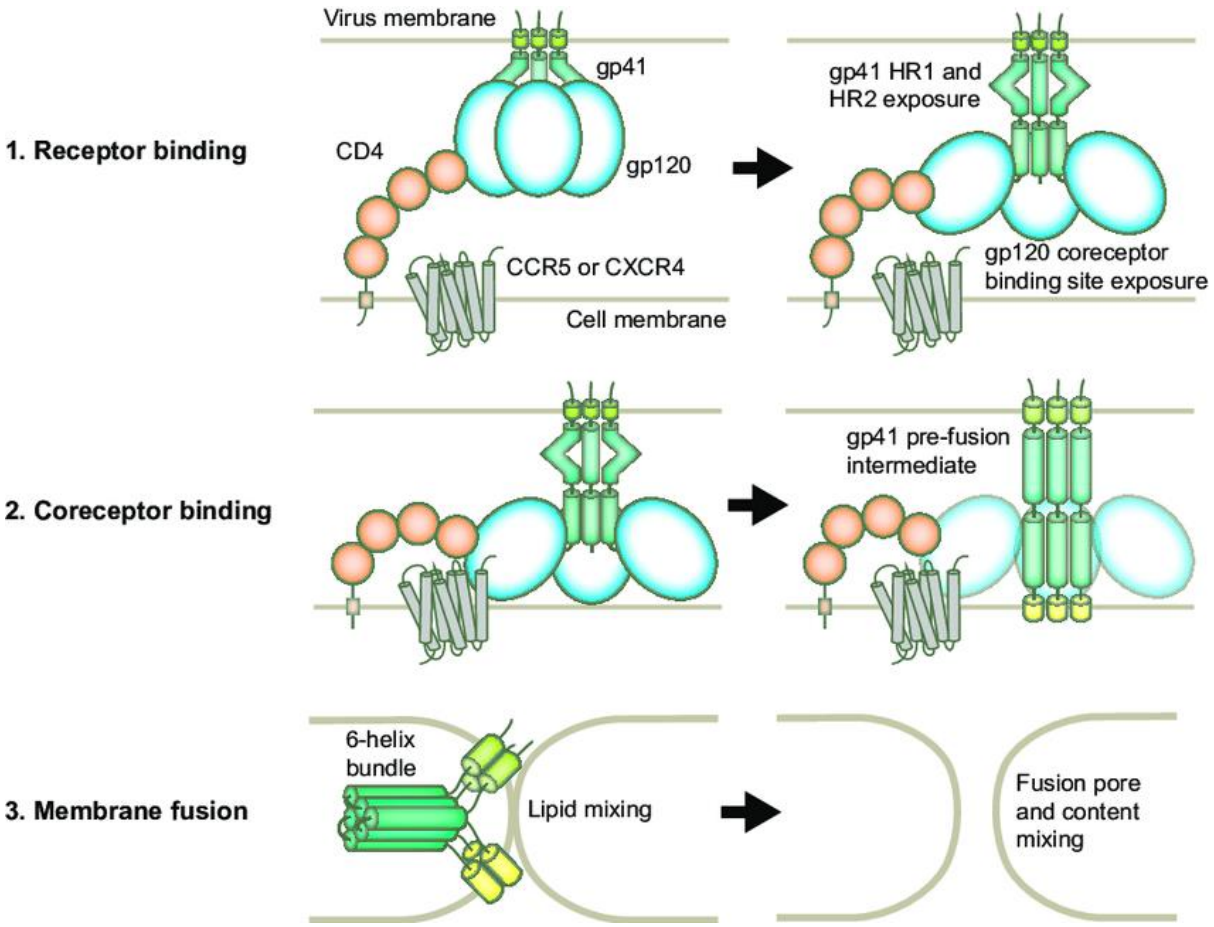


Figure 2.13 A schematic diagram of HIV entry mechanism. The gp41 HR1, HR2, and FP are depicted in light green, dark green, and yellow, respectively (Image adapted from (Falkenhagen & Joshi, 2018)).

2.3.3 Advances in drug discovery towards CCR5 in HIV therapy

The validation of CCR5 as a promising drug target in HIV-1 treatment resulted in the development of several approaches to block HIV-1 interaction with CCR5, including covalently modified natural CCR5 ligands, small molecule antagonists, and monoclonal antibodies (mAbs) (Berro et al., 2011; Henrich & Kuritzkes, 2013). However, several small-molecule CCR5 antagonists have advanced to phase 2 or 3 clinical trials, with many proving effective inhibition of HIV-1 replication (Figure 2.14). CCR5 small molecules have shown excellent *in vitro* synergy when combined with other CCR5 inhibitors (including CCR5 mAbs)

to significantly block HIV-1 entry (Ji et al., 2007; Westby et al., 2007; Lalezari et al., 2008; S Latinovic et al., 2016).

Maraviroc (MVC) was the first CCR5 small-molecule antagonist to be approved and marketed as a CCR5 HIV-1 entry inhibitor (FDA, 2007; Carter, 2013). The oral administration of Maraviroc has resulted in a substantial decrease in viral loads (Westby et al., 2007; Schlecht et al., 2008); however, the prescription of Maraviroc is limited due to identified factors such as its drug-drug interactions (especially when co-administrated with CYP3A4 inhibitors), CYP450 inhibition, and viral resistance (Garcia-Perez et al., 2015; Peng et al., 2018). The prodrug Fostemsavir (formerly BMS-663068/BMS-626529 and marketed as Rukobia) has emerged as the first-in-class HIV attachment inhibitor approved by the US FDA, mainly for HIV positive patients with multidrug-resistant HIV-1 and intolerance or safety issues with other therapies (FDA, 2020). Fostemsavir binds to HIV envelope glycoprotein 120 (gp120) when hydrolysed to its active form temsavir, which prevents the conformational change needed by gp120 for attachment to CD4 cell surface receptor of the host (Lalezari et al., 2015; Thompson et al., 2017; Kozal et al., 2020).

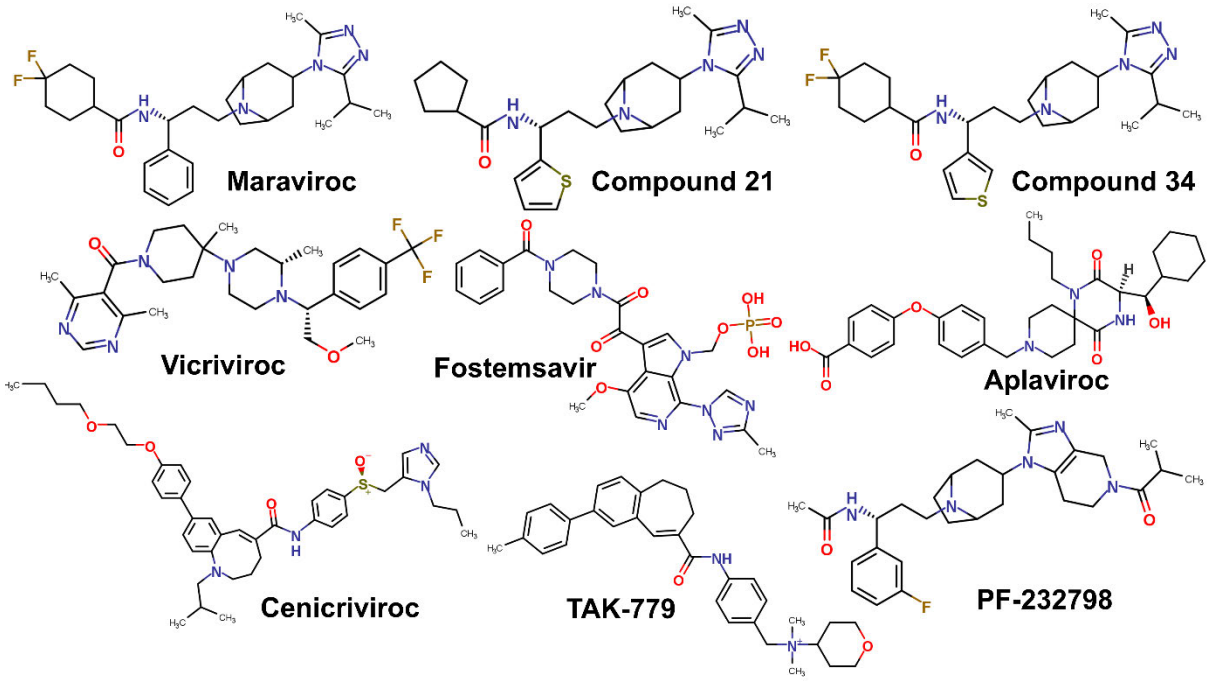


Figure 2.14 Representative approved drugs and clinical antagonists as HIV-1 entry inhibitors targeting CCR5 (Image prepared by author).

Aplaviroc (APL) and Vicriviroc (VCV) were two promising CCR5 small molecule antagonist but were both discontinued for varied reasons. Aplaviroc showed a substantial decrease of plasma HIV-1 RNA copies when administered in the first ten days of treatment (Lalezari et al.,

2005); however, development was discontinued due to reversible drug-induced hepatitis observed during phase II and III trials in five subjects (Nichols et al., 2008). Vicriviroc (VCV) also displayed substantial HIV-1 suppression but was terminated prior to phase II study due to increasing virologic failure rates compared with the control group (Gulick et al., 2007; Landovitz et al., 2008; Schürmann et al., 2007; Caseiro et al., 2012). Cenicriviroc (CVC) is another small molecule under development for inhibiting the CCR5 receptor in phase III trials (Klibanov et al., 2010). Cenicriviroc has a longer half-life than Maraviroc and shows a substantial decrease in plasma HIV-1 RNA load among HIV positive patients. Cenicriviroc further displayed favourable efficacy and safety in treating HIV-1 naive infected patients and has made progress to phase III trials (Thompson et al., 2016). Additionally, INCB009471 is also a CCR5 coreceptor investigational antagonist being developed as a selective HIV-1 entry inhibitor (Shin et al., 2011). INCB009471 showed limited adverse effect with no identified dose-limiting toxicity (Troy et al., 2007).

Genome editing techniques such as the clustered regularly interspaced short palindromic repeats (CRISPR), the zinc-finger nucleases (ZFN), and CRISPR associated nuclease 9 (CRISPR-Cas9) are emerging as an alternative approach in blocking CCR5 by deleting the CCR5 gene ex vivo. Recent results suggest that the concurrent gene modification of CXCR4 and CCR5 by CRISPR-Cas9 protect $CD4^+$ T cells from HIV-1 infection and holds the potential to provide a safe and effective functional cure for HIV-1 infection (Liu et al., 2017).

Recent structure-based optimisation of the functional and binding properties of Maraviroc lead to the design of novel 1-Heteroaryl-1, 3-Propanediamine analogues (Peng et al., 2018). The compounds (Compound 21 and Compound 34) have proven to be potential drug candidates with improved anti-HIV-1 activity and tolerable pharmacokinetic profile compared with Maraviroc (Peng et al., 2018). The binding of CCR5 inhibitors, including Maraviroc, Aplaviroc and Vicriviroc, has been considered an allosteric inhibitor (Arts & Hazuda, 2012). However, the recent crystallisation of the full-length CCR5 in complex with gp120 V3 loop suggests that Maraviroc blocks gp120 binding to CCR5 via direct competitive inhibition in contrast to earlier views of noncompetitive allosteric inhibition through conformational availability restriction (Figure 2.12) (Shaik et al., 2019).

References

Akimoto, M., Selvaratnam, R., McNicholl, E.T., Verma, G., Taylor, S.S. & Melacini, G. 2013. Signaling through dynamic linkers as revealed by PKA. *Proceedings of the National Academy of Sciences*, 110(35): 14231–14236.

Allen, J.A. & Roth, B.L. 2011. Strategies to discover unexpected targets for drugs active at G protein–coupled receptors. *Annual review of pharmacology and toxicology*, 51: 117–144.

Allers, K., Hütter, G., Hofmann, J., Loddenkemper, C., Rieger, K., Thiel, E. & Schneider, T. 2011. Evidence for the cure of HIV infection by CCR5Δ32/Δ32 stem cell transplantation. *Blood*, 117(10): 2791–2799.

Álvarez, Y., Pérez-Mañá, C., Torrens, M. & Farré, M. 2013. Antipsychotic drugs in cocaine dependence: a systematic review and meta-analysis. *Journal of substance abuse treatment*, 45(1): 1–10.

Andreoli, M., Tessari, M., Pilla, M., Valerio, E., Hagan, J.J. & Heidbreder, C.A. 2003. Selective antagonism at dopamine D 3 receptors prevents nicotine-triggered relapse to nicotine-seeking behavior. *Neuropsychopharmacology*, 28(7): 1272–1280.

Apel, A.-K., Cheng, R.K.Y., Tautermann, C.S., Brauchle, M., Huang, C.-Y., Pautsch, A., Hennig, M., Nar, H. & Schnapp, G. 2019. Crystal structure of CC chemokine receptor 2A in complex with an orthosteric antagonist provides insights for the design of selective antagonists. *Structure*, 27(3): 427–438.

Arts, E.J. & Hazuda, D.J. 2012. HIV-1 antiretroviral drug therapy. *Cold Spring Harbor perspectives in medicine*, 2(4): a007161.

Askew, D., Su, C.A., Barkauskas, D.S., Dorand, R.D., Myers, J., Liou, R., Nthale, J. & Huang, A.Y. 2016. Transient Surface CCR5 Expression by Naive CD8+ T Cells within Inflamed Lymph Nodes Is Dependent on High Endothelial Venule Interaction and Augments Th Cell–Dependent Memory Response. *The Journal of Immunology*, 196(9): 3653–3664.

Attwood, T.K. & Findlay, J.B.C. 1994. Fingerprinting G-protein-coupled receptors. *Protein Engineering, Design and Selection*, 7(2): 195–203.

Azam, S., Haque, M., Jakaria, M., Jo, S.-H., Kim, I.-S. & Choi, D.-K. 2020. G-Protein-Coupled Receptors in CNS: A Potential Therapeutic Target for Intervention in Neurodegenerative Disorders and Associated Cognitive Deficits. *Cells*, 9(2): 506.

Baik, J.-H. 2013. Dopamine signaling in reward-related behaviors. *Frontiers in neural circuits*, 7: 152.

Ballesteros, J.A. & Weinstein, H. 1995. [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In *Methods in neurosciences*. Elsevier: 366–428.

Ballon, J.S., Pajvani, U., Freyberg, Z., Leibel, R.L. & Lieberman, J.A. 2014. Molecular pathophysiology of metabolic effects of antipsychotic medications. *Trends in Endocrinology & Metabolism*, 25(11): 593–600.

Barst, R.J., Langleben, D., Frost, A., Horn, E.M., Oudiz, R., Shapiro, S., McLaughlin, V., Hill, N., Tapson, V.F. & Robbins, I.M. 2004. Sitaxsentan therapy for pulmonary arterial hypertension. *American journal of respiratory and critical care medicine*, 169(4): 441–447.

Beaulieu, J.-M. & Gainetdinov, R.R. 2011. The physiology, signaling, and pharmacology of

dopamine receptors. *Pharmacological reviews*, 63(1): 182–217.

Berger, E.A., Doms, R.W., Fenyö, E.-M., Korber, B.T.M., Littman, D.R., Moore, J.P., Sattentau, Q.J., Schuitemaker, H., Sodroski, J. & Weiss, R.A. 1998. A new classification for HIV-1. *Nature*, 391(6664): 240.

Berger, E.A., Murphy, P.M. & Farber, J.M. 1999. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Immunol Today*, 17(1): 657–700. https://doi.org/10.1146/annurev.immunol.17.1.657.

Berro, R., Klasse, P.J., Lascano, D., Flegler, A., Nagashima, K.A., Sanders, R.W., Sakmar, T.P., Hope, T.J. & Moore, J.P. 2011. Multiple CCR5 conformations on the cell surface are used differentially by human immunodeficiency viruses resistant or sensitive to CCR5 inhibitors. *Journal of virology*, 85(16): 8227–8240.

Bibb, J.A., Snyder, G.L., Nishi, A., Yan, Z., Meijer, L., Fienberg, A.A., Tsai, L.-H., Kwon, Y.T., Girault, J.-A. & Czernik, A.J. 1999. Phosphorylation of DARPP-32 by Cdk5 modulates dopamine signalling in neurons. *Nature*, 402(6762): 669–671.

Blanpain, C., Migeotte, I., Lee, B., Vakili, J., Doranz, B.J., Govaerts, C., Vassart, G., Doms, R.W. & Parmentier, M. 1999. CCR5 binds multiple CC-chemokines: MCP-3 acts as a natural antagonist. *Blood, The Journal of the American Society of Hematology*, 94(6): 1899–1905.

Brelot, A. & Chakrabarti, L.A. 2018. CCR5 revisited: how mechanisms of HIV entry govern AIDS pathogenesis. *Journal of molecular biology*, 430(17): 2557–2589.

Brooks, D.J. 2000. Dopamine agonists: their role in the treatment of Parkinson's disease.

Burg, J.S., Ingram, J.R., Venkatakrishnan, A.J., Jude, K.M., Dukkipati, A., Feinberg, E.N., Angelini, A., Waghray, D., Dror, R.O. & Ploegh, H.L. 2015. Structural basis for chemokine recognition and activation of a viral G protein–coupled receptor. *Science*, 347(6226): 1113–1117.

Caffrey, M. 2015. A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes. *Acta Crystallographica Section F: Structural Biology Communications*, 71(1): 3–18.

Caffrey, M. & Cherezov, V. 2009. Crystallizing membrane proteins using lipidic mesophases. *Nature protocols*, 4(5): 706–731.

Carlsson, J., Yoo, L., Gao, Z.-G., Irwin, J.J., Shoichet, B.K. & Jacobson, K.A. 2010. Structure-based discovery of A2A adenosine receptor ligands. *Journal of medicinal chemistry*, 53(9): 3748–3755.

Carter, P.H. 2013. Progress in the discovery of CC chemokine receptor 2 antagonists, 2009–2012. *Expert opinion on therapeutic patents*, 23(5): 549–568.

Caseiro, M.M., Nelson, M., Diaz, R.S., Gathe, J., de Andrade Neto, J.L., Slim, J., Solano, A., Netto, E.M., Mak, C., Shen, J., Greaves, W., Dunkle, L.M., Vilchez, R.A. & Zeinecker, J. 2012. Vicriviroc plus optimized background therapy for treatment-experienced subjects with CCR5 HIV-1 infection: Final results of two randomized phase III trials. *Journal of Infection*, 65(4): 326–335.

Cerchio, R. & Chen, S. 2020. Role of GPCRs in cancer. In *GPCRs*. Elsevier: 463–474.

Chan, D.C. & Kim, P.S. 1998. HIV entry and its inhibition. *Cell*, 93(5): 681–684.

Cherezov, V., Rosenbaum, D.M., Hanson, M.A., Rasmussen, S.G.F., Thian, F.S., Kobilka, T.S., Choi, H.-J., Kuhn, P., Weis, W.I. & Kobilka, B.K. 2007. High-resolution crystal

structure of an engineered human β2-adrenergic G protein–coupled receptor. *science*, 318(5854): 1258–1265.

Chien, E.Y.T., Liu, W., Zhao, Q., Katritch, V., Han, G.W., Hanson, M.A., Shi, L., Newman, A.H., Javitch, J.A. & Cherezov, V. 2010. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science*, 330(6007): 1091–1095.

Christopher, J.A., Aves, S.J., Bennett, K.A., Dore, A.S., Errey, J.C., Jazayeri, A., Marshall, F.H., Okrasa, K., Serrano-Vega, M.J. & Tehan, B.G. 2015. Fragment and structure-based drug discovery for a class C GPCR: discovery of the mGlu5 negative allosteric modulator HTL14242 (3-chloro-5-[6-(5-fluoropyridin-2-yl) pyrimidin-4-yl] benzonitrile). *Journal of medicinal chemistry*, 58(16): 6653–6664.

Chun, E., Thompson, A.A., Liu, W., Roth, C.B., Griffith, M.T., Katritch, V., Kunken, J., Xu, F., Cherezov, V. & Hanson, M.A. 2012. Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure*, 20(6): 967–976.

Combadiere, C., Ahuja, S.K., Lee Tiffany, H. & Murphy, P.M. 1996. Cloning and functional expression of CC CKR5, a human monocyte CC chemokine receptor selective for MIP-1α, MIP-1β, and RANTES. *Journal of leukocyte biology*, 60(1): 147–152.

Congreve, M. & Marshall, F. 2010. The impact of GPCR structures on pharmacology and structure-based drug design. *British journal of pharmacology*, 159(5): 986–996.

Cormier, E.G. & Dragic, T. 2002. The crown and stem of the V3 loop play distinct roles in human immunodeficiency virus type 1 envelope glycoprotein interactions with the CCR5 coreceptor. *Journal of virology*, 76(17): 8953–8957.

Dal Toso, R., Sommer, B., Ewert, M., Herb, A., Pritchett, D.B., Bach, A., Shivers, B.D. & Seeburg, P.H. 1989. The dopamine D2 receptor: two molecular forms generated by alternative splicing. *The EMBO journal*, 8(13): 4025–4034.

Deng, H., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhart, M., Marzio, P. Di, Marmon, S., Sutton, R.E. & Hill, C.M. 1996. Identification of a major co-receptor for primary isolates of HIV-1. *Nature*, 381(6584): 661–666.

Deupi, X., Standfuss, J. & Schertler, G. 2012. Conserved activation pathways in G-protein-coupled receptors.

Elemam, N.M., Hannawi, S. & Maghazachi, A.A. 2020. Role of Chemokines and Chemokine Receptors in Rheumatoid Arthritis. *ImmunoTargets and Therapy*, 9: 43.

Falkenhagen, A. & Joshi, S. 2018. HIV entry and its inhibition by bifunctional antiviral proteins. *Molecular Therapy-Nucleic Acids*, 13: 347–364.

Fan, L., Tan, L., Chen, Z., Qi, J., Nie, F., Luo, Z., Cheng, J. & Wang, S. 2020. Haloperidol bound D 2 dopamine receptor structure inspired the discovery of subtype selective ligands. *Nature communications*, 11(1): 1–11.

FDA, U.S. 2020. FDA Approves New HIV Treatment for Patients With Limited Treatment Options. *FDA NEWS RELEASE*. https://www.fda.gov/news-events/press-announcements/fda-approves-new-hiv-treatment-patients-limited-treatment-options 19 October 2020.

FDA, U.S. 2007. FDA notifications. Maraviroc approved as a CCR5 co-receptor antagonist. *AIDS Alert*, 22: 103.

Fredriksson, R., Lagerström, M.C., Lundin, L.-G. & Schiöth, H.B. 2003. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis,

paralogon groups, and fingerprints. *Molecular pharmacology*, 63(6): 1256–1272.

Gad, A.A. & Balenga, N. 2020. The Emerging Role of Adhesion GPCRs in Cancer. *ACS Pharmacology & Translational Science*, 3(1): 29–42.

Galaj, E., Manuszak, M., Babic, S., Ananthan, S. & Ranaldi, R. 2015. The selective dopamine D3 receptor antagonist, SR 21502, reduces cue-induced reinstatement of heroin seeking and heroin conditioned place preference in rats. *Drug and alcohol dependence*, 156: 228–233.

García-Nafría, J. & Tate, C.G. 2020. Cryo-electron microscopy: Moving beyond X-ray crystal structures for drug receptors and drug development. *Annual Review of Pharmacology and Toxicology*, 60: 51–71.

Garcia-Perez, J., Staropoli, I., Azoulay, S., Heinrich, J.-T., Cascajero, A., Colin, P., Lortat-Jacob, H., Arenzana-Seisdedos, F., Alcami, J. & Kellenberger, E. 2015. A single-residue change in the HIV-1 V3 loop associated with maraviroc resistance impairs CCR5 binding affinity while increasing replicative capacity. *Retrovirology*, 12(1): 1–20.

Garland, S.L. 2013. Are GPCRs still a source of new targets? *Journal of biomolecular screening*, 18(9): 947–966.

Ghosh, E., Kumari, P., Jaiman, D. & Shukla, A.K. 2015. Methodological advances: the unsung heroes of the GPCR structural revolution. *Nature Reviews Molecular Cell Biology*, 16(2): 69–81.

Giros, B., Sokoloff, P., Martres, M.-P., Riou, J.-F., Emorine, L.J. & Schwartz, J.-C. 1989. Alternative splicing directs the expression of two D2 dopamine receptor isoforms. *Nature*, 342(6252): 923–926.

Gomes, P. & Soares-da-Silva, P. 2002. Na+/H+ exchanger activity and dopamine D1-like receptor function in two opossum kidney cell clonal sublines. *Cellular Physiology and Biochemistry*, 12(5–6): 259–268.

Griffith, J.W., Sokol, C.L. & Luster, A.D. 2014. Chemokines and chemokine receptors: positioning cells for host defense and immunity. *Annual review of immunology*, 32: 659–702.

Guimarães, T.R. & Thathiah, A. 2020. G protein-coupled receptors in neurodegenerative diseases. In *GPCRs*. Elsevier: 335–365.

Gulick, R.M., Su, Z., Flexner, C., Hughes, M.D., Skolnik, P.R., Wilkin, T.J., Gross, R., Krambrink, A., Coakley, E., Greaves, W.L., Zolopa, A., Reichman, R., Godfrey, C., Hirsch, M. & Kuritzkes, D.R. 2007. Phase 2 Study of the Safety and Efficacy of Vicriviroc, a CCR5 Inhibitor, in HIV-1–Infected, Treatment-Experienced Patients: AIDS Clinical Trials Group 5211. *The Journal of Infectious Diseases*.

Gupta, M.K. & Vasudevan, N.T. 2019. GPCRs and insulin receptor signaling in conversation: Novel avenues for drug discovery. *Current topics in medicinal chemistry*, 19(16): 1436–1444.

Hauser, A.S., Attwood, M.M., Rask-Andersen, M., Schiöth, H.B. & Gloriam, D.E. 2017. Trends in GPCR drug discovery: new agents, targets and indications. *Nature reviews Drug discovery*, 16(12): 829–842.

Hauser, A.S., Chavali, S., Masuho, I., Jahn, L.J., Martemyanov, K.A., Gloriam, D.E. & Babu, M.M. 2018. Pharmacogenomics of GPCR drug targets. *Cell*, 172(1–2): 41–54.

Heidbreder, C.A. & Newman, A.H. 2010. Current perspectives on selective dopamine D3

receptor antagonists as pharmacotherapeutics for addictions and related disorders. *Annals of the New York Academy of Sciences*, 1187: 4.

Heifetz, A., Morao, I., Babu, M.M., James, T., Southey, M.W.Y., Fedorov, D.G., Aldeghi, M., Bodkin, M.J. & Townsend-Nicholson, A. 2020. Characterizing Interhelical Interactions of G-Protein Coupled Receptors with the Fragment Molecular Orbital Method. *Journal of chemical theory and computation*, 16(4): 2814–2824.

Henrich, T.J. & Kuritzkes, D.R. 2013. HIV-1 entry inhibitors: recent development and clinical use. *Current opinion in virology*, 3(1): 51–57.

Higley, A.E., Kiefer, S.W., Li, X., Gaál, J., Xi, Z.-X. & Gardner, E.L. 2011. Dopamine D3 receptor antagonist SB-277011A inhibits methamphetamine self-administration and methamphetamine-induced reinstatement of drug-seeking in rats. *European journal of pharmacology*, 659(2–3): 187–192.

Hollingsworth, S.A. & Dror, R.O. 2018. Molecular dynamics simulation for all. *Neuron*, 99(6): 1129–1143.

Holmes, A., Lachowicz, J.E. & Sibley, D.R. 2004. Phenotypic analysis of dopamine receptor knockout mice; recent insights into the functional specificity of dopamine receptor subtypes. *Neuropharmacology*, 47(8): 1117–1134.

Huang, Y., Todd, N. & Thathiah, A. 2017. The role of GPCRs in neurodegenerative diseases: avenues for therapeutic intervention. *Current opinion in pharmacology*, 32: 96–110.

Hughes, C.E. & Nibbs, R.J.B. 2018. A guide to chemokines and their receptors. *The FEBS journal*, 285(16): 2944–2971.

Isberg, V., Mordalski, S., Munk, C., Rataj, K., Harpsøe, K., Hauser, A.S., Vroling, B., Bojarski, A.J., Vriend, G. & Gloriam, D.E. 2016. GPCRdb: an information system for G protein-coupled receptors. *Nucleic acids research*, 44(D1): D356–D364.

Jacoby, E., Bouhelal, R., Gerspacher, M. & Seuwen, K. 2006. The 7 TM G-protein-coupled receptor target family. *ChemMedChem: Chemistry Enabling Drug Discovery*, 1(8): 760–782.

Jaeger, K., Bruenle, S., Weinert, T., Guba, W., Muehle, J., Miyazaki, T., Weber, M., Furrer, A., Haenggi, N. & Tetaz, T. 2019. Structural basis for allosteric ligand recognition in the human CC chemokine receptor 7. *Cell*, 178(5): 1222–1230.

Ji, C., Zhang, J., Dioszegi, M., Chiu, S., Rao, E., derosier, A., Cammack, N., Brandt, M. & Sankuratri, S. 2007. CCR5 small-molecule antagonists and monoclonal antibodies exert potent synergistic antiviral effects by cobinding to the receptor. *Molecular pharmacology*, 72(1): 18–28.

Kaar, S.J., Natesan, S., McCutcheon, R. & Howes, O.D. 2020. Antipsychotics: mechanisms underlying clinical response and side-effects and novel treatment approaches based on pathophysiology. *Neuropharmacology*, 172: 107704.

Kang, D.S., Tian, X. & Benovic, J.L. 2014. Role of β-arrestins and arrestin domain-containing proteins in G protein-coupled receptor trafficking. *Current opinion in cell biology*, 27: 63–71.

Katritch, V., Jaakola, V.-P., Lane, J.R., Lin, J., IJzerman, A.P., Yeager, M., Kufareva, I., Stevens, R.C. & Abagyan, R. 2010. Structure-based discovery of novel chemotypes for adenosine A2A receptor antagonists. *Journal of medicinal chemistry*, 53(4): 1799–1809.

Klein, M.O., Battagello, D.S., Cardoso, A.R., Hauser, D.N., Bittencourt, J.C. & Correa, R.G.

2019. Dopamine: functions, signaling, and association with neurological diseases. *Cellular and molecular neurobiology*, 39(1): 31–59.

Klibanov, O.M., Williams, S.H. & Iler, C.A. 2010. Cenicriviroc, an orally active CCR5 antagonist for the potential treatment of HIV infection. *Curr Opin Investig Drugs*, 11(8): 940–950.

Kolakowski Jr, L.F. 1994. GCRDb: a G-protein-coupled receptor database. *Receptors & channels*, 2(1): 1.

Kolb, P., Rosenbaum, D.M., Irwin, J.J., Fung, J.J., Kobilka, B.K. & Shoichet, B.K. 2009. Structure-based discovery of beta2-adrenergic receptor ligands. *Proceedings of the National Academy of Sciences of the United States of America*, 106(16): 6843–6848. https://pubmed.ncbi.nlm.nih.gov/19342484.

Kotake, T., Usami, M., Akaza, H., Koiso, K., Homtna, Y., Kawabe, K., Aso, Y., Orikasa, S., Shimazaki, J. & Isaka, S. 1999. Goserelin acetate with or without antiandrogen or estrogen in the treatment of patients with advanced prostate cancer: a multicenter, randomized, controlled trial in Japan. *Japanese journal of clinical oncology*, 29(11): 562–570.

Kozal, M., Aberg, J., Pialoux, G., Cahn, P., Thompson, M., Molina, J.-M., Grinsztejn, B., Diaz, R., Castagna, A. & Kumar, P. 2020. Fostemsavir in Adults with Multidrug-Resistant HIV-1 Infection. *New England Journal of Medicine*, 382(13): 1232–1243.

Kraneveld, A.D., Braber, S., Overbeek, S., de Kruijf, P., Koelink, P. & Smit, M.J. 2010. Chemokine receptors in inflammatory diseases. *Chemokine Receptors as Drug Targets*, 46: 105–150.

Lalezari, J., Thompson, M., Kumar, P., Piliero, P., Davey, R., Patterson, K., Shachoy-Clark, A., Adkison, K., Demarest, J., Lou, Y., Berrey, M. & Piscitelli, S. 2005. Antiviral activity and safety of 873140, a novel CCR5 antagonist, during short-term monotherapy in HIV-infected adults. *AIDS*, 19(14): 1443–1448.

Lalezari, J., Yadavalli, G.K., Para, M., Richmond, G., DeJesus, E., Brown, S.J., Cai, W., Chen, C., Zhong, J. & Novello, L.A. 2008. Safety, Pharmacokinetics, and Antiviral Activity of HGS004, a Novel Fully Human IgG4 Monoclonal Antibody against CCR5, in HIV-1—zInfected Patients. *The Journal of infectious diseases*, 197(5): 721–727.

Lalezari, J.P., Latiff, G.H., Brinson, C., Echevarría, J., Treviño-Pérez, S., Bogner, J.R., Thompson, M., Fourie, J., Pena, O.A.S. & Urbina, F.C.M. 2015. Safety and efficacy of the HIV-1 attachment inhibitor prodrug BMS-663068 in treatment-experienced individuals: 24 week results of AI438011, a phase 2b, randomised controlled trial. *The lancet HIV*, 2(10): e427–e437.

Landovitz, R.J., Angel, J.B., Hoffmann, C., Horst, H., Opravil, M., Long, J., Greaves, W. & Fätkenheuer, G. 2008. Phase II Study of Vicriviroc versus Efavirenz (both with Zidovudine/Lamivudine) in Treatment-Naive Subjects with HIV-1 Infection. *The Journal of Infectious Diseases*, 198(8): 1113–1122.

Latorraca, N.R., Venkatakrishnan, A.J. & Dror, R.O. 2017. GPCR dynamics: structures in motion. *Chemical reviews*, 117(1): 139–155.

Li, P., L Snyder, G. & E Vanover, K. 2016. Dopamine targeting drugs for the treatment of schizophrenia: past, present and future. *Current topics in medicinal chemistry*, 16(29): 3385–3403.

Liang, Y.-L., Khoshouei, M., Radjainia, M., Zhang, Y., Glukhova, A., Tarrasch, J., Thal, D.M.,

Furness, S.G.B., Christopoulos, G. & Coudrat, T. 2017. Phase-plate cryo-EM structure of a class B GPCR–G-protein complex. *Nature*, 546(7656): 118–123.

Lieberman, J.A., Stroup, T.S., McEvoy, J.P., Swartz, M.S., Rosenheck, R.A., Perkins, D.O., Keefe, R.S.E., Davis, S.M., Davis, C.E. & Lebowitz, B.D. 2005. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *New England journal of medicine*, 353(12): 1209–1223.

Lin, H.-H., Hsiao, C.-C., Pabst, C., Hébert, J., Schöneberg, T. & Hamann, J. 2017. Adhesion GPCRs in regulating immune responses and inflammation. In *Advances in immunology*. Elsevier: 163–201.

Liu, J., Bartesaghi, A., Borgnia, M.J., Sapiro, G. & Subramaniam, S. 2008. Molecular architecture of native HIV-1 gp120 trimers. *Nature*, 455(7209): 109–113.

Liu, R., Paxton, W.A., Choe, S., Ceradini, D., Martin, S.R., Horuk, R., MacDonald, M.E., Stuhlmann, H., Koup, R.A. & Landau, N.R. 1996. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell*, 86(3): 367–377.

Liu, W., Ishchenko, A. & Cherezov, V. 2014. Preparation of microcrystals in lipidic cubic phase for serial femtosecond crystallography. *Nature protocols*, 9(9): 2123–2134.

Liu, Z., Chen, S., Jin, X., Wang, Q., Yang, K., Li, C., Xiao, Q., Hou, P., Liu, S. & Wu, S. 2017. Genome editing of the HIV co-receptors CCR5 and CXCR4 by CRISPR-Cas9 protects CD4+ T cells from HIV-1 infection. *Cell & bioscience*, 7(1): 47.

Luttrell, L.M. & Lefkowitz, R.J. 2002. The role of β-arrestins in the termination and transduction of G-protein-coupled receptor signals. *Journal of cell science*, 115(3): 455–465.

Maggio, R., Scarselli, M., Capannolo, M. & Millan, M.J. 2015. Novel dimensions of D3 receptor function: Focus on heterodimerisation, transactivation and allosteric modulation. *European Neuropsychopharmacology*, 25(9): 1470–1479.

Manglik, A., Lin, H., Aryal, D.K., McCorvy, J.D., Dengler, D., Corder, G., Levit, A., Kling, R.C., Bernat, V. & Hübner, H. 2016. Structure-based discovery of opioid analgesics with reduced side effects. *Nature*, 537(7619): 185–190.

Miles, T.F., Spiess, K., Jude, K.M., Tsutsumi, N., Burg, J.S., Ingram, J.R., Waghray, D., Hjorto, G.M., Larsen, O. & Ploegh, H.L. 2018. Viral GPCR US28 can signal in response to chemokine agonists of nearly unlimited structural degeneracy. *Elife*, 7: e35850.

Missale, C., Nash, S.R., Robinson, S.W., Jaber, M. & Caron, M.G. 1998. Dopamine receptors: from structure to function. *Physiological reviews*.

Mollica Poeta, V., Massara, M., Capucetti, A. & Bonecchi, R. 2019. Chemokines and chemokine receptors: new targets for cancer immunotherapy. *Frontiers in immunology*, 10: 379.

Moritz, A.E., Free, R.B. & Sibley, D.R. 2018. Advances and challenges in the search for D2 and D3 dopamine receptor-selective compounds. *Cellular signalling*, 41: 75–81.

Murphy, P.M. 2002. International Union of Pharmacology. XXX. Update on chemokine receptor nomenclature. *Pharmacological reviews*, 54(2): 227–229.

Nagasawa, T., Hirota, S., Tachibana, K., Takakura, N., Nishikawa, S., Kitamura, Y., Yoshida, N., Kikutani, H. & Kishimoto, T. 1996. Defects of B-cell lymphopoiesis and bone-marrow myelopoiesis in mice lacking the CXC chemokine PBSF/SDF-1. *Nature*, 382(6592): 635–

638.
Nichols, W.G., Steel, H.M., Bonny, T., Adkison, K., Curtis, L., Millard, J., Kabeya, K. & Clumeck, N. 2008. Hepatotoxicity observed in clinical trials of aplaviroc (GW873140). *Antimicrobial Agents and Chemotherapy*, 52(3): 858–865.
Nishi, A., Fisone, G., Snyder, G.L., Dulubova, I., Aperia, A., Nairn, A.C. & Greengard, P. 1999. Regulation of Na+, K+-ATPase isoforms in rat neostriatum by dopamine and protein kinase C. *Journal of neurochemistry*, 73(4): 1492–1501.
Nishi, A., Snyder, G.L., Fienberg, A.A., Fisone, G., Aperia, A., Nairn, A.C. & Greengard, P. 1999. Requirement for DARPP-32 in mediating effect of dopamine D2 receptor activation. *European Journal of Neuroscience*, 11(7): 2589–2592.
Novick, D., Haro, J.M., Suarez, D., Perez, V., Dittmann, R.W. & Haddad, P.M. 2010. Predictors and clinical consequences of non-adherence with antipsychotic medication in the outpatient treatment of schizophrenia. *Psychiatry research*, 176(2–3): 109–113.
Odoemelam, C.S., Percival, B., Wallis, H., Chang, M.-W., Ahmad, Z., Scholey, D., Burton, E., Williams, I.H., Kamerlin, C.L. & Wilson, P.B. 2020. G-Protein coupled receptors: structure and function in drug discovery. *RSC Advances*, 10(60): 36337–36348.
Onuffer, J.J. & Horuk, R. 2002. Chemokines, chemokine receptors and small-molecule antagonists: recent developments. *Trends in pharmacological sciences*, 23(10): 459–467.
Oswald, C., Rappas, M., Kean, J., Doré, A.S., Errey, J.C., Bennett, K., Deflorian, F., Christopher, J.A., Jazayeri, A., Mason, J.S., Congreve, M., Cooke, R.M. & Marshall, F.H. 2016. Intracellular allosteric antagonism of the CCR9 receptor. *Nature*, 540(7633): 462–465. https://doi.org/10.1038/nature20606.
Palczewski, K., Kumasaka, T., Hori, T., Behnke, C.A., Motoshima, H., Fox, B.A., Le Trong, I., Teller, D.C., Okada, T. & Stenkamp, R.E. 2000. Crystal structure of rhodopsin: AG protein-coupled receptor. *science*, 289(5480): 739–745.
Park, S.H., Das, B.B., Casagrande, F., Tian, Y., Nothnagel, H.J., Chu, M., Kiefer, H., Maier, K., De Angelis, A.A. & Marassi, F.M. 2012. Structure of the chemokine receptor CXCR1 in phospholipid bilayers. *Nature*, 491(7426): 779–783.
Paxton, W.A., Liu, R., Kang, S., Wu, L., Gingeras, T.R., Landau, N.R., Mackay, C.R. & Koup, R.A. 1998. Reduced HIV-1 infectability of CD4+ lymphocytes from exposed-uninfected individuals: association with low expression of CCR5 and high production of β-chemokines. *Virology*, 244(1): 66–73.
Peng, P., Chen, H., Zhu, Y., Wang, Z., Li, J., Luo, R.-H., Wang, J., Chen, L., Yang, L.-M. & Jiang, H. 2018. Structure-Based Design of 1-Heteroaryl-1, 3-propanediamine Derivatives as a Novel Series of CC-Chemokine Receptor 5 Antagonists. *Journal of medicinal chemistry*, 61(21): 9621–9636.
Qin, L., Kufareva, I., Holden, L.G., Wang, C., Zheng, Y., Zhao, C., Fenalti, G., Wu, H., Han, G.W. & Cherezov, V. 2015. Crystal structure of the chemokine receptor CXCR4 in complex with a viral chemokine. *Science*, 347(6226): 1117–1122.
Qu, X., Wang, D. & Wu, B. 2020. Progress in GPCR structure determination. In *GPCRs*. Elsevier: 3–22.
Rangel-Barajas, C., Coronel, I. & Florán, B. 2015. Dopamine receptors and neurodegeneration. *Aging and disease*, 6(5): 349.
Rask-Andersen, M., Masuram, S. & Schiöth, H.B. 2014. The druggable genome: evaluation of

drug targets in clinical trials suggests major shifts in molecular class and indication. *Annual review of pharmacology and toxicology*, 54: 9–26.

Rasmussen, S.G.F., Choi, H.-J., Rosenbaum, D.M., Kobilka, T.S., Thian, F.S., Edwards, P.C., Burghammer, M., Ratnala, V.R.P., Sanishvili, R. & Fischetti, R.F. 2007. Crystal structure of the human β 2 adrenergic G-protein-coupled receptor. *Nature*, 450(7168): 383–387.

Reiter, E., Ahn, S., Shukla, A.K. & Lefkowitz, R.J. 2012. Molecular mechanism of β-arrestin-biased agonism at seven-transmembrane receptors. *Annual review of pharmacology and toxicology*, 52: 179–197.

Reiter, E. & Lefkowitz, R.J. 2006. GRKs and β-arrestins: roles in receptor silencing, trafficking and signaling. *Trends in endocrinology & metabolism*, 17(4): 159–165.

Reynes, J., Portales, P., Segondy, M., Baillat, V., André, P., Réant, B., Avinens, O., Couderc, G., Benkirane, M. & Clot, J. 2000. CD4+ T cell surface CCR5 density as a determining factor of virus load in persons infected with human immunodeficiency virus type 1. *The Journal of infectious diseases*, 181(3): 927–932.

Riddy, D.M., Delerive, P., Summers, R.J., Sexton, P.M. & Langmead, C.J. 2018. G protein–coupled receptors targeting insulin resistance, obesity, and type 2 diabetes mellitus. *Pharmacological reviews*, 70(1): 39–67.

de Roda Husman, A.-M., Koot, M., Cornelissen, M., Keet, I.P.M., Brouwer, M., Broersen, S.M., Bakker, M., Roos, M.T.L., Prins, M. & de Wolf, F. 1997. Association between CCR5 genotype and the clinical course of HIV-1 infection. *Annals of internal medicine*, 127(10): 882–890.

Rodríguez, D., Brea, J., Loza, M.I. & Carlsson, J. 2014. Structure-based discovery of selective serotonin 5-HT1B receptor ligands. *Structure*, 22(8): 1140–1151.

Rovati, G.E., Capra, V. & Neubig, R.R. 2007. The highly conserved DRY motif of class AG protein-coupled receptors: beyond the ground state. *Molecular pharmacology*, 71(4): 959–964.

S Latinovic, O., Zhang, J., Tagaya, Y., L DeVico, A., R Fouts, T., Schneider, K., R Lakowicz, J., Heredia, A. & R Redfield, R. 2016. Synergistic inhibition of R5 HIV-1 by the fusion protein (FLSC) IgG1 Fc and Maraviroc in primary cells: Implications for prevention and treatment. *Current HIV research*, 14(1): 24–36.

Samson, M., Labbe, O., Mollereau, C., Vassart, G. & Parmentier, M. 1996. Molecular cloning and functional expression of a new human CC-chemokine receptor gene. *Biochemistry*, 35(11): 3362–3367.

Santos, R., Ursu, O., Gaulton, A., Bento, A.P., Donadi, R.S., Bologa, C.G., Karlsson, A., Al-Lazikani, B., Hersey, A. & Oprea, T.I. 2017. A comprehensive map of molecular drug targets. *Nature reviews Drug discovery*, 16(1): 19–34.

Schiöth, H.B. & Fredriksson, R. 2005. The GRAFS classification system of G-protein coupled receptors in comparative perspective. *General and comparative endocrinology*, 142(1–2): 94–101.

Schlecht, H.P., Schellhorn, S., Dezube, B.J. & Jacobson, J.M. 2008. New approaches in the treatment of HIV/AIDS–focus on maraviroc and other CCR5 antagonists. *Therapeutics and clinical risk management*, 4(2): 473.

Schürmann, D., Fätkenheuer, G., Reynes, J., Michelet, C., Raffi, F., Van Lier, J., Caceres, M., Keung, A., Sansone-Parsons, A., Dunkle, L.M. & Hoffmann, C. 2007. Antiviral activity,

pharmacokinetics and safety of vicriviroc, an oral CCR5 antagonist, during 14-day monotherapy in HIV-infected adults. *AIDS*, 21(10): 1293–1299.

Schwab, R.S. 1951. Apomorphine in Parkinson's disease. *Trans Am Neurol Assoc*, 76: 251–253.

Sebastiani, G., Ceccarelli, E., Castagna, M.G. & Dotta, F. 2018. G-protein-coupled receptors (GPCRs) in the treatment of diabetes: Current view and future perspectives. *Best Practice & Research Clinical Endocrinology & Metabolism*, 32(2): 201–213.

Shaik, M.M., Peng, H., Lu, J., Rits-Volloch, S., Xu, C., Liao, M. & Chen, B. 2019. Structural basis of coreceptor recognition by HIV-1 envelope spike. *Nature*, 565(7739): 318.

Shin, N., Solomon, K., Zhou, N., Wang, K.H., Garlapati, V., Thomas, B., Li, Y., Covington, M., Baribaud, F. & Erickson-Viitanen, S. 2011. Identification and characterization of INCB9471, an allosteric noncompetitive small-molecule antagonist of CC chemokine receptor 5 with potent inhibitory activity against monocyte migration and HIV-1 infection. *Journal of Pharmacology and Experimental Therapeutics*, 338(1): 228–239.

Shioda, N. 2017. Dopamine D2L receptor-interacting proteins regulate dopaminergic signaling. *Journal of Pharmacological Sciences*, 135(2): 51–54.

Sriram, K. & Insel, P.A. 2018. G protein-coupled receptors as targets for approved drugs: how many targets and how many drugs? *Molecular pharmacology*, 93(4): 251–258.

Sunahara, R.K. & Taussig, R. 2002. Isoforms of mammalian adenylyl cyclase: multiplicities of signaling. *Molecular interventions*, 2(3): 168.

Tan, Q., Zhu, Y., Li, Jian, Chen, Z., Han, G.W., Kufareva, I., Li, T., Ma, L., Fenalti, G., Li, Jing, Zhang, W., Xie, X., Yang, H., Jiang, H., Cherezov, V., Liu, H., Stevens, R.C., Zhao, Q. & Wu, B. 2013. Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science*, 341(6152): 1387–1390.

Tate, C.G. 2012. A crystal clear solution for determining G-protein-coupled receptor structures. *Trends in biochemical sciences*, 37(9): 343–352.

Thal, D.M., Glukhova, A., Sexton, P.M. & Christopoulos, A. 2018. Structural insights into G-protein-coupled receptor allostery. *Nature*, 559(7712): 45–53.

Thal, D.M., Vuckovic, Z., Draper-Joyce, C.J., Liang, Y.-L., Glukhova, A., Christopoulos, A. & Sexton, P.M. 2018. Recent advances in the determination of G protein-coupled receptor structures. *Current opinion in structural biology*, 51: 28–34.

Thompson, M., Lalezari, J.P., Kaplan, R., Pinedo, Y., Pena, O.A.S., Cahn, P., Stock, D.A., Joshi, S.R., Hanna, G.J. & Lataillade, M. 2017. Safety and efficacy of the HIV-1 attachment inhibitor prodrug fostemsavir in antiretroviral-experienced subjects: week 48 analysis of AI438011, a Phase IIb, randomized controlled trial. *Antivir Ther*, 22(3): 215–223.

Thompson, M., Saag, M., DeJesus, E., Gathe, J., Lalezari, J., Landay, A.L., Cade, J., Enejosa, J., Lefebvre, E. & Feinberg, J. 2016. A 48-week randomized phase 2b study evaluating cenicriviroc versus efavirenz in treatment-naive HIV-infected adults with CC chemokine receptor type 5-tropic virus. *AIDS (London, England)*, 30(6): 869.

Tian, X., Kang, D.S. & Benovic, J.L. 2014. β-arrestins and G protein-coupled receptor trafficking. In *Arrestins-Pharmacology and Therapeutic Potential*. Springer: 173–186.

Tran, E.E.H., Borgnia, M.J., Kuybeda, O., Schauder, D.M., Bartesaghi, A., Frank, G.A., Sapiro, G., Milne, J.L.S. & Subramaniam, S. 2012. Structural mechanism of trimeric HIV-

1 envelope glycoprotein activation. *PLoS Pathog*, 8(7): e1002797.

Tritsch, N.X. & Sabatini, B.L. 2012. Dopaminergic modulation of synaptic transmission in cortex and striatum. *Neuron*, 76(1): 33–50.

Troy, S., Emm, T., Yeleswaram, S., Sides, E., Tayyeb, M., Solomon, K., Erickson-Viitanen, S., Levy, R. & Williams, W. 2007. Single and Multiple Dose Pharmacokinetics of INCB9471, a Potent Antagonist of CCR5 Co-Receptor. *47th Annual ICAAC*.

Tyndall, J.D.A. & Sandilya, R. 2005. GPCR agonists and antagonists in the clinic. *Medicinal chemistry*, 1(4): 405–421.

Undieh, A.S. 2010. Pharmacology of signaling induced by dopamine D1-like receptor activation. *Pharmacology & therapeutics*, 128(1): 37–60.

Vallone, D., Picetti, R. & Borrelli, E. 2000. Structure and function of dopamine receptors. *Neuroscience & biobehavioral reviews*, 24(1): 125–132.

Wang, J., Gareri, C. & Rockman, H.A. 2018. G-protein–coupled receptors in heart disease. *Circulation research*, 123(6): 716–735.

Wang, S., Che, T., Levit, A., Shoichet, B.K., Wacker, D. & Roth, B.L. 2018. Structure of the D2 dopamine receptor bound to the atypical antipsychotic drug risperidone. *Nature*, 555: 269. http://dx.doi.org/10.1038/nature25758.

Wang, S., Wacker, D., Levit, A., Che, T., Betz, R.M., McCorvy, J.D., Venkatakrishnan, A.J., Huang, X.-P., Dror, R.O. & Shoichet, B.K. 2017. D4 dopamine receptor high-resolution structures enable the discovery of selective agonists. *Science*, 358(6361): 381–386.

Westby, M., Smith-Burchnell, C., Mori, J., Lewis, M., Mosley, M., Stockdale, M., Dorr, P., Ciaramella, G. & Perros, M. 2007. Reduced maximal inhibition in phenotypic susceptibility assays indicates that viral strains resistant to the CCR5 antagonist maraviroc utilize inhibitor-bound receptor for entry. *Journal of virology*, 81(5): 2359–2371.

Wilen, C.B., Tilton, J.C. & Doms, R.W. 2012a. HIV: cell binding and entry. *Cold Spring Harbor perspectives in medicine*, 2(8): a006866.

Wilen, C.B., Tilton, J.C. & Doms, R.W. 2012b. Molecular mechanisms of HIV entry. In *Viral Molecular Machines*. Springer: 223–242.

Wu, B., Chien, E.Y.T., Mol, C.D., Fenalti, G., Liu, W., Katritch, V., Abagyan, R., Brooun, A., Wells, P. & Bi, F.C. 2010. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science*, 330(6007): 1066–1071.

Yakar, R. & Akten, E.D. 2014. Discovery of high affinity ligands for β2-adrenergic receptor through pharmacophore-based high-throughput virtual screening and docking. *Journal of Molecular Graphics and Modelling*, 53: 148–160.

Yin, J., Chen, K.-Y.M., Clark, M.J., Hijazi, M., Kumari, P., Bai, X., Sunahara, R.K., Barth, P. & Rosenbaum, D.M. 2020. Structure of a D2 dopamine receptor–G-protein complex in a lipid membrane. *Nature*: 1–5.

Zheng, Y., Han, G.W., Abagyan, R., Wu, B., Stevens, R.C., Cherezov, V., Kufareva, I. & Handel, T.M. 2017. Structure of CC chemokine receptor 5 with a potent chemokine antagonist reveals mechanisms of chemokine recognition and molecular mimicry by HIV. *Immunity*, 46(6): 1005–1017.

Zheng, Y., Qin, L., Zacarías, N.V.O., de Vries, H., Han, G.W., Gustavsson, M., Dabros, M., Zhao, C., Cherney, R.J. & Carter, P. 2016. Structure of CC chemokine receptor 2 with orthosteric and allosteric antagonists. *Nature*, 540(7633): 458–461.

Zhou, Y., Cao, C., He, L., Wang, X. & Zhang, X.C. 2019. Crystal structure of dopamine receptor D4 bound to the subtype selective ligand, L745870. *Elife*, 8: e48822.

Żuk, J., Bartuzi, D., Matosiuk, D. & Kaczor, A.A. 2020. Preferential Coupling of Dopamine D2S and D2L Receptor Isoforms with Gi1 and Gi2 Proteins—In Silico Study. *International journal of molecular sciences*, 21(2): 436.

CHAPTER 3

Membrane-Protein Simulations and Computational Methods in Drug Discovery

3.1 Introduction

The traditional approach to drug discovery is costly, time-consuming, risky, and based on multidisciplinary approaches in producing effective and safer medicines. Computational drug design and discovery, and biomolecular simulation techniques, are used widely in modern drug discovery to study drugs and bioactive compounds interactions with their biological targets. As powerful tools for studying receptor-inhibitor interactions, structure-activity relationships (SAR) and conformational dynamics of complex molecular systems, pharmaceutical research has increasingly used modern medicinal chemistry methods, including computational and molecular modelling. Also, *in silico* predictions of physicochemical properties (e.g., molecular weight, polar surface area, solubility, hydrogen bond donor and acceptor, rotatable bonds etc.), and pharmacokinetic properties (absorption, distribution, metabolism, excretion and toxicity) have significantly contributed to the success rate of modern drug discovery.

3.2 Molecular Mechanics (force fields) in Biomolecular Simulation of GPCRs

The application of quantum mechanical (QM) approaches, though more accurate, remains suitable only for performing computations on small systems consisting of few atoms (Steinbrecher & Elstner, 2013). The investigation of the structure and dynamics of a biomolecular system of interest comes at an extremely high computational cost that cannot be handled with QM methods (Habgood et al., 2020; Steinbrecher & Elstner, 2013). On the other hand, the accuracy of QM and the speed (less expensive) of MM are combined by the quantum mechanics/molecular mechanics (QM/MM) method (Saura et al., 2019). The QM/MM approach treats the inhibitor and few active site residues (where the chemical process takes place) with QM level of theory and the remaining macromolecular system treated with Molecular Mechanics (MM) (Saura et al., 2019) (Figure 3.1).

Molecular Mechanics (MM) are mathematical expressions that describe the potential energy of a system particle coordinates. MM methods (also called force field methods) utilises the Born-Oppenheimer approximation to estimate the energy of a system based only on the motions/positions while ignoring the degrees of freedom of electron (Monticelli & Tieleman, 2013).

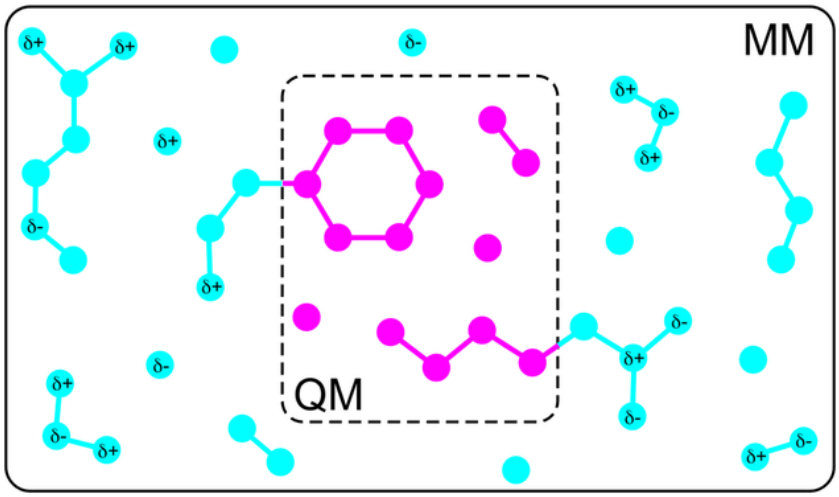


Figure 3.1 A QM/MM model for treating part of an enzyme with QM theory (ligand and few active site residues) and the remaining part of the macromolecular system with MM theory (Image adapted from https://bioexcel.eu/software/cp2k/).

The realistic simulation of a biomolecular complex system largely depends on the availability of reliable and accurate force fields. Most classical force fields consist of five different terms describing a molecular system and are represented as the bonded (bond stretching, bending of angles and rotation of the dihedral/torsional angle) and nonbonded (electrostatics and Van der Waals) interactions (Monticelli & Tieleman, 2013) (Figure 3.2)

A slight variation may exist in the force field terms of the available different MM packages. The force field terms of the AMBER molecular dynamics package used in this thesis are described in Equation 3.1.

$$E_{total} = \sum_{\text{bonds}} \frac{1}{2} k_r (r - r_0)^2 + \sum_{\text{angles}} \frac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} \frac{V_n}{2} [1 + \cos(n\emptyset - \delta)] + \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_o r_{ij}} \right) \quad \textbf{Equation 3.1}$$

The first term of equation 3.1 defines bond stretching based on Hooke's law, and the energetic penalty for the bond stretching from its reference bond length r_0 determine by force constant k_r. Similarly, the second term defines angle bending based on Hooke's law, and the energetic penalty for the distortion of the angle from its reference angle θ_0 determine by force constant k_θ. The third term also denotes the dihedral or torsion angle rotation signified by a dihedral angle $\emptyset$, a cosine series expansion with periodicity *n*, a barrier height V_n, and an offset δ.

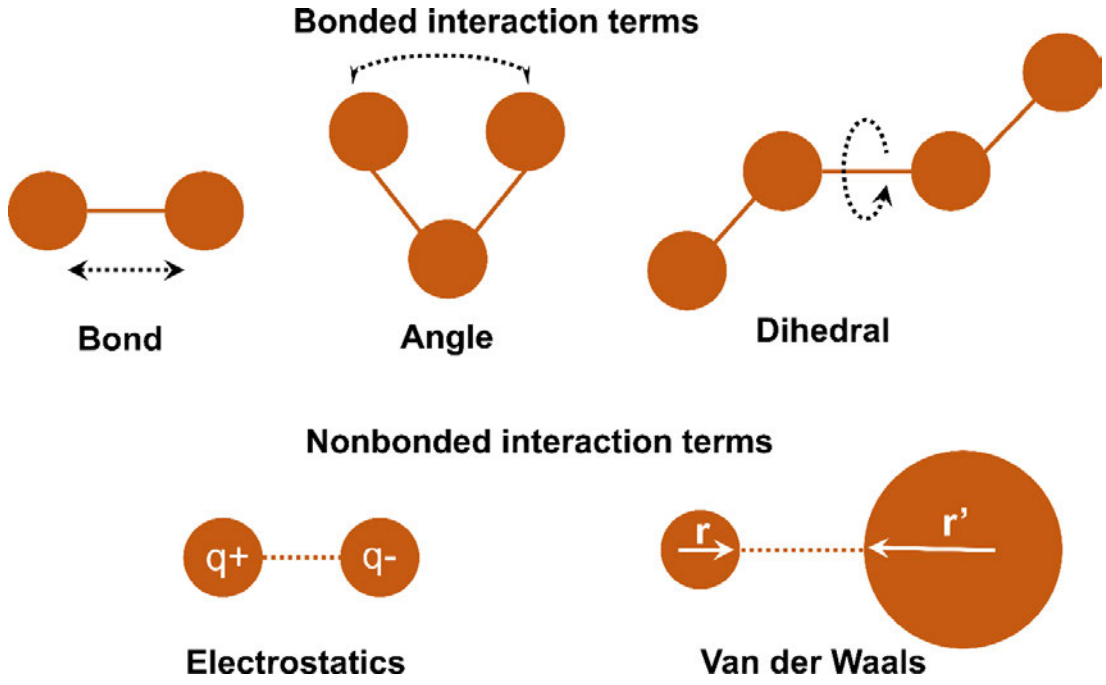


Figure 3.2 The bonded interactions and nonbonded interactions underlying molecular mechanics force field in the potential energy determination of a system (Image prepared by author).

The fourth and fifth parameters consist of the nonbonded interactions terms – the van der Waals (vdWs) interactions depicted by a 6-12 Lennard-Jones potential and the electrostatic interactions by the Coulombic potential. Molecular dynamics force fields development has made a significant stride in the past decades to accurately describe proteins, nucleic acids, lipids, and small molecules (Nerenberg & Head-Gordon, 2018).

3.2.1 Protein Force Fields

The pairwise additive approximation underlies the existing all-atom fixed-charge protein force fields for biomolecular simulation (Demerdash et al., 2014). Earlier protein force fields such as OPLS-AA, AMBER ff94 and CHARMM22 were successfully employed in short peptide and globular protein simulations. However, deficiencies in these force fields such as their inability to simulate intrinsically disordered proteins (Henriques et al., 2015; Rauscher et al., 2015; Levine & Shea, 2017), and incorrect identification of protein folding intermediates/pathways (McKiernan et al., 2017) were revealed by detailed experimental data. The backbone torsion potential modifications, which were subsequently included in the Amber ff99SB (Hornak et al., 2006) and the CMAP backbone energy correction to CHARMM22 (Mackerell Jr et al., 2004), improved their accuracy compared with the previous versions, as was demonstrated by comparing simulation results with experimental data (Best et al., 2008).

The torsion potential that regulates both side-chain dihedral angles and backbone behaviour has been a common approach to improving protein force fields. The methods used to derive these potentials included matching structural database data or only experimental NMR as used in AMBER ff99sb*/ff03*/ff03w (Best & Hummer, 2009; Best & Mittal, 2010) and ff99sbnmr1 (Li & Brüschweiler, 2010) or matching to only *ab initio* quantum chemistry data such as in AMBER ff14SBonlysc (Maier et al., 2015), AMBER ff99SB-ILDN (Lindorff-Larsen et al., 2010), AMBER-FB15 (Wang et al., 2017) and OPLS-AA/M (Robertson et al., 2015). However, the CHARMM 36 (Best, Zhu, et al., 2012; Best, Mittal, et al., 2012) and AMBERff14SB (Maier et al., 2015) utilised a combination of both data sources.

3.2.2 Lipid Force Fields

The all-atom lipid force field explicitly describes in detail all the atoms of a molecule. Some of the most frequently used lipid force fields include the AMBER Lipid11 (Skjevik et al., 2012), AMBER Lipid14 (Dickson et al., 2014), and AMBER Lipid17 (Gould et al., 2018), the CHARMM36 (C36) (Klauda et al., 2010) and CHARMM36 united atom FF (C36-UA) (Lee et al., 2014), the OPLS OPLS-UA and OPLS-AA, the GROMOS (53A6-CPK) (Piggot et al., 2012) and GROMOS54A7/54B7 (Schmid et al., 2011), and the MARTINI force field (Marrink et al., 2007). Loschwitz and colleagues have reviewed the detailed applications of these force fields and other emerging force field. (Loschwitz et al., 2020). Before the specific parameter set for lipids in AMBER was introduced, AMBER force fields were rarely used in the simulation of membrane proteins (Jójárt & Martinek, 2007). The AMBER lipid force fields have been tested for lipids bilayers such as 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (POPC), 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC), 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE), 2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC), 1,2-dilauroyl-*sn*-glycero-3-phosphocholine (DLPC), and 1,2-dimyristoyl-d54-*sn*-glycero-3-phosphocholine (DMPC) (Dickson et al., 2014).

3.2.3 Small Molecule Force Fields

The application of force fields in parameterising small molecules has played an essential role in the simulation of both natural and synthetic inhibitors. Presently, the force fields widely used for small molecules are the General AMBER Force Field (GAFF/GAFF2) (Wang et al., 2004; Wang et al., 2006), the CHARMM General force field (CGenFF) (Vanommeslaeghe et al., 2010; Vanommeslaeghe & MacKerell Jr, 2012; Vanommeslaeghe et al., 2012; Yu et al., 2012), GROMOS (Schuler et al., 2001; Oostenbrink et al., 2004; Horta et al., 2011; Horta et al., 2016),

OPLS-All-Atom (OPLS-AA) (Jorgensen et al., 1996), OPLS3 (Harder et al., 2016), and the Merck Molecular Force Field (MMFF) (Halgren, 1996a; Halgren, 1996b; Halgren, 1996c; Halgren & Nachbar, 1996; Halgren, 1996d).

The AMBER GAFF charges for small molecules are computed from the AM1-BCC semi-empirical model (Jakalian et al., 2000; Jakalian et al., 2002) or QM *ab initio* such as HF/6-31G* RESP charge (Bayly et al., 1993) charge models are used in the AMBER GAFF. The recent reparameterization of the Lennard-Jones and bonded parameters in the GAFF yielded the GAFF2. The default charge model of the OPLS3 is the CM1A-BCC or 1.14*CM1A-LBCC, which is similar to the GAFF AM1-BCC charge model (Dodda et al., 2015; Dodda et al., 2017). The CGenFF utilises a charge model known as the moiety-specific charge deviation protocol (Vanommeslaeghe et al., 2010). The GROMOS charge model is considerably different compared to the above charge models as used in the most recent release known as the 2016H66 (Horta et al., 2016).

The above force fields for small molecules have frequently been improved and actively maintained to involve new parameters for a broader range of chemical entities. The program for generating the CGenFF parameters and CHARMM topologies is accessible via the ParamChem website (Vanommeslaeghe & MacKerell Jr, 2012; Vanommeslaeghe et al., 2012). The Antechamber module (Wang et al., 2006) was designed by the AMBER developers and is used to generate GAFF and AMBER topologies for MD simulations. Most of the existing force fields used in the parameterisation of small molecules depend on non-polarisable or additive force fields for empirical potential energy functions. The common characteristics shared between these force fields are their potential energy function and the energy function parameters.

3.3 Molecular dynamics simulations of biomolecules

Molecular dynamics (MD) simulation has increasingly become an essential tool in providing deeper microscopic mechanistic insights into the structure, function, and dynamics of biological molecules (Cournia et al., 2015; Torrens-Fontanals et al., 2020). Experimental methods, including nuclear magnetic resonance (NMR) spectroscopy, cryo-electron microscopy (cryo-EM), and X-ray crystallography, have often been used to provide atomistic insight into the interactions of protein receptors with lipids. However, the single and static protein conformation obtained from these experimental methods offers little information about

the dynamics of macromolecules due to the highly dynamic nature of drug binding and molecular recognition of macromolecules (Torrens-Fontanals et al., 2020; Saurabh et al., 2020; X. Liu et al., 2018). Thus, MD simulation obtain extrapolated dynamical information on the biomolecular structure such as proteins and DNA.

3.3.1 Principles of MD Simulations

Molecular dynamics is a theoretical technique that depends on Newtonian mechanics. Solving Newton's movement equations makes it possible to model the physical motions of atoms in a molecular system over time. The forces between the system's potential energy and the particles are then measured using the MM force field (Monticelli & Tieleman, 2013). MD simulations generate dynamical trajectories of the interacting particles of a biomolecular structure over the desired time frame allowing the dynamics of interacting particles to be studied. Figure 3.3 shows the basic MD algorithm simplified.

3.3.1.1 Energy minimisation

Poor interaction sometimes characterises the starting structure of biomolecules to be used in MD simulation. Large repulsive interaction, for example, may arise when two atoms are too close to each other. Thus, the energy minimisation of the starting structure is required to remove bad atom contacts before MD simulations. The three commonly used energy minimisation methods are the conjugate gradient (CG), the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS), and the steepest descent (SD) algorithm. The robust and efficient steepest descent algorithm is sufficient for most minimisation processes. The SD algorithm uses potential energy and forces in updating the positions of the atomic particles in an iterative manner and halts when it has iterated for the number of cycles specified by the user or when the specified value is larger than the maximum force. Compared with the steepest descent algorithm, the conjugate gradient is more efficient when the system is near the energy minimum. The L-BFGS algorithm has been observed to have a faster convergence than the conjugate gradient algorithm and is more appropriate for alchemical free energy calculations.

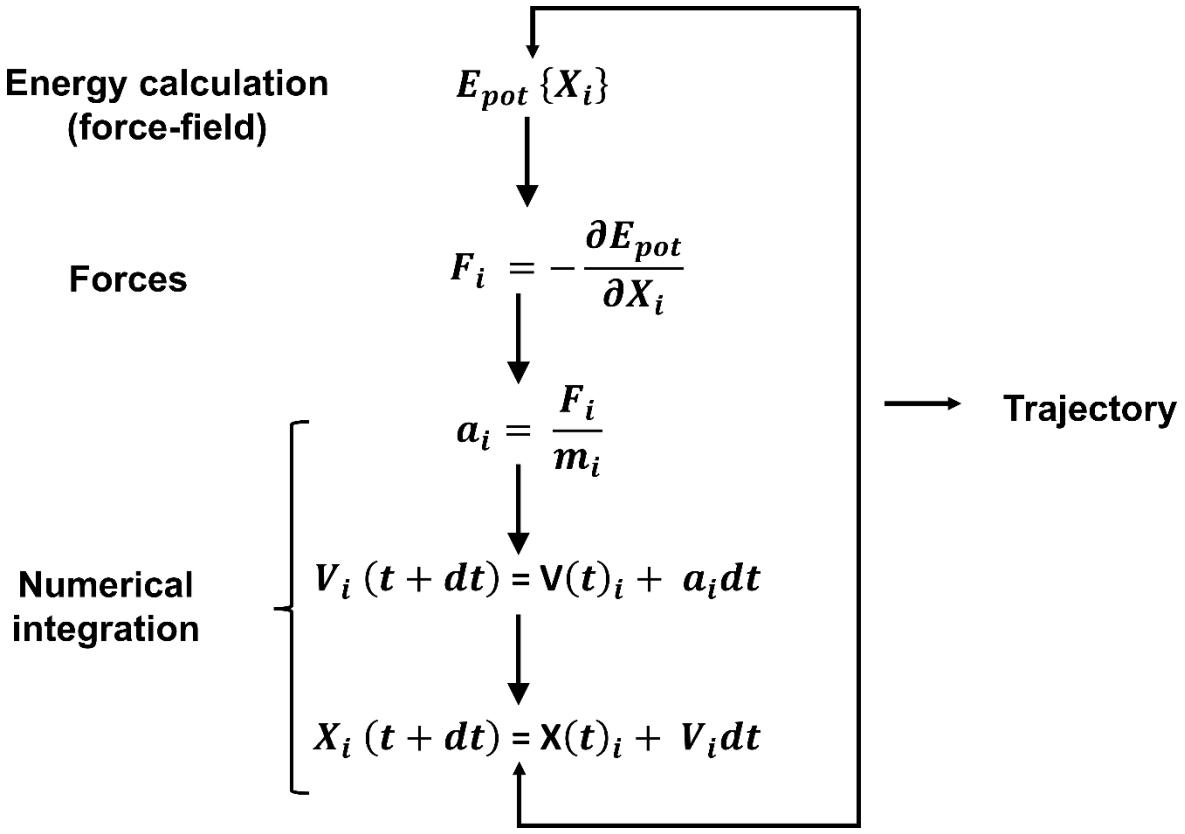


Figure 3.3 Simplified MD simulation algorithm. Where potential energy = Epot; time of iterations = dt; simulation time = t; For N simulated atoms of each spatial coordinates (i): atom coordinates = x; force component = F; a = acceleration, atom mass = m; and velocity = v (Image modified from (Hospital et al., 2015)).

3.3.1.2 Integration Algorithm

Integration methods used in MD simulation codes include the leap-frog algorithm, the Verlet algorithm, and velocity Verlet. The leap-frog algorithm is the often-used integration algorithm updating the position r and velocity v of a simulation particle by:

$$v\left(t + \frac{1}{2}\Delta t\right) = v\left(t + \frac{1}{2}\Delta t\right) + \frac{\Delta t}{m}\mathrm{F(t)}, \qquad \textbf{Equation 3.2}$$

$$r(t + \Delta t) = r(t) + \Delta t v\left(t + \frac{1}{2}\Delta t\right). \qquad \textbf{Equation 3.3}$$

3.3.1.3 Regulation of Temperature and Pressure

The regulation of pressure only applies when a system is simulated under constant pressure periodic boundary condition. A “barostat” or a “pressure bath” is utilised to regulate the pressure of the simulating system. Available pressure coupling algorithms include the Berendsen barostat (which scales the atomic coordinates and box vector every time step), the

Monte Carlo barostat (that rigorously samples from the isobaric-isothermal ensemble (NPT)) and the Parrinello-Rahman barostat.

Temperature coupling is applied under a constant temperature (NVT ensemble) MD simulation. The commonly used temperature regulation thermostats include the Berendsen thermostat (Berendsen et al., 1984), Andersen thermostat (Andersen, 1980), and the Nosé-Hoover thermostat (Nosé, 1984; Hoover, 1985).

3.3.1.4 Constraint algorithms

Constraints are introduced to hold bonds in fixed length in classical MD simulation. The SHAKE algorithm approach has traditionally been used to apply constraints in numerous MD simulation codes (Ryckaert et al., 1977). Using a reference set, the SHAKE algorithm works by modifying a set of unconstrained coordinates to a new set that satisfies a list of distance constraints. The LINear Constraint Solver (LINCS) is another approach available besides the SHAKE algorithm (Hess et al., 1997). After an unconstrained update in a non-iterative fashion, the LINCS algorithm resets the bond lengths.

3.3.1.5 Periodic Boundary Conditions

Periodic boundary conditions (PBC) are used to minimise the interface with the vacuum of a finite molecular system in an MD simulation (Schultz, 1999). To measure short-range nonbonded interactions with other particles, PBC is often applied following the minimum image convention, which regulates that only the closest image of a particle is selected. Usually, for short-range interactions, a cut-off is set. Beyond the cut-off, long-range van der Waals interactions are typically ignored. In contrast, long-range electrostatic interactions are usually recovered using the Particle-Mesh-Ewald (PME) method (Darden et al., 1993). The system is first put into a box when using PBC, which is then copied to fill the entire space in every direction. The most frequently used box types include rhombic dodecahedrons, cubes, and truncated octahedrons.

3.3.2 System Setup and MD Simulations of GPCRs in Lipid Bilayer

There are various computational approaches for membrane protein simulations that seek to balance accuracy and the size of the molecular system under investigation (Cui, 2014; Chavent et al., 2016). Generally, MD simulation studies begin with an experimental or homology modelled protein structure and simulated over time based on classical mechanics principle. However, the starting protein structures of membrane proteins are almost in all cases not

available in their native lipid bilayer. Hence, presenting one of the most significant challenges in constructing and embedding membrane proteins in lipid bilayer for use in MD simulation. This shortcoming gives an additional responsibility for appropriate modelling and insertion of the protein structure in a native lipid bilayer. Thus, membrane protein simulation compared to globular protein simulation system setup for simulation may be challenging, especially for the novice.

Biomolecular simulation software packages such as GROMACS and CHARMM have their membrane builder. The membrane builder for GROMACS, whereas MemBuilder (Ghahremanpour et al., 2014) and CHARMM-GUI (Wu et al., 2014) are for CHARMM. However, the AMBER MD package lacks its membrane builder and relies on external programs for embedding membrane proteins in a lipid bilayer. The Visual Molecular Dynamics (VMD) (Humphrey et al., 1996) package also has a plugin for setting up a simple membrane system. Other online servers packages such as PACKMOL (Martínez et al., 2009), MemProtMD (Newport et al., 2019) and PACKMOL-memgen (Schott-Verdugo & Gohlke, 2019) allows the user to build a membrane system to suit their system's needs. Recently, Khanna and colleagues (Khanna, 2018) developed an "in-house" membrane builder code known as AMBAT-*Amber Membrane Builder and Analysis Tool* specifically for the AMBER MD simulation package and has been employed in a recent study (Toroz et al., 2019). However, AMBAT has not yet been integrated into the AMBER MD simulation package.

Embedding a membrane protein in lipid bilayer using the CHARMM-GUI Membrane Builder remote server arguable appears to be more user-friendly; nevertheless, the user will need to perform some modifications for residue and atom naming compatibility before using it in the AMBER MD package. The *charmmlipid2amber.py* introduced in AMBER, therefore, provides a vigorous atom and residue renaming of CHARMM-GUI PDB structure format into Leap and Lipid14 PDB readable format. The CHARMM-GUI Membrane Builder was initially used to embed the membrane proteins used in this thesis in a lipid bilayer, prepared with the *charmmlipid2amber.py* and simulated with the AMBER MD package. Figure 3.4 represents the general procedure for the initial system setup with CHARMM-GUI.

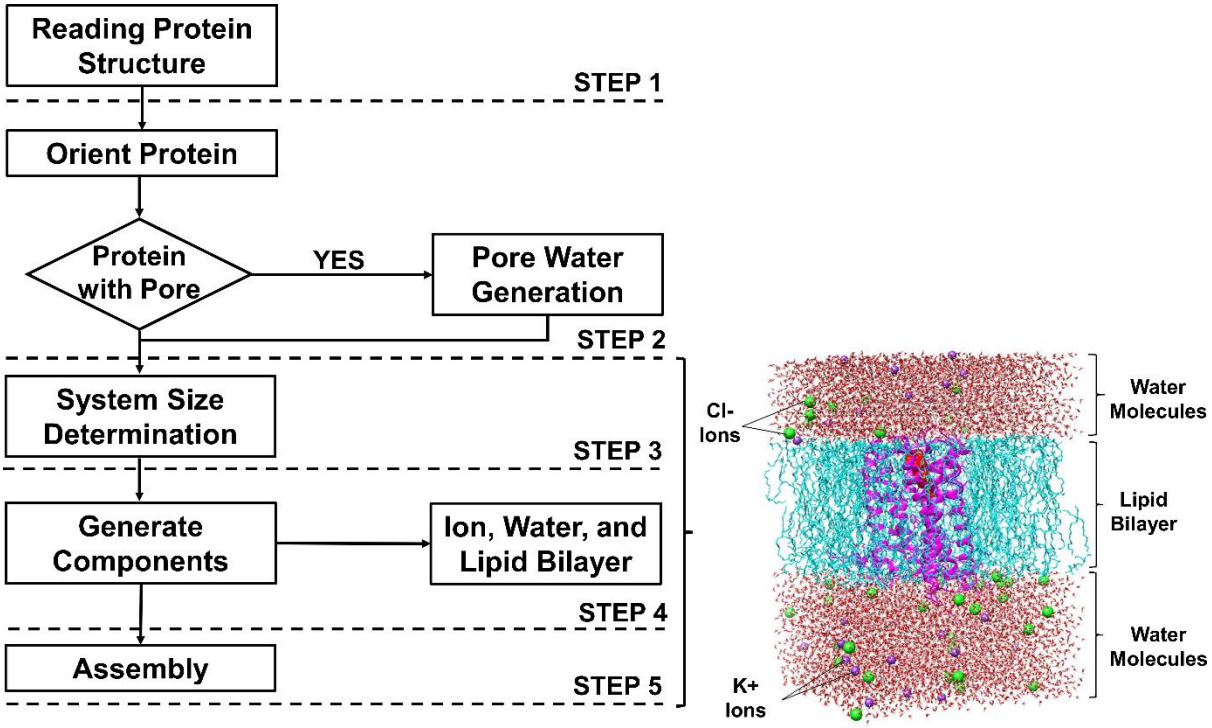


Figure 3.4 The systematic workflow in the overall embedding of membrane protein in a lipid bilayer with water molecules and ions using the CHARMM-GUI membrane-builder for AMBER MD simulation (Image modified by author from (Jo et al., 2008)).

3.3.3 Molecular Dynamics Trajectory Analyses

3.3.3.1 Atomic displacement analysis

The average change in the displacement of selected atoms for a trajectory frame relative to a reference frame is measured using the root mean square deviation (RMSD). To test the convergence/stability of the MD simulation, RMSD is computed for all the trajectory frames. The RMSD computation for trajectory frame x is:

$$\mathrm{RMSD_x} = \sqrt{\frac{1}{\mathrm{N}} \sum_{\mathrm{i=1}}^{\mathrm{N}} (r_i'(t_x) - r_i\,(t_{ref}))^2} \qquad \textbf{Equation 3.4}$$

where N is the selected number of atoms, $r_i\,(t_{ref})$ denotes the average referenced position of an atomic particle r_i at referenced time t_{ref} and $r_i'(t_x)$describes the position of specified atoms in the frame x recorded at the time t_x after reference frame superposition. This calculation is repeated for each frame of the simulated trajectory.

3.3.3.2 Residue positional mobility analysis

The root mean square fluctuation (RMSF) for specified atoms (e.g. Cα atoms) is a measure of atomic deviation between particle n position relative to a reference position:

$$\mathrm{RMSF_i} = \sqrt{\frac{1}{\mathrm{T}} \sum_{\mathrm{t=1}}^{\mathrm{T}} (r_i'(t) - r_i\,(t_{ref}))^2} \qquad \textbf{Equation 3.5}$$

where T is the specified trajectory time to computed for RMSF, $r_i\,(t_{ref})$ denotes the average referenced position of an atomic particle r_i at referenced time t_{ref} and $r_i'(t)$ describes the position of a specified atom in residue i at time t after reference frame superposition. The RMSF is computed to observe the mobility/fluctuations of residues in the macromolecular system.

3.3.3.3 Structural compactness analysis

The radius of gyration (RoG or rGyr), which estimates the compactness of a molecular structure is computed as:

$$Rg = \left(\frac{\sum_i |r_i|^2 m_i}{\sum_i m_i}\right)^2 \qquad \textbf{Equation 3.6}$$

where m_i denotes the mass of atom *i* and r_i describes the position of atom *i* relative to the molecule's centre of mass.

3.3.3.4 Principal Component Analysis and Conformational Clustering

Principal Component Analysis (PCA) is an advanced technique of trajectory analysis used to describe a given macromolecule's collective motion. PCA is performed by constructing a covariance matrix $\mathrm{C_i}$ defined by an averaged MD trajectories ensemble as:

$$\mathrm{C_i} = < (\mathrm{q_i} - < \mathrm{q_i} >)(\mathrm{q_j} - < \mathrm{q_j} >) > \quad (\mathrm{i, j} = 1,2 \ldots\ldots,3\mathrm{N}) \qquad \textbf{Equation 3.7}$$

Where N signifies the number of C-α atom, q_i denoting the mass-weighted cartesian coordinate of the *i*th N atom.

To construct the covariance matrix, the translation and rotation movements are initially excluded, which allows the 3N directions along which most of the protein motion to be identified. The covariance matrix is diagonalised to extracting a set of eigenvectors and eigenvalues. Using an orthogonal coordinate transformation matrix T, the diagonalisation is

carried out, which transforms the covariance matrix C into a diagonal matrix of the eigenvalues λi.:

$$\Lambda = T^{T} C_{ij} T \qquad \textbf{Equation 3.8}$$

Extracting the essential principal modes provides insight into the conformational dynamics of biological molecules. The principal component analysis in the manuscript chapters that make up this thesis was computed with the Bio3D package in R (Grant et al., 2006).

3.3.4 Estimation of binding free energies

The application of computational approaches in binding free energy estimation has emerged as an essential tool in the drug design process (Pearlman & Charifson, 2001; Srivastava & Sastry, 2012; Steinbrecher et al., 2017). These methods provide a guide in structure-based drug discovery and protein-ligand interaction analysis. Several of these computational approaches rely on molecular dynamics (MD) in providing a cost-effective conformational ensemble that is statistically relevant for thermodynamic calculations. The most popular MD-based binding free energy methods include thermodynamic integration (TI) (Bhati et al., 2017), free energy perturbation (FEP) (Lenselink et al., 2016; Deflorian et al., 2020), linear interaction energy (LIE) (Rifai et al., 2018), and Molecular Mechanics Poisson–Boltzmann/Generalized Born Surface Area (MM-PB/GBSA) (Aldeghi et al., 2017; Miller III et al., 2012). However, the first two methods are quite computationally demanding/expensive, while the MM-PB/GBSA approaches have been mainly used in virtual screening protocols due to their computational efficiency (Yau et al., 2020; Manhas et al., 2019). The AMBER MM-PB/GBSA calculations have successfully has been used to calculate the binding free energies in protein-protein complexes (Contini et al., 2012), DNA-ligand complexes (Ferri et al., 2011), and protein-ligand complexes (Manhas et al., 2019).

3.3.4.1 AMBER MM/PBSA.py for binding free energy calculation

The *MM/PBSA.py* is an AMBER package program for end-point free energy estimation which uses the Molecular Mechanics Poisson–Boltzmann/Generalized Born Surface Area (MM-PB/GBSA) methods (Miller III et al., 2012). For the prediction of binding affinity, the MM-PB/GBSA-based approaches are commonly used as they provide an intermediate compromise of accuracy and speed between the more robust free energy perturbation techniques and the empirical scoring functions used in docking.

MM/PBSA is more effective in computing absolute binding free energies, whereas MM/GBSA effectively computes relative binding energies. MM/GBSA is, therefore, the most appropriate approach compared to MM/PBSA to be employed in the correct ranking of inhibitors (Hou et al., 2011).

The estimation of binding free energy (ΔG) by MM/PBSA or MM/GBSA is represented by Figure 3.5 and the following equations:

$$\Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand}) \qquad \textbf{Equation 3.9}$$

$$\Delta G_{bind} = \Delta H - T\Delta S \qquad \textbf{Equation 3.10}$$

The individual free energy components of Equation 3.9 is obtained from the molecular mechanical energy (E_{MM}), entropy contribution ($-T\Delta S$), and solvation energy (G_{sol}) terms:

$$\Delta H = \Delta E_{MM} + \Delta G_{sol} - T\Delta S \qquad \textbf{Equation 3.11}$$

The molecular mechanical energy term E_{MM} of the total enthalpy contribution is expressed as:

$$\Delta E_{MM} = \Delta E_{internal} + \Delta E_{vdw} + \Delta E_{ele} \qquad \textbf{Equation 3.12}$$

Where the ΔE_{MM} is comprised of the intramolecular energy term ($\Delta E_{internal}$; involving the system's dihedral, angle, and bond energies), the van der Waals interactions (ΔE_{vdw}), and the electrostatic energy (ΔE_{ele}) in the gas phase.

The free energy of solvation is described by G_{sol} composing of the nonpolar $\Delta G_{nonpolar}$ and polar ΔG_{polar} solvation free energy terms:

$$\Delta G_{sol} = \Delta G_{polar(PB/GB)} + \Delta G_{nonpolar} \qquad \textbf{Equation 3.13}$$

The ΔG_{polar} term is generally calculated using either the Generalized Born (GB) model in MM/GBSA and the Poisson-Boltzmann (PB) model in MM/PBSA.

The solvent-accessible surface area (SASA) method is used to calculate the nonpolar solvation energy component ($G_{nonpolar}$):

$$G_{nonpolar} = \gamma * (SAS) + \beta \qquad \textbf{Equation 3.14}$$

The β and γ denote the offset and surface tension values, respectively. The conformational entropy (-TΔS) to the binding free energy is generally estimated by the AMBER *normal mode analysis (nmode*) program.

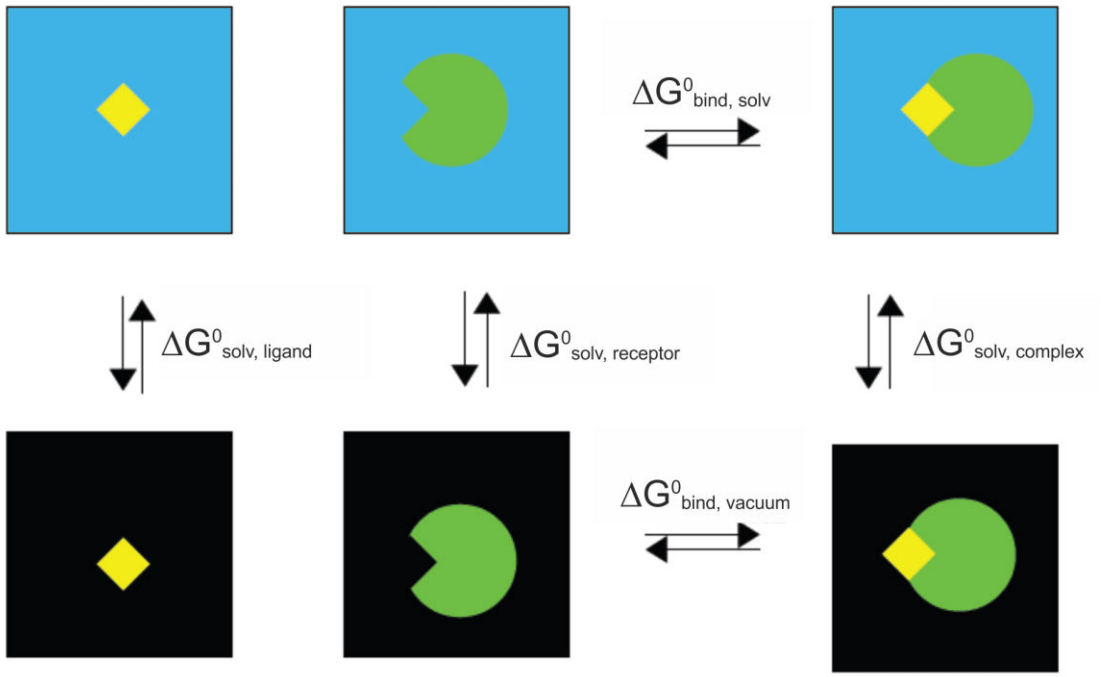


Figure 3.5 Schematic illustration of the thermodynamic cycle according to which free energy of binding is calculated. (Image adapted from (Walker, 2008)).

The MM-PBSA/GBSA approaches have been documented to predict experimental binding free energies successfully for diverse biological systems. (Chéron & Shakhnovich, 2017; Rastelli et al., 2010; Ferrari et al., 2007; Bonnet & Bryce, 2005; Wang et al., 2001). However, the MM-PBSA/GBSA approach has also been shown to produce substandard results in predicting binding free energies (Singh & Warshel, 2010; Kuhn et al., 2005; Pearlman, 2005). MM-PBSA/GBSA performance in predicting binding affinity was recently evaluated and reported for 934 known inhibitors against the crystal structures of twenty class A GPCRs (Yau et al., 2019). The authors concluded that MM-PBSA/GBSA performance in GPCR ligands binding free energy prediction is highly system-specific (Yau et al., 2019).

3.4 Computational Drug Design and Discovery

Drug design, discovery and development require a laborious inter-disciplinary approach and are considered time consuming and expensive (DiMasi et al., 2003; DiMasi et al., 2016). The application of computational drug design and discovery (CDDD) techniques in drug design projects have assisted in expediting the process. The available computational methods in drug design and discovery can be categorized into structure-based drug design (SBDD) and ligand-

based drug design (LBDD) methods. Figure 3.6 summarises the position of CDDD methods in the drug design process.

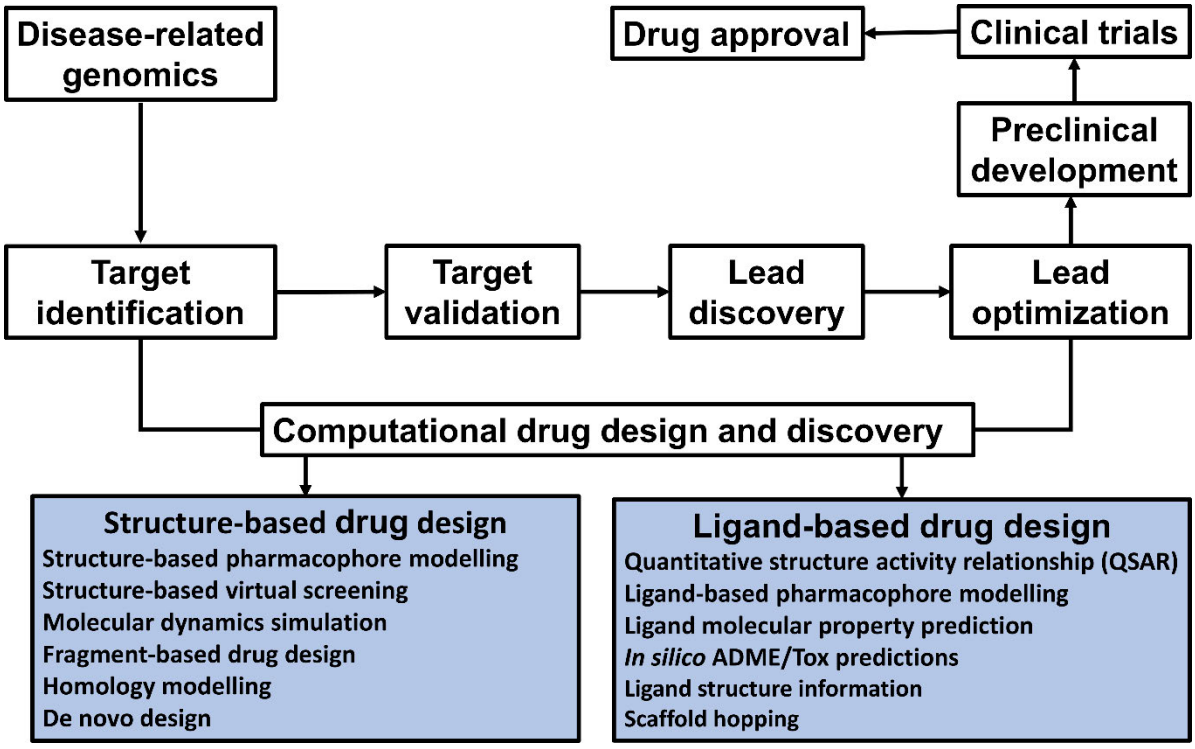


Figure 3.6 Computational drug design and discovery position in the drug discovery process (Image prepared by author).

3.4.1 Structure-based drug design and discovery

Structure-based drug design (SBDD) is a target-based computational method used to design or identify novel inhibitor that relies on the structural details of the biological target of interest (Batool et al., 2019). cryo-electron microscopy (cryo-EM), X-ray crystallography, and nuclear magnetic resonance (NMR) spectroscopy are experimental methods used to determine the three-dimensional (3D) structures of the macromolecules (Krishnan & Rupp, 2012; García-Nafría & Tate, 2020). Homology/comparative models can be used based on the available structures of closely related proteins in the absence of an experimentally defined structure (Dong et al., 2013). SBDD is an efficient, more specific, and quick process for lead identification/discovery and lead optimisation since it utilises molecular level knowledge about the disease and the protein target structure (Lionta et al., 2014; Congreve & Marshall, 2010). The use of SBDD methods in academic research and the pharmaceutical industry have helped identify several FDA-approved medicines, including HIV-1 inhibitors (Wlodawer & Vondrasek, 1998) and the antibiotic norfloxacin (Rutenber & Stroud, 1996).

3.4.1.1 Structure-based virtual screening

The *in silico* high throughput screening (HTS) technique known as structure-based virtual screening (SBVS) is a technique used to screen large chemical compound libraries or compound databases against a molecular target (Wang et al., 2020; Liu & Jockers, 2020). The process requires the protein target structure to predict the best binding interaction between the compounds and the target to form a complex. The ligands or compound libraries are then ranked based on their binding affinity to the target, with the highly promising compound displayed at the top of the list (S. Liu et al., 2018). Molecular docking is a well-known technique applied in SBVS as it is fast, less computationally expensive and achieves good results (Meng et al., 2011).

3.4.1.2 Fragment-based drug discovery

Fragment-based drug discovery (FBDD) is an effective target-based technique used to develop potent small molecules with chemical fragments as starting points and optimised toward drug-like leads. FBDD has become an attractive strategy in structure-based drug discovery with successes, even where other approaches failed for challenging targets (Erlanson et al., 2016; Murray et al., 2012). FBDD begins with screening very weak affinity low molecular weight compounds libraries (fragments) to identify "hits" against the target of interest. The identified hits are further optimised based on the target structural information to higher affinity small molecules using robust methods (Li, 2020; Erlanson et al., 2019).

3.4.2 Ligand-based drug design

Ligand-based drug design (LBDD) is a computational technique applied in the absence of a target 3D structural information. It depends on the knowledge of the molecules that bind to the biomolecular target of interest. The 3D quantitative structure-activity relationships (3D QSAR) and Ligand-based pharmacophore modelling are the most relevant and highly used methods in LBDD. These tools assist in lead identification and lead optimisation with appropriate predictive models (Acharya et al., 2011).

3.4.3 Molecular docking

Molecular docking is an extensively applicable computational technique in the process of structure-based drug design and discovery. Molecular docking is mainly used to predict the most favourable binding conformation and affinity of a ligand at the binding pocket of a macromolecular target (Fan et al., 2019) (Figure 3.7). The prediction of the binding modes and

the quantitative estimation of binding affinities through molecular docking protocols assist in the investigation of vital protein-ligand interactions. This then assists in analysing the structure-activity relationship (SAR) for rational drug design and prioritising compounds/drug candidates that can be synthesised or experimentally tested.

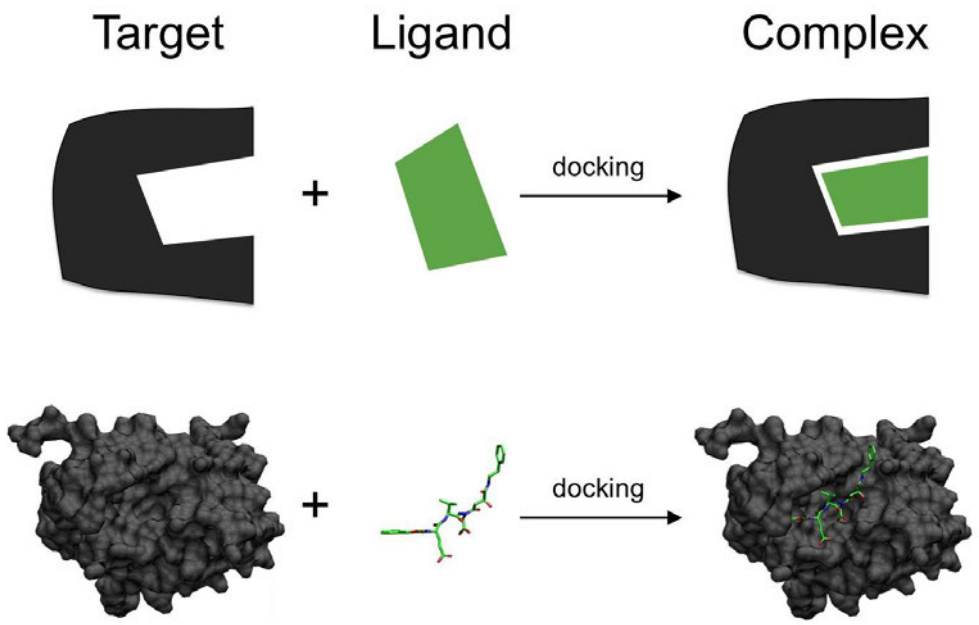


Figure 3.7 A schematic representation of the molecular docking of a ligand (green) into a protein target binding pocket (black) to produce a ligand-receptor complex (Image adapted from (Wikipedia, 2015)).

The application of large drug-like compound libraries in virtual high-performance screening to obtain leads for further drug development is paramount for molecular docking. The molecular docking process can be divided into two separate stages: ligand conformational space search within the binding pocket and the binding affinity estimation for every predicted conformation (Shen, Ding, et al., 2020).

Table 3.1 Commonly used molecular docking software.

Name	*Search Algorithm	Scoring Method	Speed	License	Applicable Areas and Features
AutoDock (Morris et al., 2009)	GA, LGA	Semi-empirical free energy estimation	Medium	Free for academic use	Flexible-rigid docking program. AutoDock is free academic software that works in combination with Autodock-Tools.
Autodock Vina (Trott & Olson, 2010)	GA	Semi-empirical free energy estimation	Fast	Open-source	Flexible-rigid docking program. AutoDock Vina employs a sophisticated local gradient optimisation procedure. It is accurate and faster than the AutoDock 4.
Dock (Lang et al., 2009)	FA	Molecular mechanic force field	Fast	Free for academic use	Flexible docking program. Applicable for flexible ligands and flexible receptor docking.
Flex X (Kramer et al., 1999)	FA	Semi-empirical free energy estimation	Fast	Commercial	Flexible-rigid docking. It uses incremental construction algorithms and is suitable for small molecule database virtual screening.
Glide (Friesner et al., 2004)	ESS	Semi-empirical free energy estimation	Medium	Commercial	A flexible program for docking. This software uses a knowledge base to decrease the search range. It has modes of extra accuracy (XP), standard accuracy (SP), and high throughput virtual display (HTVS).
GOLD (Verdonk et al., 2003)	GA	Semi-empirical free energy estimation	Fast	Commercial	Flexible GA-based docking program. The software's reliability and accuracy have been highly appraised.
LeDOCK (Shen, Wang, et al., 2020)	GA, SA	Molecular force field	Fast	Free for academic use	A new flexible molecular docking software. It is highly applicable for virtual high throughput screening due to its high accuracy and being fast (Shen, Wang, et al., 2020).
RDOCK (Li et al., 2003)	GA, MC, MIN	Molecular force field	Medium	Open-source	A rigid docking program. For refinement and scoring, it uses the CHARMm-based protocol. It is primarily designed to predict high-throughput virtual screening (HTVS) campaigns aside from predicting binding mode.
ZDOCK (Chen et al., 2003)	SC, FFT	Molecular force field	Medium		A rigid molecular docking program. It employs a novel scoring function that unites electrostatic and desolvation with pairwise shape complementarity (PSC). A webserver version available as ZDOCK server.

*FA= Fragmentation algorithm; GA= Genetic Algorithm; LGA= Lamarckian Genetic Algorithm; ESS= Exhaustive systematic search; SA = Simulated annealing; MC = Monte Carlo; MIN = Simplex Minimization; FFT = Fast Fourier Transform algorithm, shape complementarity

3.4.3.1 Ligand Conformational Search

Conformational search algorithms are utilised in molecular docking programs to search systematically for the orientations and conformations of a ligand at the binding pocket on a receptor target (Shen, Ding, et al., 2020). Thus, several poses of the ligand-receptor complexes are generated in the conformation space search. Translational and Rotational degrees of freedom are used in the case of rigid docking. In contrast, a conformational degree of freedom is added to the ligands' rotations and translations in a flexible docking protocol. (Maia et al., 2020).

3.4.3.2 Binding Affinity Estimation

Molecular docking programs employ different scoring functions to compute the molecular interaction force between a ligand and its molecular target. For each ligand-receptor complex conformation, a scoring function predicts the ligand-binding affinity at the receptor-binding pocket (Shen, Ding, et al., 2020). In modern docking programs, the molecular docking scoring functions used can be characterised as knowledge-based potential, physics-based methods (e.g. force-field or QM-based), descriptor-based scoring functions and empirical scoring functions (Liu & Wang, 2015).

3.4.4 In silico physicochemical, and pharmacokinetic predictions

In their safety and efficacy profiles, the absorption, delivery, metabolism, excretion, and toxicity (ADMET) properties of the prospective drug molecule are critical. However, it is time-consuming and expensive to apply experimental methods that also require animal testing. Due to poor druggability, the disconnection between compound optimisation and ADMET assessments may sometimes lead to the dismissal of compound candidates, although they may have demonstrated excellent *in vitro* efficacy (Wang et al., 2015). The ability to early ascertain the druggability of candidate compounds improve drug development efficiency and productivity. The application of *in silico* physicochemical and ADMET profiling provides a cost-effective approach in prioritising the candidates. Thus, in silico ADMET profiling has emerged as a tool of choice in early drug discovery projects (Wang et al., 2015). The availability of high quality physicochemical and ADMET.predictive models provide avenues for the optimisation of druggability properties and compound efficacy (Huang et al., 2013). Applying these *in silico* models increases the probability of drug candidate success with reduced overall cost due to decreased attrition rate. The intelligent integration of *in silico*, *in vivo*, and *in vitro* ADMET data maximise the outcome of ADMET models in guiding drug

discovery (Wang & Collis, 2011). Various *in silico* tools exist to predict pharmacokinetic profiles and other molecular properties such as clogP, topological polar surface area (TPSA), drug score values, fragment-based drug-likeness, etc., of candidate molecules. A representative of these tools includes SWISSADME (Daina et al., 2017), pkCSM (Pires et al., 2015), and QikProp (Ioakimidis et al., 2008).

References

Acharya, C., Coop, A., E Polli, J. & D MacKerell, A. 2011. Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. *Current computer-aided drug design*, 7(1): 10–22.

Aldeghi, M., Bodkin, M.J., Knapp, S. & Biggin, P.C. 2017. Statistical analysis on the performance of Molecular Mechanics Poisson–Boltzmann Surface Area versus absolute binding free energy calculations: Bromodomains as a case study. *Journal of chemical information and modeling*, 57(9): 2203–2221.

Andersen, H.C. 1980. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of chemical physics*, 72(4): 2384–2393.

Batool, M., Ahmad, B. & Choi, S. 2019. A structure-based drug discovery paradigm. *International journal of molecular sciences*, 20(11): 2783.

Bayly, C.I., Cieplak, P., Cornell, W. & Kollman, P.A. 1993. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *The Journal of Physical Chemistry*, 97(40): 10269–10280.

Berendsen, H.J.C., Postma, J.P.M. van, van Gunsteren, W.F., DiNola, A. & Haak, J.R. 1984. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8): 3684–3690.

Best, R.B., Buchete, N.-V. & Hummer, G. 2008. Are current molecular dynamics force fields too helical? *Biophysical journal*, 95(1): L07–L09.

Best, R.B. & Hummer, G. 2009. Optimized molecular dynamics force fields applied to the helix− coil transition of polypeptides. *The journal of physical chemistry B*, 113(26): 9004–9015.

Best, R.B. & Mittal, J. 2010. Protein simulations with an optimized water model: cooperative helix formation and temperature-induced unfolded state collapse. *The journal of physical chemistry B*, 114(46): 14916–14923.

Best, R.B., Mittal, J., Feig, M. & MacKerell Jr, A.D. 2012. Inclusion of many-body effects in the additive CHARMM protein CMAP potential results in enhanced cooperativity of α-helix and β-hairpin formation. *Biophysical journal*, 103(5): 1045–1051.

Best, R.B., Zhu, X., Shim, J., Lopes, P.E.M., Mittal, J., Feig, M. & MacKerell Jr, A.D. 2012. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain $\chi 1$ and $\chi 2$ dihedral angles. *Journal of chemical theory and computation*, 8(9): 3257–3273.

Bhati, A.P., Wan, S., Wright, D.W. & Coveney, P. V. 2017. Rapid, accurate, precise, and reliable relative free energy prediction using ensemble based thermodynamic integration. *Journal of chemical theory and computation*, 13(1): 210–222.

Bonnet, P. & Bryce, R.A. 2005. Scoring binding affinity of multiple ligands using implicit solvent and a single molecular dynamics trajectory: application to influenza neuraminidase. *Journal of Molecular Graphics and Modelling*, 24(2): 147–156.

Chavent, M., Duncan, A.L. & Sansom, M.S.P. 2016. Molecular dynamics simulations of membrane proteins and their interactions: from nanoscale to mesoscale. *Current opinion*

in structural biology, 40: 8–16.
Chen, R., Li, L. & Weng, Z. 2003. ZDOCK: an initial-stage protein-docking algorithm. *Proteins: Structure, Function, and Bioinformatics*, 52(1): 80–87.
Chéron, N. & Shakhnovich, E.I. 2017. Effect of sampling on BACE-1 ligands binding free energy predictions via MM-PBSA calculations. *Journal of Computational Chemistry*, 38(22): 1941–1951.
Congreve, M. & Marshall, F. 2010. The impact of GPCR structures on pharmacology and structure-based drug design. *British journal of pharmacology*, 159(5): 986–996.
Contini, A., Cappelletti, G., Cartelli, D., Fontana, G. & Gelmi, M.L. 2012. Molecular dynamics and tubulin polymerization kinetics study on 1, 14-heterofused taxanes: evidence of stabilization of the tubulin head-to-tail dimer–dimer interaction. *Molecular BioSystems*, 8(12): 3254–3261.
Cournia, Z., Allen, T.W., Andricioaei, I., Antonny, B., Baum, D., Brannigan, G., Buchete, N.-V., Deckman, J.T., Delemotte, L. & Del Val, C. 2015. Membrane protein structure, function, and dynamics: a perspective from experiments and theory. *The Journal of membrane biology*, 248(4): 611–640.
Cui, Q. 2014. QM/MM Methods: Recent Developments and Application to Membrane Proteins and Molecular Motors. *Biophysical Journal*, 106(2): 445a.
Daina, A., Michielin, O. & Zoete, V. 2017. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific reports*, 7: 42717.
Darden, T., York, D. & Pedersen, L. 1993. Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. *The Journal of chemical physics*, 98(12): 10089–10092.
Deflorian, F., Perez-Benito, L., Lenselink, E.B., Congreve, M., van Vlijmen, H.W.T., Mason, J.S., Graaf, C. de & Tresadern, G. 2020. Accurate prediction of GPCR ligand binding affinity with free energy perturbation. *Journal of Chemical Information and Modeling*.
Demerdash, O., Yap, E.-H. & Head-Gordon, T. 2014. Advanced potential energy surfaces for condensed phase simulation. *Annual review of physical chemistry*, 65: 149–174.
Dickson, C.J., Madej, B.D., Skjevik, Å.A., Betz, R.M., Teigen, K., Gould, I.R. & Walker, R.C. 2014. Lipid14: the amber lipid force field. *Journal of chemical theory and computation*, 10(2): 865–879.
DiMasi, J.A., Grabowski, H.G. & Hansen, R.W. 2016. Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics*, 47: 20–33. https://www.sciencedirect.com/science/article/pii/S0167629616000291.
DiMasi, J.A., Hansen, R.W. & Grabowski, H.G. 2003. The price of innovation: new estimates of drug development costs. *Journal of Health Economics*, 22(2): 151–185. https://www.sciencedirect.com/science/article/pii/S0167629602001261.
Dodda, L.S., Vilseck, J.Z., Cutrona, K.J. & Jorgensen, W.L. 2015. Evaluation of CM5 charges for nonaqueous condensed-phase modeling. *Journal of chemical theory and computation*, 11(9): 4273–4282.
Dodda, L.S., Vilseck, J.Z., Tirado-Rives, J. & Jorgensen, W.L. 2017. 1.14* CM1A-LBCC: localized bond-charge corrected CM1A charges for condensed-phase simulations. *The Journal of Physical Chemistry B*, 121(15): 3864–3870.
Dong, X., Zhao, Y., Huang, X., Lin, K., Chen, J., Wei, E., Liu, T. & Hu, Y. 2013. Structure-based drug design using GPCR homology modeling: toward the discovery of novel selective CysLT2 antagonists. *European Journal of Medicinal Chemistry*, 62: 754–763.
Erlanson, D.A., Davis, B.J. & Jahnke, W. 2019. Fragment-based drug discovery: advancing fragments in the absence of crystal structures. *Cell chemical biology*, 26(1): 9–15.
Erlanson, D.A., Fesik, S.W., Hubbard, R.E., Jahnke, W. & Jhoti, H. 2016. Twenty years on: the impact of fragments on drug discovery. *Nature reviews Drug discovery*, 15(9): 605.

Fan, J., Fu, A. & Zhang, L. 2019. Progress in molecular docking. *Quantitative Biology*: 1–7.
Ferrari, A.M., Degliesposti, G., Sgobba, M. & Rastelli, G. 2007. Validation of an automated procedure for the prediction of relative free energies of binding on a set of aldose reductase inhibitors. *Bioorganic & medicinal chemistry*, 15(24): 7865–7877.
Ferri, N., Radice, T., Antonino, M., Beccalli, E.M., Tinelli, S., Zunino, F., Corsini, A., Pratesi, G., Ragg, E.M. & Gelmi, M.L. 2011. Synthesis, structural, and biological evaluation of bis-heteroarylmaleimides and bis-heterofused imides. *Bioorganic & medicinal chemistry*, 19(18): 5291–5299.
Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., Repasky, M.P., Knoll, E.H., Shelley, M. & Perry, J.K. 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry*, 47(7): 1739–1749.
García-Nafría, J. & Tate, C.G. 2020. Cryo-electron microscopy: Moving beyond X-ray crystal structures for drug receptors and drug development. *Annual Review of Pharmacology and Toxicology*, 60: 51–71.
Ghahremanpour, M.M., Arab, S.S., Aghazadeh, S.B., Zhang, J. & van der Spoel, D. 2014. MemBuilder: a web-based graphical interface to build heterogeneously mixed membrane bilayers for the GROMACS biomolecular simulation program. *Bioinformatics*, 30(3): 439–441.
Gould, I., Skjevik, A., Dickson, C., Madej, B. & Walker, R. 2018. Lipid17: A Comprehensive AMBER Force Field for the Simulation of Zwitterionic and Anionic Lipids. *Manuscript in preparation*.
Grant, B.J., Rodrigues, A.P.C., ElSawy, K.M., McCammon, J.A. & Caves, L.S.D. 2006. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics*, 22(21): 2695–2696.
Habgood, M., James, T. & Heifetz, A. 2020. Conformational Searching with Quantum Mechanics. In *Quantum Mechanics in Drug Discovery*. Springer: 207–229.
Halgren, T.A. 1996a. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of computational chemistry*, 17(5-6): 490–519.
Halgren, T.A. 1996b. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *Journal of Computational Chemistry*, 17(5-6): 520–552.
Halgren, T.A. 1996c. Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94. *Journal of computational chemistry*, 17(5-6): 553–586.
Halgren, T.A. 1996d. Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules. *Journal of Computational Chemistry*, 17(5-6): 616–641.
Halgren, T.A. & Nachbar, R.B. 1996. Merck molecular force field. IV. Conformational energies and geometries for MMFF94. *Journal of computational chemistry*, 17(5-6): 587–615.
Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J.Y., Wang, L., Lupyan, D., Dahlgren, M.K. & Knight, J.L. 2016. OPLS3: a force field providing broad coverage of drug-like small molecules and proteins. *Journal of chemical theory and computation*, 12(1): 281–296.
Henriques, J., Cragnell, C. & Skepö, M. 2015. Molecular dynamics simulations of intrinsically disordered proteins: force field evaluation and comparison with experiment. *Journal of chemical theory and computation*, 11(7): 3420–3431.
Hess, B., Bekker, H., Berendsen, H.J.C. & Fraaije, J.G.E.M. 1997. LINCS: a linear constraint solver for molecular simulations. *Journal of computational chemistry*, 18(12): 1463–1472.

Hoover, W.G. 1985. Canonical dynamics: Equilibrium phase-space distributions. *Physical review A*, 31(3): 1695.

Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A. & Simmerling, C. 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, 65(3): 712–725.

Horta, B.A.C., Fuchs, P.F.J., van Gunsteren, W.F. & Hünenberger, P.H. 2011. New interaction parameters for oxygen compounds in the GROMOS force field: Improved pure-liquid and solvation properties for alcohols, ethers, aldehydes, ketones, carboxylic acids, and esters. *Journal of chemical theory and computation*, 7(4): 1016–1031.

Horta, B.A.C., Merz, P.T., Fuchs, P.F.J., Dolenc, J., Riniker, S. & Hünenberger, P.H. 2016. A GROMOS-compatible force field for small organic molecules in the condensed phase: The 2016H66 parameter set. *Journal of chemical theory and computation*, 12(8): 3825–3850.

Hospital, A., Goñi, J.R., Orozco, M. & Gelpí, J.L. 2015. Molecular dynamics simulations: advances and applications. *Advances and applications in bioinformatics and chemistry: AABC*, 8: 37.

Hou, T., Wang, J., Li, Y. & Wang, W. 2011. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of Chemical Information and Modeling*, 51(1): 69–82.

Huang, S.-M., Abernethy, D.R., Wang, Y., Zhao, P. & Zineh, I. 2013. The utility of modeling and simulation in drug development and regulatory review. *Journal of pharmaceutical sciences*, 102(9): 2912–2923.

Humphrey, W., Dalke, A. & Schulten, K. 1996. VMD: visual molecular dynamics. *Journal of molecular graphics*, 14(1): 33–38.

Ioakimidis, L., Thoukydidis, L., Mirza, A., Naeem, S. & Reynisson, J. 2008. Benchmarking the reliability of QikProp. Correlation between experimental and predicted values. *QSAR & Combinatorial Science*, 27(4): 445–456.

Jakalian, A., Bush, B.L., Jack, D.B. & Bayly, C.I. 2000. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. *Journal of computational chemistry*, 21(2): 132–146.

Jakalian, A., Jack, D.B. & Bayly, C.I. 2002. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *Journal of computational chemistry*, 23(16): 1623–1641.

Jo, S., Kim, T., Iyer, V.G. & Im, W. 2008. CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of computational chemistry*, 29(11): 1859–1865.

Jójárt, B. & Martinek, T.A. 2007. Performance of the general amber force field in modeling aqueous POPC membrane bilayers. *Journal of computational chemistry*, 28(12): 2051–2058.

Jorgensen, W.L., Maxwell, D.S. & Tirado-Rives, J. 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118(45): 11225–11236.

Khanna, T. 2018. Development and application of AMBER molecular mechanics force field for herbicide interaction in plants. : 1–240. https://spiral.imperial.ac.uk:8443/handle/10044/1/75143.

Klauda, J.B., Venable, R.M., Freites, J.A., O'Connor, J.W., Tobias, D.J., Mondragon-Ramirez, C., Vorobyov, I., MacKerell Jr, A.D. & Pastor, R.W. 2010. Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *The journal of physical chemistry B*, 114(23): 7830–7843.

Kramer, B., Rarey, M. & Lengauer, T. 1999. Evaluation of the FLEXX incremental

construction algorithm for protein–ligand docking. *Proteins: Structure, Function, and Bioinformatics*, 37(2): 228–241.
Krishnan, V. V & Rupp, B. 2012. Macromolecular structure determination: comparison of X-ray crystallography and NMR spectroscopy. *eLS*.
Kuhn, B., Gerber, P., Schulz-Gasch, T. & Stahl, M. 2005. Validation and use of the MM-PBSA approach for drug discovery. *Journal of medicinal chemistry*, 48(12): 4040–4048.
Lang, P.T., Brozell, S.R., Mukherjee, S., Pettersen, E.F., Meng, E.C., Thomas, V., Rizzo, R.C., Case, D.A., James, T.L. & Kuntz, I.D. 2009. DOCK 6: Combining techniques to model RNA–small molecule complexes. *Rna*, 15(6): 1219–1230.
Lee, S., Tran, A., Allsopp, M., Lim, J.B., Hénin, J. & Klauda, J.B. 2014. CHARMM36 united atom chain model for lipids and surfactants. *The journal of physical chemistry B*, 118(2): 547–556.
Lenselink, E.B., Louvel, J., Forti, A.F., van Veldhoven, J.P.D., de Vries, H., Mulder-Krieger, T., McRobb, F.M., Negri, A., Goose, J. & Abel, R. 2016. Predicting binding affinities for GPCR ligands using free-energy perturbation. *ACS omega*, 1(2): 293–304.
Levine, Z.A. & Shea, J.-E. 2017. Simulations of disordered proteins and systems with conformational heterogeneity. *Current opinion in structural biology*, 43: 95–103.
Li, D. & Brüschweiler, R. 2010. NMR-based protein potentials. *Angewandte Chemie*, 122(38): 6930–6932.
Li, L., Chen, R. & Weng, Z. 2003. RDOCK: refinement of rigid-body protein docking predictions. *Proteins: Structure, Function, and Bioinformatics*, 53(3): 693–707.
Li, Q. 2020. Application of fragment-based drug discovery to versatile targets. *Frontiers in Molecular Biosciences*, 7.
Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J.L., Dror, R.O. & Shaw, D.E. 2010. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function, and Bioinformatics*, 78(8): 1950–1958.
Lionta, E., Spyrou, G., K Vassilatis, D. & Cournia, Z. 2014. Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Current topics in medicinal chemistry*, 14(16): 1923–1938.
Liu, J. & Wang, R. 2015. Classification of current scoring functions. *Journal of chemical information and modeling*, 55(3): 475–482.
Liu, L. & Jockers, R. 2020. Structure-Based Virtual Screening Accelerates GPCR Drug Discovery. *Trends in Pharmacological Sciences*.
Liu, S., Alnammi, M., Ericksen, S.S., Voter, A.F., Ananiev, G.E., Keck, J.L., Hoffmann, F.M., Wildman, S.A. & Gitter, A. 2018. Practical model selection for prospective virtual screening. *Journal of chemical information and modeling*, 59(1): 282–293.
Liu, X., Shi, D., Zhou, S., Liu, Hongli, Liu, Huanxiang & Yao, X. 2018. Molecular dynamics simulations and novel drug discovery. *Expert opinion on drug discovery*, 13(1): 23–37.
Loschwitz, J., Olubiyi, O.O., Hub, J.S., Strodel, B. & Poojari, C.S. 2020. Computer simulations of protein–membrane systems. *Progress in Molecular Biology and Translational Science*, 170: 273.
Mackerell Jr, A.D., Feig, M. & Brooks III, C.L. 2004. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of computational chemistry*, 25(11): 1400–1415.
Maia, E.H.B., Assis, L.C., de Oliveira, T.A., da Silva, A.M. & Taranto, A.G. 2020. Structure-based virtual screening: from classical to artificial intelligence. *Frontiers in Chemistry*, 8.
Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E. & Simmerling, C. 2015. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of chemical theory and computation*, 11(8): 3696–3713.

Manhas, A., Patel, D., Lone, M.Y. & Jha, P.C. 2019. Identification of natural compound inhibitors against PfDXR: A hybrid structure-based molecular modeling approach and molecular dynamics simulation studies. *Journal of cellular biochemistry*, 120(9): 14531–14543.

Marrink, S.J., Risselada, H.J., Yefimov, S., Tieleman, D.P. & De Vries, A.H. 2007. The MARTINI force field: coarse grained model for biomolecular simulations. *The journal of physical chemistry B*, 111(27): 7812–7824.

Martínez, L., Andrade, R., Birgin, E.G. & Martínez, J.M. 2009. PACKMOL: a package for building initial configurations for molecular dynamics simulations. *Journal of computational chemistry*, 30(13): 2157–2164.

McKiernan, K.A., Husic, B.E. & Pande, V.S. 2017. Modeling the mechanism of CLN025 beta-hairpin formation. *The Journal of chemical physics*, 147(10): 104107.

Meng, X.-Y., Zhang, H.-X., Mezei, M. & Cui, M. 2011. Molecular docking: a powerful approach for structure-based drug discovery. *Current computer-aided drug design*, 7(2): 146–157.

Miller III, B.R., McGee Jr, T.D., Swails, J.M., Homeyer, N., Gohlke, H. & Roitberg, A.E. 2012. MMPBSA. py: an efficient program for end-state free energy calculations. *Journal of chemical theory and computation*, 8(9): 3314–3321.

Monticelli, L. & Tieleman, D.P. 2013. Force fields for classical molecular dynamics. In *Biomolecular simulations*. Springer: 197–213.

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S. & Olson, A.J. 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry*, 30(16): 2785–2791.

Murray, C.W., Verdonk, M.L. & Rees, D.C. 2012. Experiences in fragment-based drug discovery. *Trends in pharmacological sciences*, 33(5): 224–232.

Nerenberg, P.S. & Head-Gordon, T. 2018. New developments in force fields for biomolecular simulations. *Current opinion in structural biology*, 49: 129–138.

Newport, T.D., Sansom, M.S.P. & Stansfeld, P.J. 2019. The MemProtMD database: a resource for membrane-embedded protein structures and their lipid interactions. *Nucleic acids research*, 47(D1): D390–D397.

Nosé, S. 1984. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of chemical physics*, 81(1): 511–519.

Oostenbrink, C., Villa, A., Mark, A.E. & Van Gunsteren, W.F. 2004. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *Journal of computational chemistry*, 25(13): 1656–1676.

Pearlman, D.A. 2005. Evaluating the molecular mechanics Poisson− Boltzmann surface area free energy method using a congeneric series of ligands to p38 MAP kinase. *Journal of medicinal chemistry*, 48(24): 7796–7807.

Pearlman, D.A. & Charifson, P.S. 2001. Are free energy calculations useful in practice? A comparison with rapid scoring functions for the p38 MAP kinase protein system. *Journal of medicinal chemistry*, 44(21): 3417–3423.

Piggot, T.J., Piñeiro, A. & Khalid, S. 2012. Molecular dynamics simulations of phosphatidylcholine membranes: a comparative force field study. *Journal of chemical theory and computation*, 8(11): 4593–4609.

Pires, D.E. V, Blundell, T.L. & Ascher, D.B. 2015. pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *Journal of medicinal chemistry*, 58(9): 4066–4072.

Rastelli, G., Rio, A. Del, Degliesposti, G. & Sgobba, M. 2010. Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *Journal of computational chemistry*, 31(4): 797–810.

Rauscher, S., Gapsys, V., Gajda, M.J., Zweckstetter, M., de Groot, B.L. & Grubmüller, H. 2015. Structural ensembles of intrinsically disordered proteins depend strongly on force field: a comparison to experiment. *Journal of chemical theory and computation*, 11(11): 5513–5524.

Rifai, E.A., van Dijk, M., Vermeulen, N.P.E. & Geerke, D.P. 2018. Binding free energy predictions of farnesoid X receptor (FXR) agonists using a linear interaction energy (LIE) approach with reliability estimation: application to the D3R Grand Challenge 2. *Journal of computer-aided molecular design*, 32(1): 239–249.

Robertson, M.J., Tirado-Rives, J. & Jorgensen, W.L. 2015. Improved peptide and protein torsional energetics with the OPLS-AA force field. *Journal of chemical theory and computation*, 11(7): 3499–3509.

Rutenber, E.E. & Stroud, R.M. 1996. Binding of the anticancer drug ZD1694 to E. coli thymidylate synthase: assessing specificity and affinity. *Structure*, 4(11): 1317–1324.

Ryckaert, J.-P., Ciccotti+, G. & Berendsen, H.J.C. 1977. *Numerical integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes*.

Saura, P., Röpke, M., Gamiz-Hernandez, A.P. & Kaila, V.R.I. 2019. Quantum Chemical and QM/MM Models in Biochemistry. In *Biomolecular Simulations*. Springer: 75–104.

Saurabh, S., Sivakumar, P.M., Perumal, V., Khosravi, A., Sugumaran, A. & Prabhawathi, V. 2020. Molecular Dynamics Simulations in Drug Discovery and Drug Delivery. In *Integrative Nanomedicine for New Therapies*. Springer: 275–301.

Schmid, N., Eichenberger, A.P., Choutko, A., Riniker, S., Winger, M., Mark, A.E. & van Gunsteren, W.F. 2011. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *European biophysics journal*, 40(7): 843–856.

Schott-Verdugo, S. & Gohlke, H. 2019. PACKMOL-memgen: a simple-to-use, generalized workflow for membrane-protein–lipid-bilayer system building. *Journal of chemical information and modeling*, 59(6): 2522–2528.

Schuler, L.D., Daura, X. & Van Gunsteren, W.F. 2001. An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase. *Journal of computational chemistry*, 22(11): 1205–1218.

Schultz, P.A. 1999. Local electrostatic moments and periodic boundary conditions. *Physical Review B*, 60(3): 1551.

Shen, C., Ding, J., Wang, Z., Cao, D., Ding, X. & Hou, T. 2020. From machine learning to deep learning: Advances in scoring functions for protein–ligand docking. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 10(1): e1429.

Shen, C., Wang, Z., Yao, X., Li, Y., Lei, T., Wang, E., Xu, L., Zhu, F., Li, D. & Hou, T. 2020. Comprehensive assessment of nine docking programs on type II kinase inhibitors: prediction accuracy of sampling power, scoring power and screening power. *Briefings in bioinformatics*, 21(1): 282–297.

Singh, N. & Warshel, A. 2010. Absolute binding free energy calculations: On the accuracy of computational scoring of protein–ligand interactions. *Proteins: Structure, Function, and Bioinformatics*, 78(7): 1705–1723.

Skjevik, Å.A., Madej, B.D., Walker, R.C. & Teigen, K. 2012. LIPID11: a modular framework for lipid simulations using amber. *The Journal of Physical Chemistry B*, 116(36): 11124–11136.

Srivastava, H.K. & Sastry, G.N. 2012. Molecular dynamics investigation on a series of HIV protease inhibitors: assessing the performance of MM-PBSA and MM-GBSA approaches. *Journal of chemical information and modeling*, 52(11): 3088–3098.

Steinbrecher, T. & Elstner, M. 2013. QM and QM/MM simulations of proteins. In *Biomolecular Simulations*. Springer: 91–124.

Steinbrecher, T., Zhu, C., Wang, L., Abel, R., Negron, C., Pearlman, D., Feyfant, E., Duan, J. & Sherman, W. 2017. Predicting the effect of amino acid single-point mutations on protein stability—large-scale validation of MD-based relative free energy calculations. *Journal of molecular biology*, 429(7): 948–963.

Toroz, D., Khanna, T. & Gould, I.R. 2019. Modeling the Effect of BSEP Inhibitors in Lipid Bilayers by Means of All-Atom Molecular Dynamics Simulation. *ACS Omega*, 4(2): 3341–3350.

Torrens-Fontanals, M., Stepniewski, T.M., Aranda-García, D., Morales-Pastor, A., Medel-Lacruz, B. & Selent, J. 2020. How Do Molecular Dynamics Data Complement Static Structural Data of GPCRs. *International Journal of Molecular Sciences*, 21(16): 5933.

Trott, O. & Olson, A.J. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2): 455–461.

Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P. & Vorobyov, I. 2010. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of computational chemistry*, 31(4): 671–690.

Vanommeslaeghe, K. & MacKerell Jr, A.D. 2012. Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *Journal of chemical information and modeling*, 52(12): 3144–3154.

Vanommeslaeghe, K., Raman, E.P. & MacKerell Jr, A.D. 2012. Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *Journal of chemical information and modeling*, 52(12): 3155–3168.

Verdonk, M.L., Cole, J.C., Hartshorn, M.J., Murray, C.W. & Taylor, R.D. 2003. Improved protein–ligand docking using GOLD. *Proteins: Structure, Function, and Bioinformatics*, 52(4): 609–623.

Walker, R.C. 2008. Molecular Mechanics with a Poisson-Boltzmann/Surface Area solvent: MM-PBSA. http://ambermd.org/tutorials/advanced/tutorial3/index.php 16 November 2020.

Wang, J. & Collis, A. 2011. Maximizing the outcome of early ADMET models: strategies to win the drug-hunting battles? *Expert Opinion on Drug Metabolism & Toxicology*, 7(4): 381–386.

Wang, J., Morin, P., Wang, W. & Kollman, P.A. 2001. Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. *Journal of the American Chemical Society*, 123(22): 5221–5230.

Wang, J., Wang, W., Kollman, P.A. & Case, D.A. 2006. Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of molecular graphics and modelling*, 25(2): 247–260.

Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. & Case, D.A. 2004. Development and testing of a general amber force field. *Journal of computational chemistry*, 25(9): 1157–1174.

Wang, L.-P., McKiernan, K.A., Gomes, J., Beauchamp, K.A., Head-Gordon, T., Rice, J.E., Swope, W.C., Martínez, T.J. & Pande, V.S. 2017. Building a more predictive protein force field: a systematic and reproducible route to AMBER-FB15. *The Journal of Physical Chemistry B*, 121(16): 4023–4039.

Wang, Y., Xing, J., Xu, Y., Zhou, N., Peng, J., Xiong, Z., Liu, X., Luo, X., Luo, C. & Chen, K. 2015. In silico ADME/T modelling for rational drug design. *Quarterly reviews of biophysics*, 48(4): 488–515.

Wang, Z., Sun, H., Shen, C., Hu, X., Gao, J., Li, D., Cao, D. & Hou, T. 2020. Combined

strategies in structure-based virtual screening. *Physical Chemistry Chemical Physics*, 22(6): 3149–3159.

Wikipedia. 2015. Docking (molecular). https://en.wikipedia.org/wiki/Docking_(molecular) 1 October 2020.

Wlodawer, A. & Vondrasek, J. 1998. Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. *Annual review of biophysics and biomolecular structure*, 27(1): 249–284.

Wu, E.L., Cheng, X., Jo, S., Rui, H., Song, K.C., Dávila-Contreras, E.M., Qi, Y., Lee, J., Monje-Galvan, V. & Venable, R.M. 2014. CHARMM-GUI membrane builder toward realistic biological membrane simulations. *Journal of computational chemistry*, 35(27): 1997–2004.

Yau, M.Q., Emtage, A.L., Chan, N.J.Y., Doughty, S.W. & Loo, J.S.E. 2019. Evaluating the performance of MM/PBSA for binding affinity prediction using class A GPCR crystal structures. *Journal of computer-aided molecular design*, 33(5): 487–496.

Yau, M.Q., Emtage, A.L. & Loo, J.S.E. 2020. Benchmarking the performance of MM/PBSA in virtual screening enrichment using the GPCR-Bench dataset. *Journal of computer-aided molecular design*, 34(11): 1133–1145.

Yu, W., He, X., Vanommeslaeghe, K. & MacKerell Jr, A.D. 2012. Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *Journal of computational chemistry*, 33(31): 2451–2468.

CHAPTER 4

Published Article

Probing binding landscapes and molecular recognition mechanisms of atypical antipsychotic drugs towards the selective targeting of D_2 dopamine receptor

Patrick Appiah-Kubi[1], Fisayo Andrew Olotu[1], and Mahmoud E.S Soliman[1*]

[1]Molecular Bio-computation and Drug Design Laboratory

School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

Telephone: +27 (0) 31 260 8048

Fax: +27 (0) 31 260 78

Email: soliman@ukzn.ac.za

Website: http://soliman.ukzn.ac.za

Abstract

Dopamine receptors constitute a unique class of G-protein coupled receptors that mediate the activities of dopamine, a neurotransmitter implicated in diverse neurological diseases when dysregulated. Over the years, antipsychotic drugs have been primarily directed towards D_2 dopamine receptor (DRD2) while associable adverse effects have been centred on non-selective targeting. The recent crystal structure of DRD2 in complex with atypical antipsychotic could further aid the structure-based design of highly DRD2-selective antipsychotics. Therefore, in this study, we comprehensively investigate the molecular recognition and differential binding landscapes of class-I and II DRD2 atypical antipsychotics, using membrane-bilayer molecular dynamics simulation and binding free energy techniques. Findings revealed that selected class-I antipsychotics exhibited binding dynamics and poses dissimilar to the class-II types with different interactive mechanisms at the binding cavity of DRD2. More interestingly, the class-II drugs established a highly coordinated binding at the DRD2 active site with a pertinent and recurrent involvement of Asp114 via strong hydrogen interactions. Furthermore, while these compounds exert distinct effects on DRD2 structure, findings revealed that the class-II types favourably engaged the deep hydrophobic pocket of DRD2 compared to the class-I drugs. We speculate that these findings will be fundamental to the discovery of highly selective DRD2 antipsychotics.

Keywords: Atypical antipsychotics, D_2 dopamine receptor, G-protein coupled receptor, Membrane-bound protein, lipid bilayer Molecular dynamics simulation.

1. Introduction

G protein-coupled receptors (GPCRs) play an important role in the cell signalling process, particularly in response to neurotransmitters and hormones (Congreve & Marshall, 2010). Dopamine is an important neurotransmitter exerting its effect via the activation of the five dopamine receptors subtypes belonging to the GPCR family, namely; D_1, D_2, D_3, D4, and D_5 (Kebabian & Calne, 1979). These receptors are further grouped into two main subfamilies, D_1-like (D_1 and D_5) and D_2-like (D_2, D_3, and D_4) receptors, based on their sequence and structural similarities (Fuxe et al., 2014). Generally, these class of receptors mediate cell proliferation, and differentiation, cyclic adenosine monophosphate release and neurotransmission of other neurotransmitters (Mishra et al., 2018).

Amongst these subtypes, D_2 dopamine receptor (DRD2) is an important GPCR drug target that mediates signal transduction of neurotransmitters in the central nervous system as well as in neurological processes such as memory, attention, emotion, pleasure, lust, and love (Greengard, 2001). The dysfunction of the dopaminergic system has been associated with various neurodegenerative disorders such as depression, Parkinson's disease, attention deficit hyperactivity disorder, and schizophrenia. DRD2 has been the main target for atypical and typical antipsychotic drugs (Creese et al., 1976),(Meltzer et al., 1989) as well as drugs employed in the treatment of Parkinson disease. Unfortunately, many of these antipsychotic drugs targeting DRD2 are accompanied by a plethora of severe side effects, due to off-target interactions with other related targets (Roth et al., 2004),(Roth, 2007).

The crystal structures of D_4 dopamine receptor (DRD4) in complex with antipsychotic nemonapride (Wang et al., 2017) and D_3 dopamine receptor (DRD3) in complex with eticlopride (Chien et al., 2010) have earlier been resolved. However, the lack of crystalised structures of the DRD2 ligand complexes over the past years impeded the molecular understanding of receptor function and ligand recognition. The recent crystallisation of the DRD2 in complex with the widely prescribed antipsychotic drug risperidone in 2018 provided novel insights (Wang et al., 2018) (Figure 1). Previous studies of DRD2 have been based on homology modelled structures using either DRD3 or DRD4 crystal structures as templates (Montgomery et al., 2018; Podder et al., 2016; Salmas et al., 2017).

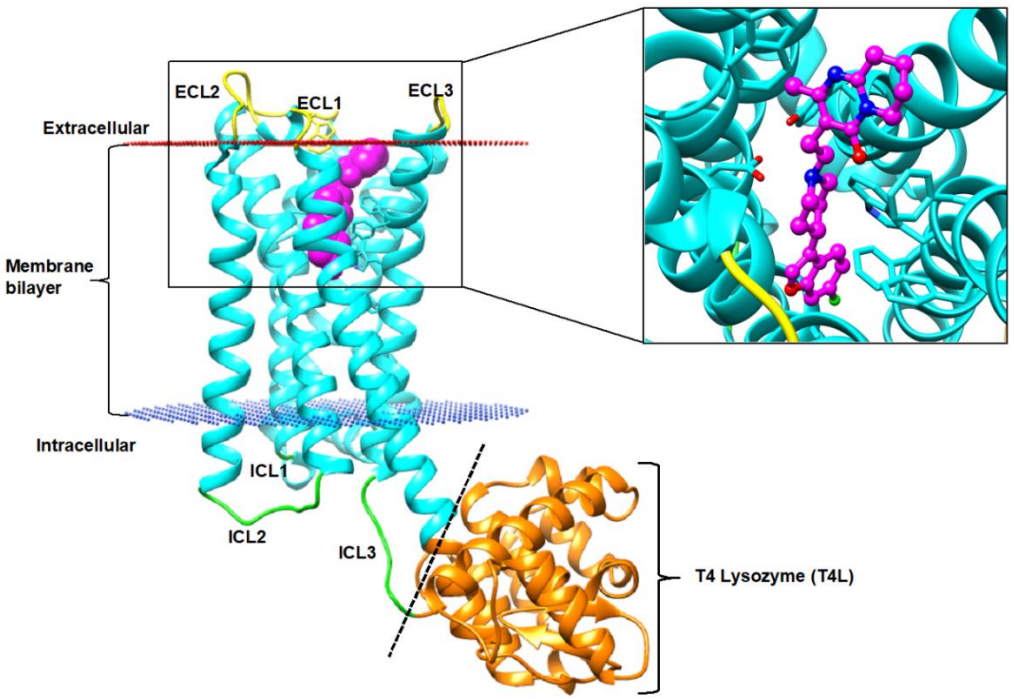


Figure 1. Crystal structure of dopamine 2 receptor in complex with risperidone binding deep into the active site. Extracellular loops (ECL), intracellular loops (ICL), T4-Lysozyme (T4L) and risperidone are depicted in yellow, green, orange, and magenta, respectively (Image prepared by author).

The solved crystal structure of DRD2 revealed an unexpected binding mode of risperidone which was contrary to findings from previous molecular docking studies that employed a non-crystal homology model of DRD2 and used DRD3/DRD4 as structural templates (Duan et al., 2015; Salmas et al., 2017). This approach failed to replicate the unique binding mode of risperidone as observed in the recent DRD2 crystal structure (Wang et al., 2018), which revealed that risperidone binds differently to DRD2. In the risperidone-DRD2 crystalised complex, the fluorobenzisoxazol ring orients deep in the hydrophobic cleft interacting with Trp386, Phe390, Phe382, Phe198, Ser197, Thr119 and Cys118 whereas the tetrahydropyridopyrimidinone ring interacts with residues Phe110, Thr412, Trp100, Tyr408 and Val191 (Wang et al., 2018). Interestingly, in the recent crystal structure of serotonin 2A receptor (5-HT_{2A}R) in complex with risperidone substantiates the observed binding mode in DRD2; the fluorobenzisoxazol ring of risperidone binds in the bottom of the hydrophobic cleft forming interactions with Trp336, Ile163, Phe340, Phe243, and Phe332 whereas the basic nitrogen of risperidone forms a salt bridge with the conserved residue Asp155 of 5-HT_{2A}R (Kimura et al., 2019). These, however, contradict the reported reversed binding mode of risperidone using DRD2 homology model (Duan et al., 2015; Salmas et al., 2017). Therefore,

the binding modes of related atypical antipsychotic DRD2-targeting drugs remain unclear, and to make significant advancements in the structure-based design of these class of drugs. There is a need to identify definite binding modes and mechanisms for risperidone and related atypical antipsychotics at the binding cavity of DRD2.

Understanding the binding modes and activation mechanisms of atypical antipsychotics at the binding cavity of DRD2 may assist in the rational-based drug design of highly selective drugs with minimal off-target tendencies. To further clarify the binding mechanisms of DRD2 antagonists, we performed a comparative study of six (6) well known DRD2 atypical antipsychotic drugs to elucidate the mechanisms of ligand recognition, binding theme and conformational changes induced by these antipsychotics. The present study utilised molecular docking, explicit membrane-bound protein lipid bilayer molecular dynamics simulation, protein-ligand interaction fingerprint and binding energy analyses approach. Findings from this study will provide a further understanding of DRD2 ligand recognition and to guide future development of novel DRD2 selective as well as safer antagonists.

2. Computational Methods

2.1 Ligand structure preparations

The 2D structures of DRD2 antagonists; aripiprazole (ZINC01851149), clozapine (ZINC19796155), olanzapine (ZINC52957434), quetiapine (ZINC19632628), ziprasidone (ZINC00538550) and risperidone (ZINC00538312) were downloaded from the ZINC database for molecular docking (Irwin et al., 2012). Risperidone was also extracted from the dopamine crystalised structure (PDB ID: 6CM4) (Wang et al., 2018) from the Protein Data Bank (http://www.rcsb.org/) for the simulation. The retrieved compounds were prepared and minimised at physiological pH (7.4) using LigPrep module of Schrodinger suite software (Schrödinger, 2015). Figure 2 shows the 2D structures of the compounds employed in this study.

2.2 Molecular Docking Analysis

Molecular docking experiment was performed using Autodock Vina software package (Trott & Olson, 2010). The software's default settings were used to rank the ligand conformation based on their predicted binding affinities during the docking experiments. Molecular docking was performed following receptor grid generation around the protein active site residues. The grid box was defined around the crystallised bound ligand (risperidone) active site residues. The grid box's X, Y and Z dimensions were defined as 17.4531, 16.8919, and 19.9398

respectively whereas the X, Y, and Z centres were defined by 9.69473, 5.99816, and -8.89723 respectively. Polar hydrogens and Gasteiger charge were assigned for all compounds. The best docking pose was output based on the best binding score. The 2D structures of risperidone and the related antipsychotics used in this study were docked to the crystal structure using the same grid box. The docked complex structures were visualised by the UCSF chimera (Pettersen et al., 2004) graphical user interface, and Molecular interactions were also visualised using Molegro Molecular Viewer (MMV) software program (http:www.clcbio.com).

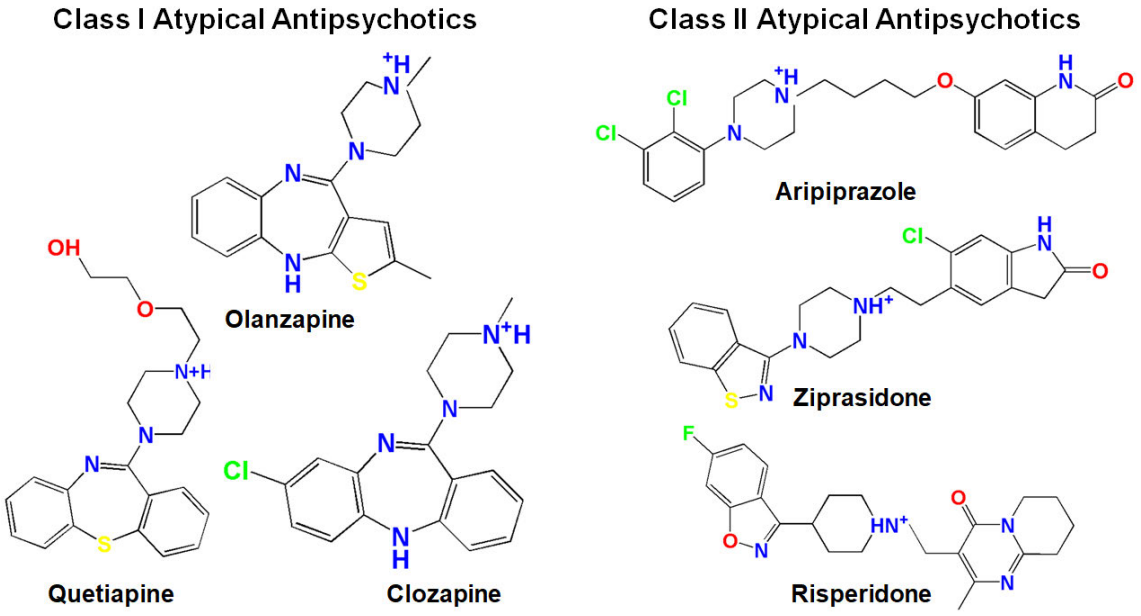


Figure 2. 2D structures of dopamine 2 receptor atypical antipsychotic drugs employed in the molecular dynamics simulations (Image prepared by author).

2.3 Protein-Membrane System Setup

The initial coordinates of the protein were taken from the Protein Data Bank (PDB ID: 6cm4) (Wang et al., 2018). The protonation state of the protein was determined using PROPKA (Dolinsky et al., 2007) at a 7.4 physiological pH. Missing loops were modelled using Modeller 9.19 (Webb & Sali, 2014) plugins in chimera software. The T4-Lysozyme (T4L) residues (Val223-Arg361) fused in the ICL3 were deleted and the exposed residues Arg222 and Lys362 capped. The protein-membrane bilayer and all the ligand-protein-membrane bilayers were assembled using the membrane builder module at the CHARMM-GUI website (http://www.charmm-gui.org/). The protein-membrane and the ligand-protein-membrane systems were aligned in dipalmitoylphosphatidylcholine (DPPC) lipid bilayer with knowledge of the protein orientation in the bilayer obtained from the Orientation of Protein Membranes (OPM) database (Lomize et al., 2006). The complex systems were solvated with TIP3P water

molecules (Jorgensen et al., 1983) and neutralised with KCl counterions at a concentration of 0.15M (Figure 3). The dimensions of the final simulated systems approximately measured 75 x 75 x 100 $Å^3$ and approximately contained 150 DPPC lipid molecules, 26 potassium ions (K^+), 37 chloride ions (Cl^-), and 10,000 TIP3P water molecules.

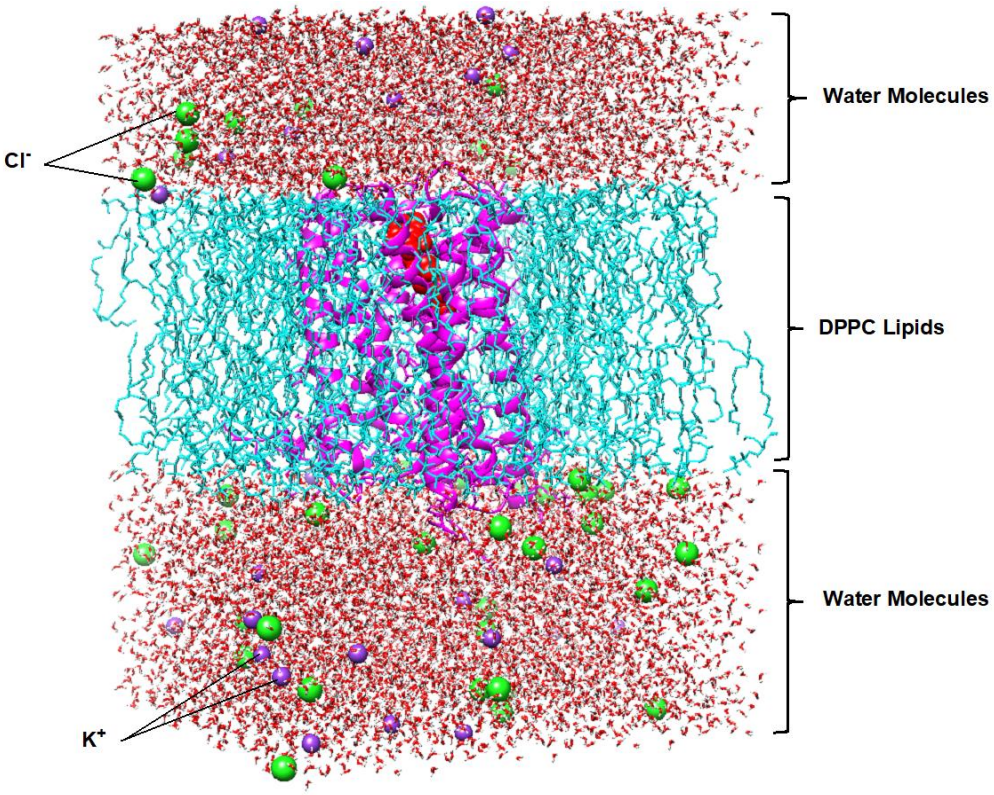


Figure 3. D_2 dopamine receptor (purple) in complex with atypical antagonist risperidone (red) embedded in DPPC lipid bilayer (cyan) (Image prepared by author).

2.4 Molecular Dynamic (MD) Simulations

The AMBER18 package (Case et al., 2018) was used to perform all MD simulations. The solvated complex systems were used as the starting structures for the MD simulations. The charmmlipid2amber.py program in AmberTools 18 was used to convert the CHARMM lipid residue names to Amber format. The general Amber force field (GAFF) (Sprenger et al., 2015), Amber force field ff14B (Maier et al., 2015) and the Lipid14 (Dickson et al., 2014) force field were employed for the ligands, protein, and lipid molecules parameterisation, respectively. K^+ and Cl ions were parameterised using the Joung-Cheatham TIP3P ion parameters (Joung & Cheatham III, 2008). The RESP model (Bayly et al., 1993) was used to derive the partial atomic charges of the ligands and electrostatic potentials generated at the HF/6–31G* level in the

GAUSSIAN 16 software package. The critical Cys182 and Cys107 disulfide bond were maintained during the simulations.

All simulations were performed on Amber18 using the Particle Mesh Ewald Molecular Dynamics (PMEMD) program. The prepared systems were energy minimised and equilibrated as follow: Energy minimisation of 10000 steps by hybrid methods of 5000 steps of steepest descent and remaining of conjugate gradient algorithm to relax the systems to remove any possible steric clashes. Individual systems were heated from 0 K to 100 K in isothermal-isochoric (NVT) ensemble using the Langevin thermostat (Larini et al., 2007) for 12.5 ps applying a 10.0 kcal mol^{-1} $Å^{-2}$ harmonic restraint on non-hydrogen atoms of protein, ligand, and lipid with a collision frequency of 1.0 ps^{-1}. Each system was further heated from 100 K to 310 K for 125 ps with anisotropic pressure scaling and a pressure of 1 bar in isothermal-isobaric (NPT) ensemble. Equilibration was further performed for a total of 10 ns at 310 K in NPT ensemble. A 2fs time step was used for heating and equilibration runs. Long-range electrostatic interactions were handled by particle mesh Ewald (PME) (Darden et al., 1999) algorithm, and non-bonded interactions were cut off at 10.0 Å. Covalent bonds involving hydrogen and heavy atoms were constrained with the SHAKE algorithm (Miyamoto & Kollman, 1992). Finally, a 200ns production run was carried out for each system at 310K and 1 bar in NPT ensembles without any restraint at a time step of 2fs and a cut-off of 10Å. Simulation snapshots were saved every 1ps.

2.5 Post-MD Trajectories Analysis

Prior to the analysis of trajectories, water molecules, counter ions and DPPC lipid molecules were stripped from the trajectories. The CPPTRAJ module (Roe & Cheatham, 2013) of AMBER18 was used to perform the analysis of the trajectories. Figures were prepared using UCSF Chimera (Pettersen et al., 2004).

2.6 MM-GBSA Binding Free Energy Analysis

The estimation of binding free energy for atypical antipsychotic was obtained from the stable MD trajectories using MM-GBSA method implemented in Amber18 package (Case et al., 2018). 4000 snapshots were evenly extracted over the 200 ns of the MD trajectories at an interval of 50ps. Prior to MM-GBSA calculations, all water molecules, counterions, and lipids were stripped. Conceptual summary of the MM-GBSA method can be found in our previous paper (Appiah-Kubi & Soliman, 2016).

2.7 Residue interaction energy analysis

The energy contribution of the active site residues to the overall binding energy of the studied antipsychotics was also decomposed from their total binding energies.

3. Results and Discussion

3.1 Binding poses of atypical drugs to DRD2

Molecular docking analysis was performed to predict the binding poses and affinities of aripiprazole, clozapine, olanzapine, quetiapine, risperidone, and ziprasidone (Figure 2 & 4) at the binding site of the crystalised D_2 dopamine receptor. Auto Dock Vina (Trott & Olson, 2010), a molecular docking tool has successfully been employed in finding the appropriate binding poses/modes and affinities of ligand molecules inside the binding/active site of proteins (Nisha et al., 2016; Kumar et al., 2018; Seyedi et al., 2016). Prior to the docking of related atypical antipsychotics, the downloaded risperidone structure (ZINC00538312) from the ZINC database was initially docked into the DRD2 receptor (Figure 4b).

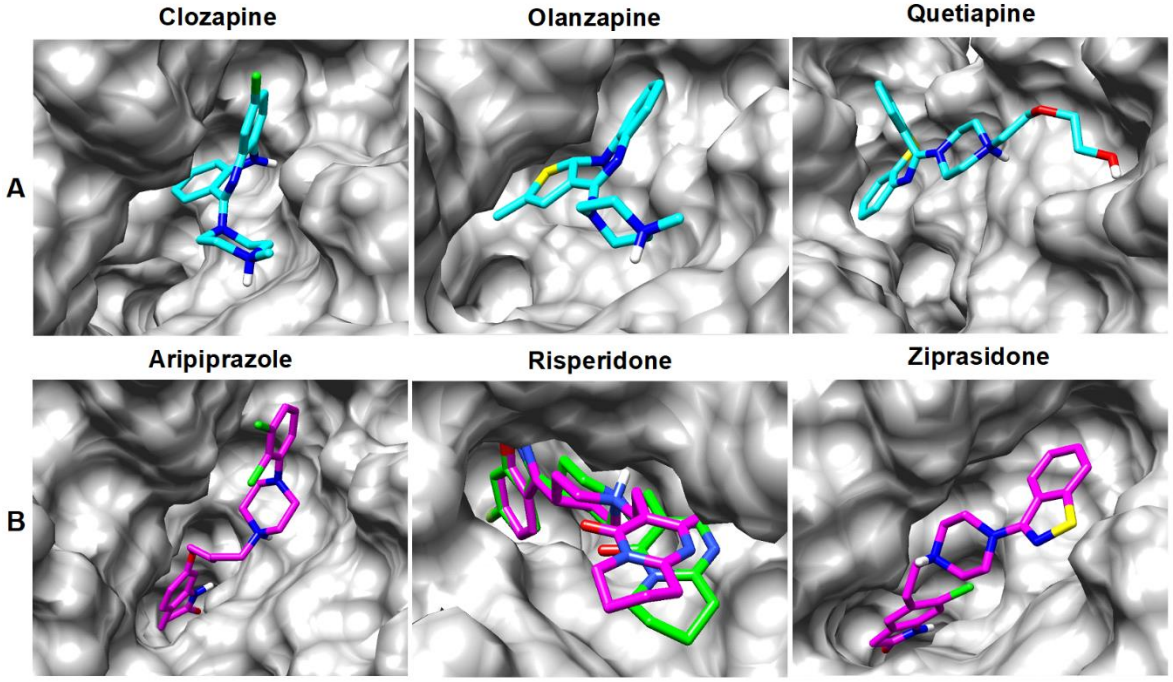


Figure 4. Binding poses of (**A**) class-I atypical antipsychotics and (**B**) class-II atypical antipsychotics docked to DRD2 receptor binding site. Docked risperidone (ZINC00538312; magenta) superimposed over the crystallographic risperidone pose (PDB ID: 6CM4; green) in DRD2 receptor. The observed RMSD difference was 0.07 Å. Interacting residues were similar for both poses (Image prepared by author).

The docked risperidone complex reproduced the vital contact residues in DRD2 binding site in reference to the Protein Data Bank crystal structure complex (Figure 4b). Thus, the co-crystalised risperidone DRD2 complex was used in the MD simulation. The docked poses of each of the studied antagonists were ranked based on their binding scores in the active site and results shown in Table 1. The docking results showed that the binding pose of class-I antipsychotics was different from those of class-II (Figure 4). The extended chemical structures of class-II atypical antipsychotics bind deeper into the hydrophobic pocket whereas the bulky class-I atypical antipsychotics bind within the orthosteric and extended binding pocket.

Table 1. Comparison of binding affinities and interacting amino acid residues of atypical antagonists against dopamine 2 receptor upon molecular docking.

Compounds	Docking Score (kcal/mol)	Interacting residues within 4Å
*Clozapine	-8.50	Leu94, Trp100, Cys107, Glu181, Cys182, Ile184, Phe110, Phe136, Tyr408, Thr412, Phe186, Ser193, His393, Phe389, Trp386, Pro405, Tyr417, Asn396, Asn402.
*Olanzapine	-8.40	Leu94, Trp100, Phe110, Tyr416, Ile184, Asp114, Phe189, Val111, Val115, Ser193, Val190, Phe390, His393, Phe389, Asn396, Pro405, Tyr408, Ser409, Thr412.
*Quetiapine	-8.60	Trp100, Phe102, Arg104, Cys107, Val111, Cys182, Trp134, Glu181, Ile184, Ala185, Ser197, Phe390, Trp386, Phe389, Thr412, His393, Tyr408, Pro405.
**Aripiprazole	-10.40	Trp100, Lys101, Phe110, Asp114, Val115, Cys118, Thr119, Ile122, Ile184, Ala185, Phe189, Ser193, Ser197, Trp386, Phe389, Phe390, His393, Pro405, Tyr408, Thr412, Trp413.
**Risperidone	-12.20	Ser409, Tyr408, Thr412, Phe389, Trp386, Phe390, Phe198, Phe382, Ile383, Ile122, Ser197, Thr119, Cys118, Val115, Ser193, Phe189, Asp114, Tyr416, Trp100, Val91, Leu94, Ile184, His414, Phe110.
**Ziprasidone	-10.70	Leu94, Val91, Trp100, Ile184, Val190, Asp114, Tyr416, Phe189, Cys118, Val115, Ser193, Ser194, Ser197, Phe390, His393, Trp386, Phe389, Thr412, Trp413, Thr119, Pro405, Val406, Tyr408.

Bolded residues are those forming important H-bonds; * represents class I atypical antipsychotics;
** represents class I atypical antipsychotics.

3.2 Dynamic interaction patterns of atypical DRD2 antipsychotics

To obtain ensemble insights into the complementary associations between atypical antipsychotics and DRD2, the binding mechanisms of the complex systems were evaluated using molecular dynamics (MD) simulations in explicit lipid bilayer.

To effectively determine the global stability of protein-ligand interactions, the root mean square deviation (RMSD) of Cα atoms with respect to the starting structures were determined. Figure

5a and 5b depict the average RMSD of the Cα atoms for atypical-receptor complexes over the entire 200ns MD simulations. In general, all complexed systems achieved conformational stability after 40ns simulations. The average apo and ligand-complex system RMSD values were 2.29Å (Apo), 1.95Å (clozapine), 1.93Å (quetiapine), 1.99 Å (olanzapine), 2.04Å (aripiprazole), 2.28Å (risperidone) and 1.82Å (ziprasidone). The ziprasidone-DRD2 complex relatively represents the most stable conformer. The above RMSD values seem to suggest that the DRD2 protein structure is less stable without inhibitor than with inhibitor. The relative dynamic stability of class-I and class-II atypical antagonists was also calculated to effectively assess the stability of these inhibitors at DRD2 binding pocket (Figure 5c & 5d). The observed averaged RMSD of the inhibitors was 0.92 Å, 1.69 Å, 0.91 Å, 1.98 Å, 1.06 Å, and 1.32 Å for clozapine, quetiapine, olanzapine, aripiprazole, risperidone, and ziprasidone, respectively.

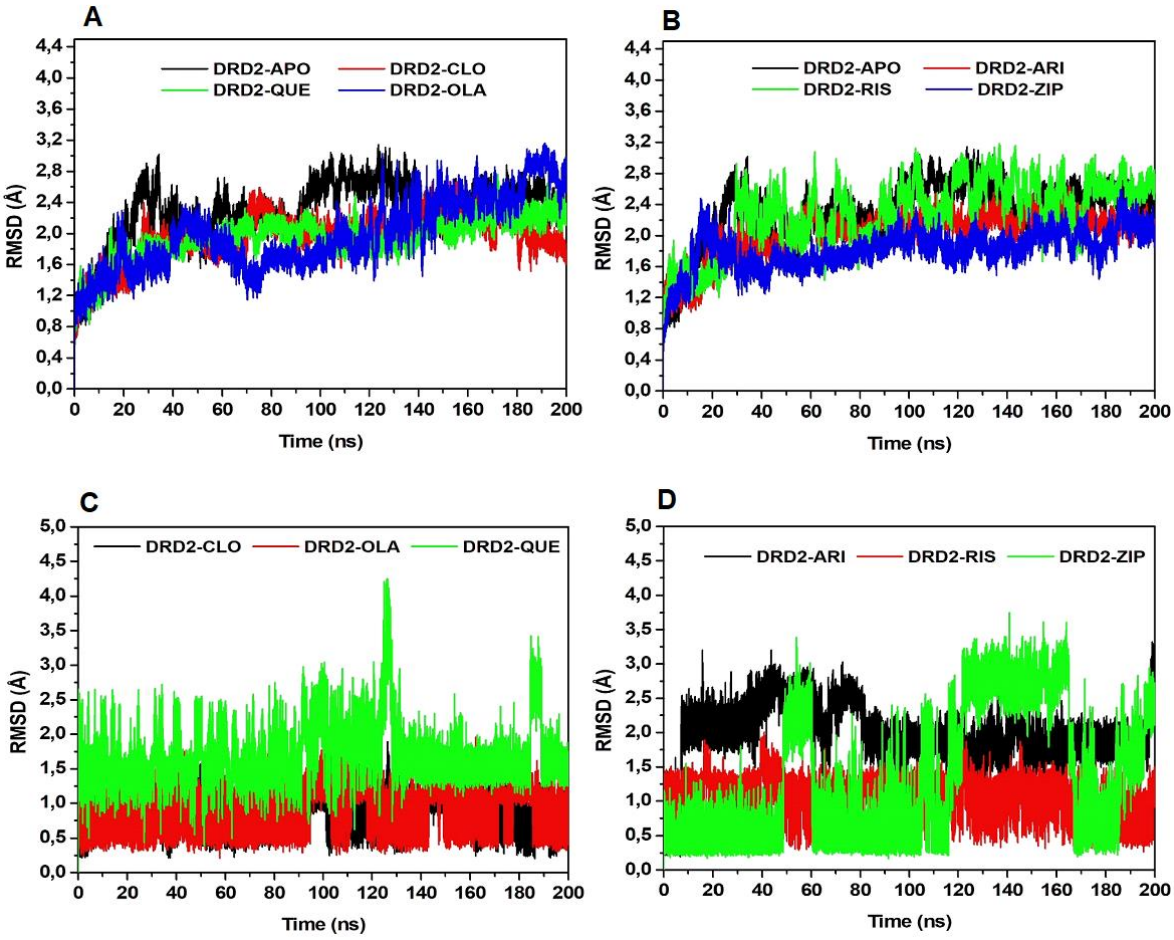


Figure 5. Conformational stability of protein-atypical drug complexes. The Cα RMSD of (**a**) atypical class-I drug-receptor complexes, (**b**) atypical class-II drug-receptor complexes, (**c**) atypical class-I drugs only and (**d**) atypical class-II drugs only relative to the starting minimised over 200ns (Image prepared by author).

To determine the regions of the protein exhibiting higher structural flexibility, the average root mean square fluctuations (RMSF) per residue of Cα atoms were determined (Fig 6a & 6b). Relatively, minimal alterations were observed at the transmembrane (TM) domain regions of the unbound (apo) and ligand-bound compared to the loops regions where a considerable degree of residue fluctuations was observed as shown in Figure 6a & b. The difference in RMSF values was also identified for class-I and class-II inhibitors for ECL1 and ECL2 residues. Risperidone and aripiprazole bound systems showed higher residue flexibility in the ECL2 compared with clozapine, olanzapine and quetiapine bound systems suggesting that the interactions of these different classes of inhibitors with DRD2 produced different restrictions on the motions of these residues. The above-observed differences in RMSF values may suggest the changes in interaction intensities and internal dynamics.

To obtain insights into the different degrees of solvent accessibility, the hydrophobic regions of DRD2 in complex with class-I and class-II atypical antipsychotics were observed and compared over the 200 ns simulations (Figure 6c & 6d). The mean solvent accessible surface area (SASA) values for class-I antipsychotics were 16162.34 $Å^2$, 16188.13 $Å^2$, and 16211.95 $Å^2$ for clozapine, quetiapine, and olanzapine respectively compared with 16365.43 $Å^2$, 16568.53 $Å^2$ and 16800.84 $Å^2$ for class-II antipsychotics aripiprazole, ziprasidone, and risperidone, respectively. Evidently, the binding poses of class-II antipsychotics engage deeper into the hydrophobic pockets which may explain their relatively larger volume of solvent accessible surface area compared to class-I antipsychotics.

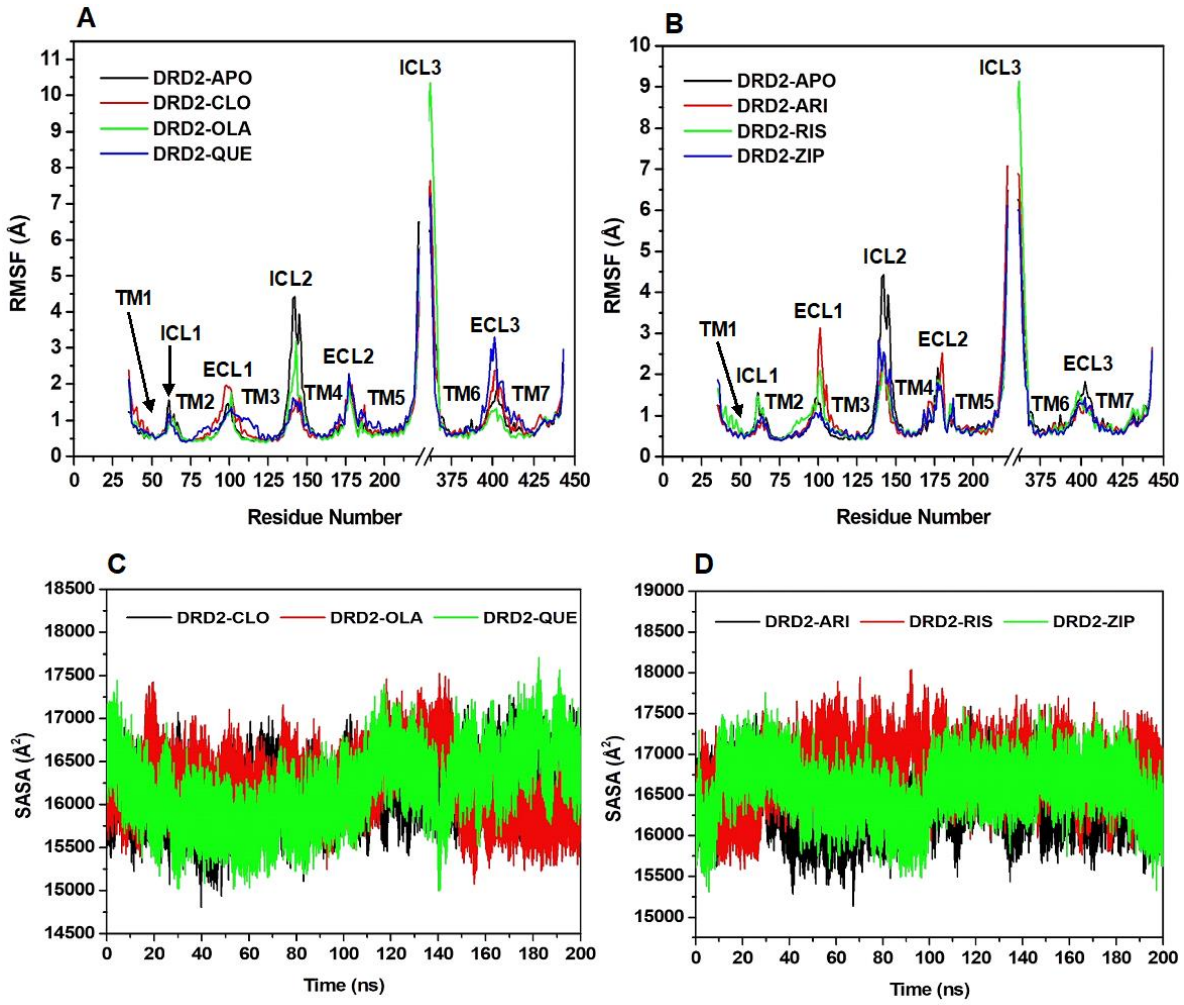


Figure 6. RMSF plot of Cα atoms in DRD2 systems for (**a**) class-I atypical and (**b**) class-II atypical over 200 ns of simulation. Solvent accessible surface area (SASA) for (**c**) class-I atypical and (**d**) class-II atypical over 200 ns of simulation (Image prepared by author).

3.3 Analysis of binding free energy interaction.

To further dissect contributions to the protein-inhibitor binding, the molecular mechanics-generalised born surface area (MM-GBSA) method was employed to predict the binding free energy of atypical antipsychotics to DRD2 using snapshots from stabilised MD trajectories (Table 2). The predicted binding free energies for class-I atypical antipsychotics were -37.22 ± 0.06, -34.21 ± 0.09, and -34.05 ± 0.09 kcal/mol respectively for olanzapine, clozapine and quetiapine (Ki values of 30.75, 147.0 and 437.0 nM, respectively). In class-II atypical antipsychotics, the predicted binding free energies were -49.87 ± 0.09, -49.27 ± 0.06, and -37.18 ± 0.08 kcal/mol respectively for aripiprazole, risperidone, and ziprasidone (Ki values of 2.30, 3.70, and 4.75 nM, respectively). The above binding energy rankings provide a good correlation with the experimental binding affinities compared to the previously reported rankings by Salmas *et al.* (Salmas et al., 2017) using the modelled DRD2 structure. From the

energy components of the total binding free energies, the van der Waals energy term (ΔE_{vdW}) favoured the binding of both atypical class-I and class-II antipsychotics, with class-II drugs depicting higher van der Waals interaction energies compared to class-I drugs. Polar solvation energy term (ΔE_{polar}) was also observed to favour the association of all inhibitors to DRD2 except quetiapine. Electrostatic interactions (ΔE_{elec}) were also observed to provide unfavourable force to all inhibitors except for quetiapine. Furthermore, non-polar solvation energy ($\Delta E_{nonpolar}$) was noted to be less strong than van der Waals interactions. Hence, it can be concluded that the van der Waals interaction energy term mainly controls the binding of atypical antipsychotic drugs to DRD2 receptor.

Table 2. MMGB/SA binding free energy components for DRD2–Antagonist complexes (expressed in kcal/mol with standard error of mean [SEM]) and their respective experimental binding affinity in nM (Stahl, 2017).

Atypical Class I Drugs								
Ligand	**ΔE_{MM}**	**ΔE_{solv}**	**ΔE_{vdW}**	**ΔE_{elec}**	**ΔE_{polar}**	**$\Delta E_{nonpolar}$**	**$\Delta E_{bind} \pm SEM$**	**$Exp_{(Ki)}$**
Olanzapine	-18.23	-18.98	-39.47	21.24	-14.51	-4.47	-37.22 ± 0.06	30.75
Clozapine	-12.78	-21.43	-36.26	23.48	-17.50	-3.93	-34.21 ± 0.09	147.0
Quetiapine	-52.06	18.01	-35.79	-16.27	22.49	-4.48	-34.05 ± 0.09	437.0
Atypical Class II Drugs								
Ligand	**ΔE_{MM}**	**ΔE_{solv}**	**ΔE_{vdW}**	**ΔE_{elec}**	**ΔE_{polar}**	**$\Delta E_{nonpolar}$**	**ΔE_{bind}**	**$Exp_{(Ki)}$**
Aripiprazole	-18.78	-31.09	-54.89	36.11	-24.57	-6.53	-49.87 ± 0.09	2.30
Risperidone	59.83	-109.1	-53.86	113.7	-102.8	-6.29	-49.27 ± 0.06	3.70
Ziprasidone	-11.95	-25.22	-44.44	32.49	-20.11	-5.11	-37.18 ± 0.08	4.75

ΔE_{ele}, electrostatic energy; ΔE_{vdw}, van der Waals energy; ΔE_{polar}, polar solvation energy; $\Delta E_{nonpolar}$, nonpolar solvation energy; ΔE_{solv}, total solvation energy term.
All the energies were averaged over 4000 snapshots at time intervals of 50 ps from the entire 200 ns MD simulations and are expressed in kcal/mol.

3.4 Hydrogen bond interaction analysis

To further identify interactions responsible for ligand recognition and conformational stability, hydrogen bond analysis was performed to identify and characterise the stability of specific drug-receptor hydrogen bonds over the course of the MD simulations. The key interactions observed between atypical antipsychotics and D_2 dopamine receptor are shown in Table 3 and Figure 8 & 9.

Table 3. Hydrogen bond occupancy of antagonists in complex with DRD2 during molecular dynamics simulations.

Ligands	H-Acceptor	H-Donor	[a]Occupancy (%)	Distance (Å)
Risperidone	Asp114-OD2	Inhibitor-N06-H06	9.02	2.85
	Thr412-OG1	Inhibitor-O23-H23	1.27	2.73
Aripiprazole	Inhibitor-O1	Asp114-OD2-HD2	71.88	2.76
Ziprasidone	Asp114-OD1	Inhibitor-N3-H1	15.38	2.84
	Inhibitor-O1	Ser193-OG-HG	2.59	2.76
	Inhibitor-O1	Asn396-ND2-HD22	1.62	2.86
Clozapine	Asn396-OD1	Inhibitor-N4-H2	30.33	2.77
	Asn402-OD1	Inhibitor-N4-H2	14.44	2.77
Quetiapine	Glu181-OE2	Inhibitor-N3-HN	35.04	2.77
	Glu181-OE2	Inhibitor-O2-H1	28.96	2.72
	Glu181-OE1	Inhibitor-N3-HN	19.37	2.77
	Glu181-OE1	Inhibitor-O2-H1	12.61	2.71
Olanzapine	Thr412-OG1	Inhibitor-N1-H2	48.36	2.89

[a] Occupancy is the percentage of time of the existence of hydrogen bonding over the 200ns simulation time.

From the obtained results it can be observed that all atypical class-II antipsychotics formed hydrogen bonds with Asp114. In the risperidone-DRD2 complex, hydrogen bond occurred between Asp114 and the tertiary amine of the piperidine ring in risperidone at a 9% occupancy and 2.85 Å distance. This hydrogen bond/salt-bridge was observed in the co-crystalised structure of risperidone (Wang et al., 2018). Additionally, a weaker hydrogen bond not observed in the co-crystalised structure occurred between Thr412 and the oxygen atom in the tetrahydropyridopyrimidinone moiety of risperidone with a 1.27% occupancy. In the aripiprazole complex, a stronger and stable hydrogen bond is formed between its butoxy oxygen and the hydroxyl of Asp114 with a 71.88% occupancy. The dihydro-indol-2-one moiety of ziprasidone also forms a hydrogen bond with Asp114 at a 15.38% occupancy. Additional, ziprasidone forms less stable hydrogen bonds with Ser193 (2.59 % occupancy) and Asn396 (1.62 % occupancy).

In the quetiapine-DRD2 complex, Glu181 was observed to form multiple stable hydrogen bonds with quetiapine between 12 - 35% occupancy, whereas a weaker hydrogen bond was also observed to be formed with Phe102 at a 2.6% occupancy. Also, clozapine was observed to form hydrogen bonds with Asn396 and Asn402 with a percentage occupancy of 30.33% and 14.44%, respectively. For olanzapine-DRD2 complex, the thieno-benzodiazepine group interacted with Thr412 via a stable hydrogen bond with a 48% occupancy.

The observed differences in the formation of hydrogen bond by class-I and class-II atypical antipsychotics reflect their distinctive binding mode and interaction pattern in the DRD2 crystal structure.

3.5 Residue interaction energy analysis

To identify key residues that make important intermolecular interaction contribution to the binding of atypical antipsychotics, the total binding energy of each inhibitor was decomposed to obtain the energy contribution of active site residues. This decomposition was performed at the atomic level for the overall atoms of each residue to obtain individual residue contribution to inhibitor binding. The van der Waals energy, electrostatic interaction energy and the total energy of residues with high energy contribution are presented in Figure 7. The binding mode of the two classes of atypical antipsychotics and important residues are also shown in Figure 8 & 9.

In the class-I atypical drug-receptor interactions, all three drugs strongly interacted commonly with two residues, namely Ile184 and Tyr408 with energy values between -0.75 and -3.768 kcal/mol. Olanzapine and quetiapine shared similar interactions with Trp100 and Val111, olanzapine and clozapine both interacted with Phe110 and Pro405, whereas clozapine and quetiapine made interactions Cys182 and Ala185. In the class-II drug-receptor interactions, residues Trp100, Phe389 and Tyr408 commonly interacted with all three drugs with energy values ranging from -0.512 to -2.882 kcal/mol. Aripiprazole and risperidone interacted with six similar additional residues including, Asp114, Cys118, Thr119, Trp386, Phe390, and Thr412 with energy values between -0.557 and -2.347 kcal/mol. These additional interactions pattern shared between aripiprazole and risperidone but not ziprasidone seems to suggest a similar binding mechanism exists in risperidone and aripiprazole toward DRD2 receptor.

Class-II antipsychotics generally interacted with residues within the deep hydrophobic binding pocket characterised by Cys118, Thr119, Ser197, Phe198, Phe382, Phe390, and Trp386 with stronger and similar residue interaction pattern observed for aripiprazole and risperidone. Overall, the residue interaction energy analysis further clarifies the difference in binding mode and residue interactions pattern of class-I and class-II atypical antipsychotics in DRD2.

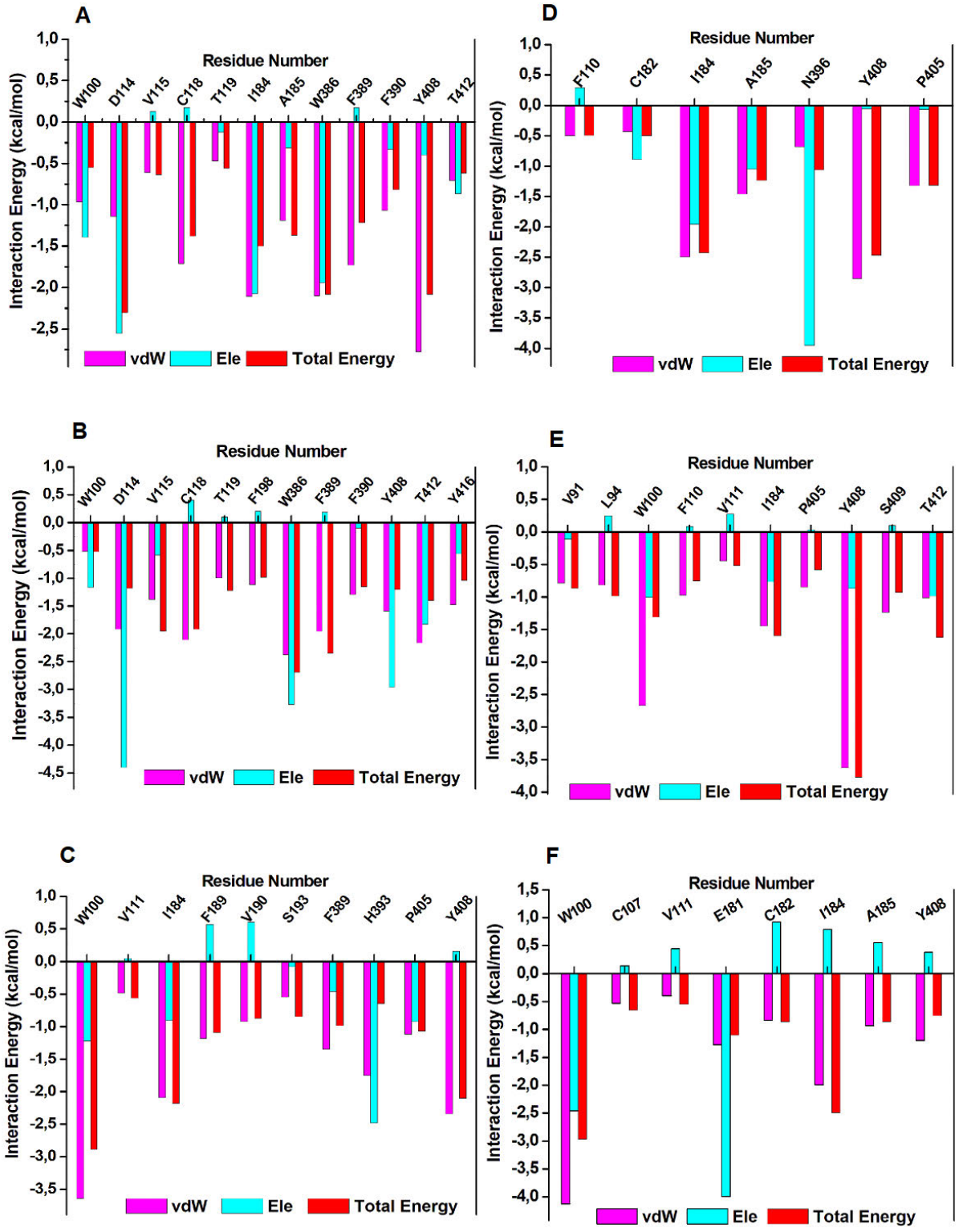


Figure 7. Per-residue binding free energy decomposition of (**a**) aripiprazole (**b**) risperidone (**c**) ziprasidone (**d**) clozapine (**e**) olanzapine and (**f**) quetiapine to DRD2 receptor (Image prepared by author).

3.6 The binding mode and interaction analysis of atypical antipsychotics

To probe the binding mode and interactions pattern of atypical antipsychotics in the crystal structure of the DRD2 receptor, these drugs were first docked into DRD2 binding pocket followed by an all-atom membrane MD simulation. The protein-ligand interaction profiler (PLIP) software (Salentin et al., 2015) and the Discovery Studio protein-ligand interaction module (Accelrys, 2013) were used to perform drug-receptor interaction fingerprint analysis to identify non-covalent interactions such as hydrophobic contacts, hydrogen bonding, pi-stacking, salt bridges, and pi-cation interactions.

In the predicted binding mode and interactions of olanzapine using the x-ray structure (Figure 8c & 9c), the protonated nitrogen atom of the thieno-benzodiazepine moiety forms a stable hydrogen bond with Thr412. Olanzapine is predicted to bind in the orthosteric and the extended binding pocket interacting with a cluster of aromatic and hydrophobic residues such as Val91, Leu94, Ile184, Trp100, Phe389 and Phe110. Additionally, olanzapine interacts with Phe389 and Tyr408 via pi-sulphur and pi-cation interactions, respectively. The residue energy interaction shows residues that are critical to the binding of olanzapine, these residues include Tyr408, Thr412, Ile184, Trp100, Leu94, Ser409, Val91, Phe110, Pro405, and Val111 with binding energies greater than -0.50 kcal/mol (Figure 7e). Tyr408 was previously observed to contribute to olanzapine binding via a hydrogen bond with the nitrogen atom of the methyl piperazine (Salmas et al., 2017). However, it was observed that Tyr408 strongly interacts with olanzapine through a pi-cation interaction (Figure 8c) which could be attributed to its higher residue energy contribution of -3.8 kcal/mol (Figure 7e).

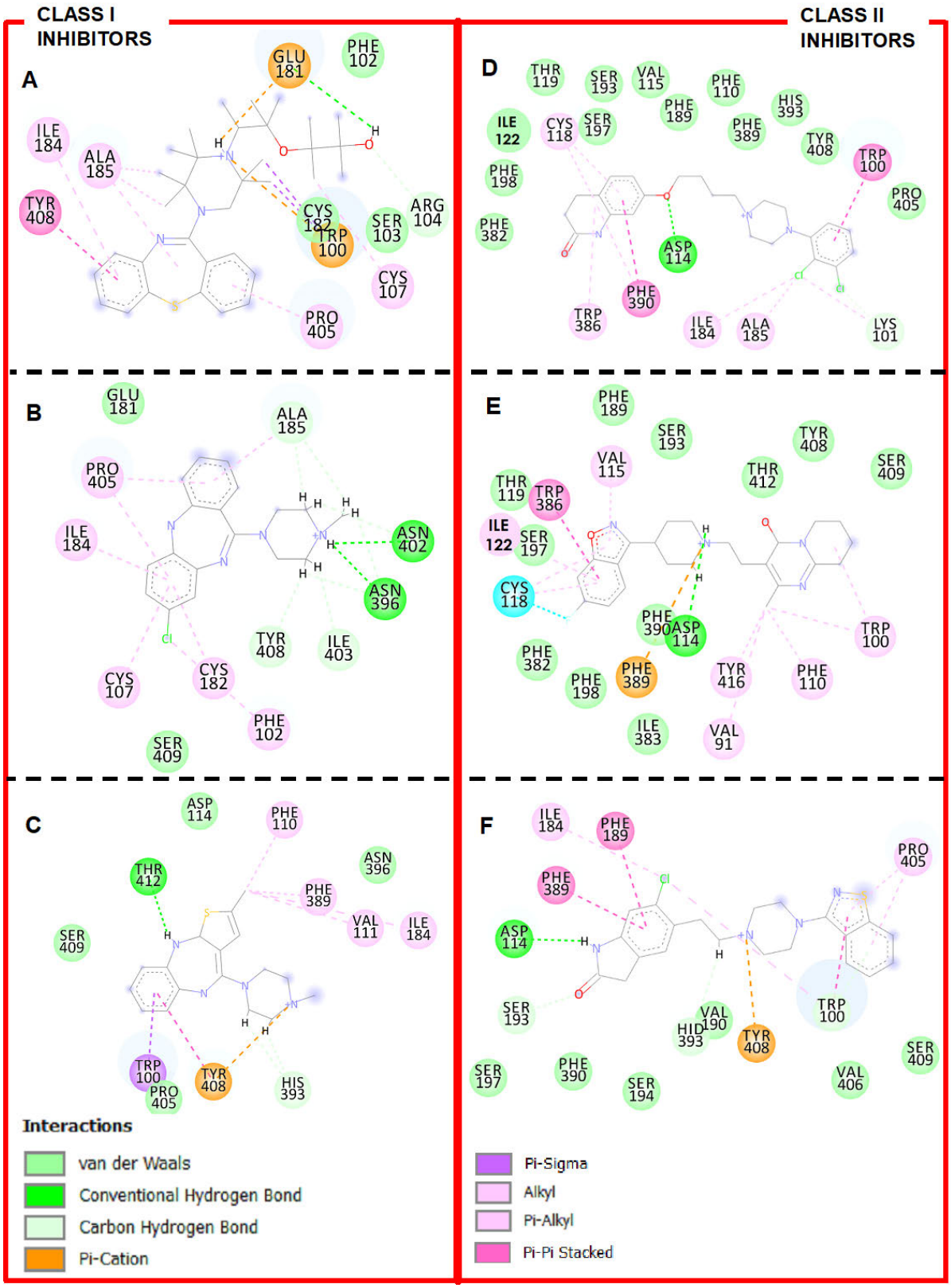


Figure 8. 2D inhibitor-complexes for (a) quetiapine (b) clozapine (c) olanzapine (d) aripiprazole (e) risperidone and (f) ziprasidone with D_2 dopamine receptor showing important protein inhibitor interactions (Image prepared by author).

In the predicted binding mode of clozapine in DRD2 receptor (Figure 8b & 9b), clozapine extends its interactions above the orthosteric binding pocket into the extended binding pocket interacting strongly with Phe110, Cys182, Ile184, Ala185, Asn396, Tyr408 and Pro405 with interaction energies greater than -0.5 kcal/mol (Figure 7d). The protonated nitrogen of the methyl-piperazine ring forms stable hydrogen bonds with Asn396 and Asn402 (Table 3 and Figure 8b & 9b). Pi-alkyl interactions are also observed between the benzo-benzodiazepine group and Phe102, Pro405, and Ile184. Clozapine additionally forms hydrophobic interactions with Phe102, Ile180, Pro405, Ile184, Glu181 and Tyr408.

The docking poses and molecular dynamics of quetiapine suggest a salt bridge interaction between Glu181 and the positively charged nitrogen of the piperazine ring. The hydroxyl in the ethoxyethanol group forms a stable hydrogen bond with Glu181 (Table 3 and Figure 8a & 9a). Quetiapine was predicted to extends its binding above the orthosteric binding site into the extended binding pocket interacting strongly with Trp100, Ile184, and Tyr408. The residue interaction energy suggests Trp100, Cys107, Val111, Glu181, Cys182, ile184, Ala185 and Tyr408 to significantly contribute to the binding of quetiapine with interaction energies higher than −0.5 kcal/mol (Figure 7f). The high interaction energy of Trp100 (-3.0 kcal/mol) with quetiapine mainly stems from the observed π-cation interaction between Trp100 and the piperazine ring of quetiapine.

The binding of risperidone in the D_2 dopamine receptor is such that the fluorobenzisoxazol ring orients deep in the hydrophobic cleft interacting with Phe390, Trp386, Phe382, Phe198, Ser197, Thr119, Cys118, Val115 whereas the tetrahydropyridopyrimidinone ring of risperidone interacts with residues Phe110, Thr412, Trp100, Tyr408 and Val191 (Figure 8e & 9e). The major residues with favourable interaction energy contribution to risperidone binding predominately originate from eleven residues (Asp114, Val115, Cys118, Thr119, Phe198, Trp386, Phe389, Phe390, Tyr408, Thr412, and Tyr416) with average energy contribution larger than −0.9 kcal/mol (Figure 7b). Asp114 forms hydrogen bond with the tertiary amine in the piperidine ring whereas a less stable hydrogen bond occurred between Thr412 and the oxygen atom in the tetrahydropyridopyrimidinone ring (Table 3). The fluorobenzisoxazol ring forms π-π interactions with Phe390, Trp386 and Phe198 whereas an alkyl-π interaction with Val115 and Cys118 were also observed. The piperidine ring also forms a π-cation interaction with Phe389.

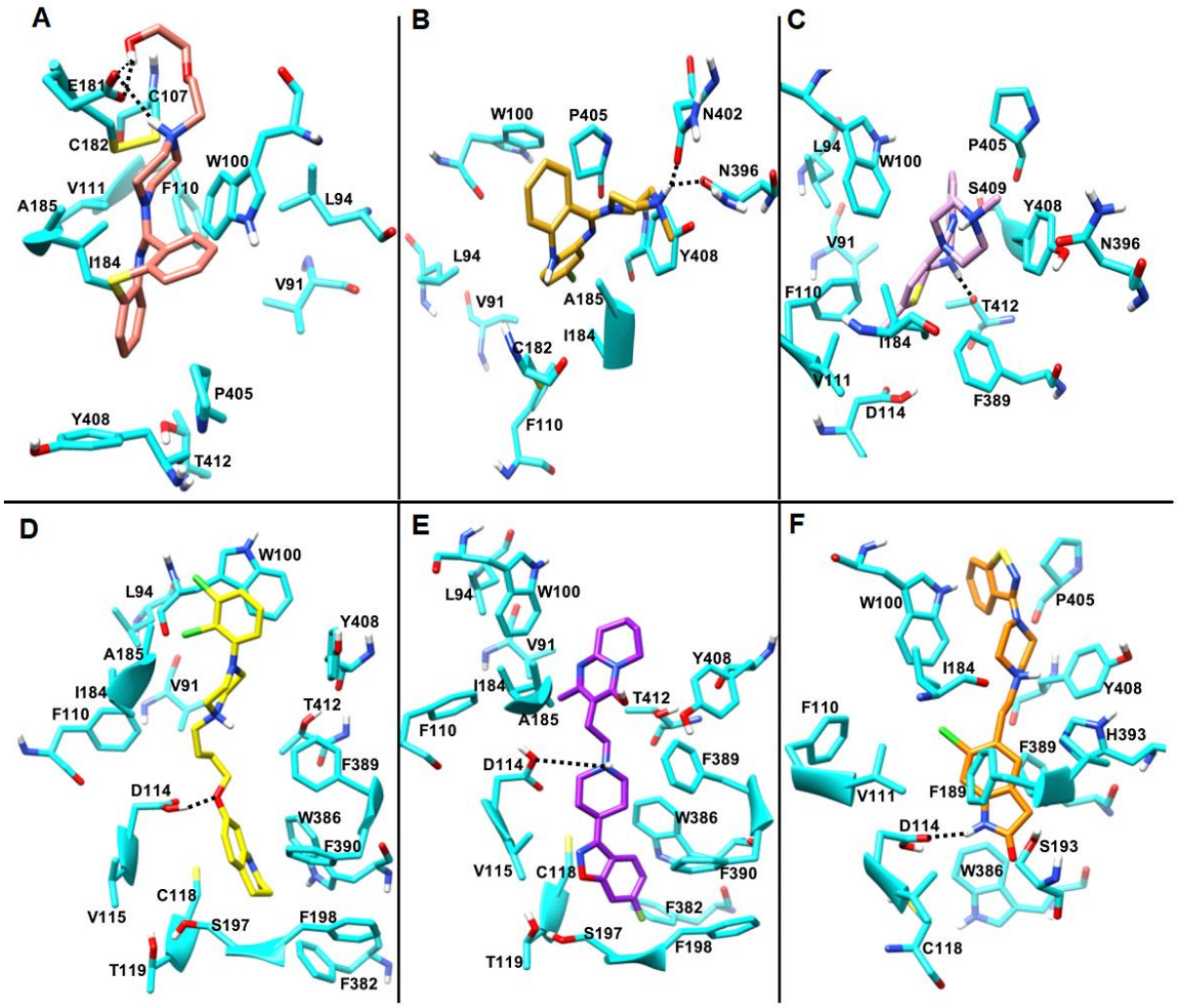


Figure 9. 3D Docked complex of (a) quetiapine (b) clozapine (c) olanzapine (d) aripiprazole (e) risperidone and (f) ziprasidone with D_2 dopamine receptor showing important interacting residues (Image prepared by author).

The predicted binding mode of aripiprazole in DRD2 active places the dihydro-quinolin-2-one moiety deep into the hydrophobic cleft below the orthosteric binding site interacting with residues Thr119, Ser197, Phe198, Phe390, Cys118, Trp386, Phe389, Phe390, and Phe382 (Figure 8d & 9d). The residue interaction analysis indicates that residues Asp114, Cys118, Ile184, Ala185, Trp386, Phe389, and Tyr408 of DRD2 contribute strongly to the binding of aripiprazole with binding energies greater than -1.0 kcal/mol (Figure 7a). Among these residues, the interaction energy of Asp114 with aripiprazole is the strongest (-2.3 kcal/mol). This stronger interaction energy from Asp114 may be due to the strong and stable hydrogen bond interaction (~72% occupancy) between Asp114 and aripiprazole (Table 3). This observed hydrogen bond interaction appears to be significant for the potency and stability of the aripiprazole-DRD2 complex. In addition, the dihydro-quinolin-2-one ring forms π-π

interactions with Phe390 and Trp386 whereas the dichlorophenyl group forms π-π stacking interactions with Trp100 and Tyr408 (Figure 8d).

The binding mode of ziprasidone (Figure 8f & 9f) suggests a salt bridge between Asp114 and the nitrogen atom of the chloro-1,3-dihydroindol-2-one moiety. Additionally, the oxygen atom of chloro-1,3-dihydroindol-2-one ring is predicted to form less stable hydrogen bonds with Ser193 and Asn396 (Table 3). The chloro-1,3-dihydroindol-2-one moiety binds in a cavity surrounded by residues Phe389, Val111, Phe189, Ser193 and Ser194, forming π-π interactions with Phe189 and Phe389. The benzothiazole ring made interactions with Ala185, Ile184, Leu94, Pro405, and Trp100 during the simulation. The residues interaction energy analysis reveals that Trp100, Val111, Ile184, Phe189, Val190, Ser193, Phe389, His393, Pro405 and Tyr408 strongly contribute to ziprasidone binding with energies greater than − 0.5 kcal/mol (Figure 6d). The high interaction energy of Trp100 (-2.9 kcal/mol) stems from the π-cation and π-sulphur interactions with the piperazine ring and benzothiazole ring, respectively whereas that of Tyr408 (-2.1 kcal/mol) stems from the π-cation interactions with the piperazine ring. These residues may contribute significantly to the stability and binding of ziprasidone to DRD2 receptor.

Asp114$^{3.32}$ (superscript denote Ballesteros Weinstein numbering system) elicits a unique interaction pattern common to all Class-II inhibitors, which further potentiates the importance of this interaction as primarily observed in the co-crystallised complex of DRD2 with risperidone (Wang et al., 2018). Asp$^{3.32}$ is a conserved residue observed in the structures of aminergic receptors and has been observed to form vital interactions with their inhibitors. The importance of this conserved residue has been reported in several aminergic crystalised structures: Asp114$^{3.32}$ forms a salt bridge with the tertiary amine of risperidone in DRD2 (Wang et al., 2018); the tertiary amine of eticlopride forms a salt bridge with Asp110$^{3.32}$ in DRD3 (Chien et al., 2010); the conserved Asp115$^{3.32}$ also interacts with nemonapride in DRD4 (Wang et al., 2017); whereas risperidone and Zotepine also form a salt bridge between the basic nitrogen of these inhibitors and Asp155$^{3.32}$ in 5-HT$_{2A}$ receptor (Kimura et al., 2019). This salt bridge appears to be vital for high-affinity inhibitor binding to aminergic subfamily of GPCRs (Shi & Javitch, 2002).

4. Conclusions

The recent availability of the crystal structure of DRD2-drug complex provides avenues for structure-based molecular dynamics simulations to be performed to uncover the mechanisms

of ligand recognition at the molecular level. In this study, molecular docking, all-atom molecular dynamics simulations of D_2 dopamine receptor in explicit lipid-bilayer were performed in complex with selected atypical antipsychotic drugs to decipher the binding mechanism and the effect of ligand binding on the conformational changes of the DRD2 receptor. The dynamic interaction patterns analysis showed that DRD2 extracellular and intracellular loop flexibility are large and binding site residues dynamic behaviour is different due to the presence of structurally different ligands. The interaction energy calculation by MM-GBSA method showed that van der Waals interaction dominates the binding of atypical antipsychotics to DRD2 receptor.

Additionally, the binding of atypical antipsychotic ligands into DRD2 binding site displayed different interactions ranging from hydrogen bonding (salt bridge), hydrophobic CH-π and π-π interactions formed between distinct subsite residues and the ligand atoms. Findings from this study further revealed a binding mode of atypical class-II drugs where the dihydro-2-quinolinone moiety of aripiprazole, the dihydro-indol-2-one moiety of ziprasidone and the fluorobenzisoxazole moiety of risperidone bind deep into DRD2 hydrophobic pocket defined by residues Thr119, Ser197, Phe198, Phe390, Cys118, Trp386, Phe389, Phe390, and Phe382 of transmembrane (TM) III, V and VI. These residues significantly contributed to the binding and stability of these drugs as evident by the residue interaction energy analysis. Asp114 was observed to form a strong and stable hydrogen bond in all class-II atypical drugs. This hydrogen bond interaction seems to be important for potency as well as for the stability of atypical class-II drugs.

Contrary to the observed binding mode and interaction pattern of class-II antipsychotics, class-I antipsychotics depict different interaction pattern and mode of action in DRD2 receptor. The binding of class-I antipsychotics extends into the extended binding pocket above the orthosteric binding pocket characterised by Trp100, Phe110, Val,91, Tyr408, Ile184 and Leu94. These residues made a significant contribution to the binding and stability of class-I antipsychotics. As reported by Wang *et al.* (Wang et al., 2018) neither DRD3 inhibitor eticlopride nor DRD4 inhibitor nemonapride occupies the deep hydrophobic pocket. Hence, the binding of class-II atypical antipsychotic drugs in the deep opening hydrophobic sub-pocket can be exploited in the design of the next generation subtype selective DRD2 antagonists.

5. Acknowledgements

The authors acknowledge the College of Health Science of University of KwaZulu-Natal for infrastructural support and the Centre for High-Performance Computing (CHPC), Cape Town,

South Africa for computational resources (www.chpc.ac.za). Ahmed A. El Rashedy is specially thanked for his technical support during the membrane protein simulation protocol development.

6. Authors contributions

P. Appiah-Kubi designed, performed the experiment, data analysis, interpretation of data and drafting the manuscript. FA. Olotu reviewed and edited the drafted manuscript. M.E.S Soliman conceptualised and supervised the entire study. The manuscript was reviewed and approved by all contributing authors.

7. Conflict of interest

The authors declare no conflict of interest.

References

Accelrys, S.I. 2013. Discovery Studio Visualizer, Release 4.0.

Appiah-Kubi, P. & Soliman, M.E.S. 2016. Dual anti-inflammatory and selective inhibition mechanism of leukotriene A4 hydrolase/aminopeptidase: insights from comparative molecular dynamics and binding free energy analyses. *Journal of Biomolecular Structure and Dynamics*, 34(11): 2418–2433. https://doi.org/10.1080/07391102.2015.1117991.

Bayly, C.I., Cieplak, P., Cornell, W. & Kollman, P.A. 1993. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *The Journal of Physical Chemistry*, 97(40): 10269–10280.

Case, D.A., Babin, V., Berryman, J., Betz, R.M., Cai, Q., Cerutti, D.S., Cheatham Iii, T.E., Darden, T.A., Duke, R.E. & Gohlke, H. 2018. Amber 18. *University of California, San Francisco*.

Chien, E.Y.T., Liu, W., Zhao, Q., Katritch, V., Han, G.W., Hanson, M.A., Shi, L., Newman, A.H., Javitch, J.A. & Cherezov, V. 2010. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science*, 330(6007): 1091–1095.

Congreve, M. & Marshall, F. 2010. The impact of GPCR structures on pharmacology and structure-based drug design. *British journal of pharmacology*, 159(5): 986–996.

Creese, I., Burt, D.R. & Snyder, S.H. 1976. Dopamine receptor binding predicts clinical and pharmacological potencies of antischizophrenic drugs. *Science*, 192(4238): 481–483.

Darden, T., Perera, L., Li, L. & Pedersen, L. 1999. New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure*, 7(3): R55–R60.

Dickson, C.J., Madej, B.D., Skjevik, Å.A., Betz, R.M., Teigen, K., Gould, I.R. & Walker, R.C. 2014. Lipid14: the amber lipid force field. *Journal of chemical theory and computation*, 10(2): 865–879.

Dolinsky, T.J., Czodrowski, P., Li, H., Nielsen, J.E., Jensen, J.H., Klebe, G. & Baker, N.A. 2007. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic acids research*, 35(Web Server issue): W522-5.

Duan, X., Zhang, M., Zhang, X., Wang, F. & Lei, M. 2015. Molecular modeling and docking

study on dopamine D2-like and serotonin 5-HT2A receptors. *Journal of Molecular Graphics and Modelling*, 57: 143–155.

Fuxe, K., Borroto-Escuela, D.O., Tarakanov, A.O., Romero-Fernandez, W., Ferraro, L., Tanganelli, S., Perez-Alea, M., Di Palma, M. & Agnati, L.F. 2014. Dopamine D2 heteroreceptor complexes and their receptor–receptor interactions in ventral striatum: novel targets for antipsychotic drugs. In *Progress in brain research*. Elsevier: 113–139.

Greengard, P. 2001. The neurobiology of dopamine signaling. *Bioscience reports*, 21(3): 247–269.

Irwin, J.J., Sterling, T., Mysinger, M.M., Bolstad, E.S. & Coleman, R.G. 2012. ZINC: A Free Tool to Discover Chemistry for Biology. *Journal of Chemical Information and Modeling*, 52(7): 1757–1768. https://doi.org/10.1021/ci3001277.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. & Klein, M.L. 1983. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2): 926–935.

Joung, I.S. & Cheatham III, T.E. 2008. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *The journal of physical chemistry B*, 112(30): 9020–9041.

Kebabian, J.W. & Calne, D.B. 1979. Multiple receptors for dopamine. *Nature*, 277(5692): 93–96.

Kimura, K.T., Asada, H., Inoue, A., Kadji, F.M.N., Im, D., Mori, C., Arakawa, T., Hirata, K., Nomura, Y. & Nomura, N. 2019. Structures of the 5-HT 2A receptor in complex with the antipsychotics risperidone and zotepine. *Nature structural & molecular biology*, 26(2): 121.

Kumar, A., Srivastava, G., Negi, A.S. & Sharma, A. 2018. Docking, molecular dynamics, binding energy-MM-PBSA studies of naphthofuran derivatives to identify potential dual inhibitors against BACE-1 and GSK-3β. *Journal of Biomolecular Structure and Dynamics*: 1–16.

Larini, L., Mannella, R. & Leporini, D. 2007. Langevin stabilization of molecular-dynamics simulations of polymers by means of quasisymplectic algorithms. *The Journal of chemical physics*, 126(10): 104101.

Lomize, M.A., Lomize, A.L., Pogozheva, I.D. & Mosberg, H.I. 2006. OPM: orientations of proteins in membranes database. *Bioinformatics (Oxford, England)*, 22(5): 623–625.

Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E. & Simmerling, C. 2015. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of chemical theory and computation*, 11(8): 3696–3713.

Meltzer, H.Y., Matsubara, S. & Lee, J.C. 1989. Classification of typical and atypical antipsychotic drugs on the basis of dopamine D-1, D-2 and serotonin2 pKi values. *Journal of Pharmacology and Experimental Therapeutics*, 251(1): 238–246.

Mishra, A., Singh, S. & Shukla, S. 2018. Physiological and functional basis of dopamine receptors and their role in neurogenesis: possible implication for Parkinson's disease. *Journal of experimental neuroscience*, 12: 1179069518779829.

Miyamoto, S. & Kollman, P.A. 1992. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of computational chemistry*, 13(8): 952–962.

Montgomery, D., Campbell, A., Sullivan, H.-J. & Wu, C. 2018. Molecular dynamics simulation of biased agonists at the dopamine D2 receptor suggests the mechanism of receptor functional selectivity. *Journal of Biomolecular Structure and Dynamics*: 1–20.

Nisha, C.M., Kumar, Ashwini, Vimal, A., Bai, B.M., Pal, D. & Kumar, Awanish. 2016. Docking and ADMET prediction of few GSK-3 inhibitors divulges 6-bromoindirubin-3-oxime as a potential inhibitor. *Journal of Molecular Graphics and Modelling*, 65: 100–

107.
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. & Ferrin, T.E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13): 1605–1612.
Podder, A., Pandey, D. & Latha, N. 2016. Investigating the structural impact of S311C mutation in DRD2 receptor by molecular dynamics & docking studies. *Biochimie*, 123: 52–64.
Roe, D.R. & Cheatham, T.E. 2013. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation*, 9(7): 3084–3095.
Roth, B.L. 2007. Drugs and valvular heart disease. *N Engl J Med*, 356(1): 6–9.
Roth, B.L., Sheffler, D.J. & Kroeze, W.K. 2004. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nature reviews Drug discovery*, 3(4): 353.
Salentin, S., Schreiber, S., Haupt, V.J., Adasme, M.F. & Schroeder, M. 2015. PLIP: fully automated protein–ligand interaction profiler. *Nucleic acids research*, 43(W1): W443–W447.
Salmas, R.E., Yurtsever, M. & Durdagi, S. 2017. Atomistic molecular dynamics simulations of typical and atypical antipsychotic drugs at the dopamine D2 receptor (D2R) elucidates their inhibition mechanism. *Journal of Biomolecular Structure and Dynamics*, 35(4): 738–754.
Schrödinger. 2015. 4: LigPrep, version 3.6. *Schrödinger, LLC, New York, NY*.
Seyedi, S.S., Shukri, M., Hassandarvish, P., Oo, A., Shankar, E.M., Abubakar, S. & Zandi, K. 2016. Computational approach towards exploring potential anti-Chikungunya activity of selected flavonoids. *Scientific reports*, 6: 24027.
Shi, L. & Javitch, J.A. 2002. The binding site of aminergic G protein-coupled receptors: the transmembrane segments and second extracellular loop. *Annual review of pharmacology and toxicology*, 42(1): 437–467.
Sprenger, K.G., Jaeger, V.W. & Pfaendtner, J. 2015. The General AMBER Force Field (GAFF) Can Accurately Predict Thermodynamic and Transport Properties of Many Ionic Liquids. *The Journal of Physical Chemistry B*, 119(18): 5882–5895.
Stahl, S.M. 2017. Drugs for psychosis and mood: unique actions at D3, D2, and D1 dopamine receptor subtypes. *CNS spectrums*, 22(5): 375–384.
Trott, O. & Olson, A.J. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2): 455–461.
Wang, S., Che, T., Levit, A., Shoichet, B.K., Wacker, D. & Roth, B.L. 2018. Structure of the D2 dopamine receptor bound to the atypical antipsychotic drug risperidone. *Nature*, 555: 269. http://dx.doi.org/10.1038/nature25758.
Wang, S., Wacker, D., Levit, A., Che, T., Betz, R.M., McCorvy, J.D., Venkatakrishnan, A.J., Huang, X.-P., Dror, R.O. & Shoichet, B.K. 2017. D4 dopamine receptor high-resolution structures enable the discovery of selective agonists. *Science*, 358(6361): 381–386.
Webb, B. & Sali, A. 2014. Protein structure modeling with MODELLER. *Protein Structure Prediction*: 1–15.

CHAPTER 5

Published Article

Exploring the structural basis and atomistic binding mechanistic of the selective antagonist blockade at D_3 dopamine receptor over D_2 dopamine receptor

Patrick Appiah-Kubi[1], Fisayo Andrew Olotu[1] and Mahmoud E. S. Soliman[1*]

[1]Molecular Bio-computation and Drug Design Laboratory

School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

Telephone: +27 (0) 31 260 8048

Fax: +27 (0) 31 260 78

Email: soliman@ukzn.ac.za

Website: http://soliman.ukzn.ac.za

Abstract

More recently, there has been a paradigm shift towards selective drug targeting in the treatment of neurological disorders, including drug addiction, schizophrenia, and Parkinson's disease mediated by the different dopamine receptor subtypes. Antagonists with higher selectivity for D_3 dopamine receptor (D3DR) over D_2 dopamine receptor (D2DR) have been shown to attenuate drug-seeking behaviour and associated side effects compared to non-subtype selective antagonists. However, high conservations among constituent residues of both proteins, particularly at the ligand-binding pockets, remains a challenge to therapeutic drug design. Recent studies have reported the discovery of two small-molecules R-VK4-40 and Y-QA31 which substantially inhibited D3DR with > 180-fold selectivity over D2DR. Therefore, in this study, we seek to provide molecular and structural insights into these differential binding mechanistic using meta-analytic computational simulation methods. Findings revealed that R-VK4-40 and Y-QA31 adopted shallow binding modes and were more surface-exposed at D3DR while on the contrary, they exhibited deep hydrophobic pocket binding at D2DR. Also, two non-conserved residues; Tyr36$^{1.39}$ and Ser182$^{45.51}$ were identified in D3DR, based on their crucial roles and contributions to the selective binding of R-VK4-40 and Y-QA31. Importantly, both antagonists exhibited high affinities in complex with D3DR compared to D2DR, while van der Waals energies contributed majorly to their binding and stability. Structural analyses also revealed the distinct stabilizing effects of both compounds on D3DR secondary architecture relative to D2DR. Therefore, findings herein pinpointed the origin and mechanistic of selectivity of the compounds which may assist in the rational design of potential small-molecules of the D_2-like dopamine family receptor subtype with improved potency and selectivity.

Keywords: D_3 dopamine receptor, D_2 dopamine receptor, Selective antagonist, R-VK4-40, Y-QA31, Molecular docking, Molecular dynamics simulation, Membrane lipid bilayer

1 Introduction

G-protein-coupled receptors (GPCRs) modulate the activity of a diverse range of stimuli signal pathways through heterotrimeric G proteins activation such as ions, peptides, small molecules, and globular proteins (Weis & Kobilka, 2018). Dopamine receptors belong to the GPCRs, which constitute seven transmembrane (TM) helical domains. The dopaminergic pathway is involved in diverse physiological processes such as motor behaviour, neuroendocrine function, cognitive function, and emotion. However, the dysfunction of the dopaminergic system and the closely related D_3 and D_2 dopamine receptors are implicated in neurological and psychiatric pathologies such as schizophrenia, Parkinson's disease, substance abuse as well as attention deficit hyperactivity disorder (ADHD)(Rangel-Barajas et al., 2015; Maggio et al., 2015; Beaulieu & Gainetdinov, 2011; Heidbreder & Newman, 2010).

The high-resolution structures of GPCRs have proven to be vital in structure-based drug design (SBDD) methods in tailoring ligand selectivity and inhibitor efficacy (Congreve et al., 2014; Liu et al., 2018). The crystal structures of the D_2 receptor bound to risperidone(Wang et al., 2018) and haloperidol (Fan et al., 2020), as well as the D_3 receptor bound to eticlopride(Chien et al., 2010), have been solved. However, the estimated 78% high degree of transmembrane (TM) sequence similarity and the near-identity of binding site residues between D_2 and D_3 dopamine receptors, poses a challenge toward the development of D_2-like receptor subtype-selective inhibitors with enhanced drug-like properties (Sibley & Monsma Jr, 1992). Current marketed antipsychotic drugs lack selectivity toward a given D_2-like receptor subtype(Li et al., 2016). They are associated with substantial side effects such as metabolic syndrome, cardiovascular hypertension, and neurological side effects, including tardive dyskinesia and extrapyramidal reactions (Kaar et al., 2020; Ballon et al., 2014; Álvarez et al., 2013). These adverse effects reduce compliance with medication and the quality of life (Lieberman et al., 2005; Novick et al., 2010).

The selective targeting of each D_2-like receptor subtypes has been posited to produce fewer side effects(Li et al., 2016; Holmes et al., 2004). Among D_2-like dopamine receptors, D3DR has received increasing pharmacotherapeutic interest in drug abuse disorders treatments (Micheli, 2011; Newman, Blaylock, et al., 2012). D_3 dopamine receptor has the highest affinity for dopamine and displays limited distribution in the mesolimbic system compared to the other dopamine receptor subtypes (Keck et al., 2015; Sokoloff & Le Foll, 2017). As such, fewer side effects are anticipated to be exerted by pharmacological agents that selectively target D3DR compared to the other dopamine receptors. Inhibitors that selectively bind to D_3 dopamine

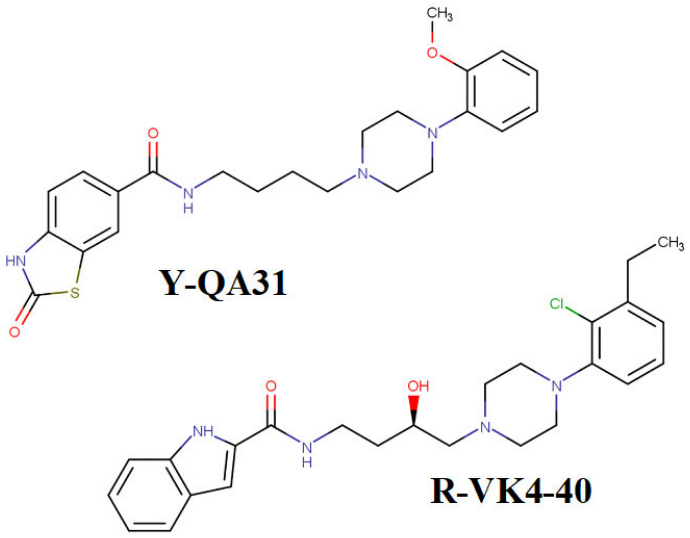
receptors may demonstrate efficacy towards a wide range of schizophrenia symptoms with fewer side effects (Sokoloff et al., 2006; Girgis et al., 2016). The D_3 dopamine receptor has been identified to improve some aspects of cognitive, negative (social withdrawal, mutism, and blunted effect) and positive (delusions and hallucinations) symptoms of schizophrenia, with reduced extrapyramidal side effects (EPS) caused by D_2 dopamine receptor blockade (Miyamoto et al., 2012).

Furthermore, the selective antagonism of D3DR has been shown to enhance the stimulant behaviour effects of cocaine in mice, which is opposite to the effect produced by selective antagonism of D2DR or nonselective D_2-like receptor antagonists (Manvich et al., 2019). Also, selective antagonist binding at the D_3 receptor has been shown to display promising results in reducing cocaine, and opioid reward and are highly effective in mitigating relapse to drug-seeking behaviour in preclinical models (Andreoli et al., 2003; Higley et al., 2011; Galaj et al., 2015). Extensive medicinal chemistry research efforts had led to the development of D_3 dopamine receptor selective partial agonists and antagonists such as BP 897, NGB 2904, SB 277,011A and GSK598,809. However, the clinical development of D3DR-selective antagonists, including SB-277,011A and GSK598,809 have been halted due to an increase in blood pressure in the presence of cocaine in freely moving dog and rat models (Appel et al., 2015; Appel & Acri, 2017; Appel & Acri, 2018). Presently, there exist no medications approved by the FDA for cocaine use disorder treatment. Moreover, currently approved drugs for opioid use disorder treatment, including buprenorphine and methadone, are opioid-based possessing several drawbacks such as respiratory suppression side effects and abuse liability (Jordan, Cao, et al., 2019).

Shaik and colleagues recently identified a novel highly selective D3DR antagonist as R-VK4-40 (Shaik et al., 2019), which is an R-enantiomer resolved from the compound (±)VK4-40 reported earlier by Kumar and colleagues as compound 40 (Kumar et al., 2016). Also, Y-QA31 is another novel D3DR selective antagonist recently identified to exhibit antipsychotic activity in preclinical studies of schizophrenia (Sun et al., 2016). These new D_3 dopamine receptor antagonists demonstrated higher selectivity (over 180-fold in ligand binding studies) toward D_3 over the D_2 dopamine receptor (Jordan, Humburg, Rice, et al., 2019; Sun et al., 2016) (Figure 1). R-VK4-40 decreases oxycodone rewarding effects while augmenting oxycodone analgesic effect via the selective inhibition of D3DR (Jordan, Humburg, Rice, et al., 2019; de Guglielmo et al., 2020). Contrary to the side effect of GSK598,809 and SB277011A in increasing blood pressure in the presence of cocaine (Appel et al., 2015; Appel & Acri, 2018),

R-VK4-40 do not exhibit adverse cardiovascular effects when combined with cocaine (Jordan, Humburg, Thorndike, et al., 2019). Also, Y-QA31 exhibit antipsychotic effects in cognitive dysfunction, negative and positive symptoms without inducing extrapyramidal side effects in preclinical models of schizophrenia (Sun et al., 2016).

However, the structural determinant and atomistic molecular mechanistic by which R-VK4-40 and Y-QA31 achieved their selectivity at D3DR over D2DR remain uncertain. To address this gap, diverse *in silico* protocols have been utilised in this study to investigate the selective binding mechanistic of R-VK4-40 and Y-QA31 D3DR over D2DR. Firstly, molecular docking was performed to predict the potential binding modes of R-VK4-40 and Y-QA31 at D_2 and D_3 dopamine receptors. Comparative molecular dynamics simulations of the four complexes were performed in a lipid bilayer environment and their dynamic properties compared. End-point binding free energy estimations were also utilized to assess the differential energetic contributions as well as critical residues driving the selectivity of R-VK4-40 and Y-QA31 towards D3DR. The results presented highlights the structural basis of selectivity and could aid in the rational design of novel potential selective inhibitors of D_3 dopamine receptor subtype.



Compound	K_i nM		selectivity fold
	D3DR	D2DR	
R-VK4-40	0.29	75.8	261
Y-QA31	0.28	52.1	186

Figure 1. Chemical structures of R-VK4-40 and Y-QA31 and their binding affinities at D_2 and D_3 dopamine receptors (Image prepared by author).

2 Materials and Methods

2.1 Protein preparation

The crystallographic structures of the inactive D_2 and D_3 dopamine receptors were retrieved from the Protein Data Bank (RCSB PDB, http://www.rcsb.org/) with entries 6CM4 (Wang et al., 2018), and 3PBL (Chien et al., 2010), respectively. The engineered proteins with T4 lysozyme (T4L) or nano antibody, co-crystallised molecular fragments, and water were deleted from the crystallographic structures. The thermostabilising crystallographic mutated residues (Ile122$^{3.40}$A, L375$^{6\,37}$A, and L379$^{6.41}$A) in D2DR and mutated residue Leu119$^{3.41}$Trp in D3DR were reverted to their respective wild types. The protein structures were further pre-processed using the Protein Preparation Wizard module from the Maestro Schrodinger suite (Schrödinger Release 2019-4, 2019). The protein preparation involved the addition of hydrogens, assigning bond orders, and the removal of water molecules beyond 5 Å of any hetero group. Missing loops were also modelled. PROPKA was used to designate the protonation state at a physiological pH. The protein was further energy minimised at a 0.50 Å root mean square deviation using the OPLS3e force field (Schrödinger Release 2019-4, 2019).

2.2 Ligand preparation

The 3D structures of R-VK4-40 and Y-QA31 antagonists were retrieved from the PubChem database with the PubChem IDs 130431314, and 46195521, respectively. The selected antagonists were pre-processed using the LigPrep module of Maestro (Schrödinger Release 2019-4, 2019) at a pH of 7.0 ± 2 to generate multiple states of possible protonation state, stereoisomers, ring conformations, and tautomer. The pre-processed ligands were then energy minimised to obtain a reasonable 3D conformation using the OPLS3e force field.

2.3 Molecular Docking

The Glide ligand docking panel (Friesner et al., 2004) implemented in the 2019 Schrödinger software package was utilized to predict the binding modes of R-VK4-40 and Y-QA31 at the active site of D_3 and D_2 dopamine receptors. The dimension of the grid box for the molecular docking was defined by the centroid of risperidone and eticlopride at D2DR and D3DR co-crystallised structures, respectively. An outer receptor grid box of 28 × 28 × 28 Å^3 with a default inner box size of 10 × 10 × 10 Å^3 were assigned for D2DR with an X, Y and Z centre grid of 9.66, 5.41 and -10.20, respectively. Similarly, an outer receptor grid box of 24 × 24 × 24 Å^3 with a default inner box size of 10 × 10 × 10 Å^3 were assigned for D3DR with an X, Y and Z centre grid of -0.03, -15.37 and 10.54, respectively. The Glide standard precision (SP) was used

in predicting various poses per compound and scored using the extra precision (XP) (Friesner et al., 2006). The docked receptor-ligand complexes with the lowest binding energies were selected as complex structures for the MD simulation setup and production. To validate the molecular docking procedure, risperidone and eticlopride were redocked as a positive control at D2DR and D3DR, respectively. The re-docked D2DR-risperidone and D3DR-eticlopride mimicked the observed crystallised binding mode (Figure S1).

2.4 D_3 and D_2 dopamine receptor complexes in lipid bilayer setup

The docked D_2 and D_3 dopamine receptor-ligand complexes were embedded in palmitoyl-oleoylphosphatidyl-choline (POPC) membrane lipid bilayer assembled with the CHARMM-GUI *membrane builder* (Wu et al., 2014). Each complex was aligned in a 180 homogeneous POPC lipid bilayer using the orientation of proteins membranes (OPM) with a rectangular box type (Lomize et al., 2006). TIP3P water molecules and 0.15M KCl were utilised to solvate and ionise the oriented systems, respectively. The *charmmlipid2amber.py* script was further used to process the solvated membrane-bound systems by renaming the POPC lipid residues according to the Amber Lipid17 force field. Molecular dynamics starting coordinate and topology files were generated with *tleap* module in Amber 18 using lipid17, ff14SB, and TIP3P forcefields for the lipid bilayer, protein, and water, respectively.

2.5 Molecular dynamics (MD) simulations

The unbound and docked D_2 and D_3 dopamine receptor-ligand complexes embedded in ionised solvated lipid bilayer were used as the starting coordinates for the unbiased molecular dynamics simulation. The simulations were run using the CUDA version of Particle Mesh Ewald Molecular Dynamics (PMEMD) in Amber18 on GPU (Salomon-Ferrer et al., 2013; Le Grand et al., 2013). Energy minimisation was performed for each system to relax and remove possible steric clashes. An initial 15,000 steps minimisations were performed with a 10.0 kcal/mol/$Å^2$ constraint on all heavy atoms of the protein and the ligand. The initial minimization consisted of 5000 steps of steepest descent and 10000 steps of a conjugate gradient. An additional 10,000 minimisation steps were performed without any restraints. The simulating systems were initially heated using the Langevin thermostat from 0 K to 100 K in isothermal-isochoric ensemble (NVT) ensemble with a 10.0 kcal/mol/$Å^2$ harmonic restraints applied on the non-hydrogen atoms of the protein and lipids for 12.5 ps. The systems were further heated to 310 K in isothermal–isobaric (NPT) ensemble with anisotropic pressure scaling and pressure scaling of 1 bar for 125 ps.

System equilibration was performed with an initial 5.0 kcal/mol/Å^2 harmonic restraint of on the protein and ligand. This was further lowered by 1.0 kcal/mol/Å^2 in a stepwise manner every 4 ns for a total of 20 ns at 310 K. An additional 5 ns unrestrained equilibration was performed before production. Finally, a duplicate all-atom MD simulation of 600 ns were run for each system, with randomized initial velocities were performed. An integration time step of 2 fs was set for each system at 310 K, and 1 bar in isothermal–isobaric (NPT) ensemble. The subsequent analysis was performed on the system where the final structure showed the utmost conformational stability. The trajectory coordinates were saved at every 50 ps intervals for subsequent analysis. All hydrogen containing bonds were constraint with the SHAKE algorithm(Ryckaert et al., 1977). For each system, the Langevin thermostat with a 1.0 ps^{-1} collision frequency, a 12.0 Å electrostatic cut-off, and the Particle Mesh Ewald method applied for long-range electrostatic interactions was set. The details of the simulation system parameters are shown in Table S1.

2.6 Analysis of MD trajectories

The simulation trajectories were analysed using the Amber 18 analysis tool CPPTRAJ(Roe & Cheatham, 2013). The overall complex root mean square deviation (RMSD) of each simulation was calculated. The radius of gyration (Rg) and the solvent-accessible surface area (SASA) values within 8 Å the binding site and the entire receptor for each system relative to the initial structural configuration were also analysed. The binding site was defined as all amino acid residues within 8 Å of the bound antagonist and the unbound binding pocket. The RMSD, SASA, and Rg values are expressed as mean ± standard deviation (Mean±SD). Snapshots of the complexes, average structures, and distances were also calculated using CPPTRAJ over the entire trajectories. Key biologically relevant protein-ligand interactions were generated with Discovery studio visualizer (Accelrys, 2013) and the Protein-Ligand Interaction Profiler (Salentin et al., 2015). The 3D visualised graphics designs of the protein-ligand interaction were done with PyMOL (Schrodinger, 2010) and Chimera (Pettersen et al., 2004).

2.7 Residue indices and numbering in transmembrane (TM) and Extracellular loops (ECL)

To enable comparison of residue among the two dopamine receptors, superscripts have been assigned to each residue denoting Ballesteros-Weinstein numbering (Ballesteros & Weinstein, 1995) for the transmembrane regions and the GPCRdb (Isberg et al., 2016) numbering for extracellular and intracellular loop regions. The Ballesteros-Weinstein nomenclature assigns

each transmembrane residue with an index number corresponding to the position of the residue in that transmembrane domain. For example, the conserved Asp residue in TM3 (which is Asp110$^{3.32}$ in D3DR and Asp114$^{3.32}$ in D2DR) is denoted as Asp$^{3.32}$ where 3 stands for TM3 and 32 stands for the position of the Asp residue in TM3.

2.8 Clustering and Principal component analysis (PCA)

The MD trajectories were subjected to clustering and principal component analysis (PCA) to probe the conformational relationships for the individual antagonist complex system. Individual frames were superposed to the original frame before the PCA calculations to eliminate rotation and displacement of structures. For each system having N atoms, the internal trajectory motion can be illustrated as 3 N_3 N of the covariance matrix C, where every column includes the cartesian coordinate X of every precise atom at each trajectory frame:

$$\mathrm{Cij} = < (X_i - < X_i >) (X_j - < Xj >) > \quad (1)$$

The decomposition of the covariance matrix C generates the principal modes as

$$\mathrm{C} = V\Lambda\mathrm{V}^{\mathrm{T}} \quad (2)$$

where V contains the eigenvector of matrix C and K denotes the eigenvalues of the diagonal matrix. In this study, the principal component analysis calculation and clustering analysis were calculated with the Amber 18 CPPTRAJ module and the R package Bio3D (Grant et al., 2006). The MD trajectories distribution was projected onto the first and second principal components.

2.9 Binding free energy calculations

The binding free energies were evaluated by the molecular mechanics/generalised Born solvent accessible surface area (MM/GBSA)(Miller III et al., 2012) method of Amber 18 to determine the interaction energies of R-VK4-40 and Y-QA31 at the D_3 and D_2 dopamine receptors. The *ante-MMPBSA.py* module (Miller III et al., 2012) of Amber 18 was utilized in generating the topologies of the *ligands, receptor*, and *complex* to be used in the MM/GBSA calculations. The MM/GBSA method estimates the relative binding free energy ($\Delta\mathrm{G}_{\mathrm{bind}}$) via the following equation:

$$\Delta\mathrm{G}_{bind} = \Delta\mathrm{E}_{MM} + \Delta\mathrm{G}_{sol} - \mathrm{T}\Delta\mathrm{S} \quad (3)$$

$$\Delta\mathrm{E}_{MM} = \Delta\mathrm{E}_{vdw} + \Delta\mathrm{E}_{ele} \quad (4)$$

$$\Delta\mathrm{G}_{sol} = \Delta\mathrm{G}_{pol} + \Delta\mathrm{G}_{nonpolar} \quad (5)$$

Where, $\Delta\mathrm{E}_{\mathrm{MM}}$ and $\Delta\mathrm{G}_{\mathrm{sol}}$ denote the molecular mechanics interaction energy and the solvation free energy, respectively. $-T\Delta S$ is the entropy contribution. The summation of the van der

Waals (ΔE_{vdw}) interaction energy and the electrostatic (ΔE_{ele}) interaction energy constitute the molecular mechanics interaction energy (ΔE_{MM}) between the receptor and ligand. The contribution of the solvation free energy is further decomposed into the polar (ΔG_{pol}) and the non-polar ($\Delta G_{nonpolar}$) terms.

In this study, the generalised Born approach was used with an implicit generalised Born solvent model and 0.15 M ionic strength. For each complex system, 2000 snapshots were extracted from the last 400ns trajectories at an interval of 200 ps. The entropy contribution was estimated using the normal mode harmonic approximation method. The decomposition of the pairwise interaction energies into active site residue contribution was also computed.

3 Results

3.1 Sequence and structural similarities between D_2 and D_3 dopamine receptors

Pairwise sequence alignment of D_2 and D_3 dopamine receptor was performed to identify the level of conservations in the amino acid sequence of the transmembrane helices (TM1-TM7), extracellular loops (ECL1-ECL3), and intracellular loops (ICL1-ICL3) (Figure 2). The amino acid sequence analysis reveals that the transmembrane domains TM2-TM7 are mainly conserved compared to the TM1 domain. The extracellular half of TM1 (1.35 to 1.50) between D2DR and D3DR contains 44% non-conserved residues (Figure 2). A substantial diversity in amino acids was also observed in the extracellular loop regions. Although the design of D_3/D_2 selective compounds has being a challenge; however, the observed differences between D_3 and D_2 receptors can be harnessed towards the development of selective antagonists.

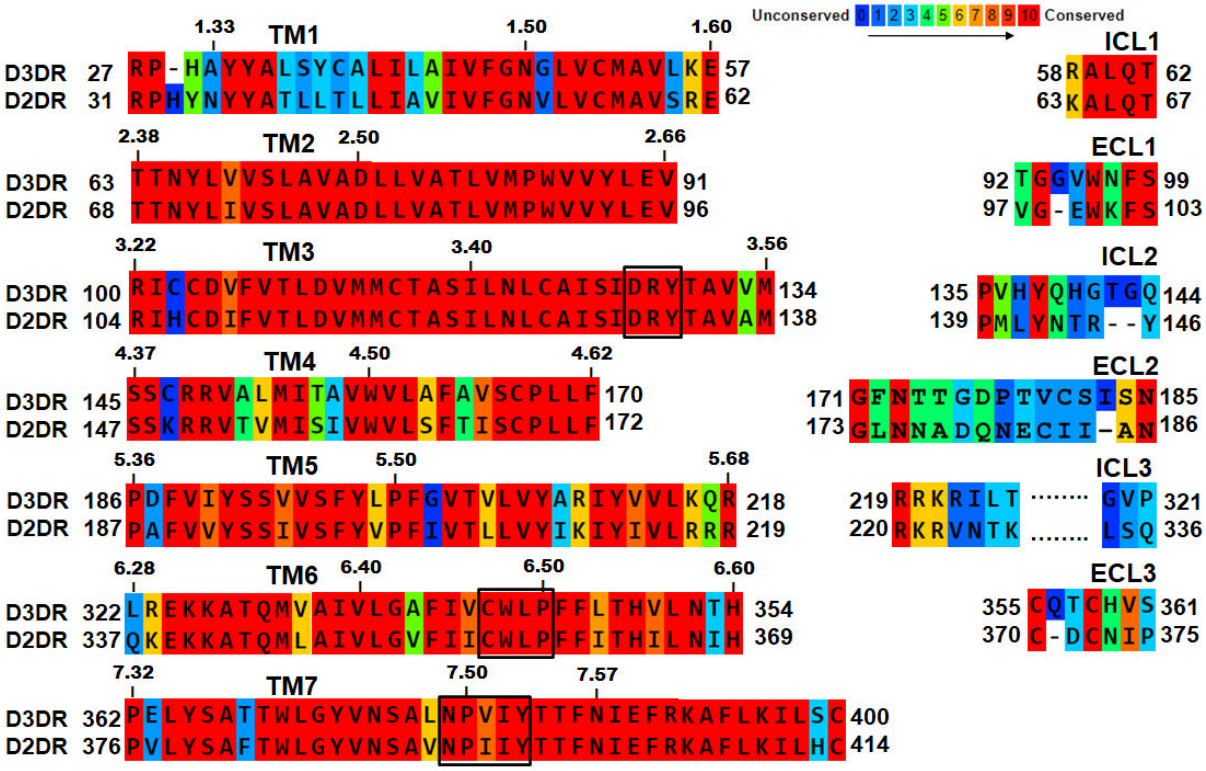


Figure 2 Comparison of amino acid sequences of D_2 and D_3 dopamine receptors showing conservation of residues and motifs. The black box markings on TM3, TM6 and TM7 highlights the conserved Class-A GPCR motifs DRY, CWXP and NPXXY, respectively. The transmembrane sections indexing is based on the Ballesteros & Weinstein system (Ballesteros & Weinstein, 1995). PRALINE multiple sequence analyser was used to construct the sequence alignment (Simossis & Heringa, 2005; Simossis & Heringa, 2003) (Image prepared by author).

3.2 Differential binding modes of R-VK4-40 and Y-QA31 to D_2 and D_3 dopamine receptors.

The potential complex structures of D2DR/R-VK4-40, D2DR/Y-QA31, D3DR/R-VK4-40, and D3DR/Y-QA31 used in this study were predicted with the Glide extra precision (XP) docking protocol. The crystallographic structures of D2DR bound with risperidone (6CM4) and D3DR in complex with eticlopride (3PBL) were used to predict the potential binding mode of R-VK4-40 and Y-QA31. To validate the predictive accuracy of the of Glide XP docking scoring, the co-crystallised inhibitors were extracted and re-docked into their respective receptors. The predicted binding poses and the experimental binding conformation overlaps in the same binding site (Figure S1). The interaction mode of R-VK4-40 and Y-QA31 with best-ranked conformers and lowest binding free energy score at D2DR and D3DR were selected for MD simulations. The predicted binding poses between the antagonists (R-VK4-40 and Y-QA31) and the receptors (D2DR and D3DR) are shown in Figure 3. In the predicted binding

mode of R-VK4-40 and Y-QA31 at D2DR, the indole moiety and the 1-(2-methoxyphenyl)piperazinyl moiety, respectively extend deeper into the orthosteric binding pocket (OBP) of D2DR defined by TM3, TM5 and TM6 side chains. In contrast, R-VK4-40 and Y-QA31 binding extend higher in the OBP, adopting a shallow binding mode at D3DR (Figure 3).

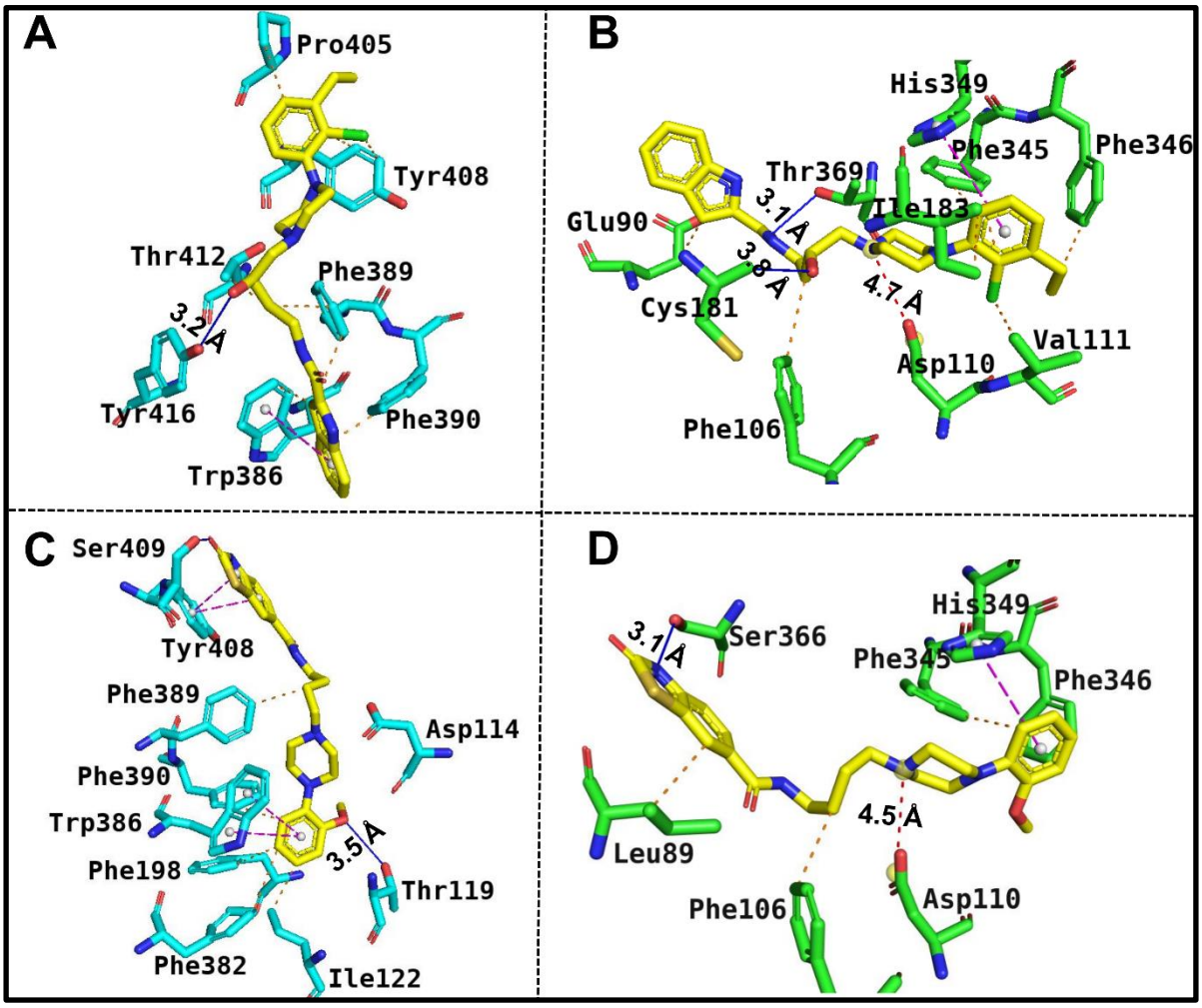


Figure 3 Intermolecular interactions for D2DR/R-VK4-40 (A), D3DR/R-VK4-40 (B), D2DR/Y-QA31 (C), and D3DR/Y-QA31 (D) from molecular docking. Hydrogen bonds, π-π stacking, hydrophobic and salt-bridge interactions are depicted in blue, magenta, orange, and red, respectively. Hydrogen bond distances are displayed in Angstrom (Å) (Image prepared by author).

3.3 Conformational dynamics of D_2 and D_3 dopamine receptors upon antagonist binding.

Molecular simulations in the past decades have enabled the detailed structural configurations and interactions of macromolecules with other molecular species to be characterised in different environments (Elber, 2016; Hosen et al., 2019). In this study, 600 ns all-atom MD

simulations were performed for each D_3 and D_2 dopamine receptor complexes and apo systems to investigate the detailed binding process and to access their dynamic structural characteristics. The dynamic stability, structural properties, as well as the energetics for all the complex systems, were assessed to obtain insights into the detailed structures of conformational change in each studied system.

3.4 Interaction energy components clarify selective binding potency at D3DR over D2DR.

The relative binding free energies were evaluated for the four complexes by the MM/GBSA approach and were decomposed into various energy contribution terms (solvation energies, gas-phase energies, and entropic contribution). Computing the binding free energy was to evaluate inhibitor binding selectivity towards D2DR and D3DR. The MM/GBSA method can provide a high correlation with experimental values and good ranking of protein-ligand complexes (Genheden & Ryde, 2015; Sun et al., 2018). Table 1 shows the estimated binding free energies (ΔG_{bind}) by MM/GBSA from the MD trajectories, and the experimental binding energy (ΔG_{exp}) obtained from the experimental inhibition constants (K_i) in Figure 1 (Sun et al., 2016; Jordan, Humburg, Rice, et al., 2019). The experimental binding energies (ΔG_{exp}) were calculated from the experimental inhibition constant (K_i) using the formula: $\Delta G_{(inhibition)} = RTlnK_i$, where T is the temperature (298.15 K), and R is the universal gas constant (1.985×10^{-3} kcal/mol/K). The predicted MM/GBSA binding affinities for R-VK4-40 and Y-QA31 at D3DR were -23.62 ± 1.37 kcal/mol and -24.65 ± 0.93 kcal/mol, whereas that of R-VK4-40 and Y-QA31 at D2DR were -11.05 ± 0.58 kcal/mol and -10.01 ± 1.36 kcal/mol, respectively. Similarly, the experimental binding free energy (ΔG_{exp}) calculated for R-VK4-40 and Y-QA31 at D3DR were -12.99 kcal/mol and -13.02 kcal/mol, whereas that of R-VK4-40 and Y-QA31 at D2DR were -9.70 kcal/mol and -9.93 kcal/mol, respectively. The above obtained quantitative theoretical and experimental binding free energies indicate that R-VK4-40 and Y-QA31 are less potent at D2DR than at D3DR.

Table 1 Individual energy component contribution to the overall binding free energy (kcal/mol) estimation using the MM-GBSA approach for D2DR and D3DR complexes.

Energetic	**R-VK4-40**		**Y-QA31**	
Terms	**D2DR**	**D3DR**	**D2DR**	**D3DR**
ΔG_{vdW}	-46.61 ± 0.06	-54.65 ± 0.08	-47.06 ± 0.10	-57.71 ± 0.08
ΔG_{elec}	-13.29 ± 0.08	-20.40 ± 0.19	-8.24 ± 0.17	-11.96 ± 0.11
ΔG_{polar}	30.25 ± 0.06	37.24 ± 0.14	24.89 ± 0.13	31.35 ± 0.08
$\Delta G_{nonpolar}$	-5.80 ± 0.01	-6.78 ± 0.01	-5.70 ± 0.01	-7.42 ± 0.01
ΔG_{MM}	-59.90 ± 0.10	-75.05 ± 0.23	-55.30 ± 0.22	-69.67 ± 0.13
ΔG_{sol}	24.46 ± 0.06	30.46 ± 0.14	19.19 ± 0.12	23.93 ± 0.08
ΔH	-35.44 ± 0.08	-44.58 ± 0.15	-36.11 ± 0.14	-45.74 ± 0.10
$-\Delta TS$	24.39 ± 0.66	20.96 ± 1.52	26.10 ± 1.50	21.09 ± 1.03
ΔG_{bind}	**-11.05 ± 0.58**	**-23.62 ± 1.37**	**-10.01 ± 1.36**	**-24.65 ± 0.93**
$^{a}\Delta G_{exp}$	**-9.70**	**-12.99**	**-9.93**	**-13.02**

ΔG_{vdW} = van der Waals energy; ΔG_{elec} = Electrostatic energy; ΔG_{MM} = molecular mechanics energy; ΔG_{sol} = solvation energy contribution; $\Delta G_{nonpolar}$ = Nonpolar solvation energy; ΔG_{polar} = Polar solvation energy; $-\Delta TS$ = Entropy contribution; ΔG_{bind} = Binding free energy. $^{a}\Delta G_{exp}$ = The calculated experimental binding free energy (ΔG) obtained from the experimental inhibition constant (Ki) in Figure 1 using the formula: $\Delta G(\text{inhibition}) = RT\ln K_i$, where T is the temperature (298.15 K), and R is the universal gas constant (1.985×10^{-3} kcal/mol/K).

From the analyses, van der Waals energy and electrostatic energy were observed to be the significant contributors to the overall binding of the antagonists. In the interactions of compounds R-VK4-40 and Y-QA31 with D2DR, the electrostatic (-13.29 and -8.24 kcal/mol, respectively) and the van der Waals (-46.61 and -47.06 kcal/mol, respectively) energy values suggest that van der Waals interactions (hydrophobic interactions) were the dominant binding forces of D2DR with its antagonists. Similarly, in the interactions of R-VK4-40 and Y-QA31 with D3DR, electrostatic (-20.40 and -11.96 kcal/mol, respectively) and van der Waals interactions (-54.65 and -57.71 kcal/mol, respectively) also made favourable interactions with D3DR binding site residues. Thus, van der Waals interaction of R-VK4-40 and Y-QA31 were more dominant at D3DR than at D2DR. For all the ligand-receptor complexes, unfavourable polar solvation energy opposed the favourable electrostatic energies in the gas phase.

Furthermore, conformational entropy ($-\Delta TS$) contribution was observed to be relatively more unfavourable for inhibitor binding at D2DR than at D3DR. As highlighted in the above descriptions, the increase in binding enthalpy with the corresponding decrease in entropy of R-VK4-40 and Y-QA31 at D3DR relative to D2DR seems to mainly drive the selective binding of these antagonists toward D3DR over D2DR. Thus, optimising the binding enthalpy of antagonists to D3DR over D2DR may be an essential aspect towards the design of D3DR selective antagonists.

3.5 Inhibitor−residue interaction network reveals the basis of antagonist selectivity

To uncover the atomistic interaction mechanistic of the antagonist selectivity and specificity at D3DR over D2DR, we investigated receptor residues that significantly interacted with the studied antagonists. The decomposition of the predicted binding affinities per active site residues was analysed to obtain further insight into the patterns of receptor-antagonist recognition and the distributions of the energy by active site residues. The residues with substantial energy contributions to antagonist binding are presented in Figure 4.

3.5.1 Binding of R-VK4-40 to D3DR versus D2DR. As shown in Figure 4A, R-VK4-40 made favourable interactions greater than −0.6 kcal/mol with ten residues in D2DR. These are Val115$^{3.33}$ (-1.43), Ile184$^{45.52}$ (-0.67), Asn186$^{5.35}$ (-0.64), Pro187$^{5.36}$ (-2.53) Phe189$^{5\,38}$ (-1.93), Val190$^{5.39}$ (-2.63), Ser193$^{5.42}$ (-1.70), Phe389$^{6.51}$ (-1.06), His393$^{6.55}$ (-0.77), and Ile397$^{6.59}$ (-1.95). The interaction energies of Phe189$^{5.38}$, Phe389$^{6.51}$ and His393$^{6\,55}$ agree structurally with the π-π stacking interactions with the indole ring and the CH-π contacts of the alkyls of Val190$^{5.39}$, Pro187$^{5.36}$ and Ile397$^{6.59}$ with the chloro-ethylbenzene ring of R-VK4-40 (Figure S2). The indole group also engaged in a π-sulphur and π-alkyl interaction with Cys118$^{3.36}$ and Val115$^{3.33}$, respectively. It is worth remarking that R-VK4-40 engages in three hydrogen bond interactions with D2DR, such as Ser193$^{5.42}$, Ile184$^{45.52}$ and His393$^{6.55}$ (Figure S2). The hydrogen bond of Ser193$^{5.42}$ is with the indole nitrogen, Ile184$^{45.52}$ with the hydroxy group in the N-((S)-3-hydroxypentyl)acetamide linker, while His393$^{6.55}$ with the oxygen atom in the acetamide unit (Figure S2).

Compared to the D2DR/R-VK4-40 complex, the binding interactions of R-VK4-40 with D3DR are relatively different (Figure 4B and Figure S3). The interaction energies of R-VK4-40 with ten residues in D3DR are observed to have energy values stronger than −0.6 kcal/mol (Figure 4B). These include Val111$^{3.33}$ (-2.47), Cys114$^{3.36}$ (-0.71), Thr115$^{3.37}$ (-0.80), Ser182$^{45.51}$ (-1.11), Ile183$^{45.52}$ (-1.50), Val189$^{5.39}$ (-0.85), Phe345$^{6.51}$ (-1.00), His349$^{6.55}$ (-0.90), Tyr365$^{7.35}$ (-0.82) and Thr369$^{7.39}$ (-1.19). The 1-(2-chloro-3-ethylphenyl)piperazinyl moiety engaged in π-π interactions with Phe197$^{5.47}$, and Tyr365$^{7.35}$, as well as π-alkyl interactions with Val111$^{3.33}$, His349$^{6.55}$, Ile183$^{45.52}$, and Phe345$^{6.51}$ (Figure S3). Additionally, π-sulphur interaction is observed with Cys114$^{3.36}$ and the 1-(2-chloro-3-ethylphenyl)piperazinyl moiety. R-VK4-40 forms a salt bridge interaction between the linker hydroxy and Asp110$^{3.32}$. A hydrogen bond is also observed between the indole nitrogen and Ser366$^{7.36}$. Additional hydrogen bonds are further formed between the carboxamide and Thr369$^{7.39}$, and Tyr365$^{7.35}$, whereas the hydroxy on the linker engage in a hydrogen bond with Ser182$^{45.51}$ (Figure S3).

The binding of R-VK4-40 at D2DR made substantially distinct interactions with residues mainly in the TM5 and TM6 domains. In contrast, R-VK4-40 made significant interactions with residues in the TM3 and TM7 domains of D3DR. However, a few common interactions are observed for residues such as Val$^{3.33}$, Ile$^{45.52}$, Phe$^{6.51}$ and His$^{6.55}$ in both receptors, which cumulatively displayed higher inhibitor interactions at D3DR over D2DR.

3.5.2 Binding of Y-QA31 to D3DR versus D2DR. As observed in Figure 4C and Figure S4, Y-QA31 made favourable interactions with energy values stronger than -0.6 kcal/mol with eight residues in D2DR. These residues include Ile184$^{45.52}$ (-0.82), Val190$^{5.39}$ (-1.35), Ser193$^{5.42}$ (-1.33), Ser194$^{5.43}$ (-1.66), Phe198$^{5.47}$ (-1.08), Phe389$^{6.51}$ (-1.54), Phe390$^{6.52}$ (-0.66) and His393$^{6.55}$ (-0.65) (Figure 4C). The deep binding pocket residues Phe198$^{5.47}$, Trp386$^{6.48}$ and Phe389$^{6.51}$ structurally make π-π stacking interactions with the 1-(2-methoxyphenyl)piperazinyl moiety of Y-QA31. Ile184$^{45.52}$ engages in both hydrogen bond and π-alkyl interactions with the carboxamide nitrogen and benzo[d]thiazol-2(3H)-one, respectively (Figure S4). Additionally, the oxygen and the nitrogen atom of the benzo[d]thiazol-2(3H)-one moiety formed hydrogen bonds with Tyr37$^{1.35}$ and Thr412$^{7.39}$, respectively. It is also noted that the benzo[d]thiazol-2(3H)-one moiety and the carboxamide oxygen engaged in π-π stacking and hydrogen bond interactions with Tyr408$^{7.35}$, respectively. Other π-alkyl interactions were observed for Val190$^{5.39}$, Pro405$^{7.32}$ and Ala185^{ECL2} (Figure S4).

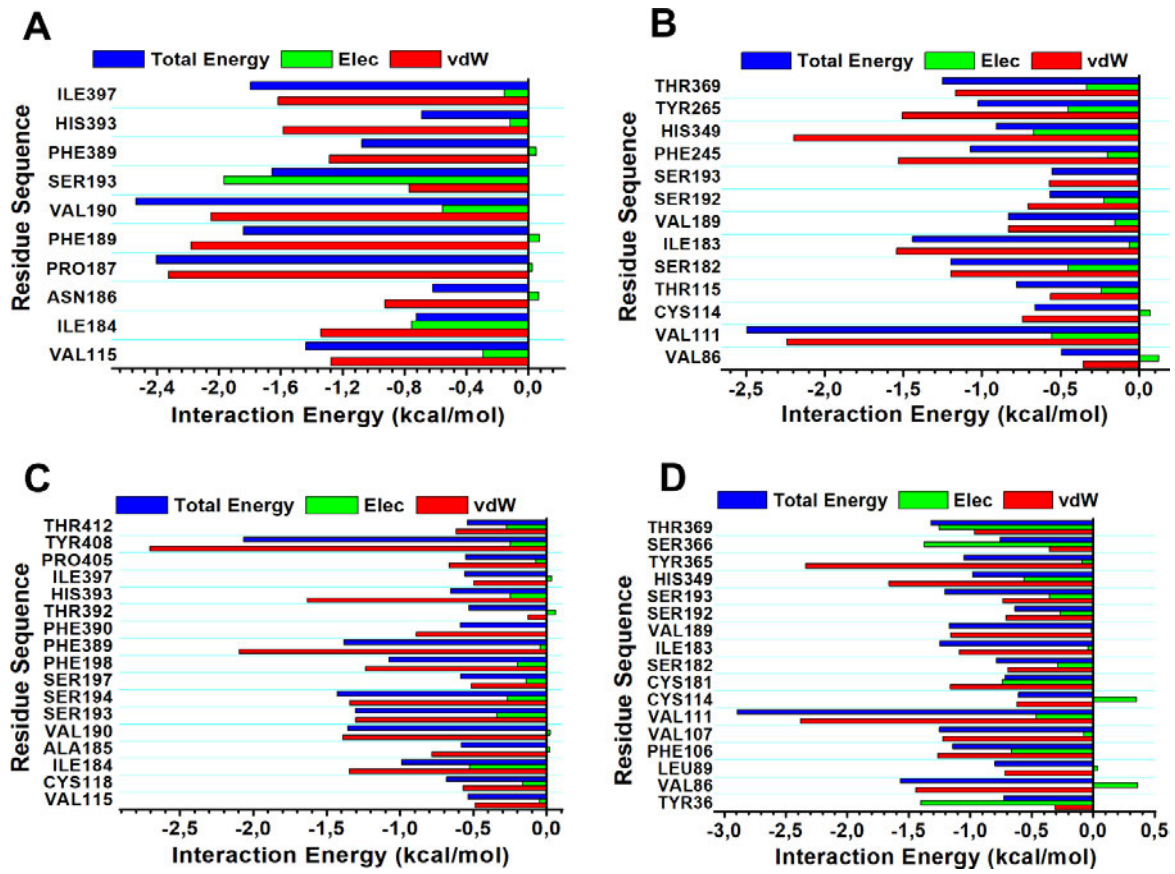


Figure 4 The interaction energy spectrum of the two antagonists with key active site residues in D2DR and D3DR. (A) D2DR/R-VK4-40 complex, (B) D3DR/R-VK4-40 complex, (C) D2DR/Y-QA31 complex, and (D) D3DR/Y-QA31complex systems (Image prepared by author).

On the other hand, the binding interactions of Y-QA31 with D3DR are quite dissimilar to D2DR. It is observed that Y-QA31 made favourable interactions with 15 residues in D3DR with energy values stronger than -0.6 kcal/mol (Figure 4D). These critical residues are Tyr36$^{1.39}$ (-0.64), Val86$^{2.61}$ (-1.60), Leu89$^{2.64}$ (-0.70), Phe106$^{3.28}$ (-1.15), Val107$^{3.29}$ (-1.28), Val111$^{3.33}$ (-2.83), Cys181$^{45.50}$ (-0.71), Ser182$^{45.51}$ (-0.85), Ile183$^{45.52}$ (-1.27), Val189$^{5\,39}$ (-1.22), Ser193$^{5.43}$ (-1.20), His349$^{6.55}$ (-1.07), Tyr365$^{7.35}$ (-0.94), Ser366$^{7.36}$ (-0.61) and Thr369$^{7.39}$ (-1.35). The interaction energies are structurally consistent with the π–π interactions of Phe106$^{3.28}$, and Tyr365$^{7.35}$ with the benzo[d]thiazol-2(3H)-one group as well as the π-alkyl interactions of Val86$^{2.61}$ Val111$^{3.33}$ Ile183$^{45.52}$ Val189$^{5.39}$ and His349$^{6.55}$ with the benzo[d]thiazol-2(3H)-one and 1-(2-methoxyphenyl)piperazinyl moieties (Figure S5). Additionally, 1-(2-methoxyphenyl)piperazinyl group further engaged in a π-sulphur and π-π interactions with Cys114$^{3.36}$ and Phe346$^{6.52}$, respectively. It is worth highlighting that Y-QA31 interacts with six D3DR residues via hydrogen bonding, including Tyr36$^{1.39}$, Asp110$^{3.32}$, Ser366$^{7.36}$, Thr369$^{7.39}$,

Tyr373$^{7.43}$, and His349$^{6.55}$ (Figure S5). The oxygen and the nitrogen atom in the benzo[d]thiazol-2(3H)-one moiety constantly engaged in hydrogen bonding with Tyr36$^{1.39}$ and Tyr373$^{7.43}$, respectively. Additionally, Ser366$^{7.36}$ and Thr369$^{7.39}$ made hydrogen bond contacts with the benzo[d]thiazol-2(3H)-one moiety (Figure S5). The carboxamide nitrogen was observed to make constant salt bridge interactions with Asp110$^{3.32,}$ which is critical for high-affinity inhibitor binding of aminergic GPCRs. The binding of Y-QA31 at D2DR engaged in more interactions with residues, mainly in the TM5 and TM6 domains. In contrast, Y-QA31 made significant interactions with residues in the TM3, ECL2, and TM7 domains of D3DR. However, a few common interactions were observed for residues such as Val$^{5.39}$, Ser$^{5.43}$, and His$^{6.55}$ in both receptors.

3.6 Conformational analyses upon ligand binding

The Cα root mean square deviation (RMSD) values were calculated from initial structures to access the dynamic stability and convergence of the MD simulations for the unbound and bound systems. The RMSD of the Cα atoms of the apo and ligand-receptor complexes for each system conducted in duplicate runs is shown in Figure S6 and Table S2. Convergence were generally observed after 250 ns of the simulation with average RMSDs less than 3.0 Å.

The radius of gyration (RoG) was computed to obtain a local (binding site residues within 8 Å of the inhibitor) and global assessment of the distribution of the receptor compactness. The radius of gyration of the receptor-ligand complex is a measure of how compact a protein is upon ligand binding (lower Rg values represent higher protein compactness) (Tanwar et al., 2017). The analysis of the simulated systems showed lower average binding pocket RoG values of 5.98 ± 0.39 Å and 5.58 ± 0.37 Å for R-VK4-40 and Y-QA31 binding at D3DR compared with the observed 10.52 ± 0.10 Å, and 10.72 ± 0.14 Å for R-VK4-40 and Y-QA31 binding at D2DR, respectively The unbound binding pocket RoG were relatively higher than the bound conformation with values of 12.77 ± 0.17 Å and 12.42 ± 0.17 Å for D2DR and D3DR, respectively (Figure 5A and Table S2). The binding pockets of the systems showed distinct compactness compared to the entire bound and unbound systems (Figure 5C and Table S2). The predicted results suggest that R-VK4-40 and Y-QA31 induced a much compact binding pocket when bound to D3DR than when bound D2DR. The comparative analyses suggest a wider configurational binding site cavity conformation at D2DR, resulting in weaker binding of R-VK4-40 and Y-QA31 compared to when bound at D3DR.

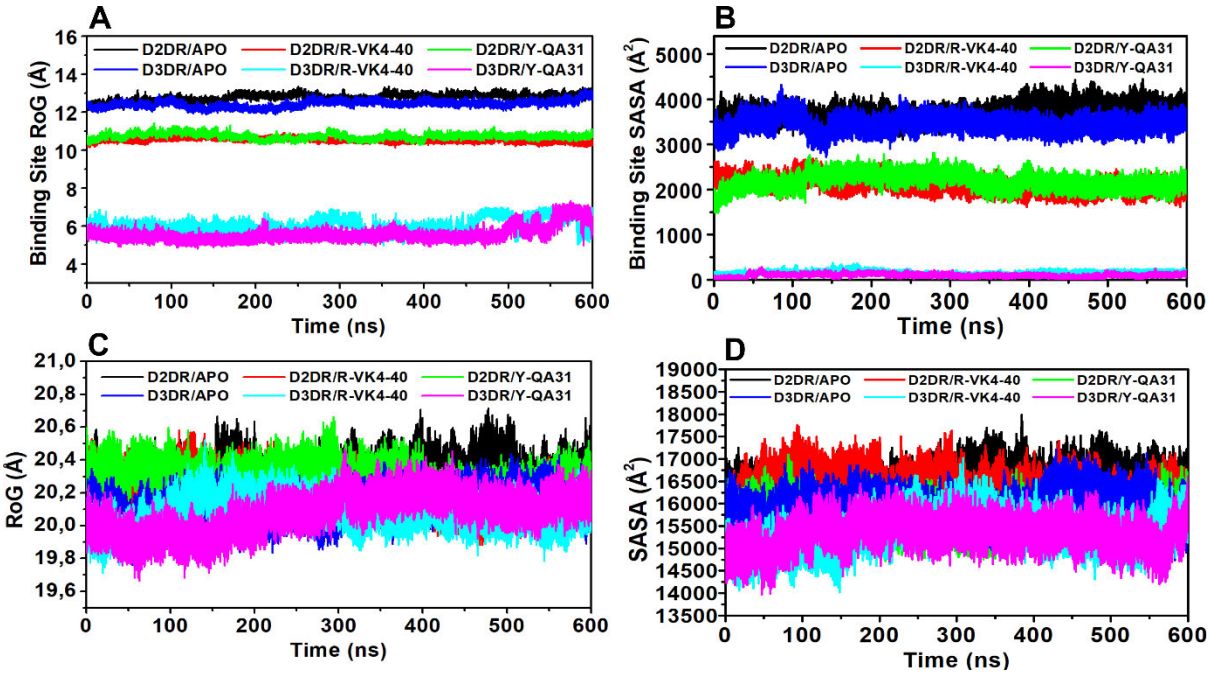


Figure 5 The radius of gyration of the binding pocket (A) and entire system (C); and the solvent accessible surface area of the binding pocket (B) and the entire system (D) for D2DR and D3DR as a function of simulation time (Image prepared by author).

The accessibility of the binding site cavity to the solvent upon antagonist binding was further analysed. In general, decreasing SASA value corresponds to increasing hydrophobic interactions. The trend of solvent accessible surface area (SASA) profiles was similar to the Rg profiles. The average D2DR binding pocket SASA values for R-VK4-40 and Y-QA31 complex were 2090.52 ± 143.14 Å^2, 2166.47 ± 157.98 Å^2, respectively; however, a significantly lower average SASA value of 153.63 ± 35.46 Å^2, 94.15 ± 35.65 Å^2 was observed for R-VK4-40 and Y-QA31 respectively at D3DR binding site. The average SASA values for the unbound D2DR and D3DR binding pocket were 3699.43 ± 166.47 Å^2 and 3439.40 ± 152.35 Å^2, respectively, which were higher than the bound systems (Figure 5B and Table S2). Thus, R-VK4-40 and Y-QA31 engage in stronger nonpolar interactions with D3DR hydrophobic binding site residues compared with D2DR binding pocket. However, the difference in the computed average SASA values for the entire receptor in the bound and unbound conformations (Figure 5D and Table S3) were not highly significant as the binding pocket.

3.7 Conformational clustering and principal component analysis (PCA)

Principal component analysis (PCA) was performed to understand the dynamic behaviour of the receptors when inhibited by the selected antagonists. PCA captured the dominant motion extracted from a large number of conformations(Sittel et al., 2014). The eigenvectors of the

covariance matrix were used to describe the overall coordinated movement of the C_α atoms, which relates to the coincident eigenvalues.

In R-VK4-40 binding at D2DR and D3DR, the top 20 principal components (PCs) accounted for 82.3% and 74.9% of the total variance in the MD simulation, respectively. The contribution of the first two PCs to the total variance in D2DR/R-VK4-40 complex were 37.5% and 10.7%, respectively, whereas in D3DR/R-VK4-40 complex they accounted for 26.4% and 8.8%, respectively (Figure 6B and 6D). Similarly, the top 20 PCs for the Y-QA31 binding at D2DR and D3DR accounted for 81.8% and 74.4% of the total covariance, respectively (Figure 7B and 7D). The contribution of PC1 and PC2 for D2DR/Y-QA31 complex were 32.8% and 9.9%, respectively, whereas accounting for 21.9% and 12.9%, respectively in D3DR/Y-QA31 complex. Thus, the selective binding of R-VK4-40 and Y-QA31 at D_3 dopamine receptor is associated with a relatively lower proportion of variance compared to their binding at the D_2 dopamine receptor.

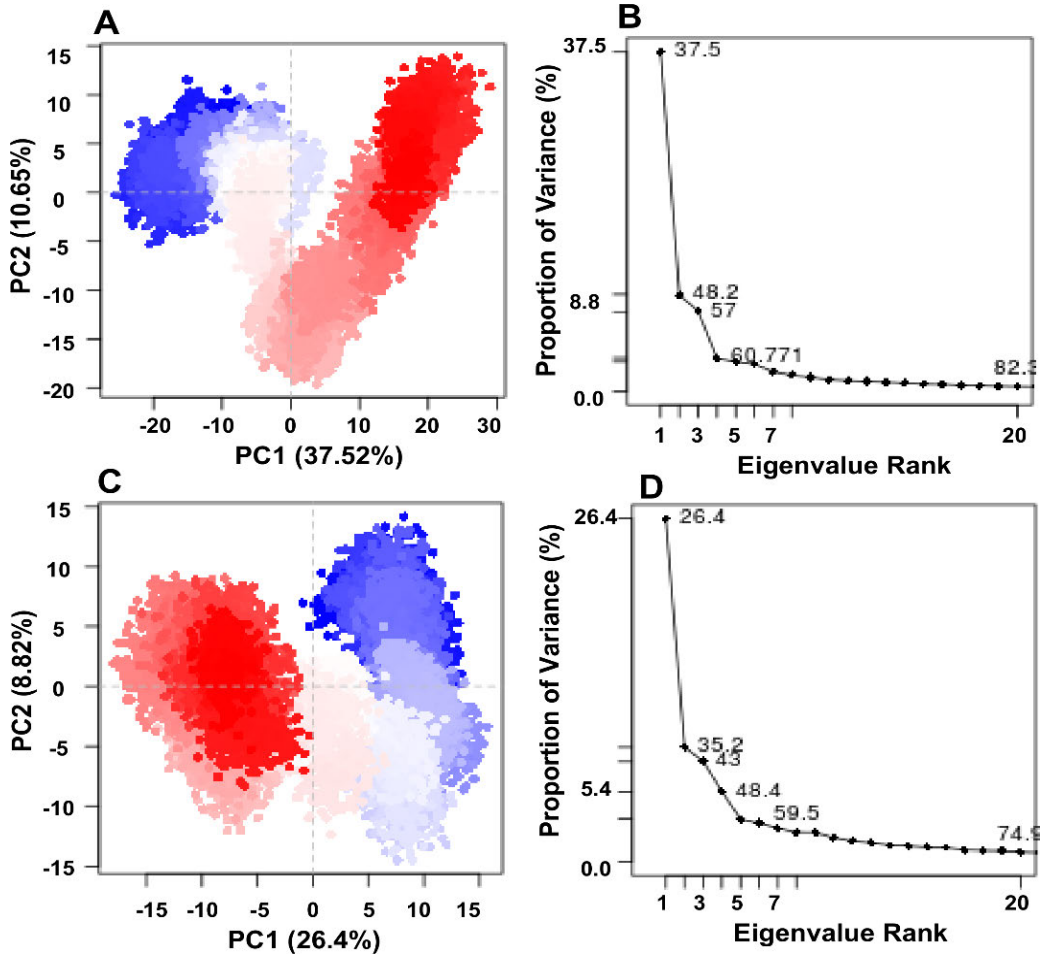


Figure 6 The projections of trajectories onto the first two principal components subspace and proportion of variance for D2DR/R-VK4-40 complex (A and B) and D3DR/R-VK4-40 complex (C and D) (Image prepared by author).

The overall contributions of the Cα atomic fluctuations variance relative to the first, second and the remaining PCs are summarised in Table S4. Hence, the first two PCs (PC1 and PC2) were sufficient in providing a valuable description of the significant fluctuations in the conformational ensembles. The conformational transitions of the studied systems were analysed based on the first and second PCs in two-dimensional subspace (Figure 6A and 6C; Figure 7A and 7C). The continuous transition of colour, which changed from red to white to blue, indicate periodic transitions between these conformations. Two distinct conformational regions were sequentially explored by D2DR/R-VK4-40 which displayed distinct periodic jumps with a substantial energy barrier while D3DR/R-VK4-40 complex showed an overlapping and uniform exploring of conformational subspace with minimal energy barrier (Figure 6). Similarly, two main conformational subspaces were visited by D2DR/Y-QA31 and D3DR/Y-QA31complexes, however wider displacement were observed along PC1 for D2DR/Y-QA31 compared to D3DR/Y-QA31complex (Figure 7). The wide displacement is an indicative of large conformational space explored by D2DR/Y-QA31. Thus, the studied D_3 selective antagonists displayed large conformational changes mainly along PC1 at D2DR than at D3DR.

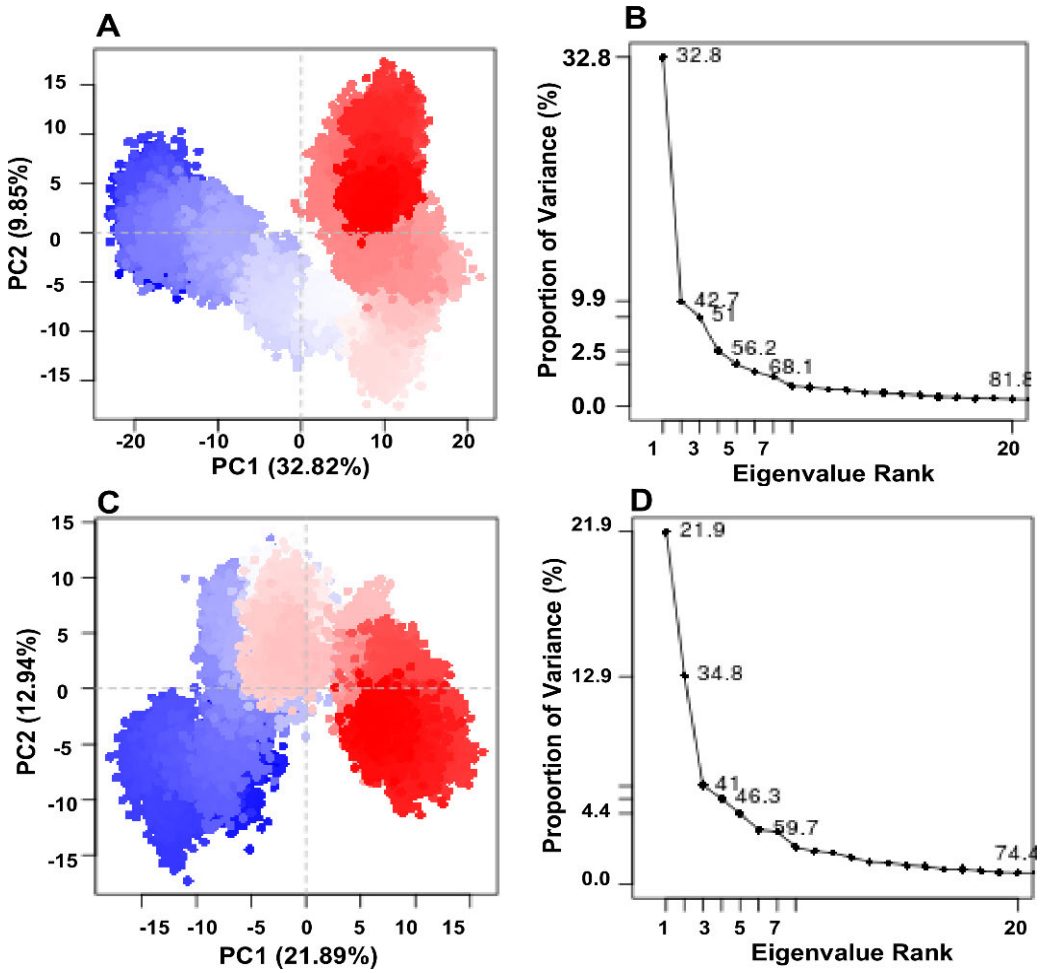


Figure 7 The projections of trajectories onto the first two principal components subspace and proportion of variance for D2DR/Y-QA31 complex (A and B) and D3DR/Y-QA31complex (C and D) (Image prepared by author).

4 Discussion

Drug binding selectivity and action are crucial in drug development that targets receptor subtypes. The availability of crystal structures provides an opportunity to identify structural differences at the atomistic molecular level between closely related GPCRs that can be exploited for novel drug design (Chien et al., 2010; Wang et al., 2018; Fan et al., 2020). However, the structural determining factors of inhibitor binding specificity at the D_3 and D_2 dopamine receptors are highly subtle due to the nearly identical binding pocket residues. The discovery of selective and potent D_2-like receptor inhibitors holds promise in the development of next-generation antipsychotic drugs in the treatment of neurological disorders (Holmes et al., 2004; Heidbreder & Newman, 2010; Xiao et al., 2014; Moritz et al., 2018). Herein, we employed the crystal structures of D_2 dopamine and D_3 dopamine receptors to investigate the

structural basis and atomistic binding mechanistic of two selective D_3 antagonists at D_3 over D_2 dopamine receptors.

The structural and physical properties of ligand binding pockets are closely associated with the ligand-binding affinities (Li et al., 2015; Smith et al., 2012). Bitopic antagonists are depicted by two pharmacophores, a primary and secondary pharmacophore connected through a linker. According to a previous study, the antagonistic activities of these compounds is more pronounced against D_3 dopamine receptor relative to D_2 dopamine receptor (Chien et al., 2010). The crystal structure of D3DR revealed an extended binding pocket (EBP) above the orthosteric binding pocket (OBP), where one of the bitopic pharmacophores interacts with TM2, TM3 and TM7 (Chien et al., 2010). In the predicted binding modes of R-VK4-40 and Y-QA31, the indole moiety and the 1-(2-methoxyphenyl)piperazinyl moiety of R-VK4-40 and Y-QA31 extended deep into the OBP of D2DR defined by TM6, TM5 and TM3 side chains, respectively.

On the other hand, the interaction patterns of R-VK4-40 and Y-QA31 extend higher above the OBP into the EBP, adopting a shallow binding mode at D3DR, making significant interactions with residues in the TM3 and TM7 domains. Newman and colleagues studied a series of D3DR bitopic selective 4-phenylpiperazine-substituted compounds. They found that the efficacy of these compounds depends on their binding mode in the OBP, whereas selectivity arises from their different interactions within the SBP (Newman, Beuming, et al., 2012). The longer linker between the binding moieties of R-VK4-40 and Y-QA31 enabled their binding at both the EBP and the OBP of D3DR (Heidbreder & Newman, 2010). The favourable flexibility of bitopic compounds seems to be conferred by the tetramethylene linker resulting in higher D3DR/D2DR selectivity (Newman, Beuming, et al., 2012).

Identifying the differences in the binding pocket dynamics of closely related receptor subtypes provides information that may assist in the design of subtype selective ligands (Latorraca et al., 2017). We further investigated the antagonist effect on the conformational dynamics (compactness and solvent accessibility) of D2DR and D3DR binding/active sites. These provide information on how the different inhibitors perturb the intrinsic dynamics of the binding site of the antagonist-bound complexes. The interactions of R-VK4-40 and Y-QA31 at D_3 dopamine receptor binding site is significantly associated with a lower average SASA and Rg values than at D2DR binding site. The radii of gyration within the binding pockets of D3DR reveals a more compact binding site with less accessibility to solvent compared to D2DR binding site. The observed binding sites conformational dynamics suggests that the binding of R-VK4-40 and Y-QA31 at D3DR yields a more compact ligand-receptor complex

with increased hydrophobic interactions compared to D2DR. This conformational change in D3DR binding site may result in the more potent inhibitor binding mode with a higher binding affinity. The relatively high hydrophobic interactions exhibited by R-VK4-40 and Y-QA31 at D3DR are consistent with an earlier report that hydrophobic interactions, particularly with TM7 are significantly involved in the binding of D3DR selective bitopic compounds at the EBP of D3DR (Newman, Beuming, et al., 2012).

The calculated binding free energies successfully predicted the higher binding affinities of R-VK4-40 and Y-QA31 at D_3 dopamine receptor over the D_2 dopamine receptor. It is worth noting that, the selectivity of R-VK4-40 and Y-QA31 at D3DR over D2DR is principally contributed by the van der Waals interactions, enthalpic and entropic energy contribution difference. The selective binding of the selective antagonists (R-VK4-40 and Y-QA31) at D3DR showed higher enthalpic interaction energy contribution while inducing a relatively favourable entropic contribution to their total binding free energies when compared to their binding at D2DR. Generally, the entropy change upon ligand binding may arise from an alteration of the conformational flexibility of the binding associates, a decrease in the rotational and translational degrees of freedom, as well as from the restructuring of their solvation shells upon binding (Bezerra et al., 2012). Desolvation effects are usually considered to drive a favourable entropic contribution to the total binding free energy as the burial of hydrophobic surfaces results in the displacement of binding site water molecules into the bulk solvent (Bezerra et al., 2012). The relatively high hydrophobic interactions (lower binding pocket solvent accessible surface area) displayed by R-VK4-40 and Y-QA31 at D3DR compared to D2DR may explain their relatively favourable entropic contribution. Also, Yildirim and Colleagues found that an increase in the conformational entropy upon ligand (pollutant) binding was associated with TPK biomolecules undergoing large conformational changes (Yildirim et al., 2016). The clustering and principal component analyses indicate that the selective binding of R-VK4-40 and Y-QA31 at D2DR is associated with a higher proportion of variance and more substantial conformational changes along PC1 relative to D3DR. The increase in conformational change at D2DR from PCA is observed to correlate with the relative increase in conformational entropy at D2DR.

The presence of water molecules in a hydrophobic enclosed pocket is less stable than the bulk water molecules due to less favourable enthalpy and/or entropy, and the ability of a ligand to displace these high-energy waters has been shown to affect selectivity (Beuming et al., 2009) and improve binding affinity (Higgs et al., 2010). In a WaterMap analysis to characterise the distribution of hydration sites in the EBP of D3DR, Newman and Colleagues found that several

high-energy water molecules were continuously positioned in the EBP of both D2DR and D3DR. In contrast, less high-energy waters were found at the TM1, TM2 and TM7 interface (Newman, Beuming, et al., 2012). The authors observed that the indole moiety of the D3DR-selective bitopic ligand (Compound 2) occupied the EBP region where several high-energy water molecules were displaced, which was consistent with compound's higher affinity for D3DR over D2DR (Newman, Beuming, et al., 2012).

The binding of R-VK4-40 and Y-QA31 made salt bridge interaction with the conserved Asp110$^{3.32}$, which is critical for inhibitors displaying high affinity to GPCR aminergic subfamily (Shi & Javitch, 2002). The ECL2 has been posited to be important in the selective targeting between D_3 and D_2 dopamine receptors since it is highly non-conserved (Heidbreder & Newman, 2010). Three critical ECL2 residues (Cys181$^{45.50}$, Ser182$^{45.51}$, and Ile183$^{45.52}$) of D_3 receptor were observed to make significant interactions with Y-QA31. Similarly, two ECL2 residues (Ser182 and Ile183) of D_3 dopamine receptor were also found to make critical interactions with R-VK4-40. The critical D3DR residue (Ser182$^{45.51}$) is non-conserved at D2DR (Ile182$^{45.51}$). Thus, Ser182$^{45.51}$ of the ECL2 region of D_3 dopamine receptor is a possible critical residue involved in the D_3/D_2 subtype selectivity of the studied antagonists. Feng and colleagues found Cys181$^{45.50}$ and Ser182$^{45.51}$ of the ECL2 of D3DR to make important interaction with the D_3 selective antagonist R-22 (Feng et al., 2012). The non-conserved residue Tyr36$^{1.39}$ of D3DR, which corresponds to Leu41$^{1.39}$ in D2DR formed a critical hydrogen bond and made more substantial interaction energy with Y-QA31 at D3DR which were not observed at D2DR. The findings highlight the essential differences in the interactions of R-VK4-40 and Y-QA31 at the binding pocket of D3DR and D2DR.

5 Conclusion

This study aimed at probing the molecular mechanistic underlying the selective binding of two antagonists (R-VK4-40 and Y-QA31) toward D3DR over D2DR using conventional molecular dynamics simulations in a POPC lipid bilayer environment. The results highlighted in this study show how R-VK4-40 and Y-QA31 achieve selective interaction at D_3 dopamine receptor over D_2 dopamine receptor. The results reveal that conformational alterations in D2DR were more prominent compared to D3DR. The dynamic binding mode and the per-residue energy decomposition enabled the identification of residues responsible for the differential binding affinities at D2DR and D3DR. The estimated binding free energies not only corroborated with the experimental binding affinities but also indicate that an increase in van der Waals interactions and a relative decrease in entropy contribution are the essential factors that underlie

high-selectivity and affinity of the antagonists for D3DR relative to D2DR. We expect that our study may assist in gaining further insight into the selective mechanistic of D3DR antagonists and provide vital information towards the design and identification of new selective antagonists of D3DR.

Authors Contributions

P. Appiah-Kubi conceived the study, performed the experiments, analysis and interpretation of data, and manuscript write up. FA. Olotu reviewed and edited the final manuscript draft. M.E.S Soliman supervised the study.

Acknowledgements

The authors acknowledge the College of Health Science, University of KwaZulu-Natal (UKZN), for their financial support and the Centre for High-Performance Computing (http://www.chpc.ac.za) for computational resources, including the license for Schrodinger Suite.

Conflict of interest

The authors declare no conflicts of interest in this work.

References

Accelrys, S.I. 2013. Discovery Studio Visualizer, Release 4.0.

Álvarez, Y., Pérez-Mañá, C., Torrens, M. & Farré, M. 2013. Antipsychotic drugs in cocaine dependence: a systematic review and meta-analysis. *Journal of substance abuse treatment*, 45(1): 1–10.

Andreoli, M., Tessari, M., Pilla, M., Valerio, E., Hagan, J.J. & Heidbreder, C.A. 2003. Selective antagonism at dopamine D 3 receptors prevents nicotine-triggered relapse to nicotine-seeking behavior. *Neuropsychopharmacology*, 28(7): 1272–1280.

Appel, N. & Acri, J. 2017. An assessment of the cardiovascular safety of the dopamine D3 receptor antagonist GSK598809 with cocaine in telemetered dogs and rats. *Journal of Pharmacological and Toxicological Methods*, 88: 220.

Appel, N.M. & Acri, J.B. 2018. Cardiovascular safety studies on dopamine D3 receptor antagonists and cocaine. *Journal of Pharmacological and Toxicological Methods*, 93: 156.

Appel, N.M., Li, S.-H., Holmes, T.H. & Acri, J.B. 2015. Dopamine D3 receptor antagonist (GSK598809) potentiates the hypertensive effects of cocaine in conscious, freely-moving dogs. *Journal of Pharmacology and Experimental Therapeutics*, 354(3): 484–492.

Ballesteros, J.A. & Weinstein, H. 1995. [19] Integrated methods for the construction of three-

dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In *Methods in neurosciences*. Elsevier: 366–428.

Ballon, J.S., Pajvani, U., Freyberg, Z., Leibel, R.L. & Lieberman, J.A. 2014. Molecular pathophysiology of metabolic effects of antipsychotic medications. *Trends in Endocrinology & Metabolism*, 25(11): 593–600.

Beaulieu, J.-M. & Gainetdinov, R.R. 2011. The physiology, signaling, and pharmacology of dopamine receptors. *Pharmacological reviews*, 63(1): 182–217.

Beuming, T., Farid, R. & Sherman, W. 2009. High-energy water sites determine peptide binding affinity and specificity of PDZ domains. *Protein science*, 18(8): 1609–1619.

Bezerra, G.A., Dobrovetsky, E., Viertlmayr, R., Dong, A., Binter, A., Abramić, M., Macheroux, P., Dhe-Paganon, S. & Gruber, K. 2012. Entropy-driven binding of opioid peptides induces a large domain motion in human dipeptidyl peptidase III. *Proceedings of the National Academy of Sciences*, 109(17): 6525–6530.

Chien, E.Y.T., Liu, W., Zhao, Q., Katritch, V., Han, G.W., Hanson, M.A., Shi, L., Newman, A.H., Javitch, J.A. & Cherezov, V. 2010. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science*, 330(6007): 1091–1095.

Congreve, M., Dias, J.M. & Marshall, F.H. 2014. Structure-based drug design for G protein-coupled receptors. In *Progress in medicinal chemistry*. Elsevier: 1–63.

Elber, R. 2016. Perspective: Computer simulations of long time dynamics. *The Journal of chemical physics*, 144(6): 60901.

Fan, L., Tan, L., Chen, Z., Qi, J., Nie, F., Luo, Z., Cheng, J. & Wang, S. 2020. Haloperidol bound D 2 dopamine receptor structure inspired the discovery of subtype selective ligands. *Nature communications*, 11(1): 1–11.

Feng, Z., Hou, T. & Li, Y. 2012. Selectivity and activation of dopamine D3R from molecular dynamics. *Journal of molecular modeling*, 18(12): 5051–5063.

Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., Repasky, M.P., Knoll, E.H., Shelley, M. & Perry, J.K. 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry*, 47(7): 1739–1749.

Friesner, R.A., Murphy, R.B., Repasky, M.P., Frye, L.L., Greenwood, J.R., Halgren, T.A., Sanschagrin, P.C. & Mainz, D.T. 2006. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein− ligand complexes. *Journal of medicinal chemistry*, 49(21): 6177–6196.

Galaj, E., Manuszak, M., Babic, S., Ananthan, S. & Ranaldi, R. 2015. The selective dopamine D3 receptor antagonist, SR 21502, reduces cue-induced reinstatement of heroin seeking and heroin conditioned place preference in rats. *Drug and alcohol dependence*, 156: 228–233.

Genheden, S. & Ryde, U. 2015. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert opinion on drug discovery*, 10(5): 449–461.

Girgis, R.R., Slifstein, M., D'Souza, D., Lee, Y., Periclou, A., Ghahramani, P., Laszlovszky, I., Durgam, S., Adham, N. & Nabulsi, N. 2016. Preferential binding to dopamine D 3 over D 2 receptors by cariprazine in patients with schizophrenia using PET with the D 3/D 2 receptor ligand [11 C]-(+)-PHNO. *Psychopharmacology*, 233(19–20): 3503–3512.

Le Grand, S., Götz, A.W. & Walker, R.C. 2013. SPFP: Speed without compromise—A mixed

precision model for GPU accelerated molecular dynamics simulations. *Computer Physics Communications*, 184(2): 374–380.

Grant, B.J., Rodrigues, A.P.C., ElSawy, K.M., McCammon, J.A. & Caves, L.S.D. 2006. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics*, 22(21): 2695–2696.

de Guglielmo, G., Kallupi, M., Sedighim, S., Newman, A.H. & George, O. 2020. Dopamine D3 receptor antagonism reverses the escalation of oxycodone self-administration and decreases withdrawal-induced hyperalgesia and irritability-like behavior in oxycodone-dependent heterogeneous stock rats. *Frontiers in Behavioral Neuroscience*, 13: 292.

Heidbreder, C.A. & Newman, A.H. 2010. Current perspectives on selective dopamine D3 receptor antagonists as pharmacotherapeutics for addictions and related disorders. *Annals of the New York Academy of Sciences*, 1187: 4.

Higgs, C., Beuming, T. & Sherman, W. 2010. Hydration site thermodynamics explain SARs for triazolylpurines analogues binding to the A2A receptor. *ACS medicinal chemistry letters*, 1(4): 160–164.

Higley, A.E., Kiefer, S.W., Li, X., Gaál, J., Xi, Z.-X. & Gardner, E.L. 2011. Dopamine D3 receptor antagonist SB-277011A inhibits methamphetamine self-administration and methamphetamine-induced reinstatement of drug-seeking in rats. *European journal of pharmacology*, 659(2–3): 187–192.

Holmes, A., Lachowicz, J.E. & Sibley, D.R. 2004. Phenotypic analysis of dopamine receptor knockout mice; recent insights into the functional specificity of dopamine receptor subtypes. *Neuropharmacology*, 47(8): 1117–1134.

Hosen, S.M.Z., Dash, R., Junaid, M., Mitra, S. & Absar, N. 2019. Identification and structural characterization of deleterious non-synonymous single nucleotide polymorphisms in the human SKP2 gene. *Computational biology and chemistry*, 79: 127–136.

Isberg, V., Mordalski, S., Munk, C., Rataj, K., Harpsøe, K., Hauser, A.S., Vroling, B., Bojarski, A.J., Vriend, G. & Gloriam, D.E. 2016. GPCRdb: an information system for G protein-coupled receptors. *Nucleic acids research*, 44(D1): D356–D364.

Jordan, C.J., Cao, J., Newman, A.H. & Xi, Z.-X. 2019. Progress in agonist therapy for substance use disorders: Lessons learned from methadone and buprenorphine. *Neuropharmacology*, 158: 107609.

Jordan, C.J., Humburg, B., Rice, M., Bi, G.-H., You, Z.-B., Shaik, A.B., Cao, J., Bonifazi, A., Gadiano, A. & Rais, R. 2019. The highly selective dopamine D3R antagonist, R-VK4-40 attenuates oxycodone reward and augments analgesia in rodents. *Neuropharmacology*, 158: 107597.

Jordan, C.J., Humburg, B.A., Thorndike, E.B., Shaik, A.B., Xi, Z.-X., Baumann, M.H., Newman, A.H. & Schindler, C.W. 2019. Newly Developed Dopamine D3 Receptor Antagonists, R-VK4-40 and R-VK4-116, Do Not Potentiate Cardiovascular Effects of Cocaine or Oxycodone in Rats. *Journal of Pharmacology and Experimental Therapeutics*, 371(3): 602–614.

Kaar, S.J., Natesan, S., McCutcheon, R. & Howes, O.D. 2020. Antipsychotics: mechanisms underlying clinical response and side-effects and novel treatment approaches based on pathophysiology. *Neuropharmacology*, 172: 107704.

Keck, T.M., John, W.S., Czoty, P.W., Nader, M.A. & Newman, A.H. 2015. Identifying

medication targets for psychostimulant addiction: unraveling the dopamine D3 receptor hypothesis. *Journal of medicinal chemistry*, 58(14): 5361–5380.

Kumar, V., Bonifazi, A., Ellenberger, M.P., Keck, T.M., Pommier, E., Rais, R., Slusher, B.S., Gardner, E., You, Z.-B. & Xi, Z.-X. 2016. Highly selective dopamine D3 receptor (D3R) antagonists and partial agonists based on eticlopride and the D3R crystal structure: new leads for opioid dependence treatment. *Journal of medicinal chemistry*, 59(16): 7634–7650.

Latorraca, N.R., Venkatakrishnan, A.J. & Dror, R.O. 2017. GPCR dynamics: structures in motion. *Chemical reviews*, 117(1): 139–155.

Li, D.-Z., Yu, G.-Q., Yi, S.-C., Zhang, Y., Kong, D.-X. & Wang, M.-Q. 2015. Structure-based analysis of the ligand-binding mechanism for DhelOBP21, a C-minus odorant binding protein, from Dastarcus helophoroides (Fairmaire; Coleoptera: Bothrideridae). *International journal of biological sciences*, 11(11): 1281.

Li, P., L Snyder, G. & E Vanover, K. 2016. Dopamine targeting drugs for the treatment of schizophrenia: past, present and future. *Current topics in medicinal chemistry*, 16(29): 3385–3403.

Lieberman, J.A., Stroup, T.S., McEvoy, J.P., Swartz, M.S., Rosenheck, R.A., Perkins, D.O., Keefe, R.S.E., Davis, S.M., Davis, C.E. & Lebowitz, B.D. 2005. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *New England journal of medicine*, 353(12): 1209–1223.

Liu, H., Hofmann, J., Fish, I., Schaake, B., Eitel, K., Bartuschat, A., Kaindl, J., Rampp, H., Banerjee, A. & Hübner, H. 2018. Structure-guided development of selective M3 muscarinic acetylcholine receptor antagonists. *Proceedings of the National Academy of Sciences*, 115(47): 12046–12050.

Lomize, M.A., Lomize, A.L., Pogozheva, I.D. & Mosberg, H.I. 2006. OPM: orientations of proteins in membranes database. *Bioinformatics (Oxford, England)*, 22(5): 623–625.

Maggio, R., Scarselli, M., Capannolo, M. & Millan, M.J. 2015. Novel dimensions of D3 receptor function: Focus on heterodimerisation, transactivation and allosteric modulation. *European Neuropsychopharmacology*, 25(9): 1470–1479.

Manvich, D.F., Petko, A.K., Branco, R.C., Foster, S.L., Porter-Stransky, K.A., Stout, K.A., Newman, A.H., Miller, G.W., Paladini, C.A. & Weinshenker, D. 2019. Selective D 2 and D 3 receptor antagonists oppositely modulate cocaine responses in mice via distinct postsynaptic mechanisms in nucleus accumbens. *Neuropsychopharmacology*, 44(8): 1445–1455.

Micheli, F. 2011. Recent advances in the development of dopamine D3 receptor antagonists: a medicinal chemistry perspective. *ChemMedChem*, 6(7): 1152–1162.

Miller III, B.R., McGee Jr, T.D., Swails, J.M., Homeyer, N., Gohlke, H. & Roitberg, A.E. 2012. MMPBSA. py: an efficient program for end-state free energy calculations. *Journal of chemical theory and computation*, 8(9): 3314–3321.

Miyamoto, S., Miyake, N., Jarskog, L.F., Fleischhacker, W.W. & Lieberman, J.A. 2012. Pharmacological treatment of schizophrenia: a critical review of the pharmacology and clinical effects of current and future therapeutic agents. *Molecular psychiatry*, 17(12): 1206–1227.

Moritz, A.E., Free, R.B. & Sibley, D.R. 2018. Advances and challenges in the search for D2

and D3 dopamine receptor-selective compounds. *Cellular signalling*, 41: 75–81.

Newman, A.H., Beuming, T., Banala, A.K., Donthamsetti, P., Pongetti, K., LaBounty, A., Levy, B., Cao, J., Michino, M. & Luedtke, R.R. 2012. Molecular determinants of selectivity and efficacy at the dopamine D3 receptor. *Journal of medicinal chemistry*, 55(15): 6689–6699.

Newman, A.H., Blaylock, B.L., Nader, M.A., Bergman, J., Sibley, D.R. & Skolnick, P. 2012. Medication discovery for addiction: translating the dopamine D3 receptor hypothesis. *Biochemical pharmacology*, 84(7): 882–890.

Novick, D., Haro, J.M., Suarez, D., Perez, V., Dittmann, R.W. & Haddad, P.M. 2010. Predictors and clinical consequences of non-adherence with antipsychotic medication in the outpatient treatment of schizophrenia. *Psychiatry research*, 176(2–3): 109–113.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. & Ferrin, T.E. 2004. UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13): 1605–1612.

Rangel-Barajas, C., Coronel, I. & Florán, B. 2015. Dopamine receptors and neurodegeneration. *Aging and disease*, 6(5): 349.

Roe, D.R. & Cheatham, T.E. 2013. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation*, 9(7): 3084–3095.

Ryckaert, J.-P., Ciccotti+, G. & Berendsen, H.J.C. 1977. *Numerical integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes*.

Salentin, S., Schreiber, S., Haupt, V.J., Adasme, M.F. & Schroeder, M. 2015. PLIP: fully automated protein–ligand interaction profiler. *Nucleic acids research*, 43(W1): W443–W447.

Salomon-Ferrer, R., Götz, A.W., Poole, D., Le Grand, S. & Walker, R.C. 2013. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *Journal of chemical theory and computation*, 9(9): 3878–3888.

Schrodinger, L.L.C. 2010. The PyMOL molecular graphics system. *Version*, 1(5): 0.

Schrödinger Release 2019-4. 2019. Maestro. https://www.schrodinger.com/Maestro.

Shaik, A.B., Kumar, V., Bonifazi, A., Guerrero, A.M., Cemaj, S.L., Gadiano, A., Lam, J., Xi, Z.-X., Rais, R. & Slusher, B.S. 2019. Investigation of novel primary and secondary pharmacophores and 3-substitution in the linking chain of a series of highly selective and bitopic dopamine D3 receptor antagonists and partial agonists. *Journal of medicinal chemistry*, 62(20): 9061–9077.

Shi, L. & Javitch, J.A. 2002. The binding site of aminergic G protein–coupled receptors: the transmembrane segments and second extracellular loop. *Annual review of pharmacology and toxicology*, 42(1): 437–467.

Sibley, D.R. & Monsma Jr, F.J. 1992. Molecular biology of dopamine receptors. *Trends in pharmacological sciences*, 13: 61–69.

Simossis, V.A. & Heringa, J. 2005. PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. *Nucleic acids research*, 33(suppl_2): W289–W294.

Simossis, V.A. & Heringa, J. 2003. The PRALINE online server: optimising progressive

multiple alignment on the web.

Sittel, F., Jain, A. & Stock, G. 2014. Principal component analysis of molecular dynamics: On the use of Cartesian vs. internal coordinates. *The Journal of chemical physics*, 141(1): 07B605_1.

Smith, R.D., Engdahl, A.L., Dunbar Jr, J.B. & Carlson, H.A. 2012. Biophysical limits of protein–ligand binding. *Journal of chemical information and modeling*, 52(8): 2098–2106.

Sokoloff, P., Diaz, J., Foll, B. Le, Guillin, O., Leriche, L., Bezard, E. & Gross, C. 2006. The dopamine D3 receptor: a therapeutic target for the treatment of neuropsychiatric disorders. *CNS & Neurological Disorders-Drug Targets (Formerly Current Drug Targets-CNS & Neurological Disorders)*, 5(1): 25–43.

Sokoloff, P. & Le Foll, B. 2017. The dopamine D3 receptor, a quarter century later. *European Journal of Neuroscience*, 45(1): 2–19.

Sun, H., Duan, L., Chen, F., Liu, H., Wang, Z., Pan, P., Zhu, F., Zhang, J.Z.H. & Hou, T. 2018. Assessing the performance of MM/PBSA and MM/GBSA methods. 7. Entropy effects on the performance of end-point binding free energy calculation approaches. *Physical Chemistry Chemical Physics*, 20(21): 14450–14460.

Sun, X., Gou, H., Li, F., Lu, G., Song, R., Yang, R., Wu, N., Su, R., Cong, B. & Li, J. 2016. Y-QA31, a novel dopamine D 3 receptor antagonist, exhibits antipsychotic-like properties in preclinical animal models of schizophrenia. *Acta Pharmacologica Sinica*, 37(3): 322–333.

Tanwar, H., Sneha, P., Kumar, D.T., Siva, R., Walter, C.E.J. & Doss, C.G.P. 2017. A computational approach to identify the biophysical and structural aspects of methylenetetrahydrofolate reductase (MTHFR) mutations (A222V, E429A, and R594Q) leading to Schizophrenia. In *Advances in protein chemistry and structural biology*. Elsevier: 105–125.

Wang, S., Che, T., Levit, A., Shoichet, B.K., Wacker, D. & Roth, B.L. 2018. Structure of the D2 dopamine receptor bound to the atypical antipsychotic drug risperidone. *Nature*, 555: 269. http://dx.doi.org/10.1038/nature25758.

Weis, W.I. & Kobilka, B.K. 2018. The molecular basis of G protein–coupled receptor activation. *Annual review of biochemistry*, 87: 897–919.

Wu, E.L., Cheng, X., Jo, S., Rui, H., Song, K.C., Dávila-Contreras, E.M., Qi, Y., Lee, J., Monje-Galvan, V. & Venable, R.M. 2014. CHARMM-GUI membrane builder toward realistic biological membrane simulations. *Journal of computational chemistry*, 35(27): 1997–2004.

Xiao, J., Free, R.B., Barnaeva, E., Conroy, J.L., Doyle, T., Miller, B., Bryant-Genevier, M., Taylor, M.K., Hu, X. & Dulcey, A.E. 2014. Discovery, optimization, and characterization of novel D2 dopamine receptor selective antagonists. *Journal of medicinal chemistry*, 57(8): 3450–3463.

Yildirim, A., Zhang, J., Manzetti, S. & van der Spoel, D. 2016. Binding of Pollutants to Biomolecules: A Simulation Study. *Chemical research in toxicology*, 29(10): 1679–1688.

CHAPTER 5

Supporting information

Exploring the structural basis and atomistic binding mechanistic of the selective antagonist blockade at D_3 dopamine receptor over D_2 dopamine receptor

Patrick Appiah-Kubi[1], Fisayo Andrew Olotu[1] and Mahmoud E. S. Soliman[1*]

[1]Molecular Bio-computation and Drug Design Laboratory

School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

Telephone: +27 (0) 31 260 8048

Fax: +27 (0) 31 260 78

Email: soliman@ukzn.ac.za

Website: http://soliman.ukzn.ac.za

Table S1 Description of the simulated systems parameters.

Simulated Systems							
Systems	**Bound ligand**	**POPC lipid molecules**	**K^+ ions**	**Cl^- ions**	**Water molecules**	**Number of atoms**	**Duration of Simulation**
D2DR/APO	-	180	41	51	15621	75477	600 ns x 2
D2DR/R-VK4-40	R-VK4-40	180	41	51	15580	75417	600 ns x 2
D2DR/Y-QA31	Y-QA31	180	41	51	15666	75671	600 ns x 2
D3DR/APO	-	180	43	50	16504	78031	600 ns x 2
D3DR/R-VK4-40	R-VK4-40	180	43	50	16529	78169	600 ns x 2
D3DR/Y-QA31	Y-QA31	180	43	50	16507	79099	600 ns x 2

Table S2 Average root mean square deviations (RMSD) of the duplicated simulation run.

Simulated Systems	**RMSD (Å)**	
	Run1	**Run2**
D2DR/APO	2.54 ± **0.78**	2.23 ± 0.23
D2DR/R-VK4-40	1.91 ± 0.31	1.62 ± 0.20
D2DR/Y-QA31	2.16 ± 0.31	2.20 ± 0.24
D3DR/APO	1.68 ± **0.23**	1.10 ± 0.24
D3DR/R-VK4-40	1.85 ± 0.21	1.76 ± 0.32
D3DR/Y-QA31	1.64 ± 0.27	1.67 ± 0.18

Table S3 The average molecular properties of the R-VK4-40 and Y-QA31 binding at D_2 and D_3 receptor binding pocket and entire system as a function of the 600 ns simulation time.

Complexes	**Entire System**		**Binding Pocket**	
	Rg ± SD (Å)	**SASA ± SD ($Å^2$)**	**Rg ± SD (Å)**	**SASA ± SD ($Å^2$)**
D2DR/APO	20.32 ± 0.09	16431.28 ± 342.10	12.77 ± 0.17	3699.43 ± 166.47
D3DR/APO	20.11 ± 0.07	15899.66 ± 286.67	12.42 ± 0.17	3439.40 ± 152.35
D2DR/R-VK4-40	20.21 ± 0.09	16349.10 ± 327.50	10.52 ± 0.10	2090.52 ± 143.14
D3DR/R-VK4-40	20.08 ± 0.08	15359.26 ± 361.13	5.98 ± 0.39	153.63 ± 35.46
D2DR/Y-QA31	20.29 ± 0.07	15707.97 ± 262.76	10.72 ± 0.14	2166.47 ± 157.98
D3DR/Y-QA31	20.07 ± 0.11	15325.34 ± 305.82	5.58 ± 0.37	94.15 ± 35.65

Table S4 The projections of the distribution of the simulated systems onto the subspace defined by the most significant principal components

Simulated systems	**Variance in Cα atomic fluctuations (%)**		
	PC1	**PC2**	**PC3-PC20**
D2DR/R-VK4-40	37.52	10.65	<8.84
D3DR/R-VK4-40	26.40	8.82	<7.74
D2DR/Y-QA31	32.82	9.85	<8.36
D3DR/Y-QA31	21.89	12.94	<6.21

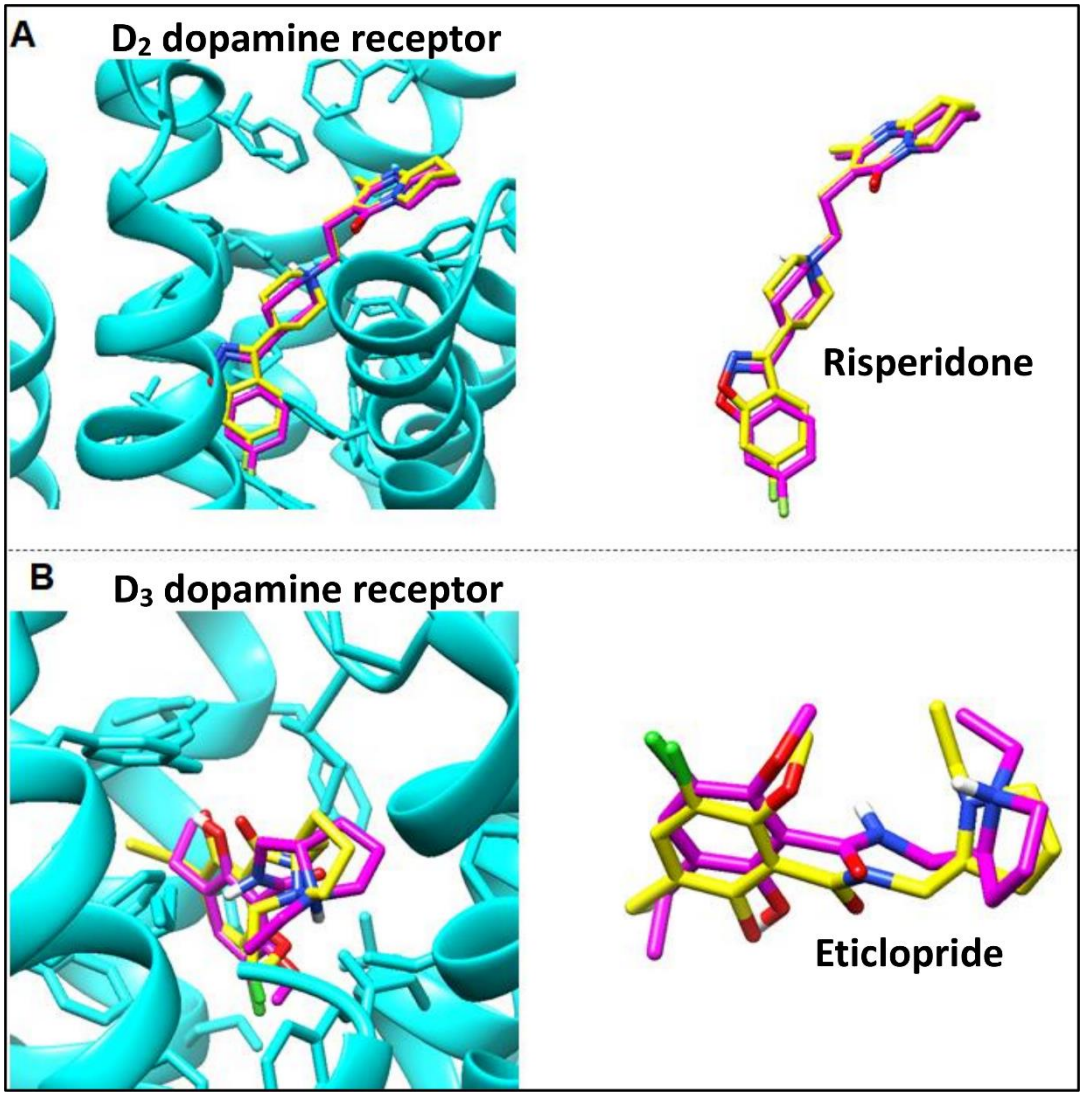


Figure S1 Redocking of Risperidone and Eticlopride at D_2 dopamine receptor (A) and D_3 dopamine receptor (B), respectively. Crystallographic and redocked ligand binding conformation in yellow and magenta colour, respectively (Image prepared by author).

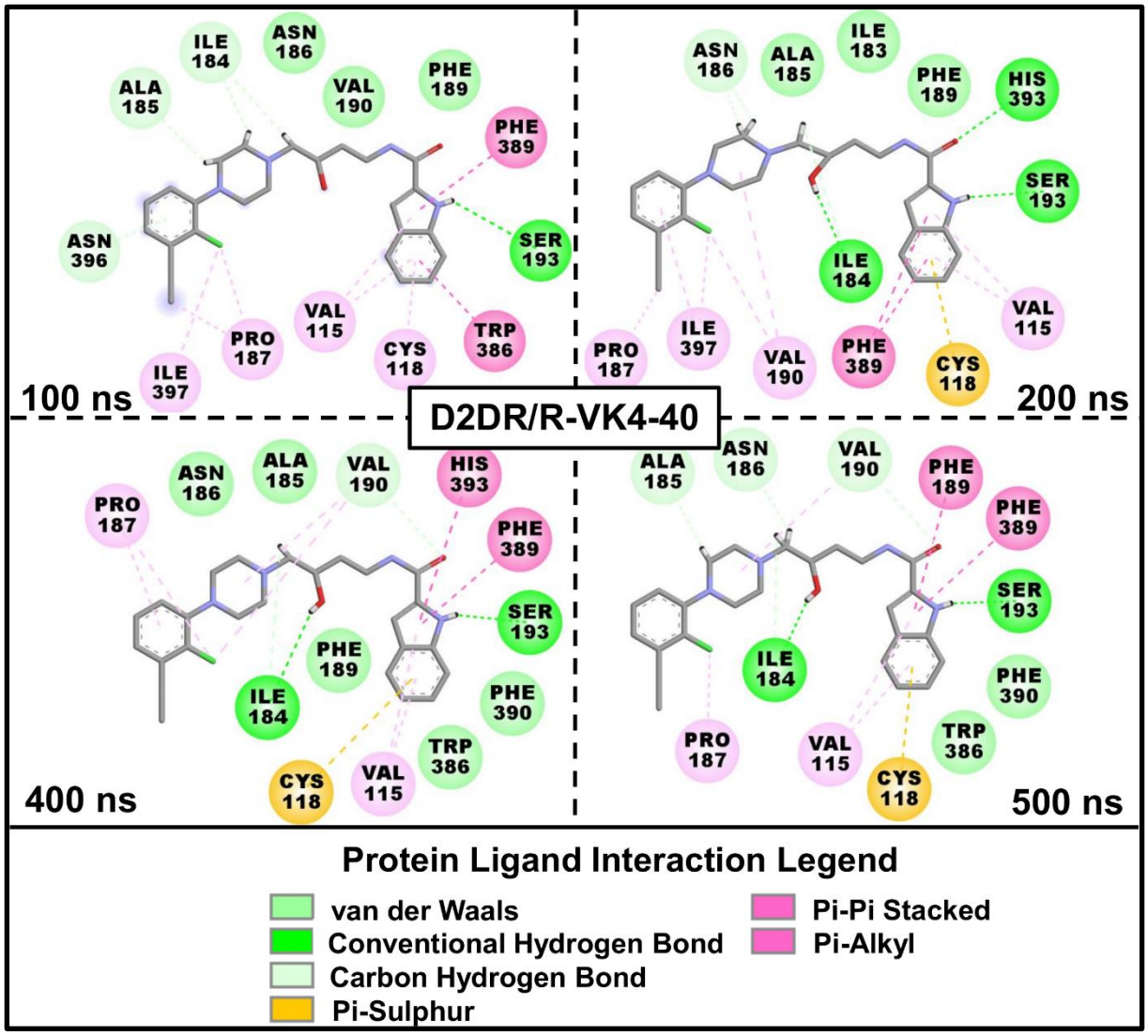


Figure S2 Representative binding mode and binding interaction of R-VK4-40 at D2DR extracted over the course of the simulation (Image prepared by author).

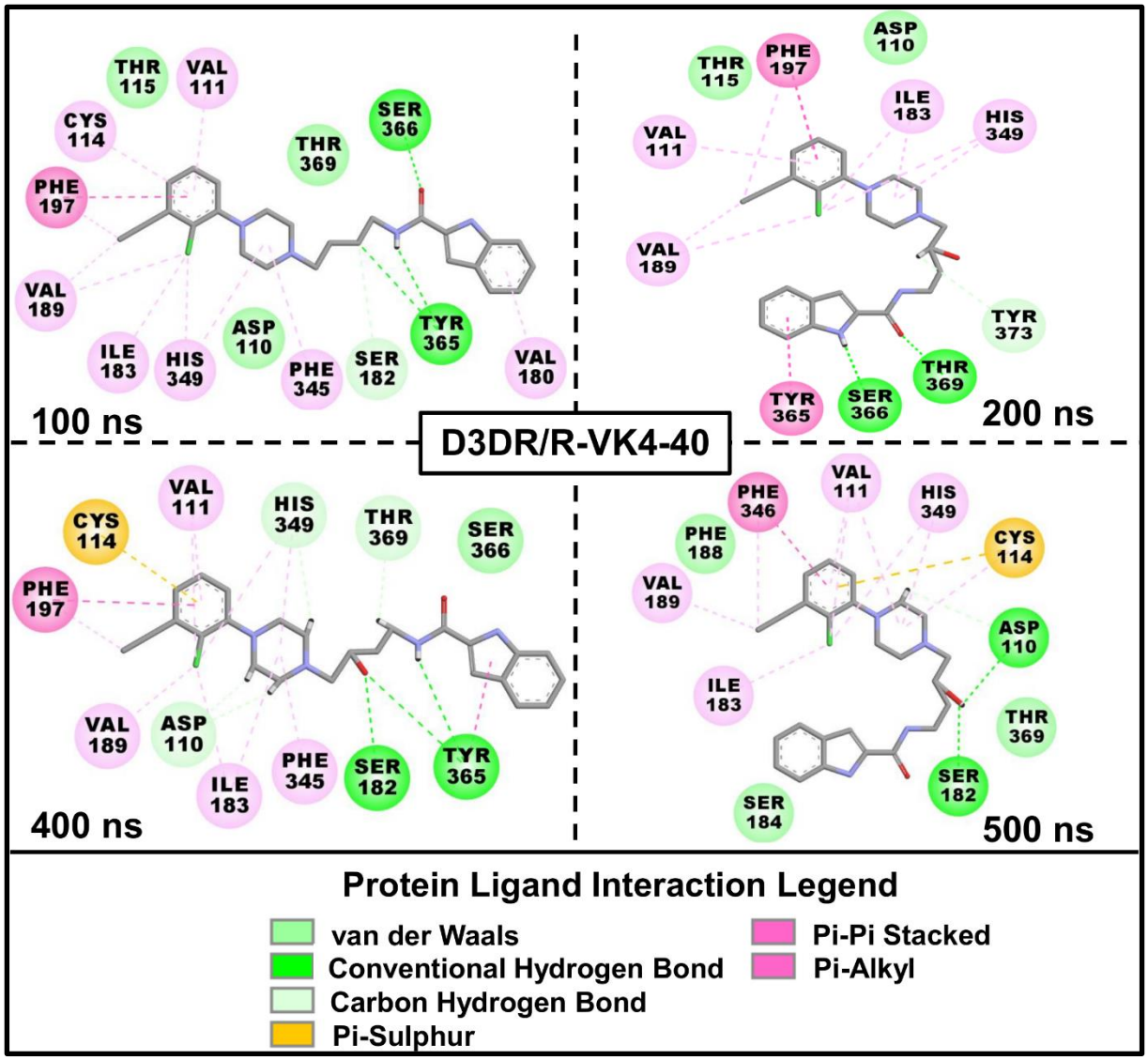


Figure S3 Representative binding mode and binding interaction of R-VK4-40 at D3DR extracted over the course of the simulation (Image prepared by author).

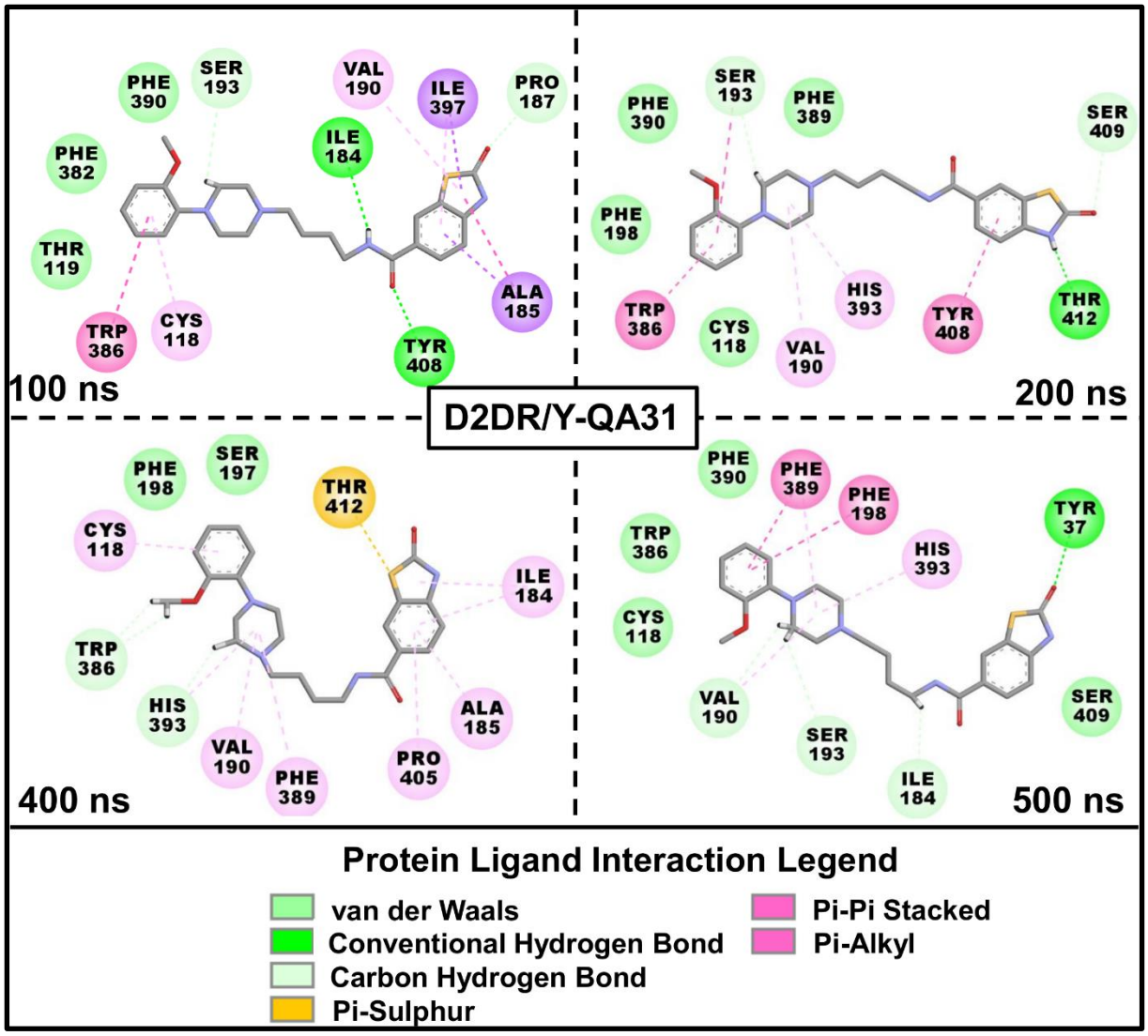


Figure S4 Representative binding mode and binding interaction of Y-QA31 at D2DR extracted over the course of the simulation (Image prepared by author).

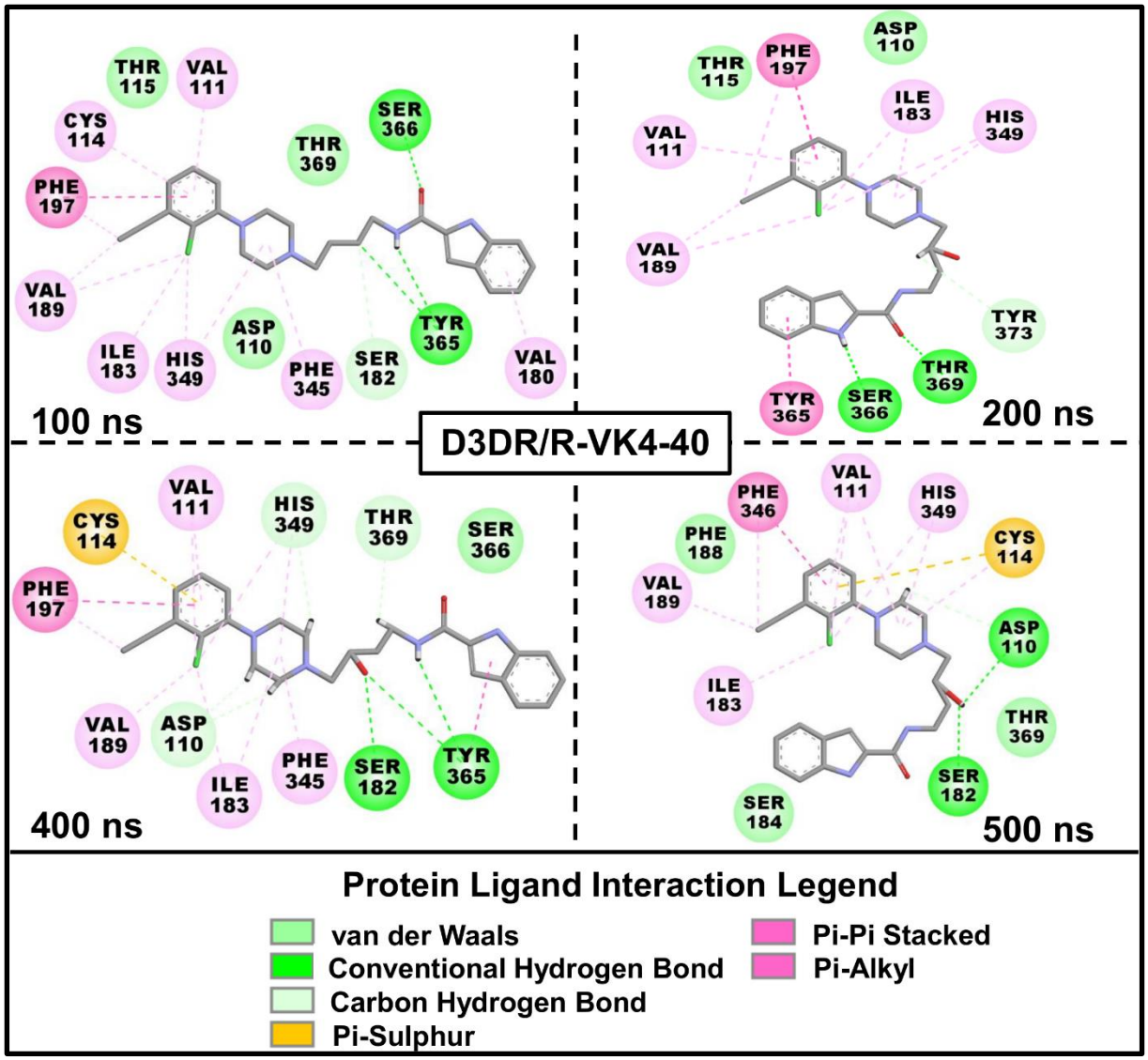


Figure S5 Representative binding mode and binding interaction of Y-QA31 at D3DR extracted over the course of the simulation (Image prepared by author).

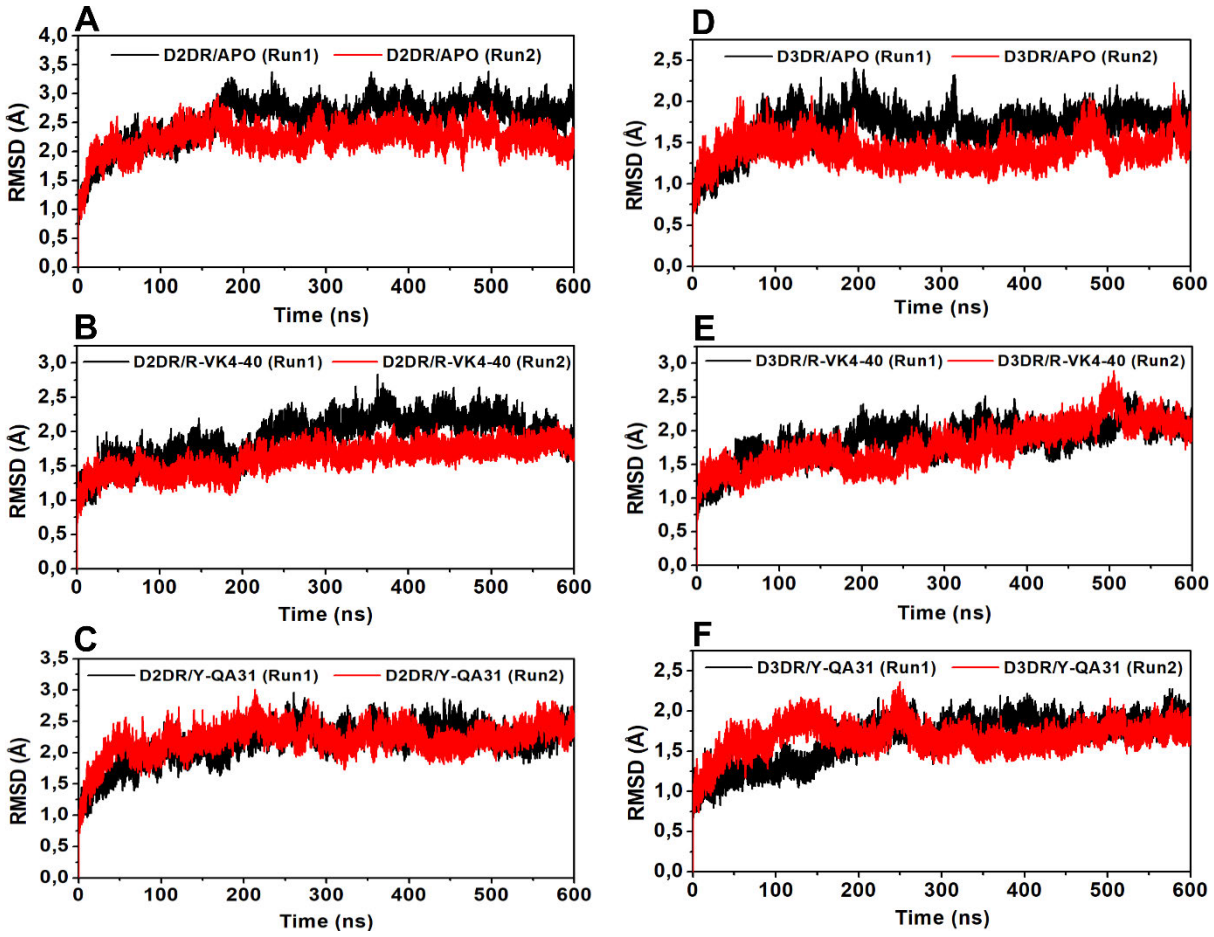


Figure S6 The Root-mean-square deviation (RMSD) of Cα atoms plots of duplicate 600 ns simulation for D2DR unbound (A), D2DR/R-VK4-40 bound (B), D2DR/Y-QA31 bound (C), D3DR unbound (D), D3DR/R-VK4-40 bound (E), and D3DR/Y-QA31(F) as a function of 600 ns simulation time (Image prepared by author).

CHAPTER 6

Published Article

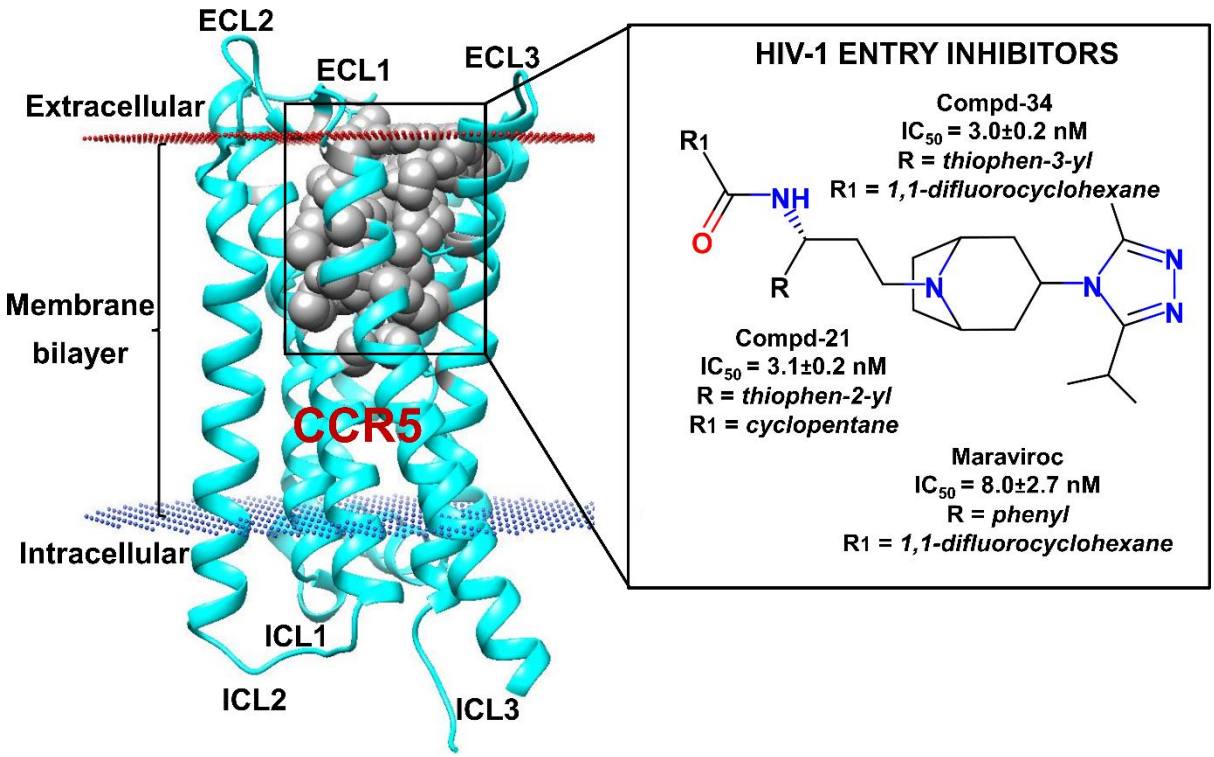
Elucidating the Disparate Inhibitory Mechanisms of Novel 1-Heteroaryl-1,3-Propanediamine Derivatives and Maraviroc towards C-C Chemokine Receptor 5: Insights for Structural Modifications in HIV-1 Drug Discovery

Patrick Appiah-Kubi[1], Fisayo Andrew Olotu[1], Mahmoud E. S. Soliman[1*]

[1]Molecular Bio-computation and Drug Design Laboratory
School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman
Telephone: +27 (0) 31 260 8048
Fax: +27 (0) 31 260 78
Email: soliman@ukzn.ac.za
Website: http://soliman.ukzn.ac.za
Patrick Appiah-Kubi appiahpat@gmail.com
Fisayo Andrew Olotu olotufisayo@gmail.com

Graphical Abstract



Abstract

Introduction: Blocking Human Immunodeficiency Virus type 1 (HIV-1) entry via C-C chemokine receptor 5 (CCR5) inhibition has remained an essential strategy in HIV drug discovery. This underlies the development of CCR5 blockers, such as Maraviroc, which, however, elicits undesirable side effects despite its potency.

Background: Recent lead optimization efforts led to the discovery of novel 1-heteroaryl-1,3-propanediamine derivatives; Compd-**21** and -**34**, which were ~3 times more potent than Maraviroc, with improved pharmacokinetics. However, atomistic molecular interaction mechanism of how slight structural variance between these inhibitors significantly affects their binding profiles have not been elucidated.

Method: This study employed explicit lipid bilayer molecular dynamics (MD) simulations, and advance analyses to explore these inhibitory discrepancies.

Results: Findings revealed that the thiophene moiety substitution common to Compd-**21** and -**34** enhanced their CCR5-inhibitory activities due to complementary high-affinity interactions with Trp86$^{2.60}$, Tyr108$^{3.32}$, Tyr251$^{6.51}$, Glu283$^{7.39}$. These cumulatively accounted for their ΔG_{bind}, which were higher than Maraviroc. Binding dynamics further revealed that the compounds mediated direct competitive inhibition at CCR5 by blocking the gp120 V3 loop. Furthermore, constituent tropane and triazole moieties in the compounds commonly engaged in interactions with Glu283$^{7.39}$ and Trp86$^{2.60}$, respectively. Structural analyses also revealed that both Compd-**21** and **-34** elicited distinct internal dynamic effect on CCR5 relative to Maraviroc.

Conclusion: Structural modifications at the thiophene substituent and the addition of new functional groups to the triazole ring may enhance inhibitor competition with gp120 V3-loop. Findings herein highlighted would contribute to future structure-based design of inhibitors of HIV-1 CCR5 with improved potencies.

Keywords: G protein-coupled receptor, HIV-1, Maraviroc, lipid bilayer, C-C chemokine receptor 5 (CCR5), Molecular dynamics simulations, 1-heteroaryl-1,3-propanediamine

1. Introduction

Human immunodeficiency virus type 1 (HIV-1) continues to threaten the quality of life, being the causal agent of acquired immunodeficiency syndrome (AIDS). Despite the progress made in the prevention of the HIV/AIDS epidemic evidenced by the development of various treatment strategies, the disease remains a global health threat. Different drugs and inhibitors have been developed over the past years toward the inhibition of HIV-1 via viral proteins targeting such as gp41, gp120, integrase, protease, and reverse transcriptase (Dyda et al., 1994; Huff, 1991; Kohlstaedt et al., 1992). However, there is the need for the development of novel HIV treatment drugs due to incidences of drug resistance easily triggered by mutations in these targets (Wensing et al., 2019).

G protein-coupled receptors (GPCRs) represent critical therapeutic intervention targets towards the development of new drugs due to their varied functions in many various cell responses (Rosenbaum et al., 2009). C-C chemokine receptor 5 (CCR5) is a member of the GPCRs family involved in immune function regulation (Flanagan, 2014; Sorce et al., 2011; Oppermann, 2004). In 1996, the chemokine receptors CXCR4 and CCR5 were found to be co-receptors for HIV-1 (Berger et al., 1999). HIV-1 envelope protein (Env) is made up of the gp160 trimeric $(gp160)_3$, which cleaves with three fragments of gp41 (fusion) and gp120 (receptor binding). The fusing of HIV-1 Env with viral and cell membrane enables host cells entry by the virus (Harrison, 2008). gp120 sequentially binds to its primary receptor CD4 and a co-receptor, thereby inducing conformational changes that result in gp120 dissociation and gp41 refolding (Harrison, 2008).

CCR5 represents the major HIV-1 co-receptor during entry into CD4+ T-cells (Berger et al., 1999). This significant role played by CCR5 in HIV-1 infection was defined when a naturally occurring mutation in the CCR5 gene (CCR5-Δ32) mediated resistance to HIV-1 disease (Allers et al., 2011). Compared with HIV-1 targets such as reverse transcriptase, protease, gp120, integrase, and gp41, the CCR5 receptor has been identified to have a low probability of mutations. Antiretroviral therapy (ART) administration slows the progressing of HIV to AIDS by decreasing viral loads in affected individuals (Detels et al., 1998; Peng et al., 2018). However, the current highly active ARTs (HAART) are associated with various setbacks, which include viral resistance, drug-drug interactions, and unwanted side effects (Günthard et al., 2014). Thus, the ability to prevent HIV-1 entry into host cells represents an attractive therapeutic approach in blocking HIV-1 infection and replication (M Gibson & J Arts, 2012). The quest to develop inhibitors capable of blocking HIV-1 entry by inhibiting the co-receptor

CCR5 lead to the identification of the first CCR5 HIV-1 entry drug Maraviroc, which was approved in 2007 by the Food and Drug Administration (FDA) (FDA, 2007). Maraviroc is the only marketed CCR5 drug but with limited prescription due to identified factors such as its drug-drug interactions (especially when co-administrated with CYP3A4 inhibitors), CYP450 inhibition, and viral resistance (Garcia-Perez et al., 2015; Peng et al., 2018). The rapid worldwide increase in patients diagnosed with HIV, therefore, necessitates the discovery of novel therapeutics for HIV treatment with fewer side effects and better efficacy.

Recently, a new series of 1-Heteroaryl-1,3-propanediamine derivatives [N-((S)-3-(exo-3-(3-Isopropyl-5-methyl-4H-1,2,4-triazol-4-yl)-8-azabicyclo[3.2.1]octan-8-yl)-1-(thiophen-2-yl)propyl)cyclopentane Carboxamide (Compd-21) and 4,4-Difluoro-N-(3-((1R,3S,5S)-3-(3-isopropyl-5-methyl-4H-1,2,4-triazol-4-yl)-8-azabicyclo[3.2.1]octan-8-yl)-1-(thiophen-3-yl)-propyl)cyclohexane-1-carboxamide (Compd-34)] (Peng et al., 2018) were synthesized as CCR5 antagonists (Fig. **1**). These inhibitors have displayed lower cytotoxicity, exceptional *in-vitro* anti-HIV-1 activity, and tolerable pharmacokinetic profile compared with Maraviroc (Peng et al., 2018). The crystal structure of Maraviroc in complex with CCR5 has previously been reported (Tan et al., 2013). Previous computational studies on CCR5 have been done to probe the interaction of Maraviroc with CCR5 through Molecular dynamics (MD) (Bai et al., 2014; Salmas et al., 2015), identify potential CCR5 inhibitors via pharmacophore-based screening and MD (Wang et al., 2016) and MD studies on CCR5 dimerization (Zhang et al., 2019). However, atomistic molecular details of the interaction mechanisms of how slight structural variance between these inhibitors (Compd-21, Compd-34, and Maraviroc) significantly affects their binding profiles at the CCR5 receptor has not been elucidated. Such an atomistic understanding would be beneficial in the identification of molecular properties and receptor interactions, which can be useful in the design of more effective HIV-1 compounds targeting CCR5 receptor.

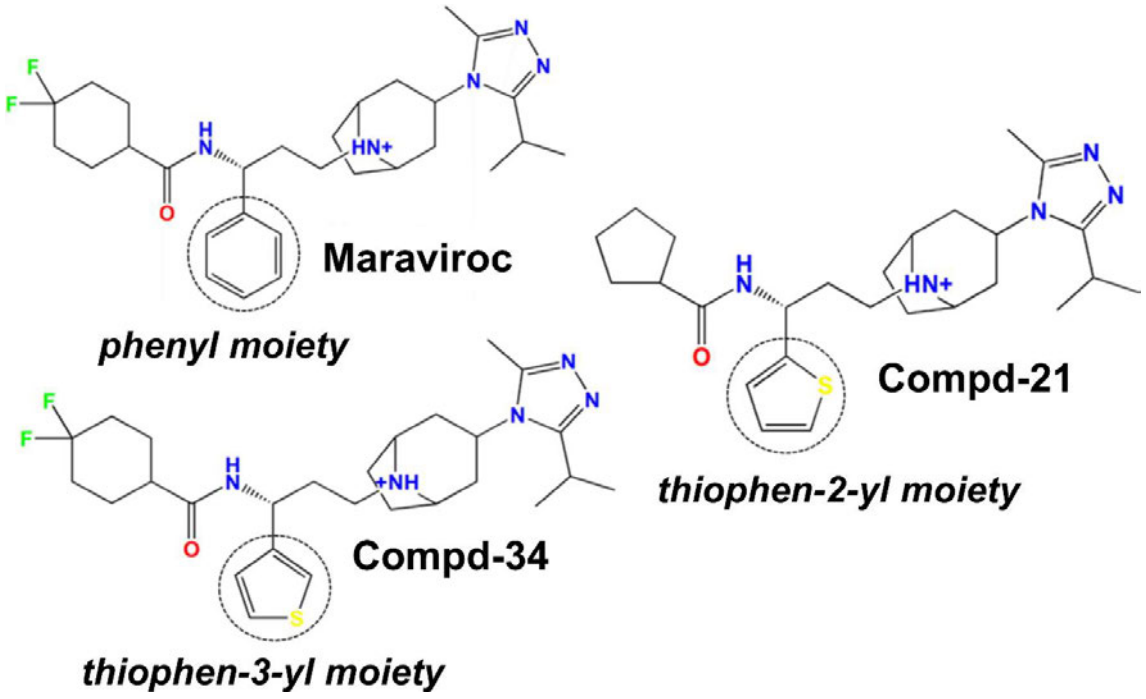


Fig. (1). 2D Chemical structures of the selected HIV-1 entry inhibitors used as co-crystallized inhibitors in the simulated complex systems. Key functional group variation highlighted in dashed circles (Image prepared by author).

In this present work, lipid bilayer molecular dynamics (MD) simulations in ionized explicit solvent, binding energy interactions, and conformational analyses have been employed to investigate the atomistic molecular basis for the higher inhibitory potency possessed by Compd-34 and Compd-21 relative to Maraviroc. Their respective molecular binding interactions provide mechanistic insights into their receptor recognition and shows how understanding ligand-receptor interaction and activation may eventually enable drug design at the CCR5 receptor for HIV treatment. This study provides novel insights into the structural determinants and interaction patterns that drive the differential binding profile of Maraviroc and the 1-heteroaryl-1,3-propanediamine derivatives (Compd-34 and Compd-21) which have vital implications for the design of improved HIV-1 agents.

2. Materials and Methods

2.1. Starting Structures

The initial coordinates for molecular dynamic simulations were obtained from the crystal of CCR5 bound to Maraviroc (PDB entry 4MBS), CCR5 bound to Compd-21 (PDB entry 6AKX) and CCR5 bound to Compd-34 (PDB entry 6AKY) (Berman et al., 2002). Maraviroc was

removed to obtain the apoprotein. The rubredoxin molecule was removed from the receptor and the intracellular loop 3 (ICL3) missing residues Cys224$^{5.68}$, Arg225, Asn226, and Glu227 were reconstructed and missing residues modelled using modeller (Webb & Sali, 2014) based on CCR5 sequence (UniProtKB ID: P51681). The thermostabilizing mutations (Cys58$^{1.60}$Tyr, Gly163$^{4.60}$Asn, Ala233$^{6.33}$Asp and Lys303$^{8.49}$Glu) in the crystal structures were reverted to their wild types.

2.2. Ligand and Protein Preparations

The Maestro LigPrep module (Schrödinger Release 2019-4, 2019) was used to generate ionization and tautomeric states of the ligands and minimized using OPLS3e force field. The Protein Preparation Wizard (Schrödinger Release 2019-4, 2019) was used to pre-process the protein using default parameters. Epik was used to generate ionization states at pH 7.0±4.0 and water molecules beyond 5 Å from the ligand deleted. The protein was subsequently refined by optimizing the hydrogen bonds and minimized using the OPLS3e force field.

2.3. Membrane-protein complex system setup

The simulated systems were assembled using the CHARMM-GUI membrane builder module (Jo et al., 2008) (http://www.charmm-gui.org/). The apo and inhibitor-complexed systems were embedded in a homogenous palmitoyl-oleoyl-phosphatidylcholine (POPC) lipid bilayer. The orientations of CCR5 Apo and CCR5 ligand complexes were aligned to the orientation of CCR5 in the membrane obtained from the Orientation of Protein Membranes (OPM) server (Lomize et al., 2006). The systems were then solvated using the TIP3P water model (Jorgensen et al., 1983), followed by a 0.15M concentration of KCl counterions for neutralization. (Fig. **2** and Table **1**). The simulated systems had an average dimension of 95 x 95 x 104 Å^3. The charmmlipid2amber.py script was used in processing structure files in renaming lipid residues according to the Amber lipid14 force field. The two essential disulphide bonds between Cys20-Cys269 and Cys101-Cys178 residues were maintained during topology and coordinate generation in *tleap* of Amber.

Table **1.** Description of the simulated CCR5-Membrane systems parameters.

Simulated Systems							
Systems	**Bound ligand**	**POPC lipid molecules**	**K^+ ions**	**Cl^- ions**	**Water molecules**	**Number of atoms**	**Duration of Simulation**
CCR5-34	Compd-34	190	47	64	17864	84039	400 ns
CCR5-21	Compd-21	190	47	64	17837	83955	400 ns
CCR5-MVR	Maraviroc	190	46	63	17498	82944	400 ns
CCR5-APO	-	190	46	60	17473	82773	400 ns

2.4. Molecular dynamics (MD) simulation protocols

To explore the stability and conformation dynamics of the systems under study, Molecular dynamics (MD) simulations were performed in an ionized explicit POPC lipid bilayer on Amber18 (Case et al., 2018). The GAFF force field (Sprenger et al., 2015) was used to generate force fields for the inhibitors. The protein and electrostatic potential distribution were obtained with the Amber force field FF14SB (Maier et al., 2015) and the Restrained Electrostatic Potential fit approach (RESP) (Bayly et al., 1993), respectively. The Lipid14 force field (Dickson et al., 2014) was used to describe the lipids. The coordinate and topology files for the simulations were generated with *tleap* software using the ff14SB, TIP3P, and lipid14 forcefields with a box dimensions of 95 x 95 x 104 $Å^3$.

The systems were initially relaxed before the MD simulations by performing 10000 minimization steps. The systems were then heated in isothermal-isochoric (NVT) ensemble from 0 K to 100 K using the Langevin thermostat (Larini et al., 2007) for 13 ps with harmonic restraint of 10 kcal mol^{-1} $Å^{-2}$ applied on non-hydrogen atoms of lipid, ligand, and protein, with a 1.0 ps^{-1} collision frequency. A 130 ps heating in isothermal-isobaric (NTP) ensemble was performed for the systems starting at 100 K to 310 K with a pressure of 1 bar. Equilibration of the system was performed with a starting harmonic restraint of 5.0 kcal/mol/ $Å^2$ on the protein and lowered by 1.0 kcal mol^{-1} $Å^{-2}$ in a stepwise manner every 4 ns for a total of 20 ns at 310 K under NTP ensemble. A subsequent 10 ns unrestrained equilibration was performed before production. Finally, 400 ns unrestrained MD simulations were run under NPT ensemble at 310 K with SHAKE constraints for bonds with hydrogen. To prevent errors during the simulations, the *skinnb* parameter value was set at 5. A 2-fs time step was used in integrating the equation of motion during the production phase. The Langevin dynamic was used to regulate the temperature with a collision frequency of 5 ps−1, whereas the anisotropic pressure coupling with a pressure relaxation time of 1.0 ps used to control the pressure of the systems. The particle

mesh Ewald was used to treating long-range electrostatic interactions (Darden et al., 1993) under periodic boundary conditions with nonbonded interactions cut-off of 12 Å.

2.4.1. Trajectory analysis

To comprehend the structural as well as functional implications of the studied systems, various structural property analyses were performed as a function of time on the obtained trajectories. The Amber CPPTRAJ module (Roe & Cheatham, 2013) was used to analyse the generated 400 ns trajectories. To assess the stability and conformational dynamics during the simulations, the root mean square fluctuation (RMSF), root mean square deviation (RMSD), the solvent-accessible surface area (SASA), the radius of gyration (Rg) and secondary structure analysis (DSSP) were computed. Similarly, hydrogen bonds formed between the inhibitors and specific residues were calculated.

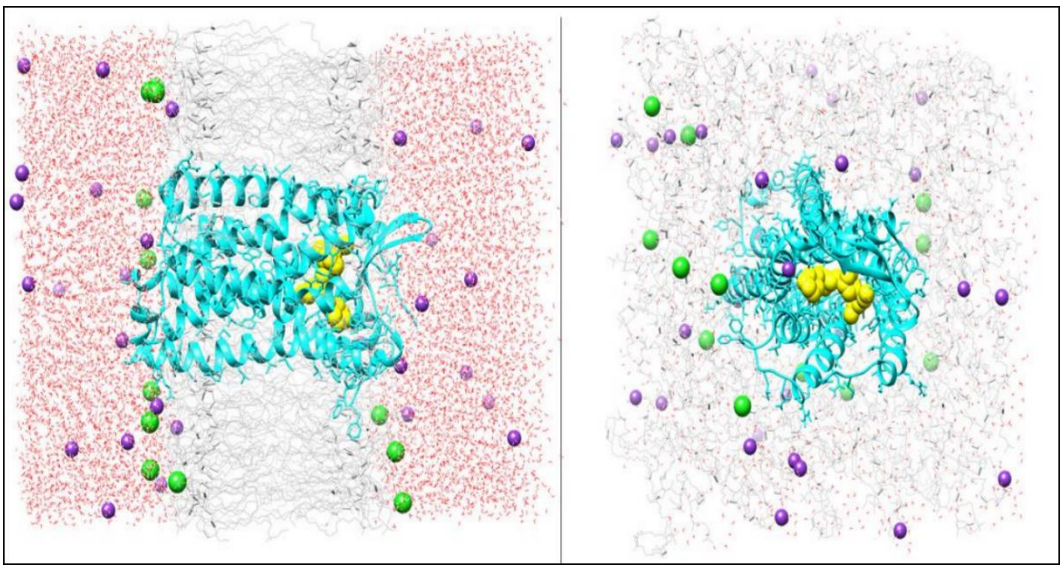


Fig. (2). CCR5-ligand lipid bilayer system setup showing side view (left) and top view (right) of CCR5-inhibitor complex (*CCR5 represented in a cyan cartoon and inhibitor in yellow sphere*) embedded in an ionized solvated POPC lipid bilayer (*POPC lipids, water, K+, and Cl- are depicted in grey wire, red, purple and green, respectively*) (Image prepared by author).

2.5. End-point interaction energy calculations

The Molecular Mechanics-Generalized-Born Surface Area (MM-GBSA) approach is a widely used method in analysing relative binding energies protein-ligand complexes (Ylilauri & Pentikäinen, 2013; Chen et al., 2016; Sun et al., 2018). Various binding energy components such as the molecular mechanics potential energy (van der Waals and electrostatic energies), polar and nonpolar interaction energies were obtained for each CCR5-inhibitor complex. A total of 500 representative snapshots were extracted over the last 300 ns conformational

ensemble. The binding free energy (ΔG_{bind}) is computed by the MM-GBSA approach via the set of equations below.

$$\Delta G_{bind} = G_{complex} - (G_{receptor} + G_{ligand}) \quad (1)$$

$$\Delta G_{bind} = E_{gas} + G_{sol} - T\Delta S \quad (2)$$

$$E_{gas} = E_{int} + E_{vdw} + E_{ele} \quad (3)$$

$$G_{sol} = G_{pol} + G_{SA} \quad (4)$$

$$G_{SA} = SASA + b \quad (5)$$

Where G_{ligand}, $G_{receptor}$, and $G_{complex}$ denote the relative free energies of the ligand, unbound protein, and protein-complex, respectively. The ΔG_{bind} was decomposed into (equations 2 to 5): the solvation energy term (G_{sol}) which is the sum of the polar (G_{pol}) and nonpolar (G_{SA}) solvation terms; the gas-phase (E_{gas}) energy contribution which is a summation of the nonbonded [van der Waals (E_{vdw}) and electrostatics (E_{ele})] and the bonded [internal energy (E_{int})] energy terms, and the entropy term ($-T\Delta S$). The conformational entropy contribution to the total binding free energy was calculated by normal mode analysis using 10 snapshots evenly extracted from the 500 snapshots.

2.5.1 Decomposition of the overall interaction energy

The total binding free energies were reduced into each residue contribution to identifying active site "hot spot" residues involved in the preferential binding of the novel inhibitors, using the MM-GBSA per-residue energy decomposition utility of Amber.

2.6. Clustering and Principal Component Analysis (PCA)

The principal component analysis (PCA) is routinely applied to MD trajectory analysis in reducing large-dimensional observations sets onto collective data. Obtaining principal components (PCs) for MD trajectories involves two key steps:

(i) Covariance matrix (C) generation,

$$C_{ij} = < (X_i - < X_i >) (X_j - < X_j >) > \quad (1)$$

(ii) The 3N x 3N covariance matrix C diagonalization that can be computed by eigenvalue decomposition (EVD) as,

$$C = V\Lambda V^T \quad (2)$$

Where V denotes a matrix, which includes eigenvectors, and Λ, describes the eigenvalues contained in the diagonal matrix. The eigenvalues display the mean squared displacements (MSD) of the Cα atoms throughout the used eigenvector. The principal component analysis in this study was applied utilizing the Bio3D package in R (Grant et al., 2006).

2.7. Graphical tools used in the study

Molecular visualizations were done with Chimera (Pettersen et al., 2004), Schrodinger Maestro(Schrödinger Release 2019-4, 2019), Origin software for plotting all graphs (Seifert, 2014), the protein-ligand interaction profiler (Salentin et al., 2015) and Bio3D package in R (Grant et al., 2006) for PCA analysis.

2.8. Transmembrane (TM) and Extracellular loops (ECL) residue numbering

The superscripts assigned to each residue denote GPCRs Ballesteros-Weinstein numbering (Ballesteros & Weinstein, 1995) for the transmembrane domains and the GPCRdb (Isberg et al., 2016) numbering for intracellular and extracellular loops.

3. Results

3.1. Phenyl → thiophen-2-yl (Compd-21) and thiophen-3-yl (Compd-34) substitutions improved binding affinities over Maraviroc.

The key structural variation between Maraviroc and the novel series of 1-heteroaryl-1,3-propanediamine derivatives (Compd-34, and Compd-21) is the substitution of the phenyl group in Maraviroc with a thiophen-2-yl and thiophen-3-yl moieties in Compd-21, and Compd-34, respectively. This substitution was observed to exhibit excellent *in vitro* anti-HIV-1 activity and improved pharmacokinetics (Peng et al., 2018). The interaction energies between the CCR5 and the selected inhibitors were evaluated to assess the mechanistic binding of Compd-34, Compd-21 and Maraviroc at the active site of CCR5 using MM-GBSA method (Hou et al., 2011; Sun et al., 2014; Chen et al., 2016). The estimated relative binding free energies and individual energy terms of Compd-34, Compd-21, and Maraviroc are listed in Table **2**. The computed binding energies (ΔG) (mean ± SEM) obtained for Maraviroc, Compd-21 and Compd-34 were -31.45 ± 2.2 kcal.mol^{-1}, -41.15 ± 1.0 kcal.mol^{-1}, and -44.94 ± 0.9 kcal.mol^{-1}, respectively. These energy values correlate in their ranking order with the experimentally reported IC_{50} values of 8.0 ± 2.7 nM for Maraviroc, 3.1 ± 0.2 nM for Compd-21, and 3.0 ± 0.2 nM for Compd-34. Compd-21 and Compd-34 relatively displayed more favourable binding than Maraviroc, which could be attributed to the stronger interactions elicited by Compd-21 and Compd-34 at CCR5 binding sites, as would be explained in subsequent sections. Aside from the ability for MM-GBSA to rank inhibitor in the order of their binding free energies, it also provides detail understanding of the inhibitor-target binding process (Yang et al., 2011) by decomposing the total binding free energies into their components such as van der Waals,

electrostatic, polar and nonpolar interactions. In all three inhibitor-complexes, a favourable van der Waals, polar, and nonpolar solvation energy terms were observed for inhibitor binding to CCR5; however, the overall electrostatic interactions disfavoured complex formation in all the complex systems. Furthermore, entropy contribution to the total binding energy was observed to be more unfavourable in Maraviroc compared to Compd-34 and Compd-21. It is evident from Table **2** that van der Waals energy (ΔE_{vdW}) most favoured the binding of Compd-21 and Compd-34 to CCR5 receptor compared with Maraviroc. Thus, van der Waals interaction is most important in stabilizing the CCR5-inhibitor complexes.

Table **2.** MM-GBSA binding free energy analysis [(mean ± SEM) kcal·mol^{-1}] of CCR5-inhibitor complexes.

Energy Composition	**CCR5 Antagonists**		
	Compd-34	**Compd-21**	**Maraviroc**
ΔE_{vdW}	-69.03 ± 0.2	-65.61 ± 0.2	-63.78 ± 0.2
ΔE_{elec}	162.44 ± 1.2	176.99 ± 1.2	181.59 ± 1.2
ΔE_{polar}	-156.48 ± 1.2	-174.57 ± 1.1	-172.47 ± 1.0
$\Delta E_{nonpolar}$	-8.10 ± 0.0	-7.47 ± 0.0	-7.79 ± 0.0
ΔE_{MM}	93.41 ± 1.2	111.39 ± 1.2	117.80 ± 1.2
ΔE_{sol}	-164.58 ± 1.2	-182.05 ± 1.1	-180.27 ± 1.1
ΔH	-71.17 ± 0.2	-70.66 ± 0.2	-62.47 ± 0.3
$-T\Delta S$	26.23 ± 1.1	29.51 ± 1.2	31.02 ± 2.5
ΔG_{bind}	**-44.94 ± 0.9**	**-41.15 ± 1.0**	**-31.45 ± 2.2**
Exp$_{(IC50)}$	**3.0 ± 0.2**	**3.1 ± 0.2**	**8.0 ± 2.7**

ΔE_{vdW} = van der Waals energy; ΔE_{ele} = Electrostatic energy; ΔE_{polar} = Polar solvation energy; $\Delta E_{nonpolar}$ = Nonpolar solvation energy; E_{MM} = Vacuum potential energy; ΔE_{sol} = Solvation energy; $\Delta G_{binding}$ = Total binding free energy; Exp = experimental inhibition (IC_{50} in nM).

3.2. High-affinity interactions at the gp120 V3 loop recognition site account for disparate binding among Compds-21, -34 and Maraviroc

To unravel the hotspot residues that may contribute to the higher potency of Compd-21 and Compd-34 over Maraviroc, the estimated ΔG_{bind} energy values were decomposed into individual interacting amino acids contributions using the MM-GBSA per-residue energy decomposition approach. The pairwise interaction energy contribution of CCR5 active site residues to the binding of the studied inhibitors is presented in Fig. **3a**. According to the residue interaction energy data, the major residues that contributed significantly to the inhibitory activity of the studied inhibitors with average energy values ≥ -1.0 kcal.mol^{-1} were: Trp86$^{2.60}$ (-2.86), Tyr108$^{3.32}$ (-3.22), Phe109$^{3.33}$ (-2.74), Ile198$^{5.42}$ (-1.31), Tyr251$^{6.51}$ (-1.0), Gln280$^{7.36}$ (-1.02), Thr284$^{7.40}$ (-1.59) and Met287$^{7.43}$ (-1.0) for Compd-34; Trp86$^{2.60}$ (-3.69), Tyr108$^{3.32}$ (-

3.51), Phe109$^{3.33}$ (-1.88), Phe112$^{3.36}$ (-1.25), Ile198$^{5.42}$ (-1.56), Tyr251$^{6.51}$ (-2.37), Leu255$^{6.55}$ (-1.04), and Met287$^{7.43}$ (1.00) for Compd-21; whereas Trp86$^{2.60}$ (-1.58), Tyr108$^{3.32}$ (-2.207), Phe109$^{3.33}$ (-2.08), Phe112$^{3.36}$ (-1.0), Ile198$^{5.42}$ (-1.48), and Gln280$^{7.36}$ (-1.0) for Maraviroc. Most of the binding site residues of CCR5 critical for the binding of gp120 V3-loop such as Trp86$^{2.60}$, Tyr108$^{3.32}$, Phe109$^{3.33}$, and Tyr251$^{6.51}$ made stronger interactions with the 1-heteroaryl-1,3-propanediamine derivatives (Compd-21 and Compd-34) than with Maraviroc.

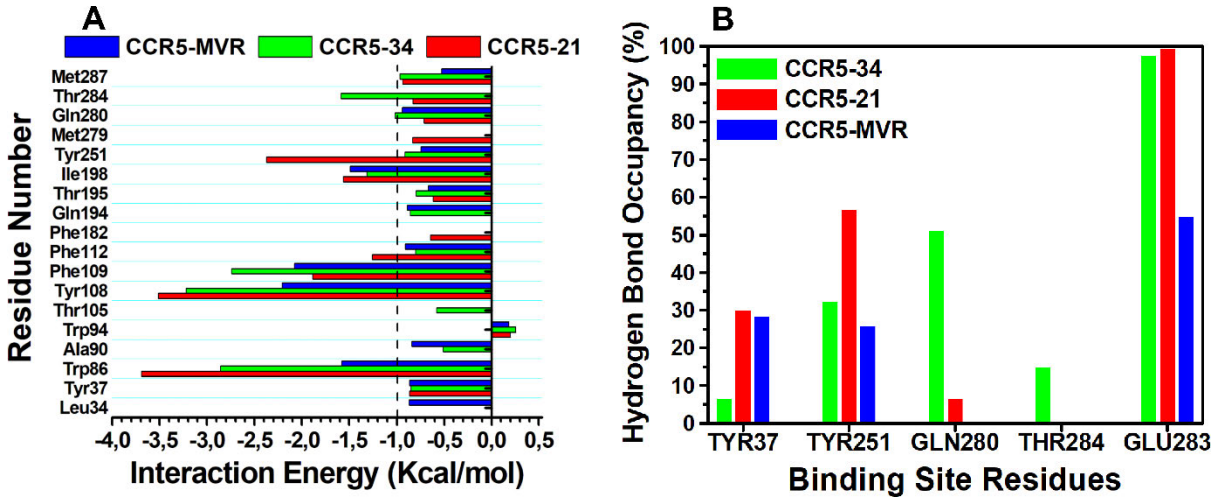


Fig. (3). Per-residue interaction energy decomposition (**a**) and hydrogen bond occupancy plot (**b**) for Compd-21, Compd-34 and Maraviroc at CCR5 binding site. The dashed line denotes residues with contribution from -1.0 kcal.mol^{-1} (Image prepared by author).

The MD trajectories were further analysed to gain an understanding of the dynamic variation in the number and strength of hydrogen bond formation throughout the simulation (Table **3** and Fig. **3b**). The inhibitors were stabilized in the active site of CCR5 receptor via hydrogen bonds involving: Glu283$^{7.39}$, Gln280$^{7.36}$, Tyr251$^{6.51}$, Thr284$^{7.40}$, and Tyr37$^{1.39}$ (with Compd-34); Glu283$^{7.39}$, Tyr251$^{6.51}$, Tyr37$^{1.39}$ and Gln280 (with Compd-21); and Glu283$^{7.39}$, Tyr37$^{1.39}$ and Tyr251$^{6.51}$, (with Maraviroc). In the CCR5-Maraviroc system, the protonated nitrogen of the tropane linker moiety maintained the most stable hydrogen bond/salt-bridge interaction with Glu283$^{7.39}$ (54.77 % occupancy). The triazole moiety also maintained a hydrogen bond with Tyr37$^{1.39}$ (26.20 % occupancy) whereas the carboxamide nitrogen formed a stable hydrogen bond with Tyr251$^{6\,51}$ (25.87 % occupancy). The dynamic hydrogen bond analysis indicates that the observed two hydrogen bond formed by Thr195$^{5.39}$ and Thr259$^{6.59}$ with one of the cyclohexane ring fluorines in the crystal structure(Tan et al., 2013) are weak interactions which disappeared at the initial stage of the MD simulation.

Table **3.** Hydrogen bond interactions between CCR5 with Compd-34, Compd-21, and Maraviroc.

Ligands	H-Acceptor	H-Donor	Donor	Occupancy (%)	Distance (Å)	Angle (°)
Compd-34	GLU283-OE1	Compd-34-HN2	Compd-34-N2	97.46	2.7	164
	GLN280-OE1	Compd-34-HN4	Compd-34-N4	50.92	2.8	158
	TYR251-OH	Compd-34-HN52	Compd-34-N5	32.42	2.9	147
	THR284OG1	Compd-34-HN3	Compd-34-N3	14.97	2.9	148
	Compd-34-N4	TYR37-HH	TYR37-OH	6.63	2.9	148
Compd-21	GLU283-OE1	Compd-21-HN1	Compd-21-N4	99.49	2.7	164
	TYR251-OH	Compd-21-H21	Compd-21-N5	56.74	2.8	160
	TYR37-OH	Compd-21-H44	Compd-21-N2	30.01	2.8	151
	Compd-21-N2	GLN280-HE21	GLN280-NE2	6.46	2.9	157
Maraviroc	GLU283-OE2	Maraviroc-HN2	Maraviroc-N2	54.77	2.7	160
	TYR37-OH	Maraviroc-HN4	Maraviroc-N4	28.34	2.9	161
	TYR251-OH	Maraviroc-HN52	Maraviroc-N5	25.87	2.8	152

The analysis of hydrogen bond occupancy in CCR5-Compd-34 complex also showed a populated and more substantial (97.46 % occupancy) salt-bridge interaction between Glu283$^{7.39}$ and the protonated nitrogen of the tropane moiety at the binding pocket. It was also observed that Gln280$^{7.36}$ and Tyr251$^{6.51}$ formed stable H-bonds with occupancy of 50.9% and 32.4%, respectively, whereas Tyr37$^{1.39}$ engaged in a weaker hydrogen bond with the triazole nitrogen at a 6.5% occupancy. The possible formation of an additional two hydrogen bonds by one of the fluorine in the cyclohexane moiety of Compd-34 with Thr259$^{6.59}$ and Thr195$^{5.39}$ as reported in the static crystallographic structure(Peng et al., 2018) were lost during the simulation. Finally, in the CCR5-Compd-21 complex system, the nitrogen in the tropane group engaged in a powerful salt-bridge interaction with Glu283$^{7.39}$ (99.5%). A stable hydrogen bond with an occupancy of 56.7% is formed between the nitrogen of the carboxamide group and Tyr251$^{6.51}$. Furthermore, the triazole ring nitrogen also interacted with Tyr37$^{1.39}$ and Gln280$^{7\,36}$ via hydrogen bonds with 30.0% and 6.5% occupancy, respectively.

In the dynamic binding mode of CCR5-34 complex structure (Fig. **4a**), the thiophen-3-yl moiety was buried into the active site making extensive contacts with Tyr108$^{3.32}$, Phe109$^{3.33}$, Phe112$^{3.36}$, Ile198$^{5.42}$, and Tyr251$^{6.51}$. The triazole moiety of Compd-34 also made hydrophobic contacts with Tyr37$^{1.39}$, Trp86$^{2.60}$, Tyr108$^{3.32}$, Gln280$^{7.36}$, and Met287$^{7.43}$. Similarly, the dynamic binding mode of Compd-21 with CCR5 indicates that thiophen-2-yl moiety made extensive contacts with Phe109$^{3.33}$, Phe112$^{3.36}$, Ile198$^{5.42}$, Trp248$^{6.48}$, and Tyr251$^{6.51}$ in the deep binding pocket. The triazole ring of Compd-21 was also involved in interactions with Tyr37$^{1.39}$, Trp86$^{2.60}$, Tyr89$^{2.63}$, Tyr108$^{3.32}$, Gln280$^{7.36}$, and Met287$^{7.43}$ (Fig. **4b**). In Fig. **4c**, Tyr108$^{3.32}$,

Phe109$^{3.33}$, Phe112$^{3.36}$, and Ile198$^{5.42}$ were also observed to engage in interactions with the phenyl group of Maraviroc whereas the triazole moiety made more contacts with Val25$^{1.27}$, Leu33$^{1.35}$, Tyr37$^{1.39}$ and Trp86$^{2.60}$.

3.3. Distinctive Structural Dynamics of CCR5 upon Inhibitors Binding Revealed by MD Simulations

Explicit MD simulations of CCR5 embedded in a hydrated ionized POPC lipid bilayer were performed to elucidate the inhibitory mechanism and the conformational dynamics associated with CCR5 by the novel 1-heteroaryl-1,3-propanediamine derivatives HIV-1 entry inhibitors. Presented in Table 1 are the different simulated systems in this study with a detailed description of system preparations outlined in the methodology section.

3.3.1. Structural Stability Evaluation

The Root Mean Squared Deviation (RMSD) of Cα atoms with respect to the minimized starting crystal structure was calculated to evaluate the structural stability of each of the studied systems throughout the simulation (Fig. **5a**). The MD simulations ultimately produced relatively stable trajectories with respect to backbone structural changes. The RMSD curves show that all the unbound and bound-complexed systems under study were evolving until 100 ns.

The overall average RMSD for the entire protein structures were 2.2 ± 0.3 Å, 2.1 ± 0.3 Å, 2.0 ± 0.3 Å, and 1.7 ± 0.2 Å for CCR5-21, CCR5-34, CCR5-MVR, and CCR5-APO, respectively. From the RMSD results, the mean deviation of all systems was lower than 2.3 Å, and a maximum RMSD value smaller than 3.0 Å. The dynamic stability of Maraviroc, Compd-21, and Compd-34 was also calculated from the RMSDs of heavy atoms of the inhibitors to adequately assess their stability at the CCR5 binding site (Fig. **S1a**). The observed average active site RMSD within 8Å of the inhibitors were 1.3 ± 0.2 Å, 1.7 ± 0.5 Å, and 2.1 ± 0.4 Å for Compd-21, Compd-34, and Maraviroc, respectively (Fig. **S1a**).

3.3.2. Alterations in CCR5 Solvent Accessibility

The Solvent Accessible Surface Area (SASA) was examined to explore the behaviour of hydrophobic and hydrophilic residues of the CCR5 complexes with water molecules during the simulation (Fig. **5c**). The SASA assessment is capable of predicting the extent of protein conformational changes occurring upon binding (Marsh & Teichmann, 2011). An average SASA value of 16329.8 ± 310 Å^2, was observed for the Apo system, while CCR5-21, CCR5-34, and CCR5-MVR complexes were 16732.1 ± 256 Å^2, 16988.3 ± 311 Å^2, and 16292.8 ± 281

Å^2, respectively. In addition, the observed average active site SASA within 8Å of the inhibitors were 118.70 ± 26.7 Å, 135.69 ± 30.4 Å, and 146.67 ± 27.4 Å for Compd-21, Compd-34, and Maraviroc, respectively (Fig. **S1c**).

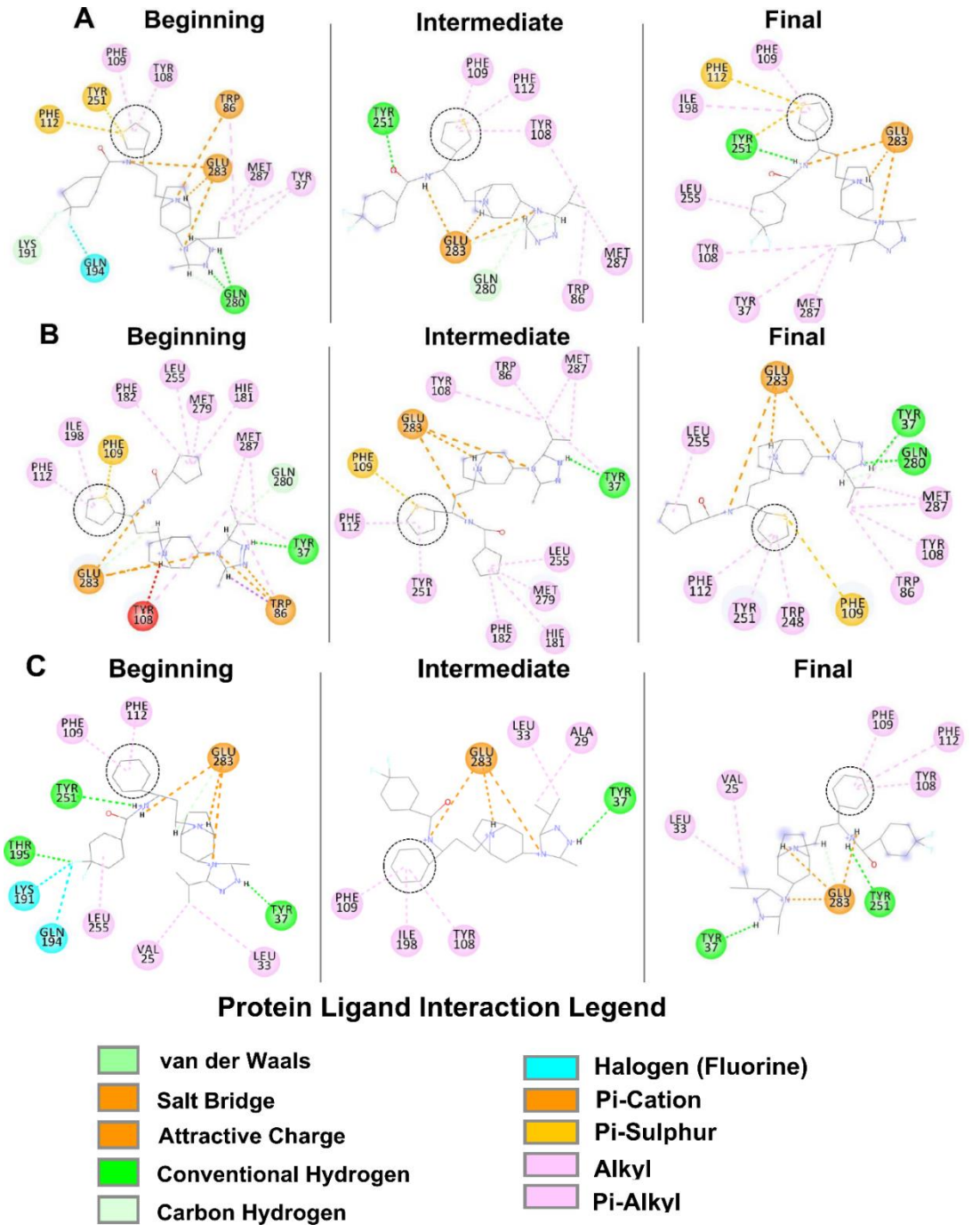


Fig. (4). Dynamic binding mode and receptor-ligand interaction fingerprint for (**a**) Compd-34, (**b**) Compd-21 and (**c**) Maraviroc at CCR5 binding site. Key functional group variations highlighted in dashed circles (Image prepared by author).

3.3.3. Protein Fold Assessment

The radius of gyration (Rg) was computed to estimate the overall change in the total compactness, folding and shape of CCR5 protein mass upon binding to the inhibitors during the simulations. The Rg graph for Cα atoms of apo and CCR5-complexed systems are shown in Fig. **5d** with average values of 21.07 ± 0.1 Å, 21.15 ± 0.1 Å, 20.98 ± 0.1 Å, and 20.89 ± 0.1 Å for CCR5-21, CCR5-34, CCR5-MVR, and CCR5-APO, respectively. Similarly, the average active site Rg values within 8 Å of the inhibitors were calculated to be 5.14 ± 0.1 Å, 5.57 ± 0.1 Å, and 5.68 ± 0.1 Å for Compd-21, Compd-34, and Maraviroc, respectively (Fig. **S1b**). The Rg results showed that the binding of the inhibitors did not significantly affect the active site and the overall conformational diversity of the CCR5 protein system.

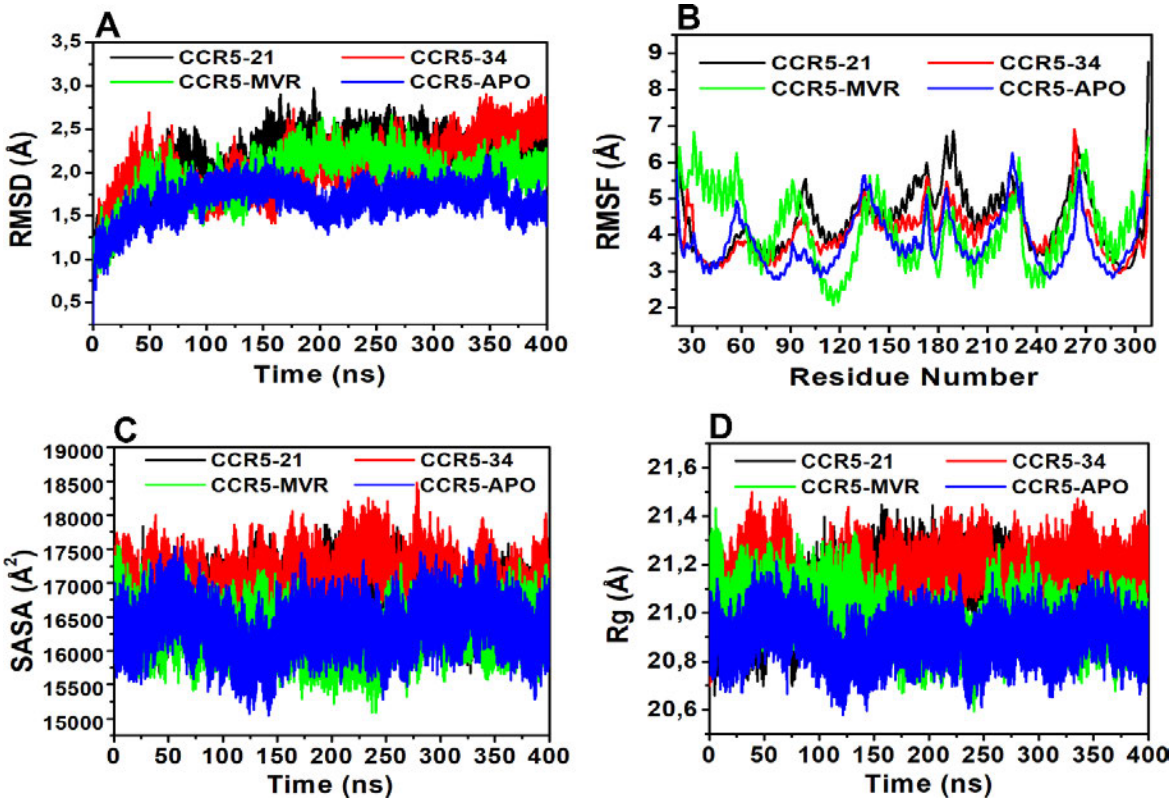


Fig. (5). Time series of (**a**) RMSDs for complexed and apo systems (**b**) RMSF for complexed and apo systems (**c**) SASA for complexed and apo systems and (**d**) Rg for complexed and apo systems over 400ns (Image prepared by author).

3.3.4. Flexibility Analysis

The Root Mean Square Fluctuation (RMSF) for each system Cα atom was calculated to gain insight into the extent to which amino acid residues mobility varies in CCR5 receptor in the bound and unbound state (Fig. **5b**). The results suggest that while transmembrane regions (TMI-TMVII) showed high stability for all systems (~ 1.0 Å RMSF), higher residual fluctuations were observed in the intracellular (ICL1, ICL2 and ICL3) and extracellular (ECL1,

ECL2, and ECL3) loop regions (up to 7 Å RMSF). This increased loop residual fluctuations were prominent in the ICL3, ECL2 and ECL3 regions. The observed higher intra- and extracellular residual mobility is expected since they belong to the outer parts of the receptor that is solvent exposed.

3.3.5. Secondary Structural Analysis

The DSSP analysis was further carried out to provide an overview of the gain and loss of secondary structure of both the CCR5 and CCR5-inhibitor complexes throughout the simulations (Fig. **6**). The results showed that the majority of the secondary structure elements were stable. All the seven-transmembrane α-helices [TMI (22-58); TMII (63-92); TMIII (97-132); TMIV (141-167); TMV (186-224); TMVI (228-265); and TMVII (268-300)] persisted throughout the simulation in the bound and unbound state. Also, the parallel and antiparallel sheet of the β-hairpin in the ECL2 (Thr167-Tyr184) were stable and unaltered throughout the entire simulation in all systems. However, slight structural changes were observed in the inhibitor bound and unbound state. For instance, the Turn and Bend in the ECL3 (266-267) of the unbound state (Fig. **6a**) extended into the TMVI (228-265) in the bound state (Fig. **6c-d**). Similarly, helix-8 (301-313) was more α-helix in Maraviroc bound than in Compd-21, Compd-34 and the unbound state.

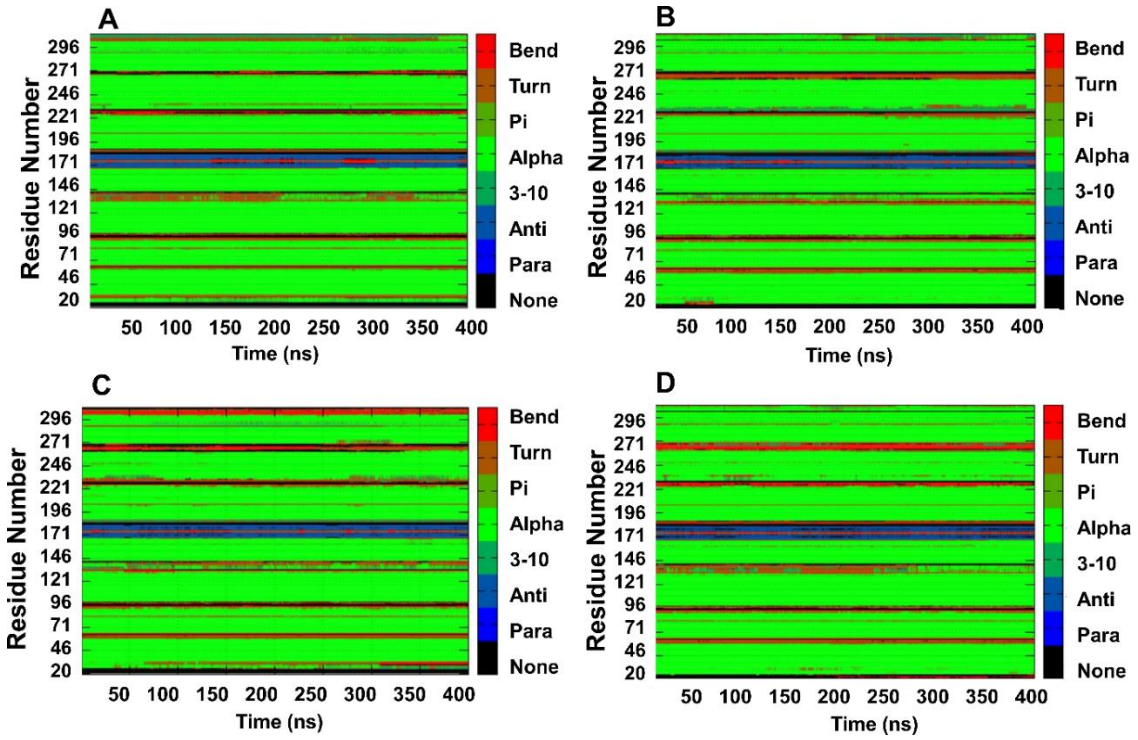


Fig. (6). DSSP analysis for secondary structure prediction for (**a**) CCR5-apo, (**b**) CCR5-21, (**c**) CCR5-34 and (**d**) CCR5-MVR (Image prepared by author).

3.3.6. Essential Conformational Dynamic Analysis

To identify dominant motion in CCR5 complexes, the principal component analysis (PCA) was performed to capture the combined movements of the Cα atoms in the protein for the first few essential eigenvectors of the covariance matrix. The principal component and clustering analysis provide considerable insight into the nature of conformational differences associated with the binding of an inhibitor to a receptor (Desdouits et al., 2015; Martínez-Archundia et al., 2019). The sum of eigenvalues increases as a function of the number of eigenvalues resulting from the collected MD trajectories is shown in Fig. **7**. The obtained spectrum of eigenvalues displays the proportion of variances against the eigenvalues for the inhibitor-complexed systems.

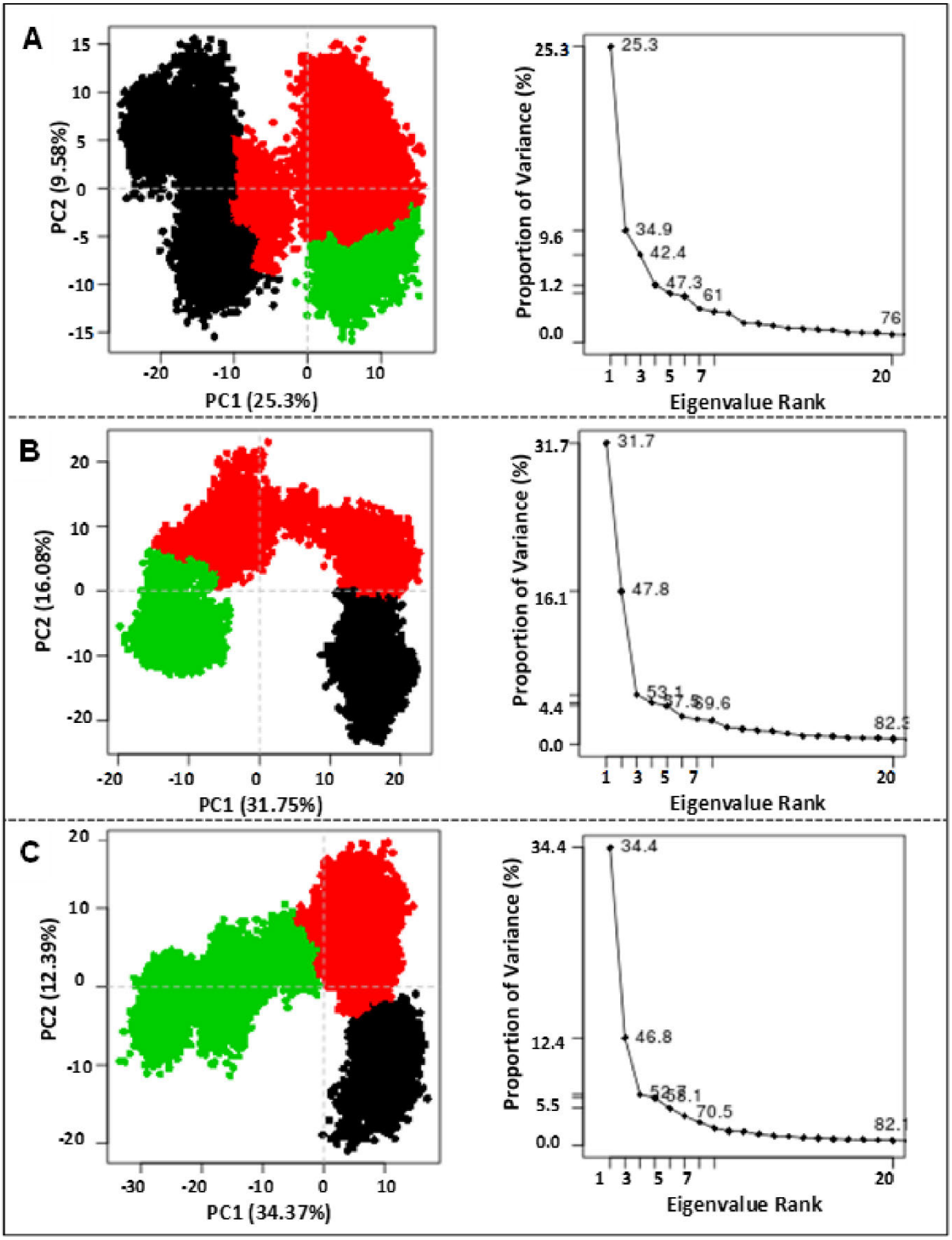


Fig. (7). Clustering and principal component projection of the trajectory conformers unto the planes formed by the first two principal components based on the dominant motion of the complex system for (**a**) Maraviroc (**b**) Compd-21, and (**c**) Compd-34 using the Bio3D package in R (Image prepared by author).

The PCA showed that the proportion of variance of the first two principal components (PCs) accounted for 25.3%, 31.8%, and 34.4% of the overall variation in the detected motion of CCR5-MVR, CCR5-Compd21, and CCR5-Compd34 complexed trajectories, respectively. The clustering of the complex protein structures for PC1 and PC2 unto two-dimensional subspace (Fig. **7**) indicates conformational distribution variance in the different CCR5 complex systems. Whereas the subspace of PCs 1 and 2 of CCR5-MVR complex showed a uniform and overlapping conformational subspace, CCR5-Compd24 and CCR5-Compd34 depict distinct periodic jump within the conformational subspace. The internal dynamics of CCR5 when bound to Compd34 and Compd21 are diverse from when bound to Maraviroc.

4. Discussion

In this study, we investigated the binding energy interactions, the structural basis and conformational changes associated with the higher binding affinity of Compd-34 and Compd-21 compared with Maraviroc at CCR5 binding site. An integrated *in silico* approach that combines explicit lipid bilayer MD simulation, principal component analysis, and MM-GBSA binding free energy calculations were employed to obtain atomistic molecular details of the interaction mechanisms of how slight structural variance between these inhibitors significantly affects their binding profiles at CCR5 receptor.

The predicted binding free energies of the studied inhibitors corroborates with the experimental bioactivity data. Furthermore, the decomposition of the total interaction energy components suggests that van der Waals interactions appear to dominate in stabilizing the CCR5-inhibitor complexes. The dynamic hydrogen bond analysis showed that hydrogen bond interactions between the antagonists and the active site residues Glu283$^{7.39}$, Gln280$^{7.36}$, Tyr251$^{6\,51}$, Thr284$^{7.40}$, and Tyr37$^{1.39}$ (with Compd-34); Glu283$^{7\,39}$, Tyr251$^{6.51}$, Tyr37$^{1.39}$ and Gln280$^{7\,36}$ (with Compd-21); and Glu283$^{7.36}$, Tyr37$^{1.39}$, and Tyr251$^{6.51}$ (with Maraviroc) are vital for the stability of the ligand-bound conformations. Bai *et al*. (2014) demonstrated that residues Tyr37$^{1.39}$, Tyr251$^{6.51}$ and Glu283$^{7.39}$ form stable hydrogen bonds with Maraviroc (Bai et al., 2014). The overall hydrogen bond strength and stability were observed to be higher for Compd-34 and Compd-21 than for Maraviroc at CCR5 active site.

The binding of gp120 to CCR5 has been reported to be sensitive to the mutations of active site residues such as Trp86$^{2.60}$, Trp94$^{23\,50}$, Tyr108$^{3.32}$, Trp248$^{6.48}$, and Tyr251$^{6.51}$. These residues were suggested to form a potential binding cavity for gp120 (Garcia-Perez et al., 2011). The recent crystallization of the full-length gp120 in complex with unmodified CCR5 receptor

reveals that the gp120 V3 loop makes extensive contacts with residues of the chemokine recognition site 2 (Phe109$^{3\,33}$, Tyr108$^{3.32}$, Trp86$^{2.60}$, Tyr89$^{2.63}$, Glu283$^{7.39}$, and Tyr251$^{6.51}$) (Shaik et al., 2019). The gp120 V3 loop of HIV-1 has been observed to predominantly occupy the minor sub-pocket formed by the transmembrane helices I-III and VII of the chemokine recognition site 2 (CRS2) of CCR5. The V3 loop Arg313 residue seems to be sandwiched between Glu283$^{7.39}$ and Tyr251$^{6.51}$ of CCR5(Shaik et al., 2019). The per-residue interaction energy decomposition reveals that Trp86$^{2.60}$, Tyr108$^{3.32}$, Tyr251$^{6.51}$ make overall stronger interactions with Compd-21 and Compd-34 compared with Maraviroc. Similarly, Tyr251$^{6.51}$ and Glu283$^{7.39}$ further engage in stronger hydrogen bond interactions with Compd-21 and Compd-34 compared to Maraviroc. These observed interactions cumulatively accounted for Compd-21 and Compd-34 higher ΔG binding energies than Maraviroc. These findings further support the experimental findings (Tan et al., 2013) and the theoretical result (Bai et al., 2014) demonstrating that Tyr37$^{1.39}$, Tyr251$^{6.51}$, Glu283$^{7.39}$ of TMI, TMVI and TMVII play a vital role in the binding of inhibitors to CCR5 binding site.

The Principal component analysis (PCs 1 and 2) revealed a subspace of CCR5-MVR complex with a uniform and overlapping conformational subspace, whereas CCR5-Compd31 and CCR5-Compd-21 complexes showed distinct periodic jump within the conformational subspace. The SASA results showed that active site amino acid residues within 8 Å of CCR5-21 and CCR5-34 inhibitors had lower SASA values compared to CCR5-MVR.

The MD simulations and atomistic interaction analysis of the studied compounds substantiate the recent findings that Maraviroc blocks gp120 binding to CCR5 via direct competitive inhibition in contrast to earlier views of allosteric inhibition via conformational availability restriction (Shaik et al., 2019). As evident by the hydrogen bond analysis, residue energy decomposition analysis, and the protein-ligand interaction, the studied inhibitors made important contacts with the residues that are critical in gp120 V-loop binding. The gp120 V3 loop has been observed to overlap mainly with Maraviroc in the minor sub pocket, which is occupied primarily by the triazole moiety of Maraviroc, Compd-31, and Compd-34. Structural modifications at the thiophene substituent with functional group(s) that may maintain strong hydrogen bond with Glu283$^{7.39}$ and Tyr251$^{6.51}$ as well as the addition of new functional groups to the triazole ring may increase inhibitor competition with gp120 V3-loop with enhanced potency (Shaik et al., 2019).

5. Conclusion

In summary, the findings highlighted in this work provide a structural understanding of the novel 1-heteroaryl-1,3-propanediamine derivatives (Compd-21 and Compd-34) and Maraviroc targeting the CCR5 receptor. The MD simulation analyses reveal that Trp86$^{2.60}$, Tyr108$^{3.32}$, Tyr251$^{6.51}$ make overall stronger interaction with Compd-21 and Compd-34 compared with Maraviroc. Similarly, Tyr251$^{6.51}$ and Glu283$^{7.39}$ further engage in stronger hydrogen bond interactions with Compd-21 and Compd-34 compared with Maraviroc. Thus, the substitution of the thiophene moieties in Compd-34 and Compd-21 cumulatively made stronger interactions with residues critical for V3-loop binding compared with the phenyl group in Maraviroc. Further structural modifications at the thiophene substituent and the addition of additional functional groups to the triazole ring may increase inhibitor competition with gp120 V3-loop with enhanced potency. This offers a foundation for the onward structural modifications and rational design of novel potent antagonists of CCR5 in HIV-1 treatment.

Ethics Approval and Consent to Participate

Not applicable

Human and Animal Rights

No humans and animals were used in the study.

Consent for Publication

Not applicable

Availability of Data and Materials

Not applicable

Funding

None

Conflict of Interest

The authors declare no conflicts of interest in this work.

Acknowledgements

The authors acknowledge the Centre for High-Performance Computing (CHPC) (http://www.chpc.ac.za), Cape Town, South Africa for computational support and Schrodinger

Maestro licence. We are further grateful to the College of Health Sciences of the University of KwaZulu-Natal for supporting the research.

ORCID IDs

Prof Mahmoud E.S. Soliman 0000-0002-8711-7783

Dr Fisayo Andrew Olotu 0000-0003-3604-5983

Patrick Appiah-Kubi 0000-0002-5904-3051

Authors Contribution

P. Appiah-Kubi conceptualized and designed the study, performed the experiment, data analysis and plots, interpretation of data, and wrote the manuscript. FA. Olotu reviewed and edited the final manuscript draft. M.E.S Soliman supervised the study.

References

Allers, K., Hütter, G., Hofmann, J., Loddenkemper, C., Rieger, K., Thiel, E. & Schneider, T. 2011. Evidence for the cure of HIV infection by CCR5Δ32/Δ32 stem cell transplantation. *Blood*, 117(10): 2791–2799.

Bai, Q., Zhang, Y., Li, X., Chen, W., Liu, H. & Yao, X. 2014. Computational study on the interaction between CCR5 and HIV-1 entry inhibitor maraviroc: insight from accelerated molecular dynamics simulation and free energy calculation. *Physical Chemistry Chemical Physics*, 16(44): 24332–24338.

Ballesteros, J.A. & Weinstein, H. 1995. [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In *Methods in neurosciences*. Elsevier: 366–428.

Bayly, C.I., Cieplak, P., Cornell, W. & Kollman, P.A. 1993. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *The Journal of Physical Chemistry*, 97(40): 10269–10280.

Berger, E.A., Murphy, P.M. & Farber, J.M. 1999. Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease. *Immunol Today*, 17(1): 657–700. https://doi.org/10.1146/annurev.immunol.17.1.657.

Berman, H.M., Battistuz, T., Bhat, T.N., Bluhm, W.F., Philip, E., Burkhardt, K., Feng, Z., Gilliland, G.L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Thanki, N., Weissig, H. & Westbrook, J.D. 2002. The Protein Data Bank. *Biological Crystallography*, 58: 899–907.

Case, D.A., Babin, V., Berryman, J., Betz, R.M., Cai, Q., Cerutti, D.S., Cheatham Iii, T.E., Darden, T.A., Duke, R.E. & Gohlke, H. 2018. Amber 18. *University of California, San Francisco*.

Chen, F., Liu, H., Sun, H., Pan, P., Li, Y., Li, D. & Hou, T. 2016. Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein–protein binding free energies and re-rank binding poses generated by protein–protein docking.

Physical Chemistry Chemical Physics, 18(32): 22129–22139.

Darden, T., York, D. & Pedersen, L. 1993. Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. *The Journal of chemical physics*, 98(12): 10089–10092.

Desdouits, N., Nilges, M. & Blondel, A. 2015. Principal component analysis reveals correlation of cavities evolution and functional motions in proteins. *Journal of Molecular Graphics and Modelling*, 55: 13–24.

Detels, R., Munoz, A., McFarlane, G., Kingsley, L.A., Margolick, J.B., Giorgi, J., Schrager, L.K., Phair, J.P. & Investigators, M.A.C.S. 1998. Effectiveness of potent antiretroviral therapy on time to AIDS and death in men with known HIV infection duration. *Jama*, 280(17): 1497–1503.

Dickson, C.J., Madej, B.D., Skjevik, Å.A., Betz, R.M., Teigen, K., Gould, I.R. & Walker, R.C. 2014. Lipid14: the amber lipid force field. *Journal of chemical theory and computation*, 10(2): 865–879.

Dyda, F., Hickman, A.B., Jenkins, T.M., Engelman, A., Craigie, R. & Davies, D.R. 1994. Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. *Science*, 266(5193): 1981–1986.

FDA, U.S. 2007. FDA notifications. Maraviroc approved as a CCR5 co-receptor antagonist. *AIDS Alert*, 22: 103.

Flanagan, C.A. 2014. Receptor conformation and constitutive activity in CCR5 chemokine receptor function and HIV infection. In *Advances in Pharmacology*. Elsevier: 215–263.

Garcia-Perez, J., Rueda, P., Alcami, J., Rognan, D., Arenzana-Seisdedos, F., Lagane, B. & Kellenberger, E. 2011. Allosteric model of maraviroc binding to CC chemokine receptor 5 (CCR5). *Journal of Biological Chemistry*, 286(38): 33409–33421.

Garcia-Perez, J., Staropoli, I., Azoulay, S., Heinrich, J.-T., Cascajero, A., Colin, P., Lortat-Jacob, H., Arenzana-Seisdedos, F., Alcami, J. & Kellenberger, E. 2015. A single-residue change in the HIV-1 V3 loop associated with maraviroc resistance impairs CCR5 binding affinity while increasing replicative capacity. *Retrovirology*, 12(1): 1–20.

Grant, B.J., Rodrigues, A.P.C., ElSawy, K.M., McCammon, J.A. & Caves, L.S.D. 2006. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics*, 22(21): 2695–2696.

Günthard, H.F., Aberg, J.A., Eron, J.J., Hoy, J.F., Telenti, A., Benson, C.A., Burger, D.M., Cahn, P., Gallant, J.E. & Glesby, M.J. 2014. Antiretroviral treatment of adult HIV infection: 2014 recommendations of the International Antiviral Society–USA Panel. *Jama*, 312(4): 410–425.

Harrison, S.C. 2008. Viral membrane fusion. *Nature structural & molecular biology*, 15(7): 690.

Hou, T., Wang, J., Li, Y. & Wang, W. 2011. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of Chemical Information and Modeling*, 51(1): 69–82.

Huff, J.R. 1991. HIV protease: a novel chemotherapeutic target for AIDS. *Journal of medicinal chemistry*, 34(8): 2305–2314.

Isberg, V., Mordalski, S., Munk, C., Rataj, K., Harpsøe, K., Hauser, A.S., Vroling, B., Bojarski, A.J., Vriend, G. & Gloriam, D.E. 2016. GPCRdb: an information system for G protein-

coupled receptors. *Nucleic acids research*, 44(D1): D356–D364.

Jo, S., Kim, T., Iyer, V.G. & Im, W. 2008. CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of computational chemistry*, 29(11): 1859–1865.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. & Klein, M.L. 1983. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2): 926–935.

Kohlstaedt, L.A., Wang, J., Friedman, J.M., Rice, P.A. & Steitz, T.A. 1992. Crystal structure at 3.5 A resolution of HIV-1 reverse transcriptase complexed with an inhibitor. *Science*, 256(5065): 1783–1790.

Larini, L., Mannella, R. & Leporini, D. 2007. Langevin stabilization of molecular-dynamics simulations of polymers by means of quasisymplectic algorithms. *The Journal of chemical physics*, 126(10): 104101.

Lomize, M.A., Lomize, A.L., Pogozheva, I.D. & Mosberg, H.I. 2006. OPM: orientations of proteins in membranes database. *Bioinformatics (Oxford, England)*, 22(5): 623–625.

M Gibson, R. & J Arts, E. 2012. Past, present, and future of entry inhibitors as HIV microbicides. *Current HIV research*, 10(1): 19–26.

Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E. & Simmerling, C. 2015. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of chemical theory and computation*, 11(8): 3696–3713.

Marsh, J.A. & Teichmann, S.A. 2011. Relative Solvent Accessible Surface Area Predicts Protein Conformational Changes upon Binding. *Structure(London, England:1993)*, 19(6): 859–867.

Martínez-Archundia, M., Correa-Basurto, J., Montaño, S. & Rosas-Trigueros, J.L. 2019. Studying the collective motions of the adenosine A2A receptor as a result of ligand binding using principal component analysis. *Journal of Biomolecular Structure and Dynamics*, 37(18): 4685–4700.

Oppermann, M. 2004. Chemokine receptor CCR5: insights into structure, function, and regulation. *Cellular signalling*, 16(11): 1201–1210.

Peng, P., Chen, H., Zhu, Y., Wang, Z., Li, J., Luo, R.-H., Wang, J., Chen, L., Yang, L.-M. & Jiang, H. 2018. Structure-Based Design of 1-Heteroaryl-1, 3-propanediamine Derivatives as a Novel Series of CC-Chemokine Receptor 5 Antagonists. *Journal of medicinal chemistry*, 61(21): 9621–9636.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. & Ferrin, T.E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13): 1605–1612.

Roe, D.R. & Cheatham, T.E. 2013. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation*, 9(7): 3084–3095.

Rosenbaum, D.M., Rasmussen, S.G.F. & Kobilka, B.K. 2009. The structure and function of G-protein-coupled receptors. *Nature*, 459(7245): 356.

Salentin, S., Schreiber, S., Haupt, V.J., Adasme, M.F. & Schroeder, M. 2015. PLIP: fully automated protein–ligand interaction profiler. *Nucleic acids research*, 43(W1): W443–W447.

Salmas, R.E., Yurtsever, M. & Durdagi, S. 2015. Investigation of inhibition mechanism of

chemokine receptor CCR5 by micro-second molecular dynamics simulations. *Scientific reports*, 5: 13180.

Schrödinger Release 2019-4. 2019. Maestro. https://www.schrodinger.com/Maestro.

Seifert, E. 2014. OriginPro 9.1: Scientific data analysis and graphing software - Software review. *Journal of Chemical Information and Modeling*, 54(5): 1552–1552.

Shaik, M.M., Peng, H., Lu, J., Rits-Volloch, S., Xu, C., Liao, M. & Chen, B. 2019. Structural basis of coreceptor recognition by HIV-1 envelope spike. *Nature*, 565(7739): 318.

Sorce, S., Myburgh, R. & Krause, K.-H. 2011. The chemokine receptor CCR5 in the central nervous system. *Progress in neurobiology*, 93(2): 297–311.

Sprenger, K.G., Jaeger, V.W. & Pfaendtner, J. 2015. The General AMBER Force Field (GAFF) Can Accurately Predict Thermodynamic and Transport Properties of Many Ionic Liquids. *The Journal of Physical Chemistry B*, 119(18): 5882–5895.

Sun, H., Duan, L., Chen, F., Liu, H., Wang, Z., Pan, P., Zhu, F., Zhang, J.Z.H. & Hou, T. 2018. Assessing the performance of MM/PBSA and MM/GBSA methods. 7. Entropy effects on the performance of end-point binding free energy calculation approaches. *Physical Chemistry Chemical Physics*, 20(21): 14450–14460.

Sun, H., Li, Y., Tian, S., Xu, L. & Hou, T. 2014. Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set. *Physical Chemistry Chemical Physics*, 16(31): 16719–16729.

Tan, Q., Zhu, Y., Li, Jian, Chen, Z., Han, G.W., Kufareva, I., Li, T., Ma, L., Fenalti, G., Li, Jing, Zhang, W., Xie, X., Yang, H., Jiang, H., Cherezov, V., Liu, H., Stevens, R.C., Zhao, Q. & Wu, B. 2013. Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science*, 341(6152): 1387–1390.

Wang, J., Shu, M., Wang, Yuanqiang, Hu, Y., Wang, Yuanliang, Luo, Y. & Lin, Z. 2016. Identification of potential CCR5 inhibitors through pharmacophore-based virtual screening, molecular dynamics simulation and binding free energy analysis. *Molecular BioSystems*, 12(11): 3396–3406.

Webb, B. & Sali, A. 2014. Protein structure modeling with MODELLER. *Protein Structure Prediction*: 1–15.

Wensing, A.M., Calvez, V., Ceccherini-Silberstein, F., Charpentier, C., Günthard, H.F., Paredes, R., Shafer, R.W. & Richman, D.D. 2019. 2019 update of the drug resistance mutations in HIV-1. *Topics in antiviral medicine*, 27(3): 111.

Yang, Y., Shen, Y., Liu, H. & Yao, X. 2011. Molecular dynamics simulation and free energy calculation studies of the binding mechanism of allosteric inhibitors with p38α MAP kinase. *Journal of chemical information and modeling*, 51(12): 3235–3246.

Ylilauri, M. & Pentikäinen, O.T. 2013. MMGBSA as a tool to understand the binding affinities of filamin-peptide interactions. *Journal of Chemical Information and Modeling*, 53(10): 2626–2633.

Zhang, F., Yuan, Y., Xiang, M., Guo, Y.-Z., Li, M.-L., Liu, Y. & Pu, X.-M. 2019. Molecular Mechanism regarding Allosteric Modulation of Ligand Binding and the Impact of Mutations on Dimerization for CCR5 Homodimer. *Journal of chemical information and modeling*.

CHAPTER 6

Supplementary Material

Elucidating the Disparate Inhibitory Mechanisms of Novel 1-Heteroaryl-1,3-Propanediamine Derivatives and Maraviroc towards C-C Chemokine Receptor 5: Insights for Structural Modifications in HIV-1 Drug Discovery

Patrick Appiah-Kubi[1], Fisayo Andrew Olotu[1], Mahmoud E. S. Soliman[1*]

[1]Molecular Bio-computation and Drug Design Laboratory

School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

Telephone: +27 (0) 31 260 8048

Fax: +27 (0) 31 260 78

Email: soliman@ukzn.ac.za

Website: http://soliman.ukzn.ac.za

Patrick Appiah-Kubi appiahpat@gmail.com

Fisayo Andrew Olotu olotufisayo@gmail.com

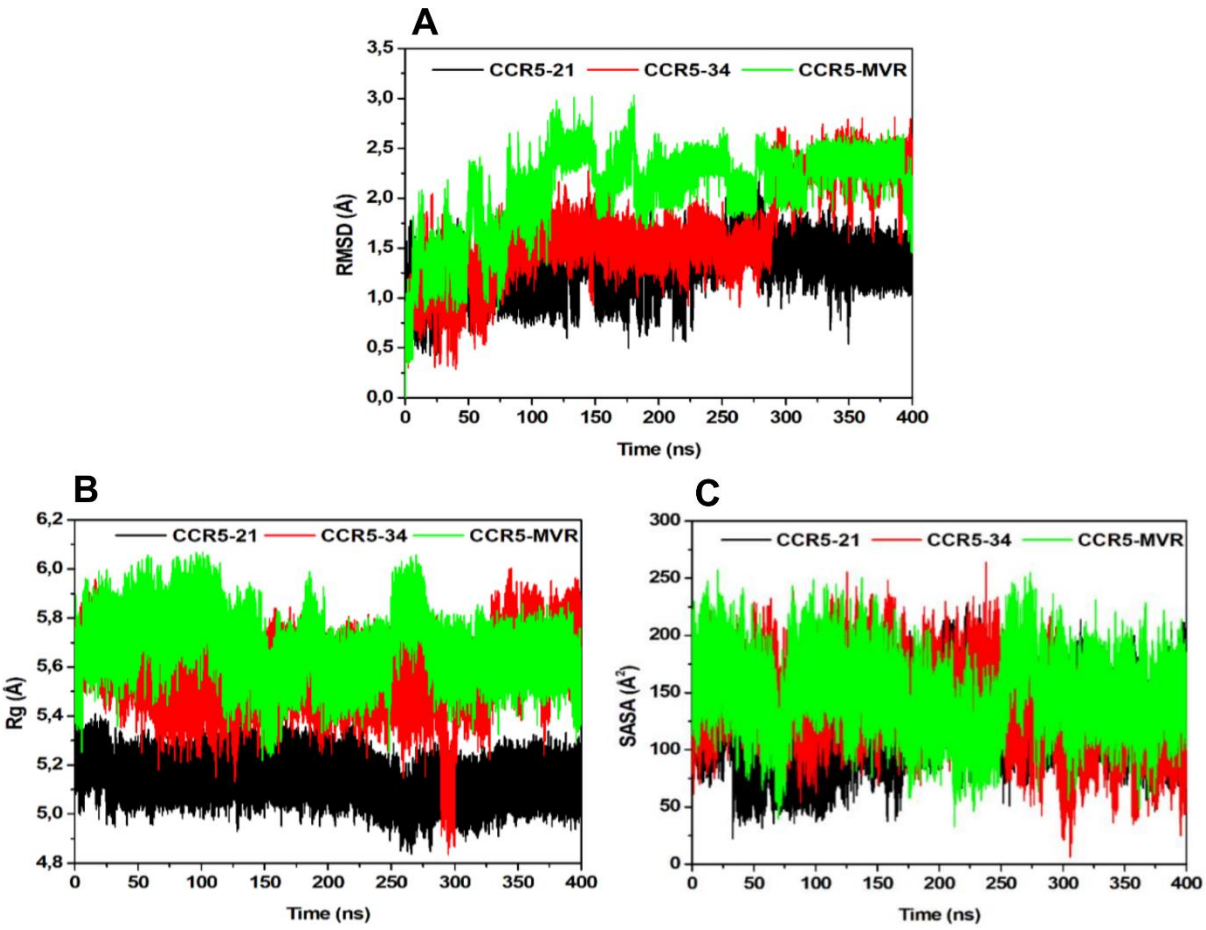


Figure S1: Time series of active site (**a**) RMSDs for complexed systems, (**b**) Rg for complexed systems and (**c**) SASA for complexed systems over 400ns (Image prepared by author).

CHAPTER 7

Submitted Manuscript

Structure-based identification of novel scaffolds as potential HIV-1 entry inhibitors towards CCR5 in HIV disease therapy

Patrick Appiah-Kubi[1], Emmanuel Amarachi Iwuchukwu [1], Mahmoud E. S. Soliman[1*]

[1]Molecular Bio-computation and Drug Design Laboratory

School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

Telephone: +27 (0) 31 260 8048

Fax: +27 (0) 31 260 78

Email: soliman@ukzn.ac.za

Website: http://soliman.ukzn.ac.za

Abstract

The C-C chemokine receptor 5 (CCR5) viral coreceptor belonging to the G protein-coupled receptor family is one of the families of chemokine receptors. The interactions of CCR5 with HIV-1 during viral entry positions it as an effective therapeutic target for the design of potent antiviral therapies. The FDA approved the small-molecule Maraviroc as CCR5 drug in 2007, while clinical trials failure has characterised many of the other CCR5 inhibitors. Thus, the continual identification of potential CCR5 inhibitors is, therefore, warranted. In this study, a structure-based discovery approach has been utilised to screen and retrieved novel potential CCR5 inhibitors from the Asinex antiviral compound (~ 8,722) database. Explicit lipid-bilayer molecular dynamics simulation, *in silico* physicochemical and pharmacokinetic analyses, were further performed for the top compounds. A total of 23 structurally diverse compounds with binding scores higher than Maraviroc were selected. Subsequent molecular dynamics (MD) simulations analysis of the top four compounds LAS 51495192, BDB 26405401, BDB 26419079, and LAS 34154543 maintained stability at the CCR5 binding site. Furthermore, these compounds made pertinent interactions with CCR5 residues critical for the HIV-1 gp120-V3 loop binding such as Trp86, Tyr89, Phe109, Tyr108, Glu283 and Tyr251. Additionally, the predicted *in silico* physicochemical and pharmacokinetic descriptors of the selected compounds were within the acceptable range for drug-likeness. The results suggest positive indications that the identified molecules may represent promising CCR5 entry inhibitors. Further structural optimisations and biochemical testing of the proposed compounds may assist in the discovery of effective HIV-1 therapy.

Keywords: C-C Chemokine Receptor 5, Maraviroc, Structure-based drug design, Virtual screening, Molecular docking, MD Simulations, Asinex database, ADMET.

1. Introduction

C-C chemokine receptor 5 (CCR5) is a member of the G protein-coupled receptors (GPCRs) and is essential in the amelioration of human immunodeficiency virus (HIV) infection (Oppermann, 2004). C-C chemokine receptor 5 plays a critical role in the early stages of HIV-1 infection and is a functionally co-receptor for HIV-1 viral entry (Tan et al., 2013). The human immunodeficiency virus-1 penetrates cells by binding its envelope glycoprotein gp120 to the CD4 receptor and co-receptors like CCR5 and CXCR4 (Woollard & Kanmogne, 2015). CCR5 is the most prominent of all chemokine co-receptors employed by HIV-1 for cell penetration, typically at the onset of infection with the R5-tropic HIV-1 strains being transmitted the most (Vangelista & Vento, 2018). CCR5 is also a potential target for ameliorating inflammatory, allergic, infectious, and autoimmune diseases such as diabetes and rheumatoid arthritis (Pereira et al., 2009; Spagnolo et al., 2005).

Extensive research in the development of inhibitors that could block the entry of HIV-1 through the target of CCR5 resulted in the approval of Maraviroc for the treatment of HIV-1 infection (FDA, 2007; Tan et al., 2013). Maraviroc is the first CCR5 drug approved by the FDA to inhibit HIV-1 entry (FDA, 2007). However, its prescription is limited due to identified factors such as its CYP450 inhibition, drug-drug interactions, particularly with CYP3A4 inhibitors, and viral resistance (Peng et al., 2018). Clinical studies on other CCR5 antagonists like Aplaviroc (Nichols et al., 2008), Vicriviroc (Schürmann et al., 2007; Gulick et al., 2007) and Cenicriviroc (Klibanov et al., 2010) failed at some stages of their clinical trials such as lack of efficacy and/or hepatotoxicity.

The limitations of the approved drug and clinical drugs targeting CCR5, as well as the reported case of clinical drug failures, necessitate the identification and development of novel effective and enhanced inhibitors of CCR5. Structure-based virtual screening incorporated in computer-aided drug design has contributed immensely to advancing the drug discovery process (Jin et al., 2020; Hughes et al., 2019; Rodríguez et al., 2014). This is because of its cost-effectiveness in the identification of novel therapeutics within a short time frame. Structure-based drug discovery enhances database screening by employing the principle of molecular docking, which predicts the best conformation for target-ligand interaction. Recent advancement in X-ray crystallography and nuclear magnetic resonance and electron cryo-microscopy (cryoEM) has promoted structure-based drug discovery (SBDD) due to the availability of diverse 3D protein crystal structures (Wang et al., 2018). Structure-based drug discovery can be employed to investigate the binding mode and ligand binding process with substantial accuracy, which provides information about their respective biological activities (Mandal et al., 2009; Wang et

al., 2018). Data like these can be harnessed toward the design of potent high-affinity ligands with basic moieties that produce the needed pharmacological and therapeutic activities (Ferreira et al., 2015; Llanos et al., 2017; Macalino et al., 2015).

Computational strategies employed in the past years toward the identification of potential CCR5 inhibitors due to the absence of CCR5 crystal structures were restricted to ligand-based approaches such as; shape-based virtual screening (Pérez-Nueno et al., 2008), quantitative structure activity relationship (QSAR) (Xu et al., 2004), and ligand-based pharmacophore modelling (Debnath, 2003). The first CCR5 crystal structure in complex with Maraviroc was reported in 2013 (Tan et al., 2013). More recent approaches following the availability of CCR5 crystal structure have focused on structure-based methods (Wang et al., 2016; Lin et al., 2019). These studies have used compounds from databases such as the National Cancer Institute (NCI), ChEMBL database, and ZINC database.

The recent crystallisation of the full-length CCR5 in complex with gp120 V3 loop suggests that Maraviroc blocks gp120 binding to CCR5 via direct competitive inhibition in contrast to earlier views of non-competitive allosteric inhibition through conformational availability restriction (Shaik et al., 2019). Herein, we screened the unexplored antiviral compound library of the Asinex database against the recently crystalised full-length CCR5 in complex with gp120 V3 loop (Shaik et al., 2019). Multistep structure-based virtual screening techniques were employed in this study to identify potential CCR5 inhibitors. The best four hits of the identified compounds in complex with CCR5 were subjected to conventional computational molecular dynamics simulations to understand their atomistic mechanism of action and conformational dynamics. The identified compounds could be optimised/tested as potent CCR5 HIV-1 entry inhibitors.

2. Materials and Methods

In this work, diverse pharmacoinformatic techniques, including molecular docking, *in silico* physicochemical and pharmacokinetic analysis, and molecular dynamics simulation, have been applied to identify novel potential CCR5 HIV1 entry inhibitors.

2.1 Compound library and protein preparation

A total of 8722 antiviral compounds were freely downloaded from the antiviral sub-class of the Asinex database (ASINEX, 2020) for the virtual high throughput screening. The Asinex database contains compounds with unique features that can aid in hit-to-lead identification,

fragment-based drug design, and structure-based drug design. The single structural data format (SDF) of the 8722 compounds were downloaded and further prepared for Autodock Vina and Schrodinger Maestro to eliminate unwanted elements and improper valency and converted into individual three-dimensional (3D) mol2 format.

2.2 Ligand and Receptor Preparations for Autodock Vina Screening

The compounds were initially prepared and converted to pbdqt with raccoon (Forli, 2010). The pdbqt format incorporated partial charges and AutoDock atom types acceptable by AutoDock Vina (Trott & Olson, 2010). The protein was prepared by adding polar hydrogen atoms and Gasteiger charges. All nonstandard residues were deleted from the crystal structure. The grid box centre coordinates were 159.28, 148.39 and 161.65 in the X, Y, and Z axes, whereas the size of 25.20 × 20.86 × 24.85 Å was assigned to the pocket X, Y, and Z axes with exhaustiveness set to 8. The receptor, ligand, and grid parameters were written in the format that is acceptable by the Autodock Vina program.

2.3 Ligands and Protein Preparation procedure for Glide docking

The LigPrep module implemented in Maestro v12 (Schrödinger Release 2019-4, 2019) was used to generate low energy 3D structures of the compounds. The protonation states of the compounds were estimated at a pH 7.4 ± 0.2 using the program EpiK (Greenwood et al., 2010) and three-dimensional conformations generated. The compounds were desalted, and the possible tautomeric conformations set to ~32 tautomer per compound. Minimisation was finally performed for the compounds using the OPLS3e force field. The structure of the full-length CCR5 receptor in complex with the gp120 V3 loop (PDB ID: 6MEO) was used in the receptor grid generation for the Glide docking. The centre of the active site grid was defined around the binding position of the gp120 V3 loop in the structure.

The full-length CCR5 receptor was prepared using the Protein Preparation Wizard implemented in Maestro v12 (Schrödinger Release 2019-4, 2019) before the Glide docking. Pre-processing and energy minimisation of the receptor with OPLS3e force field was performed. Protonation states were assigned at physiological pH, followed by the addition of hydrogen atoms using the default parameters. The side chains stereochemistry was checked to avoid significant induced perturbations during the structure preparation.

2.4 Structure-based Virtual High Throughput Screening

The current study followed the schematic workflow presented in Fig. **1**. The binding poses and affinities of the compounds were initially predicted using the AutoDock Vina molecular

docking program (Trott & Olson, 2010). Maraviroc was docked and analysed as a positive reference control (Maraviroc Autodock Vina Score = -9.0 kcal/mol). A minimum binding affinity score of -10.0 kcal/mol threshold was chosen as the cut-off, and compounds with a binding affinity lower than the selected threshold were subsequently chosen for the Maestro virtual screening workflow (VSW) (Schrödinger Release 2019-4, 2019).

To further remove possible false-positive results, the top 500 compounds obtained from the Autodock molecular docking were gradually subjected to filtering steps via the Glide virtual screening workflow of Schrodinger Glide. The Glide virtual screening workflow module (Friesner et al., 2004) possesses unique virtual screening features like the high throughput virtual screening (HTVS), standard precision (SP) and extra precision (XP) for improved and efficient docking accuracy (Friesner et al., 2006). The above mentioned three docking precision levels implement an efficient search for the orientational, conformational and positional space of the docked ligand (Friesner et al., 2004; Friesner et al., 2006). Each compound docking progressed from HTVS to SP to XP, which involved Glide predicting the binding affinity and ranking the compounds. The top 20 % of compounds from the HTVS and SP docking levels were selected for XP docking.

Finally, the top four candidate compounds showing the best possible inhibitor binding pose from both the Autodock Vina and Glide-XP dockings were selected for binding mode analyses, MD simulations and molecular trajectory analyses. The type and pattern of binding interaction were equally analysed by utilising the protein-ligand interaction module of Maestro.

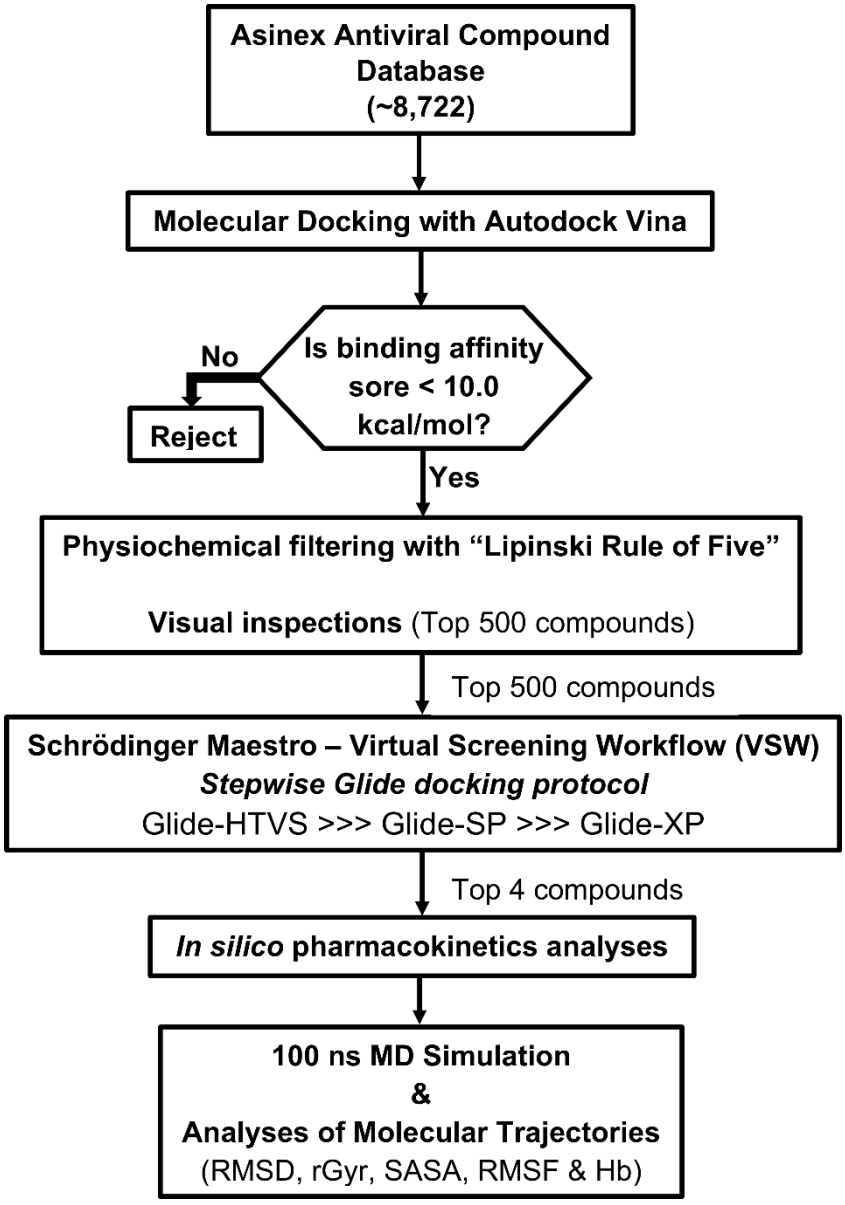


Fig. 1. Structure-based virtual screening and assessment workflow to identify potential CCR5 entry inhibitors (Image prepared by author).

2.5 Pharmacokinetics profiling and physiochemical property predictions

In silico pharmacokinetic profiling and drug-likeness estimations were further performed on the top compounds retrieved from the VSW using SwissADME (Daina et al., 2017) and pkCSM-pharmacokinetics (Pires et al., 2015) online software. Lipinski's rule of five (RO5) for empirical drug-likeness prediction was used to evaluate drug-likeness of the selected compounds (Meanwell, 2011). The rule includes: 1) hydrogen bond donors not more than five, 2) hydrogen bond acceptors not more than ten, 3) an octanol-water partition coefficient logP not more than five, 4) molecular weight less than 500 Daltons, and 5) the number of rotatable

bonds less than eight. Several vital molecular properties such as hydrophobicity (SlogP), and polar surface area (PSA) were also analysed. Additionally, pharmacokinetic parameters such as absorption, distribution, metabolism, excretion, and toxicity (ADMET) were analysed for the selected top four compounds relative to Maraviroc using the pkCSM-pharmacokinetics algorithm protocol.

2.6 Membrane-protein complex system setup.

The prepared docked complex structures were embedded in palmitoyl-oleoyl-phosphatidylcholine (POPC) lipid bilayer using the CHARMM-GUI membrane builder (http://www.charmm-gui.org/?doc=input/membrane.bilayer) (Jo et al., 2008). The orientations of the complexes were aligned on the CCR5 transmembrane helices obtained from the Orientation of Protein Membranes (OPM) server (Lomize et al., 2006). The complexes were then neutralised and solvated with 0.15M of KCl counterions in explicit solvent using the TIP3P water model (Jorgensen et al., 1983). The dimension of the final systems approximately measured 86 x 86 x 140 $Å^3$, consisting of about 60 potassium ions, 71 chloride ions, 180 lipids and 22500 water molecules.

The charmmlipid2amber.py script was used in the renaming of lipid residues according to the Amber lipid14 force field. The Lipid14 force field (Dickson et al., 2014) was used to describe the lipids. The general AMBER Force Field (GAFF) (Sprenger et al., 2015) was utilised in parameterising the inhibitors via the ANTECHAMBER (Wang et al., 2001) and LEaP module of AMBER 18. The coordinate and topology files for the simulations were generated with *tleap* software using the ff14SB, TIP3P, and lipid14 forcefields.

2.7 Molecular Dynamics (MD) Simulation Protocol

Molecular dynamic (MD) simulations were performed on the selected ligand-CCR5 complexes in an ionised explicit POPC lipid bilayer performed using the Particle Mesh Ewald Molecular Dynamics (PMEMD) on the graphical processing unit (GPU) implemented in AMBER 18(Case et al., 2005).

An initial minimisation of 6000 steps consisting of 3000 steps of steepest descent and 3000 steps of conjugate gradient method with harmonic restrain potential of 10 $kcal.mol^{-1}.Å^{-2}$ on the protein and lipids. The entire system was relaxed in a second minimisation step of 15000 cycles. The systems were heated from 0 to 100 K using the Langevin thermostat in an isothermal-isochoric (NVT) ensemble over 30 ps with 10.0 $kcal \cdot mol^{-1} \cdot Å^{-2}$ harmonic restraints on the non-hydrogen atoms of protein, lipid, and ligand with a collision frequency of 1.0 ps^{-1}.

The system was further heated to 310 K over 260 ps in an isothermal-isobaric (NPT) ensemble with an anisotropic pressure coupling and a pressure of one bar. Additional equilibration at 310 K was performed with harmonic restraints on the protein and ligand beginning at 5.0 kcal·mol^{-1}·Å^{-2}. This was decreased by 1.0 kcal·mol^{-1}·Å^{-2} in every 4 ns, for a total of 20 ns of additional restrained equilibration under the NTP ensemble. A subsequent unrestrained 5 ns equilibration was performed before production.

Finally, 100 ns unrestrained production at 310 K was run under NPT ensemble with SHAKE constraints for bonds with hydrogen. The anisotropic pressure coupling with a pressure relaxation time of 1.0 ps was used to control the systems' pressure. Long-range electrostatic interactions were treated with the particle mesh Ewald (Darden et al., 1993) under periodic boundary conditions with nonbonded interactions cut-off of 12 Å.

2.7.1 Analysis of MD simulations trajectories

Post-MD analysis such as the radius of gyration (rGyr), root mean square deviation (RMSD), root mean square fluctuation (RMSF), intermolecular hydrogen bond analysis was computed using the CPPTRAJ module (Roe & Cheatham, 2013) of the Amber18 suite. All visualisations and plots were respectively performed with UCSF Chimera molecular modelling tool and Origin data analysis software version 6 (http://www.originlab.com) (Seifert, 2014).

3. Results and Discussion

3.1 Structure-based Virtual High throughput Screening

Virtual high throughput screening of databases for novel inhibitors is a useful approach to identify potential hit-to-lead candidates for further biological testing with a higher likelihood of blocking or triggering the activity of a drug target (Jin et al., 2020; Hughes et al., 2019; Rodríguez et al., 2014). The AutoDock Vina was initially used to dock the antiviral compounds at the CCR5 binding pocket due to its high scoring function with improved speed and docking accuracy (Trott & Olson, 2010). We further performed serial docking on high-ranking molecules using the virtual screening workflow (VSW) in Maestro (Schrödinger Release 2019-4, 2019). The Schrodinger VSW, which uses the grid-based ligand docking energetics (Glide) algorithm, was further used to screen down the number of ligands. The Glide docking program incorporates high throughput virtual screening (HTVS), standard precision (SP), and extra precision (XP)

docking protocols (Friesner et al., 2006; Friesner et al., 2004). These tools are ranking filters that try to locate the precise positions of ligands in the binding pocket of CCR5 as well as predict an accurate ligand pose in the binding pocket. The top four (4) compounds with the lowest binding scores and a good binding pose compared with Maraviroc were selected for further analysis (Table **1**). Also, Table **S1** and Table **S3** present the chemical structures and binding scores, respectively, for the remaining 19 compounds that can further be optimised/tested toward CCR5. The 2D structures of the best four hits and Maraviroc are presented in Fig. **2,** and the other 19 compounds shown in Table **S1**.

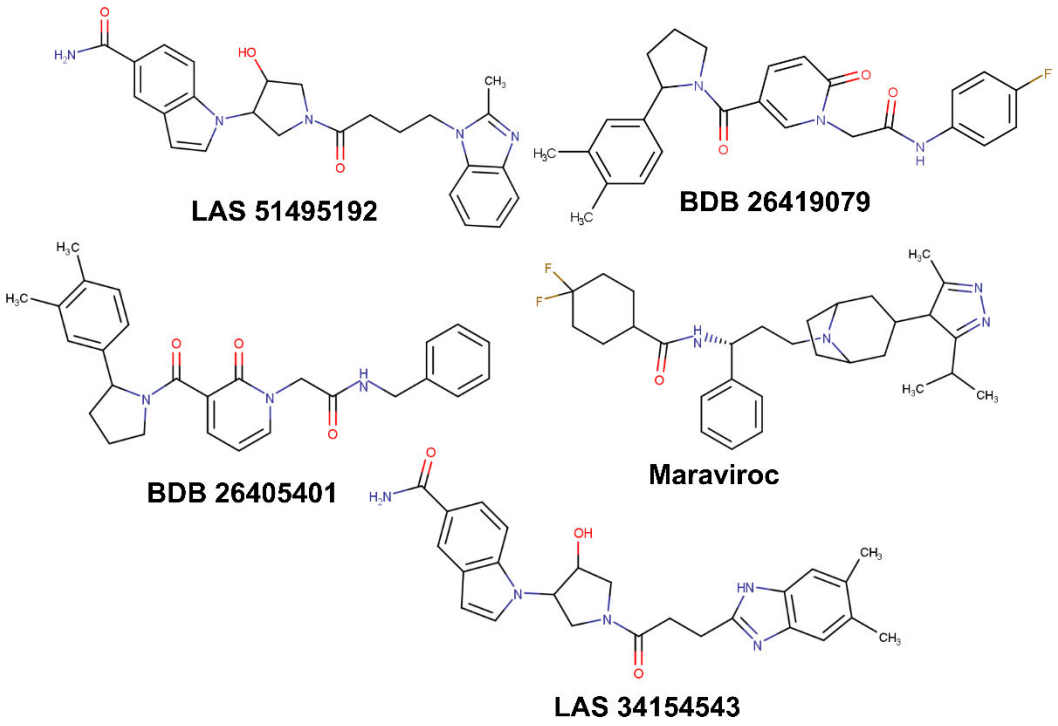


Fig. 2. The selected top four compounds following Autodock Vina, Glide HTVS, Glide SP and Glide XP docking and the reference drug (Maraviroc) used for molecular dynamics simulation (Image prepared by author).

The docking scores of selected four best-hit compounds (LAS 51495192, BDB 26405401, BDB 26419079, LAS 34154543) ranged from -10.0 kcal/mol to -10.5 kcal/mol for Autodock Vina and from -6.91 kcal/mol to -7.63 kcal/mol for Glide XP scores. These binding energies were higher than that of Maraviroc (Autodock Vina score = 9.0 kcal/mol: Glide XP score = 6.73 kcal/mol) (Table **1**). The docking results of the remaining 19 compounds that exhibited higher docking scores over Maraviroc are also presented in Supplementary Table **S3**.

Table 1. Summary of binding energy scores obtained from Autodock Vina and Glide XP for the identified top-four compounds.

Entry Number	Asinex Compound ID	Autodock Score (kcal/mol)	Glide XP Score (kcal/mol)
8356	LAS 51495192	-10.5	-7.63
287	BDB 26405401	-10.3	-7.30
579	BDB 26419079	-10.0	-7.20
7073	LAS 34154543	-10.3	-6.91
Maraviroc	N/A	-9.0	-6.73

3.2 Molecular interactions and binding modes of best-hits compounds relative to Maraviroc

To obtain insights on the binding interactions between the top four compounds from Glide XP within the CCR5 receptor binding pocket, their complexes were explored through their 2D interaction maps (Fig. **3**) and 3D structures (Fig. **4**). The molecular interactions of the best four hits at the CCR5 binding site were assessed using Maestro protein-ligand interaction. The binding mode of the top four compounds (LAS 51495192, BDB 26405401, BDB 26419079, and LAS 34154543) relative to Maraviroc bind at a similar location of the active site overlapping with the gp120 V3-loop (Fig. **4F**).

The identified top four compounds were observed to engage in hydrogen bond interactions (Fig. **3** and Table **2**) with residues such as Glu283, Gln280, Thr105, Tyr37, Ser180, Trp190, Thr195, Asn258, Lys191, Tyr251, Tyr108, and Thr167 similar to Maraviroc with Glu283, Tyr37, Tyr251, Ser180, and Trp86. Additionally, the top four compounds were also involved in π-π stacking interactions with residues such as Trp86, Tyr89, Tyr108, Phe112, and Tyr251. Maraviroc interacts with Tyr108, Phe109 and Phe112 via π-π stacking. The residues such as Trp94$^{23.50}$, Trp86$^{2.60}$, Tyr108$^{3.32}$, Tyr251$^{6.51}$, and Trp248$^{6.48}$ have been suggested to form a potential binding cavity for gp120 and has also been described to be sensitive to mutations of these residues (Garcia-Perez et al., 2011). The full-length gp120 in complex with unmodified CCR5 receptor recently reported shows that the gp120 V3 loop makes extensive contacts with residues such as Tyr108$^{3.32}$, Phe109$^{3\ 33}$, Glu283$^{7.39}$, Tyr89$^{2.63}$, Tyr251$^{6\ 51}$, and Trp86$^{2.60}$ of the chemokine recognition site 2 (Shaik et al., 2019). The residue Arg313 of the gp120 V3 loop appears to be sandwiched between Glu283$^{7.39}$ and Tyr251$^{6.51}$ of CCR5(Shaik et al., 2019). The identified compounds made important contacts with residues critical for gp120 V3 loop interaction at the CCR5 binding site, such as Trp86$^{2.60}$, Tyr89$^{2.63}$, Phe109$^{3.33}$, Tyr108$^{3\ 32}$, Glu283$^{7.39}$ and Tyr251$^{6.51}$.

Table **2**. Dynamic hydrogen bond interactions of the potential inhibitors relative to Maraviroc at CCR5 binding site over the 100ns MD simulation.

Compounds	H-Acceptor	H-Donor	Occupancy (%)	Distance (Å)	Angle (°)
LAS 51495192	THR195-OG1	LIGAND-N5	24.83	2.8	152
	ASN258-OD1	LIGAND-N5	16.95	2.8	157
	LIGAND-N2	TYR37-OH	9.17	2.9	153
	LIGAND-O2	ASN258-ND2	5.38	2.8	155
	LIGAND-O2	LYS191-NZ	2.62	2.8	153
BDB 26405401	GLU283-OE2	LIGAND-N3	72.35	2.8	162
	LIGAND-O1	GLN280-NE2	7.21	2.8	153
	LIGAND-O3	SER180-OG	4.11	2.7	160
BDB 26419079	LIGAND-O1	TYR37-OH	58.41	2.7	163
	LIGAND-O3	TYR251-OH	35.22	2.7	163
	GLU283-OE1	LIG314-N3	14.67	2.7	156
	LIGAND-O3	TYR108-OH	10.04	2.8	159
LAS 34154543	THR167-OG1	LIGAND-N3	62.73	2.8	157
	LIGAND-O2	TYR-37-OH	43.16	2.7	161
	LIGAND-O2	GLN280-NE2	31.66	2.8	159
	LIGAND-N4	SER180-OG	16.71	2.8	154
	THR105-OG1	LIGAND-O3	7.64	2.7	159
	TYR37-OH	LIGAND-N5	3.19	2.9	145
Maraviroc	GLU283-OE1	LIGAND-N2	99.48	2.7	161
	LIGAND-N4	TYR37-OH	34.32	2.8	157
	TYR_251-OH	LIGAND-N1	9.17	2.9	159
	LIGAND-O1	SER180-OG	3.56	2.7	162

The superimposed conformation of Maraviroc, the gp120 V3 loop and the top four (4) docked compounds at the active site of CCR5 are shown in Fig. **4F**. Evidently, from Fig. **4A-F**, it can be observed that the identified compounds, the gp120 V3 loop and Maraviroc overlap at the binding pocket of CCR5.

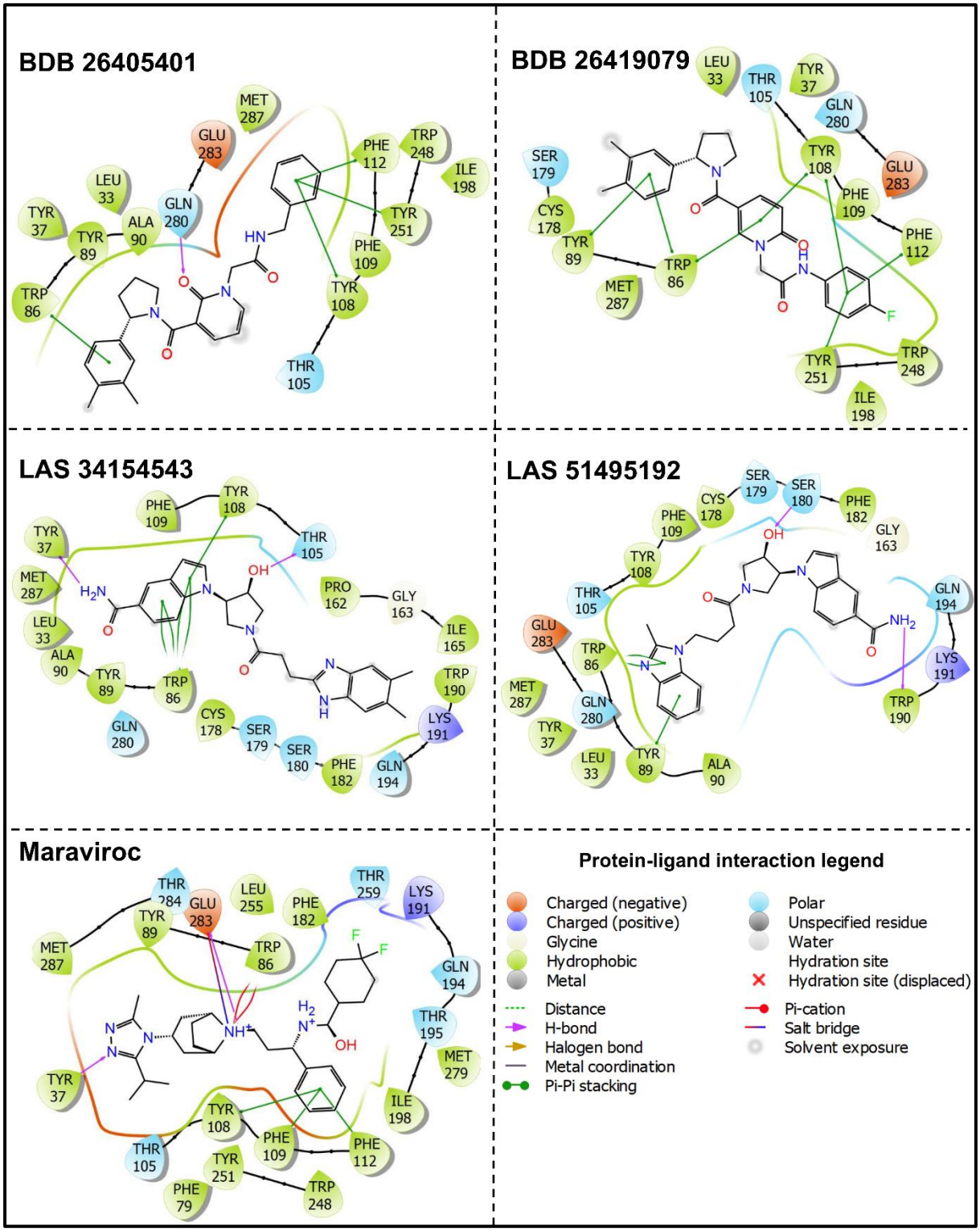


Fig. 3. The two-dimensional (2D) ligand-protein interactions diagram of Maraviroc and the best-hit compounds (LAS 51495192, BDB 26405401, BDB 26419079, and LAS 34154543) at CCR5 binding pocket (Image prepared by author).

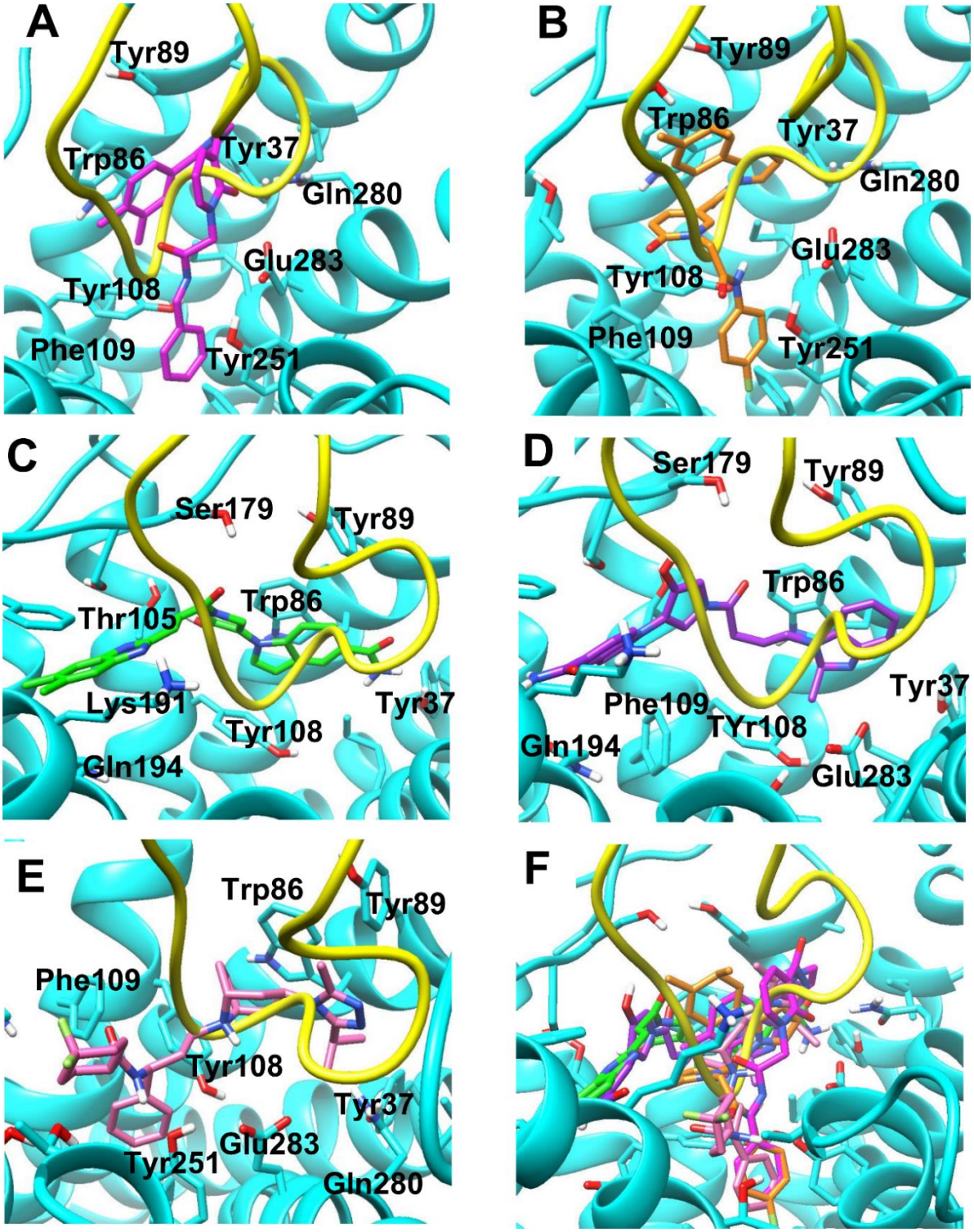


Fig. 4. Three-dimensional (3D) binding poses of the top four compounds and Maraviroc at CCR5 binding site overlapping with gp120 V3 loop (yellow). The binding poses of BDB 26405401 (**A**), BDB 26419079 (**B**), LAS 34154543 (**C**), LAS 51495192 (**D**), Maraviroc (**E**), Superposition of all top four compounds, Maraviroc and gp120 V3 loop (**F**) in the active site of CCR5 (Image prepared by author).

3.3 Conformational dynamic interaction patterns of the identified compounds revealed by molecular dynamics

Molecular dynamics simulations provide valuable information on ligand binding stability within the active site of a target as well as the dynamic behaviour of the receptor (Yang et al., 2011; Podder et al., 2016; Chaudhary & Aparoy, 2017). To obtain an ensemble understanding of the dynamic behaviour of the best-hit compounds relative to Maraviroc at CCR5, MD simulations were used to evaluate the stability of the CCR5-inhibitor complex interactions. The structural stability and residue fluctuations of the individual complex binding site were monitored for 100 ns.

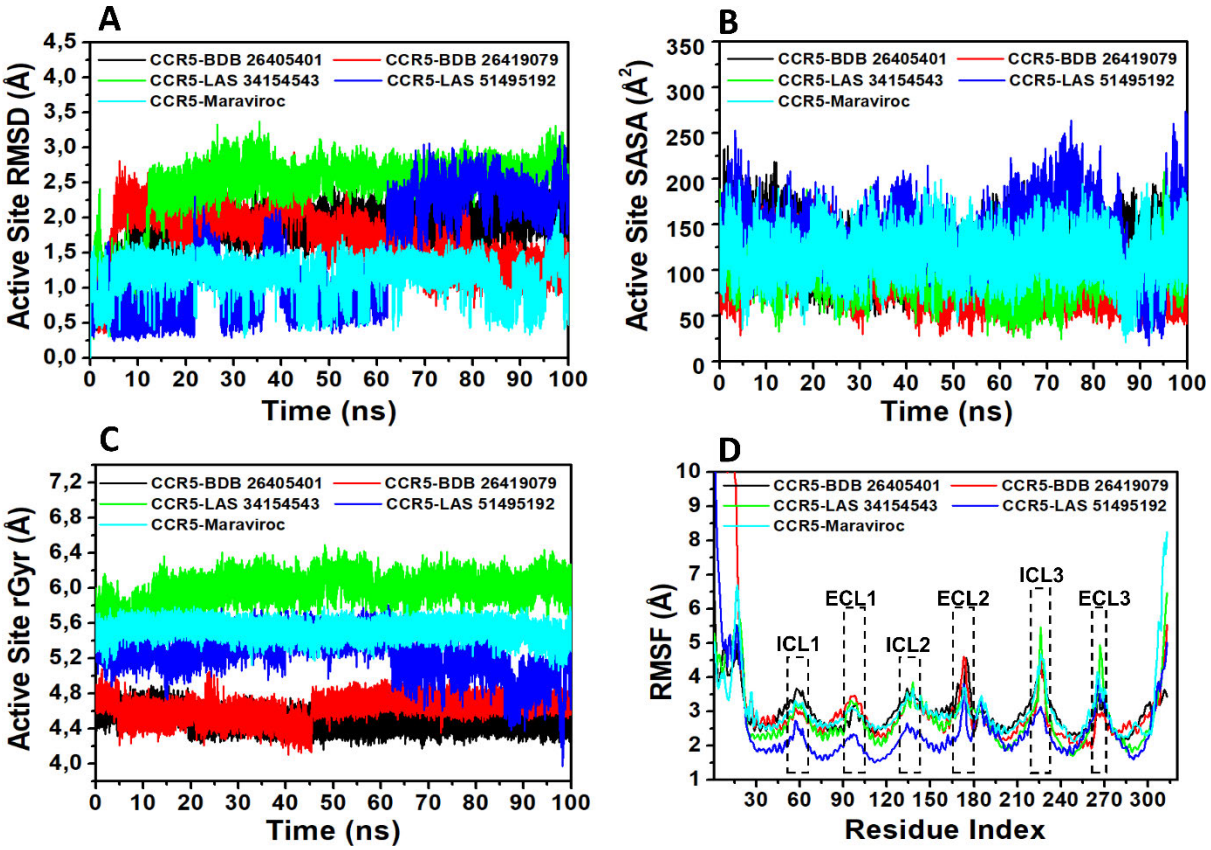


Fig. 5. The active site root mean square deviation (RMSD) (**A**), the active site solvent accessible surface area (SASA) (**B**), the binding site radius of gyration (rGyr) (**C**) and root mean square fluctuation (RMSF) (**D**) as a function of the 100 ns simulation time (Image prepared by author).

The RMSD plot for the active site showed an initial conformational fluctuation and stabilising with average values of 1.42 Å, 1.81 Å, 1.68 Å, 2.42 Å, and 1.16 Å for LAS 51495192, BDB 26405401, BDB 26419079, LAS 34154543, and Maraviroc, respectively (Fig. **5A** and Table **S5**). Except for compound LAS 34154543, all the compounds and Maraviroc displayed

averaged RMSD below 2 Å. The degree of protein binding site volume expansion of each complex system was subsequently assessed by computing the solvent accessible surface area (SASA) from individual MD trajectories (Fig. **5B**) and their average SASA values provided in Table **S5**. The SASA values of CCR5-BDB26419079 (86.37 $Å^2$) and CCR5-LAS34154543 (95.38 $Å^2$) complexes were lower than CCR5-Maraviroc (113.66 $Å^2$) system, whereas CCR5-BDB26405401 (119.92 $Å^2$) and CCR5-LAS51495192 (136.05 $Å^2$) were slightly higher than that of Maraviroc system.

The radius of gyration (rGyr), which shows the compactness of a biomolecular system, was computed within 8 Å of the ligand binding site (Fig. **5C** and Table **S5)**. Apart from LAS 34154543 complex system (5.98 Å), the remaining three compounds had relatively lower rGyr values (4.50 Å–5.25 Å) compared to Maraviroc complex system (5.52 Å). To further probe the receptor residue mobility of CCR5 during the simulation, the RMSF of the individual CCR5 residues were monitored for each complex system. As presented in Fig. **5D**, compounds LAS 51495192 and LAS 34154543 displayed lower residue fluctuation than BDB 26405401 and BDB 26419079, which showed similar residue mobility patterns like Maraviroc. Additionally, essential residues such as Tyr37, Trp86, Tyr89, Thr105, Tyr108, Phe109, Phe112, Ser180, Trp190, Tyr251, and Glu283 showed very low mobility. High residue mobilities were also observed in all inhibitors at the intracellular loop (ICL) and extracellular loop (ECL) regions compared to the transmembrane domains.

Additionally, the stability of the best-hits and Maraviroc complexes were observed throughout the simulation process for the averaged volume, pressure, temperature, the total energy (ETOT), and potential energy (EPtot) values and presented in Table **S6.** The overall simulation quality shows that these parameters were stable over the 100 ns simulation (Fig. **S1.**). Noticeably, the best-hits were observed to relatively have lower total energy and potential energy compared to Maraviroc. In contrast, the other parameters were observed to be in a similar range (Table **S6**).

3.4 *In silico* evaluation of molecular properties and pharmacokinetics of best hits

To gain insight into the drug-likeness (molecular properties) and potential pharmacokinetic properties of the final hits, a comparative *in silico* analysis was performed. The physiochemical properties of the top four compounds and the additional 19 compounds were also analysed and presented in Table **3** and Table **S4,** respectively. The *in silico* physicochemical properties and pharmacokinetic profiling were carried out using SwissADME (Daina et al., 2017) and pkCSM-pharmacokinetics (Pires et al., 2015), respectively.

An increase in the molecular weight of a therapeutic compound has been reported to decrease the amount of concentration at the intestinal epithelium surface and thus reduces absorption (Renukuntla et al., 2013). It has equally been suggested that this could obstruct bilayer membrane passive diffusion of the compound in question. Therefore, it is recommended that prospective drug candidates possess a molecular weight less than 500 g/mol in line with Lipinski RO5 (Lipinski et al.; Omran & Rauch, 2014). The final retrieved compounds were suitable as potential lead molecules as they fell within the acceptable range for a drug-like molecule (Table **3** and Table **S4**). The retrieved compounds outscored Maraviroc in terms of molecular weight, suggesting they could exhibit good intestinal absorption and cellular uptake. An additional step to examine the effect of molecular weight on bioactivity was the number of rotatable bonds, as it is an excellent descriptor of molecular flexibility (Lu et al., 2004). A report by Veber et al. (Veber et al., 2002) revealed that the number of rotatable bonds increases with an increase in molecular weight. The Lipinski RO5 suggests that a drug-like compound should have a rotatable bond count of less than 10; this promotes absorption and bioavailability (Lu et al., 2004; Di & Kerns, 2016a; Vallianatou et al., 2015). Maraviroc had nine rotatable bond count from our estimation, which was above those obtained for the new compounds. This further put the retrieved compounds as potential candidates relative to the reference drug.

Table 3. The physicochemical properties of the top four hit compound relative to Maraviroc.

Physicochemical Properties	**Compounds**				
	LAS 51495192	**BDB 26405401**	**BDB 26419079**	**LAS 34154543**	**Maraviroc**
Formula	C25H27N5O3	C27H29N3O3	C26H26FN3O3	C25H27N5O3	C29H41F2N5O
MW (g/mol)	445.51	443.54	447.50	445.51	513.67
MLog $P_{o/w}$	1.42	3.02	3.45	1.42	4.15
LogS (Ali) (mol/L)	-3.18 Soluble	-4.83 Moderately soluble	-4.50 Moderately soluble	-3.80 Soluble	-6.22 Poorly soluble
TPSA (A^2)	106.38	71.41	71.41	117.24	63.05
Molar Refractivity	130.04	132.80	129.49	130.27	145.84
HBA	4	3	4	4	6
HBD	2	1	1	3	1
Rotatable bonds	7	8	7	6	9
Lipinski violations	No; 0	No; 0	No; 0	No; 0	Yes; 2

HBD = number of hydrogen bond donors (HBD ≤ 5); HBA = number of hydrogen bond donors (HBD ≤ 10); TPSA = total polar surface area (≤140); MLOGP = predicted octanol/water partition coefficient (MLOGP < 4.15); MW = molecular weight (g/mol) (MW ≤ 500); Rotatable bonds (≤ 8); TPSA = Topological polar surface area (≤ 140)

LogP measures the hydrophilicity of chemical compounds, which is the logarithm of the coefficient of their permeation across n-octanol and water (Octanol/Water) (Lipinski, 2000). Hence, a high MLogP value indicates a reduction in aqueous solubility, leading to a decrease in absorption. All the top four hits had MLogP lower than Maraviroc (Table **3)**. Besides compound BDG 51127568 in Table **S4**, all the other compounds displayed lower MLogP values than Maraviroc.

Also, LogS was employed to estimate the aqueous solubility of the retrieved compounds relative to the approved drug. The LogS, in addition to contributing to the determination of membrane permeability, equally provide an idea of the oral bioavailability of drugs (Wang & Hou, 2011). A minimum acceptable range of 0 to -6 for aqueous solubility accounts for 95 % of the existing drugs (Wang & Hou, 2011). The obtained values in Table **3** revealed that all the chemical compounds fall within the acceptable range for being optimally absorbed in the body. The retrieved compound solubility ranged from moderately soluble to soluble, whereas poor solubility was attributed to Maraviroc. Also, except compound BDG 51127568, the remaining compounds in Table **S4** had good LogS values compared to Maraviroc.

The topological polar surface area (TPSA) puts together the surface polar atoms, most notably the oxygen and nitrogen, in connection with the hydrogen atoms that they are attached. It aids in predicting the ability of a chemical compound to pass through the cells, where a low TPSA score indicates the compound can permeate the cell (Ertl et al., 2000; Prasanna & Doerksen, 2008). It also provides a picture of the molecular size and volume, which regulates its physiological transport across the lipid bilayer, which includes the gastrointestinal tract (GIT) and blood-brain barrier (BBB) (Shityakov et al., 2013). High TPSA has been posited to disrupt the transportability of drug candidates, which in turn impact the bioactivity of these drugs (Daga et al., 2018; Di & Kerns, 2016a). From the indications in Table **3**, Maraviroc had the least TPSA; however, all the evaluated top four and the 19 compounds were within the acceptable range of TPSA of $\leq$140 $Å^2$. The compounds BDF 33909572, BDG 34130390, BDF 34027559, BDG 33694125, BDG 51127568, BDG 34042546, BDG 33691535, BDF 34035988, BDG 33694314, and BDF 33909571 in Table **S4** had TPSA values lower than Maraviroc.

Hydrogen bond has been suggested to affect the solubility of therapeutic compounds since they must get broken to facilitate the transportation of these compounds across the lipid bilayer membrane (Di & Kerns, 2016b; Sun et al., 2018). Thus, a high number of hydrogen bond influences permeation by passive diffusion due to the decline in partitioning from the aqueous region into the lipid bilayer membrane. Hydrogen bonding is a unique indicator of the number

of hydrogen bond donors (HBD) and hydrogen bond acceptors (HBA) in a molecule. This has been extensively employed in the estimation of drug-likeness of a chemical compound. The Lipinski RO5 states that a drug-like compound should have HBD count of ≤5 and HBA of ≤10; this is the basis for oral bioavailability and activity(Lipinski et al.; Lipinski, 2000). The calculation in Table **3** and Table **S4** indicate that all the retrieved top compounds passed the score for adequate hydrogen bond donation and acceptance.

The ADMET descriptors' profiling was determined using the pkCSM-pharmacokinetics algorithm protocol (Pires et al., 2015). The predictive factors for drug absorption include human intestinal absorption, skin permeability levels, water solubility and Caco-2 permeability. The factors affecting the distribution of a drug include the volume of distribution (VDss), CNS permeability, and the blood-brain barrier (logBB). The predictions of drug metabolism are based on the inhibitor or substrate CYP models (CYP3A4, CYP2D6, CYP1A2, CYP2C19, CYP2D6, CYP3A4, and CYP2C9). Also, the renal Organic Cation Transporter 2 (OCT2) substrate and the total clearance model were used to predict the excretion of the potential drug. The estimation of potential drug toxicity was based on models such as skin sensitisation, AMES toxicity, and Oral rat acute toxicity. These predictive parameters were estimated for the top four compounds and the reference drug Maraviroc (Table **4**).

Regarding the parameters for absorption, the percentage of human intestinal absorption (HIA) were higher for all the top four compounds (88.7 – 93.6%) except for LAS 34154543 (80.3 %) when compared to Maraviroc (88.2%). Also, optimal Caco-2 permeability, which is a predictive indicator for the absorption of orally administered drugs (> 0.9 suggests high Caco-2 permeability) were observed for the selected compounds (~0.9 – 1.2), which outscored Maraviroc (0.7) (Pires et al., 2015). The predictive absorption indicators of selected compounds demonstrate their potential for good oral bioavailability.

The volume of distribution in human (VDss) is a predictive parameter that describes the extent of drug distribution (VDss < -0.15 and VDss > 0.45 denote low and high distribution, respectively). The compounds LAS 51495192, BDB 26405401, and BDB 26419079 shows tolerable distribution, whereas LAS 34154543 and Maraviroc depicted high distribution in tissue.

Table 4. *In silico* pharmacokinetics prediction of the newly identified compounds relative to Maraviroc.

Pharmacokinetics Properties	Compounds				
	LAS 51495192	BDB 26405401	BDB 26419079	LAS 34154543	Maraviroc
Absorption					
Water solubility (Log mol/L)	-3.06	-4.95	-4.74	-3.07	-4.29
Caco2 permeability (log Papp in 10^{-6} cm/s)	0.99	1.22	1.19	0.88	0.74
Human intestinal absorption (% Absorbed)	88.72	93.55	92.69	80.26	88.15
Distribution					
Blood Brain Barrier Permeability (log BB)	-0.68	-0.03	-0.25	-0.84	0.19
VDss (human) (log L/kg)	-0.06	0.02	0.06	0.49	1.40
CNS permeability (log PS)	-2.71	-2.12	-2.11	-2.76	-1.96
Metabolism					
CYP2D6 substrate	Yes	No	No	No	Yes
CYP3A4 substrate	No	Yes	Yes	Yes	Yes
CYP1A2 inhibitor	No	No	No	Yes	No
CYP2C19 inhibitor	Yes	Yes	Yes	Yes	No
CYP 2C9 inhibitor	Yes	Yes	Yes	Yes	No
CYP 2D6 inhibitor	No	No	No	No	No
CYP 3A4 inhibitor	No	Yes	Yes	No	Yes
Excretion					
Total Clearance (log ml/min/kg)	1.09	1.21	0.289	1.19	0.64
Renal OCT2 substrate	Yes	No	No	Yes	Yes
Toxicity					
AMES toxicity	No	No	No	Yes	No
Oral rat acute toxicity (LD_{50}) (mol/kg)	2.51	2.52	2.59	2.20	2.81
Skin Sensitisation	No	No	No	No	No

It can also be observed from Table **4** that two of the identified compounds (BDB 26405401 and BDB 26419079) and Maraviroc were predicted to inhibit the critical enzyme of cytochrome P450 responsible for drug metabolism (CYP3A4 isoform). The selected compounds were also within acceptable/tolerable range for central nervous system (CNS) penetration. The predicted values of the total clearance (Table **4**), which assess the efficiency of a potential drug to be eliminated from the body, suggests that the reference drug and all the top four compounds possess good renal elimination. Finally, the predictive AMES toxicity test posits toxicity for the compounds LAS 34154543, whereas the rest and the reference drug passed the test. Additionally, all the compounds were predicted to display no skin sensitisation with good oral rat acute toxicity. The various physicochemical and pharmacokinetic predictions (Table **3** and Table **4)** suggest that the selected compounds could be tested or further manipulated as potential candidates toward the CCR5 receptor as HIV-1 entry inhibitors.

4. Conclusion

CCR5 inhibitors have been demonstrated to display great potential strides in HIV-1 infection treatment. However, the risk of life-threatening adverse effects, such as allergic reaction, skin reaction, liver and damage heart attack is caused by the FDA approved drug, Maraviroc (FDA, 2007; Peng et al., 2018). Hence, identifying novel potential inhibitors of CCR5 with new chemical scaffolds, relatively less adverse effects and displaying improved binding affinity are highly desired. The structure of the full-length CCR5 is significantly beneficial in the discovery and designing of potential lead candidates as HIV-1 entry inhibitors. Structure-based virtual screening remains an integral approach in the identification of new chemical scaffolds against the active site of a known protein target.

In this study, structure-based virtual screening of the Asinex antiviral database and further analyses lead to the identification of novel potential CCR5 inhibitors. In summary, the Asinex antiviral database was screened against CCR5. The applied *in silico* approaches identified compounds LAS 51495192, BDB 26405401, BDB 26419079, and LAS 34154543 with binding scores higher than Maraviroc. These compounds were also observed to make interactions with residues critical for gp120 V3 loop binding. Additional 19 compounds with similar or improved physiochemical properties and higher binding scores compared with Maraviroc that could be further optimised have also been reported. Furthermore, *in silico* pharmacokinetics and physiochemical estimations also indicated the predicted values of the physicochemical descriptors were within the acceptable range for drug-likeness, therefore suggesting positive indications that the identified potential molecules are promising drug-like entities. The successful application of multidisciplinary computational drug discovery approaches has allowed the identification of diverse potential compounds laying the groundwork for experimental exploration of the suggested compounds. Further optimisations and testing of these compounds may assist in the discovery of effective HIV-1 therapy.

Ethics Approval and Consent to Participate

Not applicable

Human and Animal Rights

No humans and animals were used in the study.

CONSENT FOR PUBLICATION

Not applicable

Availability of Data and Materials

Not applicable

Funding

None

Conflict of Interest

The authors declare no conflicts of interest in this work.

Acknowledgements

The authors would like to acknowledge the Centre for High Performance Computing (CHPC), Cape Town, South Africa, for providing access to Schrödinger Maestro v12 molecular modelling suite.

Authors Contribution

P. Appiah-Kubi: Conceptualization and study design, screening and simulation experiment, data analysis, and manuscript writing. E.A. Iwuchukwu: Data analysis, manuscript writing. M.E.S Soliman: Supervision and review.

Reference

ASINEX. 2020. ASINEX database. *ASINEX Corporation*.

Case, D.A., Cheatham, T.E., Darden, T., Gohlke, H., Luo, R., Merz, K.M., Onufriev, A., Simmerling, C., Wang, B. & Woods, R.J. 2005. The Amber biomolecular simulation programs. *Journal of Computational Chemistry*, 26(16): 1668–1688.

Chaudhary, N. & Aparoy, P. 2017. Deciphering the mechanism behind the varied binding activities of COXIBs through Molecular Dynamic Simulations, MM-PBSA binding energy calculations and per-residue energy decomposition studies. *Journal of biomolecular structure & dynamics*, 35(4): 868–882.

Daga, P.R., Bolger, M.B., Haworth, I.S., Clark, R.D. & Martin, E.J. 2018. Physiologically Based Pharmacokinetic Modeling in Lead Optimization. 2. Rational Bioavailability Design by Global Sensitivity Analysis to Identify Properties Affecting Bioavailability. *Molecular Pharmaceutics*, 15(3): 831–839.

Daina, A., Michielin, O. & Zoete, V. 2017. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific reports*, 7: 42717.

Darden, T., York, D. & Pedersen, L. 1993. Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. *The Journal of chemical physics*, 98(12): 10089–10092.

Debnath, A.K. 2003. Generation of predictive pharmacophore models for CCR5 antagonists: study with piperidine-and piperazine-based compounds as a new class of HIV-1 entry

inhibitors. *Journal of medicinal chemistry*, 46(21): 4501–4515.
Di, L. & Kerns, E.H. 2016a. Benefits of Property Assessment and Good Drug-Like Properties. In *Drug-Like Properties*.
Di, L. & Kerns, E.H. 2016b. Prediction Rules for Rapid Property Profiling from Structure. In *Drug-Like Properties*.
Dickson, C.J., Madej, B.D., Skjevik, Å.A., Betz, R.M., Teigen, K., Gould, I.R. & Walker, R.C. 2014. Lipid14: the amber lipid force field. *Journal of chemical theory and computation*, 10(2): 865–879.
Ertl, P., Rohde, B. & Selzer, P. 2000. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *Journal of Medicinal Chemistry*, 43(20): 3714–3717.
FDA, U.S. 2007. FDA notifications. Maraviroc approved as a CCR5 co-receptor antagonist. *AIDS Alert*, 22: 103.
Ferreira, L.G., Dos Santos, R.N., Oliva, G. & Andricopulo, A.D. 2015. Molecular docking and structure-based drug design strategies. *Molecules*, 20(7): 13384–13421.
Forli, S. 2010. Raccoon| AutoDock VS: an automated tool for preparing AutoDock virtual screenings.
Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., Repasky, M.P., Knoll, E.H., Shelley, M. & Perry, J.K. 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry*, 47(7): 1739–1749.
Friesner, R.A., Murphy, R.B., Repasky, M.P., Frye, L.L., Greenwood, J.R., Halgren, T.A., Sanschagrin, P.C. & Mainz, D.T. 2006. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein− ligand complexes. *Journal of medicinal chemistry*, 49(21): 6177–6196.
Garcia-Perez, J., Rueda, P., Alcami, J., Rognan, D., Arenzana-Seisdedos, F., Lagane, B. & Kellenberger, E. 2011. Allosteric model of maraviroc binding to CC chemokine receptor 5 (CCR5). *Journal of Biological Chemistry*, 286(38): 33409–33421.
Greenwood, J.R., Calkins, D., Sullivan, A.P. & Shelley, J.C. 2010. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. *Journal of computer-aided molecular design*, 24(6–7): 591–604.
Gulick, R.M., Su, Z., Flexner, C., Hughes, M.D., Skolnik, P.R., Wilkin, T.J., Gross, R., Krambrink, A., Coakley, E., Greaves, W.L., Zolopa, A., Reichman, R., Godfrey, C., Hirsch, M. & Kuritzkes, D.R. 2007. Phase 2 Study of the Safety and Efficacy of Vicriviroc, a CCR5 Inhibitor, in HIV-1–Infected, Treatment-Experienced Patients: AIDS Clinical Trials Group 5211. *The Journal of Infectious Diseases*.
Hughes, T.E.T., Del Rosario, J.S., Kapoor, A., Yazici, A.T., Yudin, Y., Fluck III, E.C., Filizola, M., Rohacs, T. & Moiseenkova-Bell, V.Y. 2019. Structure-based characterization of novel TRPV5 inhibitors. *Elife*, 8: e49572.
Jin, Z., Du, X., Xu, Y., Deng, Y., Liu, M., Zhao, Y., Zhang, B., Li, X., Zhang, L. & Peng, C. 2020. Structure of M pro from SARS-CoV-2 and discovery of its inhibitors. *Nature*: 1–5.
Jo, S., Kim, T., Iyer, V.G. & Im, W. 2008. CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of computational chemistry*, 29(11): 1859–1865.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. & Klein, M.L. 1983. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2): 926–935.

Klibanov, O.M., Williams, S.H. & Iler, C.A. 2010. Cenicriviroc, an orally active CCR5 antagonist for the potential treatment of HIV infection. *Curr Opin Investig Drugs*, 11(8): 940–950.

Lin, H.-Y., Ho, Y. & Liu, H.-L. 2019. Structure-Based Pharmacophore Modeling to Discover Novel CCR5 Inhibitors for HIV-1/Cancers Therapy. *Journal of Biomedical Science and Engineering*, 12(1): 10–30.

Lipinski, C.A. 2000. Drug-like properties and the causes of poor solubility and poor permeability. *Journal of Pharmacological and Toxicological Methods*, 44(1): 235–249.

Lipinski, C.A., Lombardo, F., Dominy, B.W. & Feeney, P.J. *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings*. Elsevier Science B.V.

Llanos, G.G., Araujo, L.M., Jiménez, I.A., Moujir, L.M., Rodríguez, J., Jiménez, C. & Bazzocchi, I.L. 2017. Structure-based design, synthesis, and biological evaluation of withaferin A-analogues as potent apoptotic inducers. *European Journal of Medicinal Chemistry*, 140: 52–64.

Lomize, M.A., Lomize, A.L., Pogozheva, I.D. & Mosberg, H.I. 2006. OPM: orientations of proteins in membranes database. *Bioinformatics (Oxford, England)*, 22(5): 623–625.

Lu, J.J., Crimin, K., Goodwin, J.T., Crivori, P., Orrenius, C., Xing, L., Tandler, P.J., Vidmar, T.J., Amore, B.M., Wilson, A.G.E., Stouten, P.F.W. & Burton, P.S. 2004. Influence of molecular flexibility and polar surface area metrics on oral bioavailability in the rat. *Journal of Medicinal Chemistry*, 47(24): 6104–6107.

Macalino, S.J.Y., Gosu, V., Hong, S. & Choi, S. 2015. Role of computer-aided drug design in modern drug discovery. *Archives of Pharmacal Research*, 38(9): 1686–1701.

Mandal, S., Moudgil, M. & Mandal, S.K. 2009. Rational drug design. *European Journal of Pharmacology*, 625(1–3): 90–100.

Meanwell, N.A. 2011. Improving drug candidates by design: a focus on physicochemical properties as a means of improving compound disposition and safety. *Chemical research in toxicology*, 24(9): 1420–1456.

Nichols, W.G., Steel, H.M., Bonny, T., Adkison, K., Curtis, L., Millard, J., Kabeya, K. & Clumeck, N. 2008. Hepatotoxicity observed in clinical trials of aplaviroc (GW873140). *Antimicrobial Agents and Chemotherapy*, 52(3): 858–865.

Omran, Z. & Rauch, C. 2014. Acid-mediated Lipinski's second rule: Application to drug design and targeting in cancer. *European Biophysics Journal*, 43(4–5): 199–206.

Oppermann, M. 2004. Chemokine receptor CCR5: insights into structure, function, and regulation. *Cellular signalling*, 16(11): 1201–1210.

Peng, P., Chen, H., Zhu, Y., Wang, Z., Li, J., Luo, R.-H., Wang, J., Chen, L., Yang, L.-M. & Jiang, H. 2018. Structure-Based Design of 1-Heteroaryl-1, 3-propanediamine Derivatives as a Novel Series of CC-Chemokine Receptor 5 Antagonists. *Journal of medicinal chemistry*, 61(21): 9621–9636.

Pereira, A.B., Rezende, N.A. de, Teixeira Junior, A.L., Teixeira, M.M. & Simões e Silva, A.C. 2009. Citocinas e quimiocinas no transplante renal. *Jornal Brasileiro de Nefrologia*,

31(4): 286–296.

Pérez-Nueno, V.I., Ritchie, D.W., Borrell, J.I. & Teixido, J. 2008. Clustering and classifying diverse HIV entry inhibitors using a novel consensus shape-based virtual screening approach: further evidence for multiple binding sites within the CCR5 extracellular pocket. *Journal of chemical information and modeling*, 48(11): 2146–2165.

Pires, D.E. V, Blundell, T.L. & Ascher, D.B. 2015. pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *Journal of medicinal chemistry*, 58(9): 4066–4072.

Podder, A., Pandey, D. & Latha, N. 2016. Investigating the structural impact of S311C mutation in DRD2 receptor by molecular dynamics & docking studies. *Biochimie*, 123: 52–64.

Prasanna, S. & Doerksen, R. 2008. Topological Polar Surface Area: A Useful Descriptor in 2D-QSAR. *Current Medicinal Chemistry*, 16(1): 21–41.

Renukuntla, J., Vadlapudi, A.D., Patel, A., Boddu, S.H.S. & Mitra, A.K. 2013. Approaches for enhancing oral bioavailability of peptides and proteins. *International Journal of Pharmaceutics*, 447(1–2): 75–93.

Rodríguez, D., Brea, J., Loza, M.I. & Carlsson, J. 2014. Structure-based discovery of selective serotonin 5-HT1B receptor ligands. *Structure*, 22(8): 1140–1151.

Roe, D.R. & Cheatham, T.E. 2013. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation*, 9(7): 3084–3095.

Schrödinger Release 2019-4. 2019. Maestro. https://www.schrodinger.com/Maestro.

Schürmann, D., Fätkenheuer, G., Reynes, J., Michelet, C., Raffi, F., Van Lier, J., Caceres, M., Keung, A., Sansone-Parsons, A., Dunkle, L.M. & Hoffmann, C. 2007. Antiviral activity, pharmacokinetics and safety of vicriviroc, an oral CCR5 antagonist, during 14-day monotherapy in HIV-infected adults. *AIDS*, 21(10): 1293–1299.

Seifert, E. 2014. OriginPro 9.1: Scientific Data Analysis and Graphing Software—Software Review. *Journal of Chemical Information and Modeling*, 54(5): 1552–1552.

Shaik, M.M., Peng, H., Lu, J., Rits-Volloch, S., Xu, C., Liao, M. & Chen, B. 2019. Structural basis of coreceptor recognition by HIV-1 envelope spike. *Nature*, 565(7739): 318.

Shityakov, S., Neuhaus, W., Dandekar, T. & Förster, C. 2013. Analysing molecular polar surface descriptors to predict blood-brain barrier permeation. *International Journal of Computational Biology and Drug Design*, 6(1–2): 146–156.

Spagnolo, P., Renzoni, E.A., Wells, A.U., Copley, S.J., Desai, S.R., Sato, H., Grutters, J.C., Abdallah, A., Taegtmeyer, A., Du Bois, R.M. & Welsh, K.I. 2005. C-C chemokine receptor 5 gene variants in relation to lung disease in sarcoidosis. *American Journal of Respiratory and Critical Care Medicine*, 172(6): 721–728.

Sprenger, K.G., Jaeger, V.W. & Pfaendtner, J. 2015. The General AMBER Force Field (GAFF) Can Accurately Predict Thermodynamic and Transport Properties of Many Ionic Liquids. *The Journal of Physical Chemistry B*, 119(18): 5882–5895.

Sun, R., Han, Y., Swanson, J.M.J., Tan, J.S., Rose, J.P. & Voth, G.A. 2018. Molecular transport through membranes: Accurate permeability coefficients from multidimensional potentials of mean force and local diffusion constants. *Journal of Chemical Physics*, 149(7): 072310.

Tan, Q., Zhu, Y., Li, Jian, Chen, Z., Han, G.W., Kufareva, I., Li, T., Ma, L., Fenalti, G., Li,

Jing, Zhang, W., Xie, X., Yang, H., Jiang, H., Cherezov, V., Liu, H., Stevens, R.C., Zhao, Q. & Wu, B. 2013. Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science*, 341(6152): 1387–1390.

Trott, O. & Olson, A.J. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*, 31(2): 455–461.

Vallianatou, T., Giaginis, C. & Tsantili-Kakoulidou, A. 2015. The impact of physicochemical and molecular properties in drug design: Navigation in the "Drug-Like" chemical space. *Advances in Experimental Medicine and Biology*, 822: 187–194.

Vangelista, L. & Vento, S. 2018. The expanding therapeutic perspective of CCR5 blockade. *Frontiers in Immunology*, 8(JAN).

Veber, D.F., Johnson, S.R., Cheng, H.Y., Smith, B.R., Ward, K.W. & Kopple, K.D. 2002. Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry*, 45(12): 2615–2623.

Wang, J. & Hou, T. 2011. Recent Advances on Aqueous Solubility Prediction. *Combinatorial Chemistry & High Throughput Screening*.

Wang, J., Shu, M., Wang, Yuanqiang, Hu, Y., Wang, Yuanliang, Luo, Y. & Lin, Z. 2016. Identification of potential CCR5 inhibitors through pharmacophore-based virtual screening, molecular dynamics simulation and binding free energy analysis. *Molecular BioSystems*, 12(11): 3396–3406.

Wang, J., Wang, W., Kollman, P.A. & Case, D.A. 2001. Antechamber, An Accessory Software Package For Molecular Mechanical Calculations Correspondence to. *Journal of Chemical Information and Computer Sciences*, 26(222): 403–408.

Wang, X., Song, K., Li, L. & Chen, L. 2018. Structure-Based Drug Design Strategies and Challenges. *Current Topics in Medicinal Chemistry*, 18(12): 998–1006.

Woollard, S.M. & Kanmogne, G.D. 2015. Maraviroc: A review of its use in hivinfection and beyond. *Drug Design, Development and Therapy*, 9: 5447–5468.

Xu, Y., Liu, H., Niu, C., Luo, C., Luo, X., Shen, J., Chen, K. & Jiang, H. 2004. Molecular docking and 3D QSAR studies on 1-amino-2-phenyl-4-(piperidin-1-yl)-butanes based on the structural modeling of human CCR5 receptor. *Bioorganic & medicinal chemistry*, 12(23): 6193–6208.

Yang, Y., Shen, Y., Liu, H. & Yao, X. 2011. Molecular dynamics simulation and free energy calculation studies of the binding mechanism of allosteric inhibitors with p38α MAP kinase. *Journal of chemical information and modeling*, 51(12): 3235–3246.

CHAPTER 7

Supporting Information

Structure-based identification of novel scaffolds as potential HIV-1 entry inhibitors towards CCR5 in HIV disease therapy

Patrick Appiah-Kubi[1], Emmanuel Amarachi Iwuchukwu [1], Mahmoud E. S. Soliman[1*]

[1]Molecular Bio-computation and Drug Design Laboratory

School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

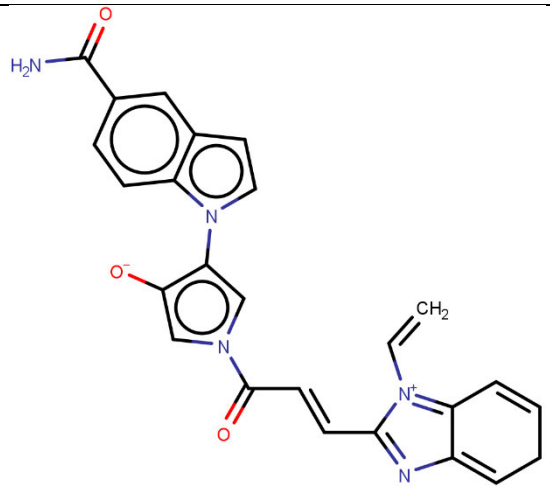
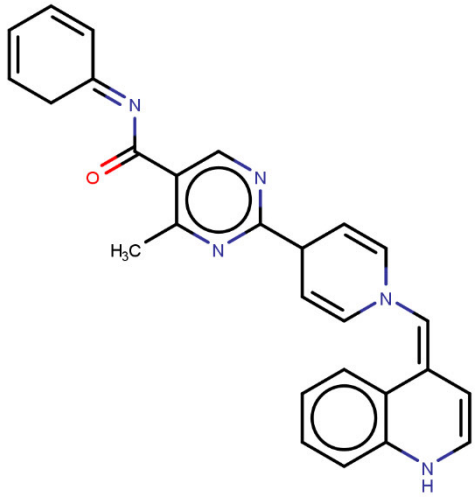
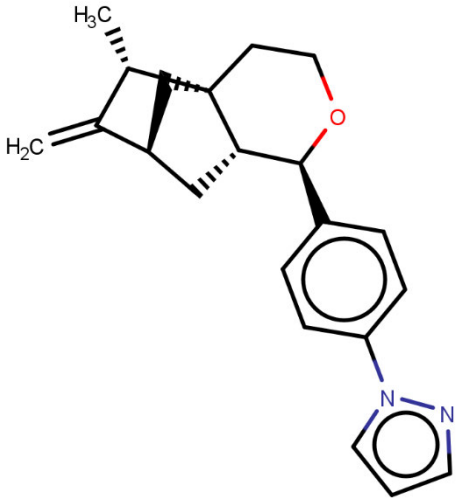
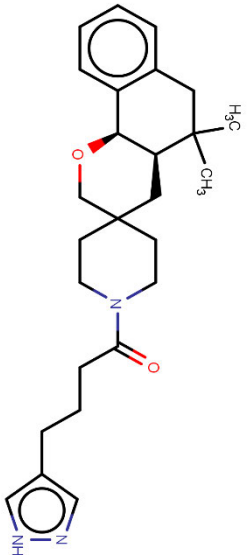
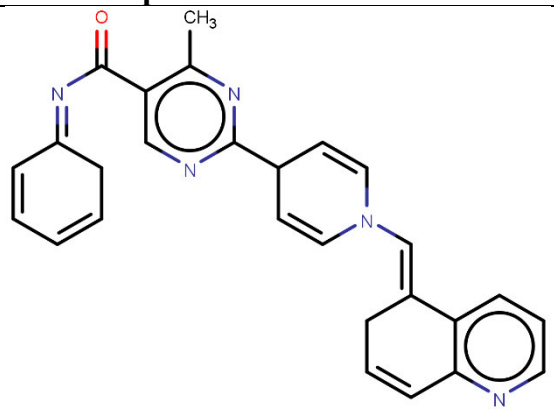
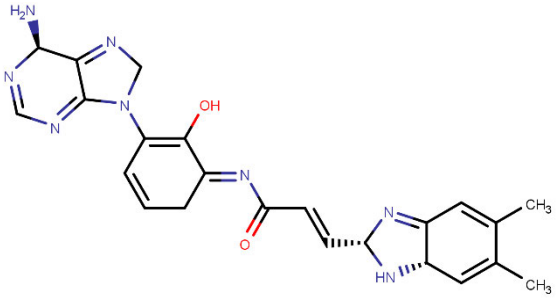
Telephone: +27 (0) 31 260 8048

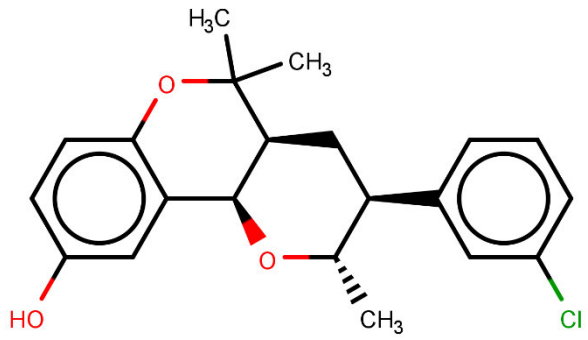
Fax: +27 (0) 31 260 78

Email: soliman@ukzn.ac.za

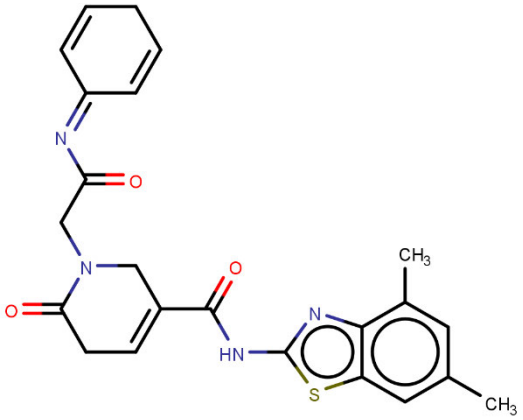
Website: http://soliman.ukzn.ac.za

Table **S1.** 2D chemical structures of the remaining 19 top hit compounds against CCR5 protein target.

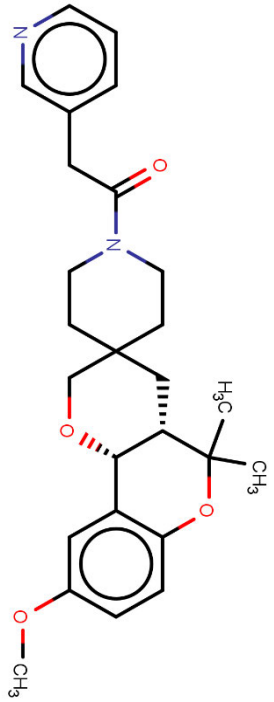
 **Entry number:** 8354 **Asinex compound ID:** LAS 51495184	 **Entry number:** 1093 **Asinex compound ID:** BDC 23205804
 **Entry number:** 3673 **Asinex compound ID:** BDF 33909572	 **Entry number:** 5477 **Asinex compound ID:** BDG 34130390
 **Entry number:** 1077 **Asinex compound ID:** BDC 23205600	 **Entry number:** 8530 **Asinex compound ID:** LAS 51502049



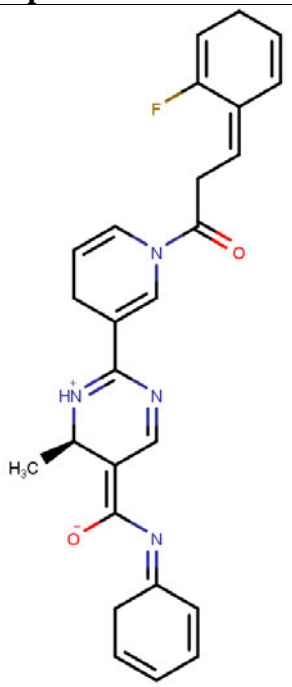
Entry number: 4856
Asinex compound ID: BDG 34037901



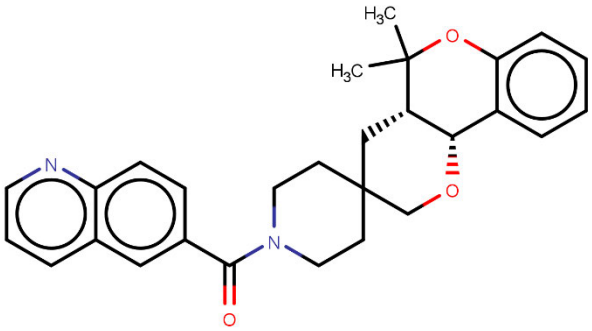
Entry number: 543
Asinex compound ID: BDB 26418354



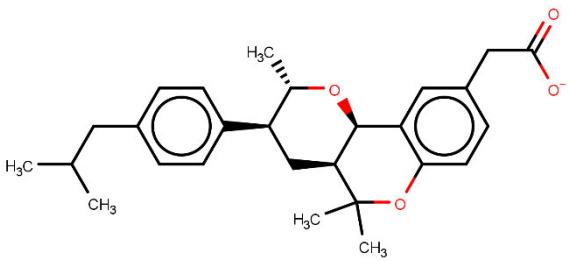
Entry number: 4055
Asinex compound ID: BDG 33691535



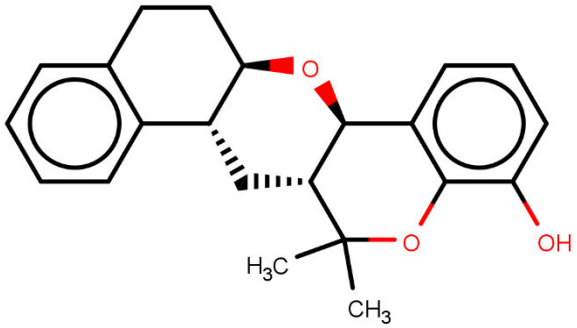
Entry number: 995
Asinex compound ID: BDC 23197464



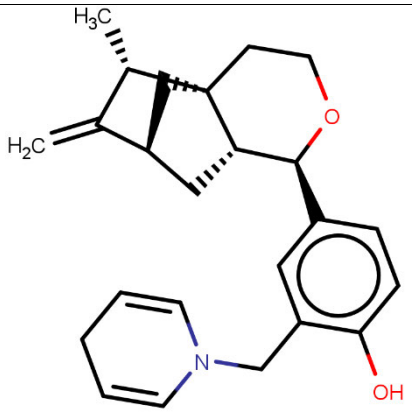
Entry number: 4156
Asinex compound ID: BDG 33694125



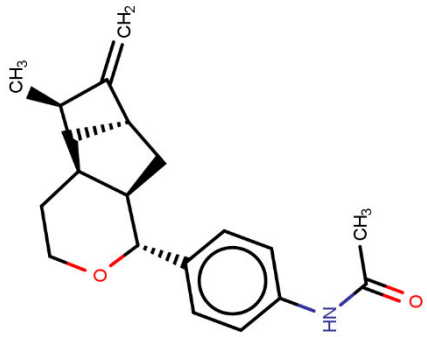
Entry number: 5573
Asinex compound ID: BDG 51127568



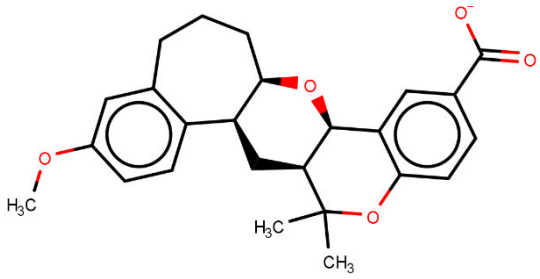
Entry number: 4946
Asinex compound ID: BDG 34042546



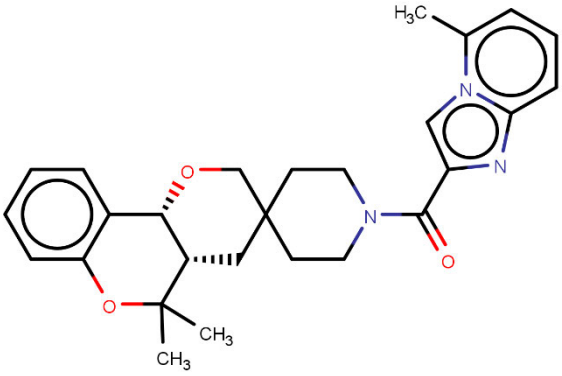
Entry number: 3742
Asinex compound ID: BDF 34027559



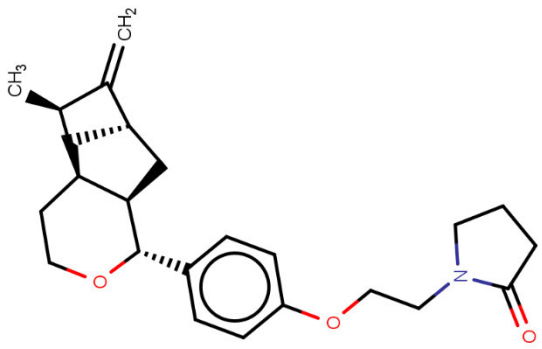
Entry number: 3746
Asinex compound ID: BDF 34035988



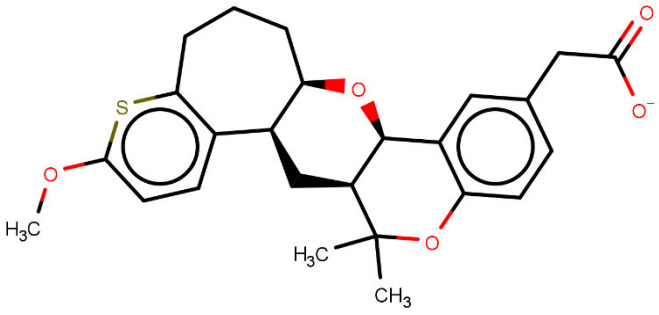
Entry number: 5582
Asinex compound ID: BDG 51129965



Entry number: 4157
Asinex compound ID: BDG 33694314



Entry number: 3672
Asinex compound ID: BDF 33909571



Entry number: 5564 **Asinex compound ID:** BDG 51126017

Table **S2.** Canonical smiles for the top four best-hit compounds.

Entry Number	Asinex Compound ID	Canonical smile
8356	LAS 51495192	Cc2nc1ccccc1n2CCCC(=O)N5CC(O)C(n4ccc3cc(C(N)=O)ccc34)C5
287	BDB 26405401	Cc4ccc(C1CCCN1C(=O)c3cccn(CC(=O)NCc2ccccc2)c3=O)cc4C
579	BDB 26419079	Cc4ccc(C1CCCN1C(=O)c3ccc(=O)n(CC(=O)Nc2ccc(F)cc2)c3)cc4C
7073	LAS 34154543	Cc5cc4nc(CCC(=O)N3CC(O)C(n2ccc1cc(C(N)=O)ccc12)C3)[nH]c4cc5C
Maraviroc	N/A	CC1=NN=C(N1C2CC3CCC(C2)N3CCC(C4=CC=CC=C4)NC(=O)C5CCC(CC5)(F)F)C(C)C

Table **S3.** Canonical smiles and molecular docking scores (Autodock Vina and HTVS Glide score) for the additional 19 compounds and Maraviroc.

Entry Number	Asinex Compound ID	Canonical smile	Autodock Score (kcal/mol)	Glide HTVS Score (kJ/mol)
8354	LAS 51495184	CCn1c(CCC(=O)N2CC(O)C(C2)n2ccc3cc(ccc23)C(N)=O)nc2ccccc12	-10.1	-7.445
1093	BDC 23205804	Cc1nc(ncc1C(=O)Nc1ccccc1)C1CCN(Cc2ccnc3ccccc23)CC1	-10.0	-6.881
3673	BDF 33909572	CC1C(=C)C2CC3C(OCCC13C2)c1ccc(cc1)-n1cccn1	-10.2	-6.082
5477	BDG 34130390	CC1(C)Cc2ccccc2C2OCC3(CCN(CC3)C(=O)CCCc3cn[nH]c3)CC12	-10.0	-5.720
1077	BDC 23205600	Cc1nc(ncc1C(=O)Nc1ccccc1)C1CCN(Cc2cccc3ncccc23)CC1	-10.6	-5.560
8530	LAS 51502049	Cc1cc2nc(CCC(=O)NC3CCCC(C3O)n3cnc4c(N)ncnc34)[nH]c2cc1C	-10.0	-5.522
4856	BDG 34037901	CC1OC2C(CC1c1cccc(Cl)c1)C(C)(C)Oc1ccc(O)cc21	-10.6	-5.452
543	BDB 26418354	Cc1cc(C)c2nc(NC(=O)c3ccc(=O)n(CC(=O)NC4CCCCC4)c3)sc2c1	-10.6	-5.599
3742	BDF 34027559	CC1C(=C)C2CC3C(OCCC13C2)c1ccc(O)c(CN2CCCCC2)c1	-10.2	-5.440
995	BDC 23197464	Cc1nc(ncc1C(=O)Nc1ccccc1)C1CCCN(C1)C(=O)CCc1ccccc1F	-10.1	-5.331
4156	BDG 33694125	CC1(C)Oc2ccccc2C2OCC3(CCN(CC3)C(=O)c3ccc4ncccc4c3)CC12	-10.6	-5.050
5573	BDG 51127568	CC(C)Cc1ccc(cc1)C1CC2C(OC1C)c1cc(CC(O)=O)ccc1OC2(C)C	-10.5	-4.975
4946	BDG 34042546	CC1(C)Oc2c(O)cccc2C2OC3CCc4ccccc4C3CC12	-11.4	-4.905
4055	BDG 33691535	COc1ccc2OC(C)(C)C3CC4(CCN(CC4)C(=O)Cc4cccnc4)COC3c2c1	-10.1	-4.868
3746	BDF 34035988	CC1C(=C)C2CC3C(OCCC13C2)c1ccc(NC(C)=O)cc1	-10.0	-4.819
5582	BDG 51129965	COc1ccc2C3CC4C(OC3CCCc2c1)c1cc(ccc1OC4(C)C)C(O)=O	-11.3	-4.777
4157	BDG 33694314	Cc1cccc2nc(cn12)C(=O)N1CCC2(CC1)COC1C(C2)C(C)(C)Oc2ccccc12	-10.7	-4.768
3672	BDF 33909571	CC1C(=C)C2CC3C(OCCC13C2)c1ccc(OCCN2CCCC2=O)cc1	-10.2	-4.708
5564	BDG 51126017	COc1ccc2C3CC4C(OC3CCCc2c1)c1cc(CC(O)=O)ccc1OC4(C)C	-11.1	-4.697
Maraviroc	N/A	CC1=NN=C(N1C2CC3CCC(C2)N3CCC(C4=CC=CC=C4)NC(=O)C5CCC(CC5)(F)F)C(C)C	-9.0	-4.695

Table **S4.** Physicochemical properties of the additional nineteen hit compounds and Maraviroc.

Entry Number	Asinex Compound ID	Formula	MW (g/mol)	MLog*P*	LogS (Ali) (mol/L)	TPSA (A^2)	Molar Refractivity	HBA	HBD	Rotatable bonds	Lipinski Drug likeness
8354	LAS 51495184	C25H27N5O3	445.51	1.42	-3.07	106.38	130.04	4	2	7	Yes; 0 violation
1093	BDC 23205804	C27H27N5O	437.54	2.70	-5.00	71.01	134.70	5	1	6	Yes; 0 violation
3673	BDF 33909572	C21H24N2O	320.43	3.85	-3.90	27.05	95.42	2	0	2	Yes; 0 violation
5477	BDG 34130390	C26H35N3O2	421.58	3.36	-4.78	58.22	126.35	3	1	5	Yes; 0 violation
1077	BDC 23205600	C27H27N5O	437.54	2.70	-5.00	71.01	134.70	5	1	6	Yes; 0 violation
8530	LAS 51502049	C23H28N8O2	448.52	0.61	-4.21	147.63	125.57	6	4	6	Yes; 0 violation
4856	BDG 34037901	C21H23ClO3	358.86	3.83	-5.36	38.69	99.96	3	1	1	Yes; 0 violation
543	BDB 26418354	C23H26N4O3S	438.54	2.67	-5.86	121.33	123.82	4	2	7	Yes; 0 violation
3742	BDF 34027559	C24H33NO2	367.52	3.89	-4.38	32.70	114.02	3	1	3	Yes; 0 violation
995	BDC 23197464	C26H27FN4O2	446.52	3.12	-4.76	75.19	129.17	5	1	8	Yes; 0 violation
4156	BDG 33694125	C28H30N2O3	442.55	3.53	-5.19	51.66	132.32	4	0	2	Yes; 0 violation
5573	BDG 51127568	C27H34O4	422.56	4.22	-6.52	55.76	123.86	4	1	5	Yes; 1 violation
4946	BDG 34042546	C22H24O3	336.42	3.56	-4.73	38.69	97.81	3	1	0	Yes; 0 violation
4055	BDG 33691535	C26H32N2O4	436.54	2.24	-3.89	60.89	125.89	5	0	4	Yes; 0 violation
3746	BDF 34035988	C20H25NO2	311.42	3.34	-3.31	38.33	92.61	2	1	3	Yes; 0 violation
5582	BDG 51129965	C25H28O5	408.49	3.52	-5.50	64.99	114.04	5	1	2	Yes; 0 violation
4157	BDG 33694314	C27H31N3O3	445.55	2.96	-5.60	56.07	130.73	4	0	2	Yes; 0 violation
3672	BDF 33909571	C24H31NO3	381.51	3.36	-3.69	38.77	113.72	3	0	5	Yes; 0 violation
5564	BDG 51126017	C26H30O5	422.51	3.45	-5.44	64.99	118.63	5	1	3	Yes; 0 violation
Maraviroc	N/A	C29H41F2N5O	513.67	4.15	-6.22	63.05	145.84	6	1	9	No; 2 violations

HBD = number of hydrogen bond donors (HBD $\leq$ 5); HBA = number of hydrogen bond donors (HBD $\leq$ 10); TPSA = total polar surface area ($\leq$140); MLOGP = predicted octanol/water partition coefficient (MLOGP < 4.15); MW = molecular weight (g/mol) (MW $\leq$ 500); Rotatable bonds ($\leq$ 8); TPSA = Topological polar surface area ($\leq$ 140)

Table **S5.** The molecular dynamics properties of the inhibitors bound within 8 Å of CCR5 active site as a function of the 100 ns simulation time for the solvent accessible surface area (SASA), the root mean square deviation (RMSD), the radius of gyration (rGyr) as well as the intramolecular hydrogen bonds (intraHB) of the entire complex.

CCR5-complex	rGyr±SD (Å)	intra HB±SD	SASA±SD ($Å^2$)	RMSD±SD (Å)
CCR5-LAS 51495192	5.25±0.22	0.87±0.80	136.05±32.42	1.42±0.76
CCR5-BDB 26405401	4.50±0.09	0.99±0.62	119.92±21.54	1.81±0.32
CCR5-BDB 26419079	4.63±0.11	1.29±0.72	86.37±19.89	1.68±0.39
CCR5-LAS 34154543	5.98±0.16	1.73±0.99	95.38±20.96	2.42±0.50
CCR5-Maraviroc	5.52±0.08	1.48±0.58	113.66±20.41	1.16±0.23

Table **S6.** Analysis of simulation quality of the top five systems for average total energy (ETOT), Potential energy (EPtot), temperature, pressure, and volume of the systems over the 100 ns simulation time.

CCR5-complex System	Parameter	Average±SD
CCR5-LAS 51495192	ETOT (kcal/mol)	-203475.4±334.53
	EPtot (kcal/mol)	-266924.4±267.89
	Pressure (bar)	0.88±84.80
	Temperature (K)	310.0±0.95
	Volume ($Å^3$)	956782.39±820.18
CCR5-BDB 26405401	ETOT (kcal/mol)	-203432.1±334.80
	EPtot (kcal/mol)	-266867.6±267.66
	Pressure (bar)	1.12±83.66
	Temperature (K)	310.0±0.96
	Volume ($Å^3$)	956622.51±817.16
CCR5-BDB 26419079	ETOT (kcal/mol)	-203136.9±334.43
	EPtot (kcal/mol)	-266503.8±267.04
	Pressure (bar)	0.81±84.53
	Temperature (K)	310.0±0.95
	Volume ($Å^3$)	955515.10±829.59
CCR5-LAS 34154543	ETOT (kcal/mol)	-203638.9±332.41
	EPtot (kcal/mol)	-267121.4±266.08
	Pressure (bar)	1.49±84.59
	Temperature (K)	310.0±0.95
	Volume ($Å^3$)	957330.56±801.62
CCR5-Maraviroc	ETOT (kcal/mol)	-194321.7±326.53
	EPtot (kcal/mol)	-254890.4±260.72
	Pressure (bar)	1.01±86.19
	Temperature (K)	310.0±0.98
	Volume ($Å^3$)	913167.63±794.6

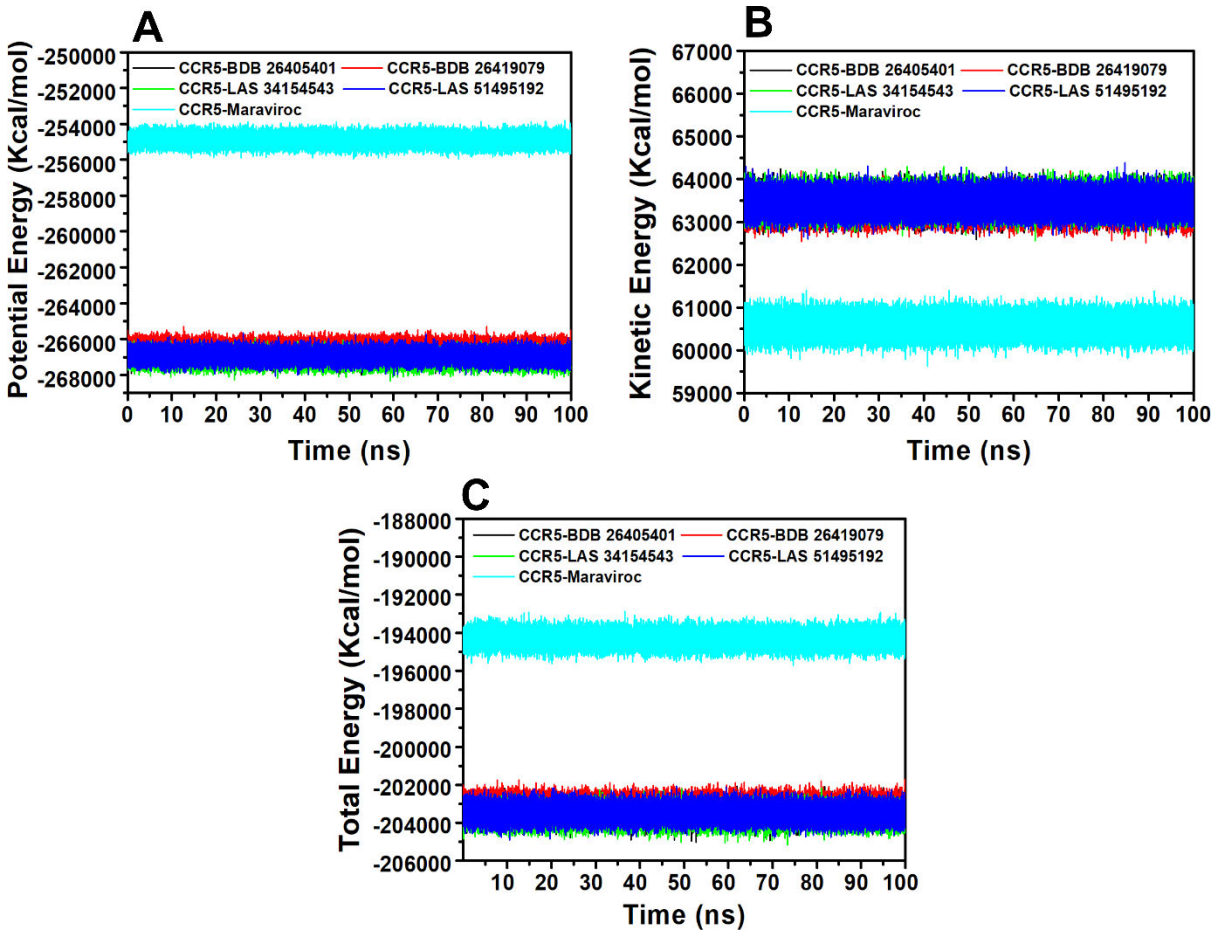


Fig. (S1). Analysis of the average energetics for potential energy (**A**), kinetic energy (**B**), and the total energy (**C**) of the systems over the 100 ns simulation time (Image prepared by author).

CHAPTER 8

8.0 Conclusions, Recommendations and Future Perspectives

In this thesis, the binding mechanistic of inhibitors targeting the human GPCRs, i.e. C-C chemokine receptor 5 (CCR5), D_2 and D_3 dopamine receptors (D2DR and D3DR), were studied to provide a multifaceted and novel understanding of inhibitor/drug-receptor interactions at the selected GPCR targets. In this concluding chapter, novel insights are highlighted, and recommendations for future investigations in this research area are outlined.

8.1 Conclusions from this thesis

The current availability of resolved drug/ligand complex structures for GPCRs such as the D_2-like dopamine receptors and CCR5 provide avenues to uncover atomic molecular level ligand recognition mechanisms and conformational dynamics. Understanding the dynamic signalling properties of GPCRs and the structural basis of their ligand recognition/binding interactions enable full leverage on the power of rational drug design. The main aim of this thesis was to provide novel mechanistic insights into how dopamine receptors and CCR5 interact with their small molecule inhibitors and apply a structure-based discovery approach to identifying new potential CCR5 inhibitor scaffolds.

The following conclusions were made in deciphering the binding mechanism of atypical antipsychotic drugs and their effect on D2DR conformational dynamics.

i. Class II atypical antipsychotics adopts a favourable binding mode and interactions at the D2DR deep hydrophobic pocket. In contrast, the binding conformations of Class I atypical antipsychotics is shallow, reaching out above the orthosteric binding pocket into the D2DR extended binding pocket.
ii. van der Waals interaction energy contribution dominates the higher binding affinity of Class II atypical antipsychotics at D2DR receptor compared to the relatively lower binding affinities of the Class I antipsychotics.
iii. A unique salt bridge interaction with $Asp114^{3.32}$ is common to all the studied Class II atypical antipsychotic drugs. This salt bridge is vital for inhibitors displaying high binding affinity at the aminergic subfamily of GPCR (Shi & Javitch, 2002).

In addition, the molecular mechanistic and conformational dynamics underlying the selective binding of the two bitopic antagonists, R-VK4-40 and Y-QA31, toward D3DR over D2DR were explored. R-VK4-40 and Y-QA31 adopt shallow binding modes at D3DR orthosteric binding pocket (OBP) and extended/second binding pocket (E/SBP), on the contrary, they display deep hydrophobic pocket binding at D2DR. The efficacy of bitopic compounds has been found to depend on their binding mode in the OBP, whereas their selectivity arises from their different interactions within the SBP (Newman et al., 2012). The favourable flexibility of bitopic compounds seems to be conferred by their linker, allowing R-VK4-40 and Y-QA31 to bind in the S/EBP of D3DR with higher selectivity. Also, two non-conserved residues (Tyr36$^{1.39}$ and Ser182$^{45.51}$) have been identified in D3DR, due to their essential role in the selective binding of R-VK4-40 and Y-QA31. The estimated binding free energies also suggest an increase in van der Waals interactions and a relative decrease in entropy contribution were crucial factors that underlie the high selectivity and affinity of the studied antagonists for D3DR relative to D2DR.

Furthermore, this thesis has also provided a structural understanding and binding mechanistic of the novel 1-heteroaryl-1,3-propanediamine derivatives (Compd-21 and Compd-34) and Maraviroc at CCR5. The findings from this study have shown the structural basis for Compd-21 and Compd-34 higher binding profile at CCR5 binding pocket compared to Maraviroc. The binding dynamics have revealed that the substitution of the phenyl group in Maraviroc with the thiophene moieties cumulatively engaged in higher affinity interactions with CCR5 binding pocket residues critical for gp120 V3-loop binding. These results corroborate the recent findings that Maraviroc blocks gp120 binding to CCR5 via direct competitive inhibition as opposed to earlier views of allosteric inhibition through conformational availability restriction (Shaik et al., 2019).

Finally, the successful application of structure-based virtual screening techniques identified potentially new inhibitor scaffolds that displayed binding modes and interaction profiles at the CCR5 binding pocket. These compounds made substantial interactions with residues vital for gp120 V3 loop binding.

8.2 Limitations of the Study

The validity of the binding affinities, the physicochemical and pharmacokinetic properties of identified potential CCR5 inhibitors is subject to experimental evaluations. Thus the results are to be interpreted as theoretical findings.

8.3 Recommendations and Future Perspectives

The deep hydrophobic sub-pocket binding of Class II atypical antipsychotic drugs may be explored in the design of the next generation of D_2 dopamine receptor subtype-selective antagonists. The role of the unique salt bridge interaction of D2DR Asp114$^{3.32}$ should be considered during the rational design of potential D2DR inhibitors.

Bitipic compound design and their optimisation for their different interactions at D3DR E/SBP (selectivity) and OBP (higher affinity) may assist in the rational design of novel D_3 dopamine receptor-selective antagonists with higher binding affinity.

Further structural modifications at the thiophene substituent with functional groups that make stronger interactions with the identified residues critical for gp120 V3 loop binding will be beneficial. Also, the addition of additional functional groups to the triazole ring may increase inhibitor competition with gp120 V3-loop with enhanced potency at CCR5 (Shaik et al., 2019). This offers a foundation for the onward structural modifications and rational design of novel potent antagonists of CCR5 in HIV-1 treatment.

The identified potential CCR5 inhibitors have displayed promising interactions with residues critical for gp120 V3 loop binding. However, biochemical investigations of these compound scaffolds obtained from the *in silico* structure-based studies are still needed to verify their antiviral activity at CCR5. Furthermore, the selected compounds toward CCR5 add to the available chemical scaffolds for structural modifications and optimisations that may assist in the lead identification of more potent antiviral drugs of HIV-1.

8.4 References

Newman, A.H., Beuming, T., Banala, A.K., Donthamsetti, P., Pongetti, K., LaBounty, A., Levy, B., Cao, J., Michino, M. & Luedtke, R.R. 2012. Molecular determinants of selectivity and efficacy at the dopamine D3 receptor. *Journal of medicinal chemistry*, 55(15): 6689–6699.

Shaik, M.M., Peng, H., Lu, J., Rits-Volloch, S., Xu, C., Liao, M. & Chen, B. 2019. Structural

basis of coreceptor recognition by HIV-1 envelope spike. *Nature*, 565(7739): 318.
Shi, L. & Javitch, J.A. 2002. The binding site of aminergic G protein-coupled receptors: the transmembrane segments and second extracellular loop. *Annual review of pharmacology and toxicology*, 42(1): 437–467.

Appendix I: Proof of Published Article

Full Paper www.molinf.com molecular informatics

DOI: 10.1002/minf.201900044

Probing Binding Landscapes and Molecular Recognition Mechanisms of Atypical Antipsychotic Drugs towards the Selective Targeting of D_2 Dopamine Receptor

Patrick Appiah-Kubi,[a] Fisayo Andrew Olotu,[a] and Mahmoud E. S. Soliman*[a]

Abstract: Dopamine receptors constitute a unique class of G-protein coupled receptors that mediate the activities of dopamine, a neurotransmitter implicated in diverse neurological diseases when dysregulated. Over the years, antipsychotic drugs have been primarily directed towards D_2 dopamine receptor (DRD2) while associable adverse effects have been centred on non-selective targeting. The recent crystal structure of DRD2 in complex with atypical antipsychotic could further aid the structure-based design of highly DRD2-selective antipsychotics. Therefore, in this study, we comprehensively investigate the molecular recognition and differential binding landscapes of class-I and II DRD2 atypical antipsychotics, using membrane-bilayer molecular dynamics simulation and binding free energy techniques. Findings revealed that selected class-I antipsychotics exhibited binding dynamics and poses dissimilar to the class-II types with different interactive mechanisms at the binding cavity of DRD2. More interestingly, the class-II drugs established a highly coordinated binding at the DRD2 active site with a pertinent and recurrent involvement of Asp114 via strong hydrogen interactions. Furthermore, while these compounds exert distinct effects on DRD2 structure, findings revealed that the class-II types favourably engaged the deep hydrophobic pocket of DRD2 compared to the class-I drugs. We speculate that these findings will be fundamental to the discovery of highly selective DRD2 antipsychotics.

Keywords: Atypical antipsychotics · D_2 dopamine receptor · G-protein coupled receptor · Membrane-bound protein · lipid bilayer Molecular dynamics simulation

1 Introduction

G protein-coupled receptors (GPCRs) play an important role in the cell signalling process, particularly in response to neurotransmitters and hormones.[1] Dopamine is an important neurotransmitter exerting its effect via the activation of the five dopamine receptors subtypes belonging to the GPCR family, namely; D_1, D_2, D_3, D4, and D_5.[2] These receptors are further grouped into two main subfamilies, D_1-like (D_1 and D_5) and D_2-like (D_2, D_3, and D_4) receptors, based on their sequence and structural similarities.[3] Generally, these class of receptors mediate cell proliferation, and differentiation, cyclic adenosine monophosphate release and neurotransmission of other neurotransmitters.[4]

Amongst these subtypes, D_2 dopamine receptor (DRD2) is an important GPCR drug target that mediates signal transduction of neurotransmitters in the central nervous system as well as in neurological processes such as memory, attention, emotion, pleasure, lust, and love.[5] The dysfunction of the dopaminergic system has been associated with various neurodegenerative disorders such as depression, Parkinson's disease, attention deficit hyperactivity disorder, and schizophrenia. DRD2 has been the main target for atypical and typical antipsychotic drugs[6,7] as well as drugs employed in the treatment of Parkinson disease. Unfortunately, many of these antipsychotic drugs targeting DRD2 are accompanied by a plethora of severe side effects, due to off-target interactions with other related targets.[8,9]

The crystal structures of D_4 dopamine receptor (DRD4) in complex with antipsychotic nemonapride[10] and D_3 dopamine receptor (DRD3) in complex with eticlopride[11] have earlier been resolved. However, the lack of crystalized structures of the DRD2 ligand complexes over the past years impeded the molecular understanding of receptor function and ligand recognition. The recent crystallization of the DRD2 in complex with the widely prescribed antipsychotic drug risperidone in 2018 provided novel insights[12] (Figure 1). Previous studies of DRD2 have been based on homology modelled structures using either DRD3 or DRD4 crystal structures as templates .[13–15]

The solved crystal structure of DRD2 revealed an unexpected binding mode of risperidone which was contrary to findings from previous molecular docking studies that employed a non-crystal homology model of

[a] *P. Appiah-Kubi, F. A. Olotu, M. E. S. Soliman*
Molecular Bio-computation and Drug Design Laboratory
School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa
Telephone: +27 (0) 31 260 8048
Fax: +27 (0) 31 260 78
E-mail: soliman@ukzn.ac.za

Appendix II: Proof of Published Article

Journal of Molecular Recognition
Online Journal:
=Free Color
=More Color
=Faster Publication

RESEARCH ARTICLE

Exploring the structural basis and atomistic binding mechanistic of the selective antagonist blockade at D_3 dopamine receptor over D_2 dopamine receptor

Patrick Appiah-Kubi, Fisayo Andrew Olotu, Mahmoud E. S. Soliman

First published: 05 January 2021 | https://doi.org/10.1002/jmr.2885

Abstract

More recently, there has been a paradigm shift toward selective drug targeting in the treatment of neurological disorders, including drug addiction, schizophrenia, and Parkinson's disease mediated by the different dopamine receptor subtypes. Antagonists with higher selectivity for D_3 dopamine receptor (D3DR) over D_2 dopamine receptor (D2DR) have been shown to attenuate drug-seeking behavior and associated side effects compared to non-subtype selective antagonists. However, high conservations among constituent residues of both proteins, particularly at the ligand-binding pockets, remain a challenge to therapeutic drug design. Recent studies have reported the discovery of two small-molecules R-VK4-40 and Y-QA31 which substantially inhibited D3DR with >180-fold selectivity over D2DR. Therefore, in this study, we seek to provide molecular and structural insights into these differential binding mechanistic using meta-analytic computational simulation methods. Findings revealed that R-VK4-40 and Y-QA31 adopted shallow binding modes and were more surface-exposed at D3DR while on the contrary, they exhibited deep hydrophobic pocket binding at D2DR. Also, two non-conserved residues; $Tyr36^{1.39}$ and $Ser182^{45.51}$ were identified in D3DR, based on their crucial roles and contributions to the selective binding of R-VK4-40 and Y-QA31. Importantly, both antagonists exhibited high affinities in complex with D3DR compared to D2DR, while van der Waals energies contributed majorly to their binding and stability. Structural analyses also revealed the distinct stabilizing effects of both compounds on D3DR secondary architecture relative to D2DR. Therefore, findings herein pinpointed the origin and mechanistic of selectivity of the compounds, which may assist in the rational design of potential small molecules of the D_2-like dopamine family receptor subtype with improved potency and selectivity.

Appendix III: Proof of Published Article

Full-Text Inquiry

Research Article

Elucidating the Disparate Inhibitory Mechanisms of Novel 1-Heteroaryl1,3-Propanediamine Derivatives and Maraviroc towards C-C Chemokine Receptor 5: Insights for Structural Modifications in HIV-1 Drug Discovery

(E-pub Abstract Ahead of Print)

Author(s): Patrick Appiah-Kubi, Fisayo Andrew Olotu, Mahmoud E. S. Soliman*

Journal Name: Medicinal Chemistry

DOI : 10.2174/1573406417666201208122110

Journal Home

Abstract:

Introduction: Blocking Human Immunodeficiency Virus type 1 (HIV-1) entry via C-C chemokine receptor 5 (CCR5) inhibition has remained an essential strategy in HIV drug discovery. This underlies the development of CCR5 blockers, such as Maraviroc, which, however, elicits undesirable side effects despite its potency.

Background: Recent lead optimization efforts led to the discovery of novel 1-heteroaryl-1,3-propanediamine derivatives; Compd-21 and -34, which were ~3 times more potent than Maraviroc, with improved pharmacokinetics. However, atomistic molecular interaction mechanism of how slight structural variance between these inhibitors significantly affects their binding profiles have not been elucidated.

Method: This study employed explicit lipid bilayer molecular dynamics (MD) simulations, and advance analyses to explore these inhibitory discrepancies.

Results: Findings revealed that the thiophene moiety substitution common to Compd-21 and -34 enhanced their CCR5- inhibitory activities due to complementary high-affinity interactions with Trp862.60, Tyr1083.32, Tyr2516.51, Glu2837.39. These cumulatively accounted for their ΔGbind which were higher than Maraviroc. Binding dynamics further revealed that the compounds mediated direct competitive inhibition at CCR5 by blocking the gp120 V3 loop. Furthermore, constituent tropane and triazole moieties in the compounds commonly engaged in interactions with Glu2837.39 and Trp862.60, respectively. Structural analyses also revealed that both Compd-21 and -34 elicited distinct internal dynamic effect on CCR5 relative to Maraviroc.

Conclusion: Structural modifications at the thiophene substituent and the addition of new functional groups to the triazole ring may enhance inhibitor competition with gp120 V3-loop. Findings herein highlighted would contribute to future structure-based design of inhibitors of HIV-1 CCR5 with improved potencies.

Keywords: G protein-coupled receptor, HIV-1, Maraviroc, lipid bilayer, C-C chemokine receptor 5 (CCR5), Molecular dynamics simulations, 1-heteroaryl-1, 3-propanediamine

Appendix IV: Proof of Submitted Manuscript

Medical Hypotheses

Structure-based identification of novel scaffolds as potential HIV-1 entry inhibitors towards CCR5 in HIV disease therapy

--Manuscript Draft--

Manuscript Number:	
Article Type:	Full length article
Keywords:	C-C Chemokine Receptor 5; Maraviroc; Structure-based Virtual screening; Molecular docking; MD Simulations; ADMET
Corresponding Author:	Patrick Appiah-Kubi, MMedSc University of KwaZulu-Natal College of Health Sciences Durban, KwaZulu-Natal SOUTH AFRICA
First Author:	Patrick Appiah-Kubi, MMedSc
Order of Authors:	Patrick Appiah-Kubi, MMedSc
	Emmanuel Amarachi Iwuchukwu, PhD
	Mahmoud E. S. Soliman, PhD
Abstract:	The C-C chemokine receptor 5 (CCR5) viral coreceptor belonging to the G protein-coupled receptor family is one of the families of chemokine receptors. The interactions of CCR5 with HIV-1 during viral entry positions it as an effective therapeutic target for the design of potent antiviral therapies. The FDA approved the small-molecule Maraviroc as CCR5 drug in 2007, while clinical trials failure has characterised many of the other CCR5 inhibitors. Thus, the continual identification of potential CCR5 inhibitors is, therefore warranted. In this study, a structure-based discovery approach has been utilised to virtually screen and retrieved novel potential CCR5 inhibitors from the Asinex antiviral compound (~ 8,722) database. Explicit lipid-bilayer molecular dynamics simulation, in silico physicochemical and pharmacokinetic analyses, were further performed for the top compounds. A total of 23 structurally diverse compounds with binding scores higher than Maraviroc were selected. Subsequent molecular dynamics (MD) simulations analysis of the top four compounds LAS 51495192, BDB 26405401, BDB 26419079, and LAS 34154543 maintained stability at CCR5 binding site. Furthermore, these compounds made pertinent interactions with CCR5 residues critical for the HIV-1 gp120-V3 loop binding such as Trp86, Tyr89, Phe109, Tyr108, Glu283 and Tyr251. Additionally, the predicted in silico physicochemical and pharmacokinetic descriptors of the selected compounds were within the acceptable range for drug-likeness. The results suggest positive indications that the identified molecules may represent promising CCR5 entry inhibitors. Further structural optimisations and biochemical testing of the proposed compounds may assist in the discovery of effective HIV-1 therapy.
Suggested Reviewers:	Shideh Montasser Kouhsari, Ph.D University of Tehran College of Science montasser20sh@khayam.ut.ac.ir Has expertise in computational molecular modeling and biomolecular simulations
	Nima Razzaghi-Asl, Ph.D Ardabil University of Medical Sciences: Ardebil University of Medical Sciences n.razzaghi@arums.ac.ir Has expertise in computer aided drug design and simulations
	Jerônimo Lameira Silva, Ph.D Federal University of Para: Universidade Federal do Para lameira@ufpa.br Expert in Computational modeling and drug design