COMPUTATIONAL CHEMISTRY STUDIES OF SUBTYPES B AND SOUTH AFRICAN C HIV PROTEASES

ZAINAB KEMI SANUSI

215079499

2016

COMPUTATIONAL CHEMISTRY STUDIES OF SUBTYPES B AND SOUTH AFRICAN C HIV PROTEASES

ZAINAB KEMI SANUSI

215079499

2016

A thesis submitted to the School of Pharmacy and Pharmacology, Faculty of Health Science, University of KwaZulu-Natal, Westville, for the degree of Master of Medical Science.

This is the thesis in which the chapters are written as a set of discrete research publications, with an overall introduction and final summary. Typically, these chapters will have been published in internationally recognized, peer-reviewed journals.

This is to certify that the content of this thesis is the original research work of Miss Zainab Kemi Sanusi.
As the candidate's supervisor, I have approved this thesis for submission.

Supervisor: Signed: ---------- Name: **Dr. B. Honarparvar** Date: 13/12/2015

Co-Supervisor: Signed: ---------- Name: **Prof. H. G. Kruger** Date: 13/12/2015

Co-Supervisor: Signed: ---------- Name: **Dr. G. E. M. Maguire** Date: 13/12/2015

ABSTRACT

HIV/AIDs is a prevalent disease infecting millions of people throughout the world. Although a lot of improvement has been achieved over the year in regard to the reduction of AIDs related deaths, a huge task lies ahead as the HIV/AIDs global epidemic keeps spreading annually. It is therefore paramount to discover and develop more and efficient drug inhibitors against HIV.

The HIV protease (HIV PR) is a C2-symmentric homodimer and consisting of 99-amino acids in each monomer and because of the important role it plays in the HIV mutation, it became a major HIV drug target for the past three decades. It is on this basis that various effective antiretroviral protease inhibitors have been designed and approved for application in HIV therapy.

The HIV subtype B strain is prominent in Europe and North America and is the most researched virus. The majority of the antiretroviral drugs were designed and tested against HIV subtype B. However, non-subtype B strains of the HIV virus makes up most of these infections in Southern and Eastern Africa, which are highly affected regions in the world. In South Africa, subtype C HIV-1 is the dominant strain and little research has been done regarding drug design for this subtype or testing of the effectiveness of the HIV approved antiretroviral drugs against these non-subtype B strains. Two potentially devastating mutations of subtype C-SA HIV PR were recently reported by our group. These were designated I36T↑T and L38L↑N↑L HIV PR. The I36T↑T PR mutant includes an extra amino acid, the mutation occurs at position 36 (isoleucine to threonine) and is followed by an insertion at the second threonine indicated by the upward arrow. The L38L↑N↑L PR mutant involves two amino acids insertions that is completely different from the usual 99-amino acids HIV PR, as well as five point mutations occur at the E35D, I36G, N37S, M46L and D60E. The two insertions occur at position 38 (asparagine and leucine) indicated by the two upward arrows. Therefore, the I36T↑T and L38L↑N↑L mutations consist of 100 and 101-amino acids in each monomer of the proteases respectively.

In this thesis, a hybrid computational model (QM: MM) using the ONIOM approach was followed. The selected FDA inhibitors were complexed with the various proteases in the active pocket interacting with Asp 25/25' catalytic residues using the same pose in the subtype B PR as a reference X-ray structure. The HIV PR inhibitors and Asp 25/25' were treated at a high-level

with quantum mechanics (QM) theory using B3LYP/6-31G(d), and the remaining HIV PR residues were considered at a low layer using molecular mechanics (MM) with the AMBER force field. This method was applied to calculate the binding free interaction energies of the selected FDA approved HIV PR drugs complexed to the HIV protease enzyme. The aim was to create and test this computational model that will reflect the experimental binding energies against subtype B, C-SA HIV PR and also a mutant from the subtype C-SA PR designated L38L↑N↑L HIV PR.

The calculated binding free interaction energies results from the subtype B follow a satisfactory trend with the experimental data. However, the C-SA HIV PR inhibitor—enzyme complexes showed some discrepancies and this was ascribed to the simplified computational model that omitted water in the active site of the enzyme. The calculated binding free interaction energies for L38L↑N↑L PR as well as experimental results, showed reduced binding affinities for all the selected FDA approved inhibitors in comparison with the subtype C-SA HIV PR. The deviation could be as a result of the insertion and mutation of the subtype C HIV-1 PR that is expected to have a significant effect in altering either the binding affinity of the HIV PR inhibitors and or characteristics of the parent protease.

The computational model used in this research will be improved by introducing water into the active pocket of the Asp 25/25' catalytic residues that will be treated at least at semi-empirical level. Optimization of the different ONIOM levels will be attempted in order to accurately predict activities of new potential HIV PR inhibitors.

DECLARATION

I, Miss Zainab Kemi Sanusi, 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 persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This thesis does not contain other persons' 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 References sections.

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

Signed:

LIST OF PUBLICATIONS

Publication **1**: Investigating the binding free energies of FDA approved drugs against subtype B and C-SA HIV PR: ONIOM approach - Zainab K. Sanusi, Thavendran Govender, Glenn E. M. Maguire, Sibusiso B Maseko, Johnson Lin, Hendrik G. Kruger and Bahareh Honarparvar.

Zainab K. Sanusi contributed to the design of the project, carried out all the calculations and wrote the paper.
The paper was submitted to Biophysical Journal 1st of December, 2016.

Publication **2**: Exploring the Binding Interaction energies of FDA approved PR Inhibitors against novel C-SA HIV-1 PR mutant L38L↑N↑L PR: ONIOM method - Zainab K. Sanusi, Thavendran Govender, Glenn E. M. Maguire, Sibusiso B Maseko, Johnson Lin, Hendrik G. Kruger and Bahareh Honarparvar.

Zainab K. Sanusi contributed to the design of the project, carried out all the calculations and wrote the paper.
The paper is in preparation for submission.

ACKNOWLEDGMENTS

My foremost appreciation goes to my Creator, Almighty Allah, who has been my help in ages past and my hope in years to come. I appreciate Him for seeing me through the thick and thin of this research work.

This thesis is the result of the dedicated effort of: Prof. Gert Kruger, Dr. Bahareh Honarparvar and Dr. Glenn E. M. Maguire, who adroitly supervised the entire research work. I also appreciate Zeynab Fakhar and Dr. Md. Rauf Shah for their love and patience in teaching me.

My overwhelming gratitude goes to my parents, Mr. L.K Sanusi and Mrs. I.Y Sanusi and my siblings (Hafeez Sanusi, Malik Sanusi, Minkahil Sanusi, Olaide Sanusi, Omolola Sanusi, Aminat Dauda) for their financial, moral and spiritual support. I love you so much and God bless you.

My appreciation also goes to Prof. Thavi Govender, Dr. Tricia Naicker, other Catalysis and Peptide Research Unit (CPRU) principled investigators and the crew at large. I hail my Niggas, Elton Chiwunze, Sifiso Mkathini, Rufaro and Soxs. My sincere thanks to: Adeola Shobo, Ishaq Lawal, Samson Akpotu, for been a worthy brother, Monsurat Lawal for been a sister, best counselor, morale booster, companion and Iffy for her care, Melissa Ramtahal and Estelle Ramchuran for their administrative advice and support.

My profound gratitude also goes to these distinct persons, my very own Sanusi families both near and far, Badru Abdulbaqi, Oni Oluwafemi, Stephen Ibijola, Fatima Ifeacho, Bode Ojugbele and group, for their love, motivation, prayers, support and care.

I also appreciate College of Health Science, University of KwaZulu-Natal and National Research Foundation, South Africa for this opportunity and financial support.

Table of Contents

ABSTRACT ..... **ii**
DECLARATION ..... **iv**
LIST OF PUBLICATIONS ..... **v**
ACKNOWLEDGMENTS ..... **vi**
LIST OF FIGURES ..... **i**
LIST OF TABLES ..... **iii**
LIST OF ACRONYMS ..... **iv**
CHAPTER ONE **1**
INTRODUCTION **1**
1.1 Preface 1
1.2 HIV Virus 2
1.2.1 HIV-1 Life-cycle 4
1.3 HIV Inhibitors 6
1.3.1 Irreversible Inhibitors 7
1.3.2 Reversible Inhibitors 7
1.4 Types of HIV Inhibitors 8
1.4.1 Entry Inhibitors 8
1.4.2 Reverse Transcriptase Inhibitors 10
1.4.3 Retroviral Integrase Inhibitors 11
1.4.4 Protease Inhibitors 12
1.5 HIV-1 Protease 14
1.5.1 Domains in HIV-1 PR Dimer 15
1.5.2 Dimer Stability of HIV-1 PR 15
1.5.3 Substrate Binding 16
1.5.4 HIV-1 PR Catalytic Mechanism 16
1.6 HIV-1 PR Mutations 18
1.7 Objectives of this Study 19
1.8 Thesis Outline 19
References **21**

CHAPTER TWO **29**
COMPUTATIONAL BACKGROUND **29**
2.1 Introduction to Computational Techniques 29
2.2 Theoretical Models 29
2.2.1 Quantum Mechanics 29
2.2.2 Molecular Mechanics (MM) 31
2.3 ONIOM Hybrid Method (QM:MM) 32
2.3.1 Calculating Binding Free Energy with ONIOM 33
References **34**
CHAPTER THREE **39**
Investigating the binding free energies of FDA approved drugs against subtype B and C-SA HIV PR: ONIOM approach **39**
Abstract **39**
1.0 Introduction **40**
2.0 Materials and Methods **42**
2.1 ONIOM binding free energies 46
3.0 Results and Discussion **47**
4.0 Conclusion **54**
References **56**
CHAPTER FOUR **64**
Exploring the Binding Free energies of FDA approved PR inhibitors against novel C-SA HIV-1 PR mutant L38L↑N↑L PR: ONIOM method **64**
Abstract **64**
1.0 Introduction **65**
2.0 Computational Methods **67**
2.1 Structural preparation of inhibitor—enzyme complexes 68
2.2 The ONIOM evaluation method 69
3.0 Results and Discussion **70**
4.0 Conclusion **76**
References **78**

CHAPTER FIVE 87
CONCLUSION 87
APPENDIX A 89
Supplementary Material for Chapter Three 89
APPENDIX B 105
Supplementary Material for Chapter Four 105

LIST OF FIGURES

CHAPTER ONE;

Figure 1. Diagram of HIV genome 4
Figure 2. Schematic representation of HIV Life cycle 6
Figure 3. Structure of a CCR5 antagonist inhibitor 9
Figure 4. Fusion inhibitor sequence ……..…….9
Figure 5. Schematic representation of binding inhibitor ……..10
Figure 6. Structure of FDA approved NRTI drugs ……..10
Figure 7. Structure of FDA approved NNRTI drugs …….....11
Figure 8. Schematic representation of HIV-1 integrase inhibitors …….....12
Figure 9. Structures of HIV-1 approved protease inhibitors ……....13
Figure 10. Ribbon representation of HIV-1 protease ……...14
Figure 11. Schematic representation of the catalytic residues triad ……..…..16
Figure 12. A representation of the PR substrate binding ……..……..….16
Figure 13. A representation of HIV-1 catalytic mechanics …….....18

CHAPTER TWO;

Figure 1. Schematic representation of two and three-layer ONIOM models 33

CHAPTER THREE;

Figure 1. Homo dimeric X-ray structure of subtype C-SA protease 43
Figure 2. The structures of HIV-1 FDA approved protease inhibitors ……..…….....45
Figure 3. Schematic diagram of the two-layered ONIOM model …….....46
Figure 4. Schematic representation of super-imposed inhibitor—enzyme complex of subtype B with C-SA PR. ……..47
Figure 5. The structure of hydrogen bond distances between subtype B and HIV PR drugs before and after optimization ……..….50
Figure 6. The electrostatic and hydrogen bond plots of the HIV subtype B PR …….….52

CHAPTER FOUR;

Figure 1. Schematic representation of super-imposed subtype C-SA and mutant L38L↑N↑L PR. 67

Figure 2. Schematic representation of two and three-layer ONIOM model ……...…….70

Figure 3. The structure of hydrogen bond distances between L38L↑N↑L and HIV PR drugs before and after optimization ……...…….73

Figure 4. The electrostatic and hydrogen bond plots of the L38L↑N↑L PR ……....74

APPENDIX A;

Figure S1. 3D Schematic representation of C-SA HIV PR complexed with **a:** APV, **b:** ATV, **c:** DRV, **d:** IDV, **e:** LPV, **f:** NFV, **g:** RTV, **h:** SQV, and **i:** TPV PIs……..89-90

Figure S2. Electrostatic and hydrogen bon interactions plots of subtype B and C-SA HIV PR respectively with the various FDA approved HIV PIs…….91-94

Figure S3. Hydrogen bond distances between the hydroxyl groups of the selected FDA approved ATV drug with the catalytic ASP25 and ASP25' residues of **a1:** subtype B HIV PR, **a2:** subtype C-SA PR before and after optimization……..95-103

APPENDIX B;

Figure S1. 3D Schematic representation of L38L↑N↑L PR complexed with **a:** APV, **b:** ATV, **c:** DRV, **d:** IDV, **e:** LPV, **f:** SQV, **g:** RTV, **h:** NFV, and **i:** TPV PIs……..105-106

Figure S2. The structures of the FDA approved HIV PR inhibitors..... 106

Figure S3. Electrostatic and hydrogen bon interactions plots of L38L↑N↑L HIV PR respectively with the various FDA approved HIV PIs ……...…….108-109

Figure S4. Hydrogen bond distances between the hydroxyl groups of the selected FDA approved ATV drug with the catalytic ASP25 and ASP25' residues of L38L↑N↑L PR before and after optimization …….110-113

LIST OF TABLES

CHAPTER THREE;

Table 1. The binding free energies (kcal mol^{-1}) for HIV PIs against Subtype B and C-SA obtained by ONIOM (B3LYP/6-31G(d):AMBER). 49

Table 2. The binding free energies, enthalpies and entropy of the various FDA approved HIV PIs against subtype B PR. 53

CHAPTER FOUR;

Table 1. The binding free energies (kcal mol^{-1}) for FDA HIV PIs against C-SA and L38L↑N↑L PRs (reported experimental results and calculated ONIOM results) ……... 70

Table 2. The binding free energies, enthalpies and entropy of the various FDA approved HIV PIs against L38L↑N↑L PR. …….……. 75

APPENDIX A;

Table S1. The Free energies, Entropy and Enthalpies contributions of FDA approved HIV PIs against subtype C-SA PR. 104

LIST OF ACRONYMS

AIDs	Acquired Immune Deficiency Syndrome
AMBER	Assisted Model Building with Energy Refinement
ASP	Aspartate
B3LYP	Becke3 Lee-Yang-Parr
C-SA PR	Subtype C-South African Protease
DFT	Density Functional Theory
FDA	Food and Drug Administration
HIV	Human Immune Virus
MM	Molecular Mechanics
ONIOM	Our own N-layered Integrated molecular Orbital and molecular Mechanics
PR	Protease
PRIs	Protease Inhibitors
PDB	Protein Data Bank
QM	Quantum Mechanics
RMS	Root Mean Square

CHAPTER ONE

INTRODUCTION

1.1 Preface

Human immune deficiency virus (HIV) has been a long-standing disease since it was first identified in the early 80's at the Institute Pasteur, Paris by Luc Montagnier[1, 2], before it was further characterized by Robert Gallo and Jay Levy in 1984[3]. This virus infects both human and animals and if not quickly treated/prevented, develops into a deadlier disease known as Acquired immune deficiency syndrome (AIDS)[1, 4, 5]. HIV is transmitted through different means including; (i) mother to child (childbirth), (ii) incision (iii) blood transfusion and (iv) sexual intercourse (which is the most common)[6].

According to statistics provided by the World Health Organization, HIV has claimed the lives of millions of people. An estimate of 36.9 million people was reported to be infected with HIV as of 2015, with women having the greater percentage (17.4 million) and an estimated 2 million new HIV infections occur globally each year. Of this, 69% people are living in sub-Sahara Africa, while in South Africa statistics indicate that 5.26 million people are living with HIV, which is an estimated 10% of the country population[7].

These figures, both worldwide and in South Africa, indicate that scientist continue to be faced with the challenge of finding a cure for the virus. Various research has been carried out by scientists all over the world with the support of the different organizations to find a lasting and more effective drug to totally inhibit this epidemic[5, 8]. This has led to the study of the HIV virus at large.

Intensive study of the life cycle, structure, of the human immune deficiency virus led to the development of unrivaled anti-viral drugs targeting the three (3) important retroviral enzymes; Integrase (IN), Reverse transcriptase (RT), Proteases (PR). The combination therapy (highly active antiretroviral therapy HARRT) that was developed in the 1990's is made up of a NRTI (Nucleoside reverse transcriptase inhibitors), NNRTI (Non-nucleoside reverse transcriptase

inhibitors) and PRIs (Protease inhibitors). This treatment was the first to give HIV/AIDS patients better prospects for the future[9, 10,11].

HIV protease inhibitors (HIV PRIs) have been found to rapidly decrease the spread of viral infections in tractable cells which makes this aspartate enzyme a predominant target in the prevention of the deadly disease known as Acquired Immune Deficiency Syndrome (AIDS)[12, 13]. Protease inhibitor drugs approved by the Food and Drug Administration (FDA) were mainly designed for HIV subtype B[14-17]. The confined data available indicate that these protease inhibitors show less activity against subtype A and C[18]. This may be due to the mutation and late stage maturation in HIV-PR that renders the protease inhibitors inactive[18, 19]. To date, only five of the approved FDA antiretroviral drugs are available in South Africa for the Human immune virus protease (HIV-1 PR C-SA)[20]. Therefore, it is imperative to study the potency of the approved FDA drugs against this strain and its mutants.

Several computational techniques ranging from Thermodynamic integration (TI), Free energy perturbation (FEP), Molecular mechanics Generalized Born surface area (MM-GBSA), linear response (LR), Molecular mechanics Poisson-Boltzmann surface area (MM-PBSA), are widely known for the investigation of binding energies of HIV protease―ligand complexes[18, 21, 22].

A great number of studies have been done using Our Own N-layered Integrated molecular Orbital and molecular Mechanics (ONIOM) process and it was found to be applicable to large molecules in different research areas[23, 24]. These includes; reaction mechanisms for organic system, determining binding interaction energies, enzyme reactions, surface reaction and cluster models of surfaces, and reactivity of organic and organometallic compounds[25-28].

In this study, we investigated the inhibitory activity of the current FDA drugs with respect to their binding affinities against subtype B, C-SA and L38L↑N↑L HIV-1 PRs using a ONIOM model.

1.2 HIV Virus

The Human immune deficiency virus belongs to the family of retrovirus (*lentivirus*)[1,8] and can be categorized into two types; HIV-1, which is similar to Simian immunodeficiency virus (SIV). It

is the most common and found worldwide, while HIV-2, is frequent in Western and Southern Central Africa and also Western India[4]. These types of HIV can be further classified into subtypes, sub-subtypes and recombinant forms[29, 30]. Of the ten subtypes of HIV-1 virus, subtype C is prevalent in Africa, mostly South Africa, while subtype A is found to be common in Central Africa, and Australia, Western Europe, and North America majorly account for subtype B HIV type 1[18, 29-31]. The major difference between the different HIV subtypes is the natural amino acids polymorphisms at various positions, but the mechanisms in which they operate remain the same[32].

The structure of the Human immune virus type 1 (**Figure 1)** is 100 nM in width, and the viral particle is enveloped with a lipid bilayer diaphragm, that cut across the uppermost layer of gp120 (glycoprotein) and binds along gp41 (transmembrane proteins). Matrix shell protein (p17) as shown in figure1 forms the linear surface directly underneath the lipid bilayer membrane. The p24 (capsid matrix) form a cone-shaped positioning that encompasses the p7 and p9 gag (nucleocapsid proteins), integrase, reverse transcriptase and protease[1, 33]. The HIV virus genes comprise of two lone RNA (Ribonucleic acid) strands that are subsequently transcribed into DNA (Deoxy-ribonucleic acid) by viral transcriptase when it penetrates the human immune cells. Susceptible cells such as macrophages, dendritic cells, and $CD4^+$ T cells are the major human cells the HIV virus infects[34].

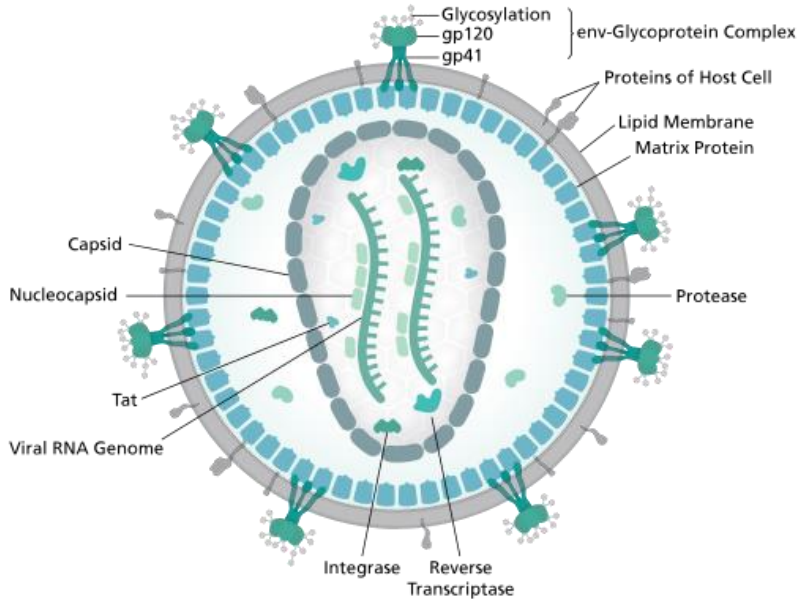


Figure 1. Diagram of HIV genome[35].

1.2.1 HIV-1 Life-cycle

The absence of antiretroviral drugs in HIV-infected patients leads to 3 phases; the acute phase is the first stage and this is when the viral replication is very fast, the chronic phase (second stage), takes a very long time estimated at a period of 9-10 years. The last stage is full-blown AIDS, and in this stage, the immune system has totally lost control over the viral replication[36].

When the HIV virus enters a human cell a process begins by interacting with the host cells CD4+ protein and co-receptor CCRS (CC Chemokine receptor 5)[37], that is recognized by a viral envelope gp120 (glycoprotein) or gp41 (transmembrane protein) in the immune system[38] (**Figure 2**). Once HIV is manifested in the T-cell, 3 enzymes; Reverse Transcriptase, Integrase Transcriptase, and Protease are important in the viral replication process[36].

The viral genetic material from RNA enters the cytoplasm and reverse transcribed into a single DNA before it is further transcribed to a double stranded deoxy-ribonucleic acid (ds DNA). These series of actions are subjected to an immense number of errors and leads to the emergence of mutations for each viral generation produced. As the double-stranded DNA viral genome enters the nucleus of the cell, by means of reverse integrase, becomes the host genome[1, 32, 39]. The provirus DNA is again transcribed into RNA and further transcribed into polyproteins that

comprise of *env*, *gag* and *gag-pol* precursors[8, 19]. Polyproteins and immature RNA virus are assembled into viral particles, the protease helps in cleaving the polyproteins to more functional protein building blocks, which is required for the mature virus[36, 40-42]. Each enzymatic stage can potentially be interfered with, thereby slowing down or stopping the replication of the virus, these essential enzymes are therefore potential drug targets. Zidovudine was the first successful synthesized drug for the treatment of cancer in 1964 but was later found to be effective against HIV in 1987, it was on this basis that other drugs were then designed and synthesized[5].

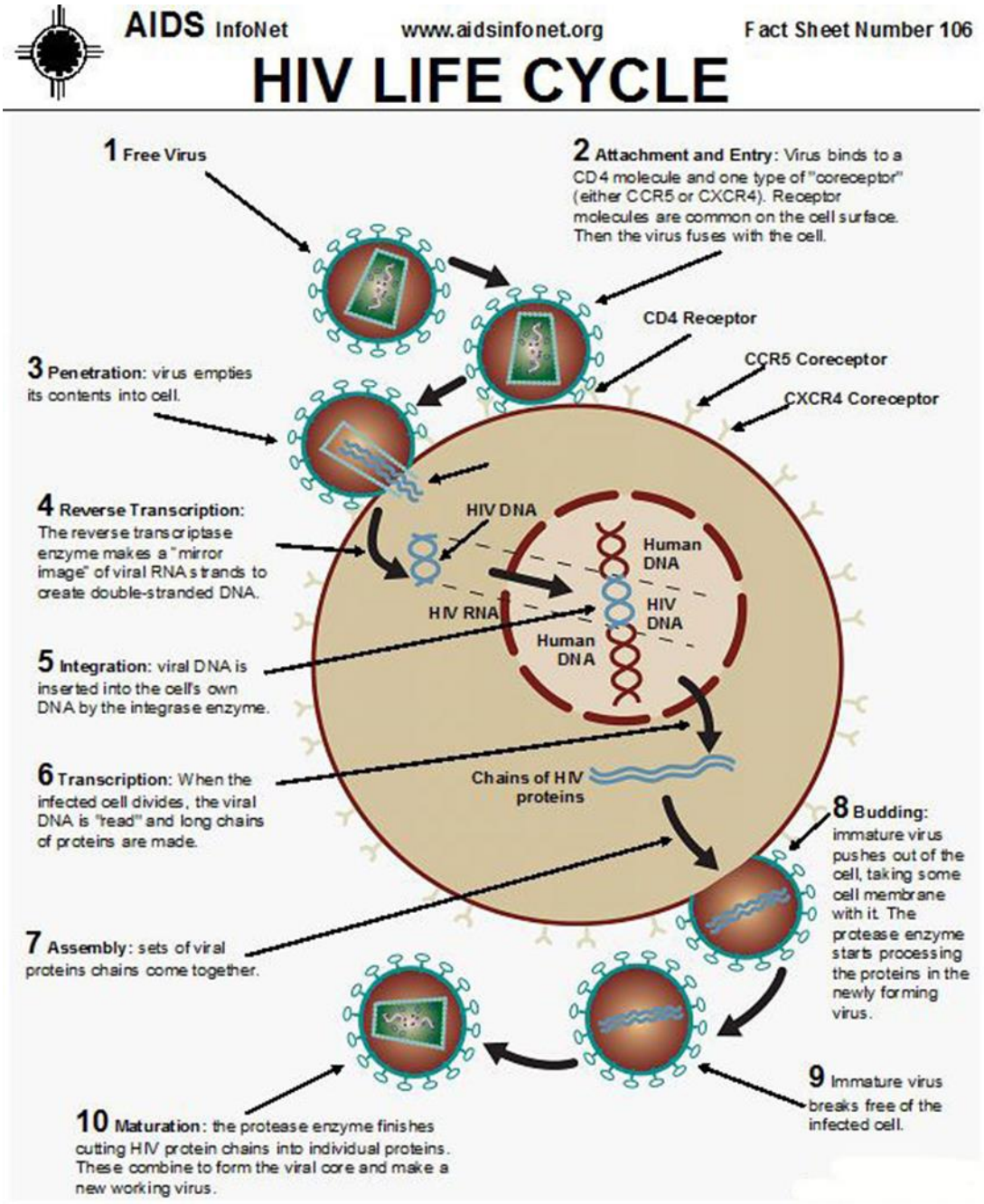


Figure 2. Schematic representation of HIV Life cycle[43]

1.3 HIV Inhibitors

Inhibitors bind onto enzymes and block or inhibit their activities[44]. Enzymes inhibitors are categorized based on their mechanism of action and it can either be specific or non-specific[45].

Specific enzyme inhibition can be further classified into irreversible and reversible inhibitors. While non-specific inhibition is irreversible in nature and it binds strongly to the enzyme and it is very unlikely to dissociate from it, only by extreme pH condition or high temperature and this denatures the enzyme[45].

1.3.1 Irreversible Inhibitors

Irreversible inhibitors covalently bind with an enzyme and because of their high affinity, once bound they do not dissociate easily from the enzyme. They can also form a non-covalent binding interaction with the active site of an enzyme[45].

1.3.2 Reversible Inhibitors

The defining properties of these inhibitors is the ease with which they non-covalently bind and can dissociate from the enzymes under certain conditions[45-47]. A discussion of reversible inhibitors is provided next.

1.3.2.1 Competitive Inhibitors

These inhibitors resemble the structure of natural substrates that bind in the active site and because of the similarity in structure, these inhibitor bind exactly in the same location as the substrates[44]. Therefore, competitive inhibitors as the name implies, will compete with the substrate for the active site and due to higher affinity than the substrate and once bound prevents the substrate binding to the active site[45, 47, 48].

1.3.2.2 Uncompetitive Inhibitors

Uncompetitive inhibitors bind in the pocket created by a change in enzyme conformation of substrate binding to the active site. This forms an enzyme-substrate inhibitor (ESI) complex which essentially blocks the activity of the enzyme[45, 47].

1.3.2.3 Non-competitive Inhibitors

Some enzymes have a permanent allosteric site (non-substrate binding site) that is generally different from the active site[48]. This allows binding of inhibitors to the allosteric site, regardless of if the enzyme is free or a substrate is bound to it. These inhibitors are non-competitive because they do not compete for the active site with the natural substrate and thus block the enzymatic activity of the enzyme[44, 47].

1.3.2.4 Mixed Inhibitors

The inhibitor in this type of inhibition not only compete for the active site with the substrate but also binds unto a free enzyme molecule which further binds a substrate or binds to a substrate-enzyme complex and reduces catalytic turnover[45-47].

1.4 Types of HIV Inhibitors

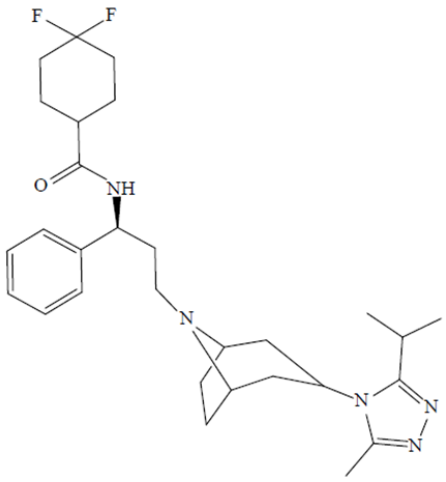
Different inhibitors at different stages of the HIV life cycle have proved to be efficacious in inhibiting enzymes responsible for reproducing the catalytic activity of the virus, acting as potential drug targets[49].

1.4.1 Entry Inhibitors

Entry inhibitors inhibit virus from entering and transmitting to the host cell. These inhibitors are broadly categorized into; co-receptor antagonist, fusion inhibitors, and attachment or binding inhibitors[1, 38].

1.4.1.1 Co-receptor Antagonists (CRAs)

The T-cell consists of two co-receptor namely CCR5 and CXCR4 and this inhibits the entry of the HIV virus that interacts with the co-receptor[50]. A principal drug that targets the C-C chemokine receptor type 5 (CCR5) co-receptor includes monoclonal antibody and Pro140 which prevents HIV type1 virus by engaging the glycoprotein (gp120) binding spot on CCR5[51]. Co-receptor antagonist capability to block HIV infection has been displayed in animal models and cell culture systems. An early example of a CCR5 antagonist (**Figure 3**) that was approved in 2007 is Maraviroc[52, 53].



Maraviroc (2003)

Figure 3. Structure of a CCR5 antagonist inhibitor[54]

1.4.1.2 Fusion Inhibitors (FIs)

Enfuvirtide was approved by the FDA in 2003 and so far is the first and only therapeutic administered fusion drug[55], the structure is a linear chain of 36-amino acids that corresponds to 127-162 residues of viral gp14 (**Figure 4**). This type of inhibitor blocks the change in formation of the glycoprotein (gp14) that is involved in the integration of the virus and host cell membrane[50].

Ac-Tyr-Thr-Ser-Leu-Ile-His-Ser-Leu-Ile-Glu-Glu-Ser-Gln-Asn-Gln-Gln-Glu-Lys-Asn-Glu-GlnGlu-Leu-Leu-Glu- Leu-Asp-Lys-Trp-Ala-Ser-Leu-Trp-Asn-Trp-Phe-NH2

Enfuvirtide (2003)

Figure 4. Fusion inhibitor sequence[56]

1.4.1.3 Binding or Attachment Inhibitors (AIs)

Attachment inhibitors interrupt the inceptive binding of the HIV-type1 glycoprotein 120 exterior to the CD4 T-cell receptor by one of the various feasible mechanisms[50]. **Figure 5** presents samples of clinically successful attachment drugs[51, 57].

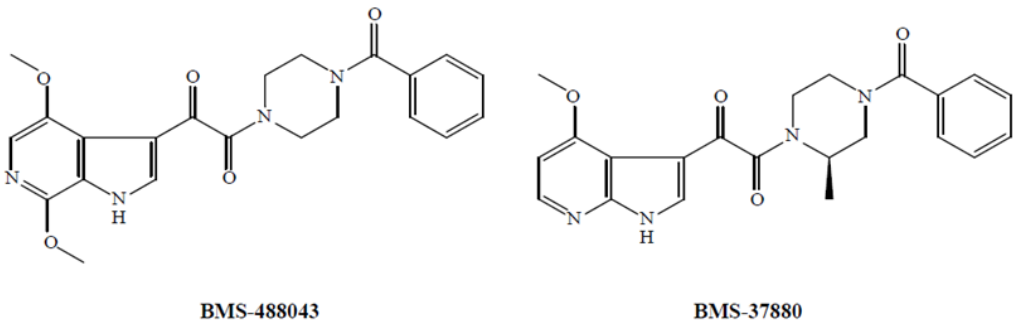


Figure 5. Schematic representation of binding inhibitors[51]

1.4.2 Reverse Transcriptase Inhibitors

The RTIs (that is, 'non-nucleoside reverse transcriptase' (NNRT) and nucleoside reverse transcriptase (NRT) inhibits reverse transcriptase enzymatic purpose and intercept the synthesis completion of the double-stranded viral DNA. Thus restrain human immune virus from multiplying[36, 58].

1.4.2.1 Nucleoside Reverse Transcriptase Inhibitors

NRTIs compete with the natural deoxynucleotides substrates even though they lack the 3-hydroxyl group on the deoxyribose moiety. Hence, the addition of NRTIs during DNA transcription prevents further covalent 5'-3' phosphodiester bond that is formed by the next incoming deoxynucleotide. Thus, resulting in viral transcription chain termination[59, 60]. Zidovudine (AZT) was the first nucleoside reverse transcriptase inhibitors (NRTIs) drug approved in 1987 for HIV treatment and presently, there are seven FDA approved NRTIs drugs in the market[11]. Some examples are shown in **Figure 6**.

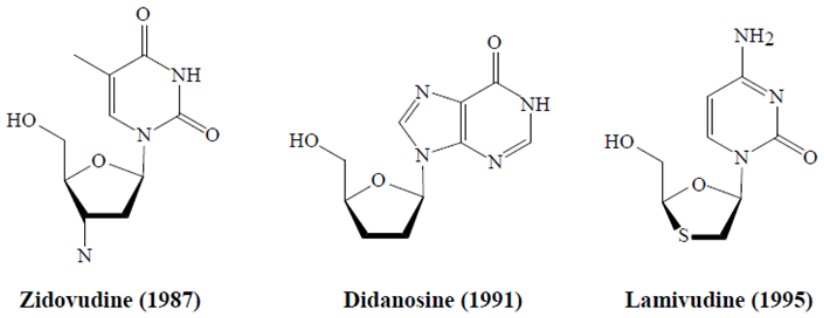


Figure 6. Structure of FDA approved NRTI drugs[11]

1.4.2.2 Non-nucleoside Reverse Transcriptase Inhibitors

NNRTIs, in contrast inhibits reverse transcriptase by binding at the allosteric site (non-substrate binding site) specifically different from the active site. Thus, blocking the movement of protein domains of HIV-1 RT-DNA viral replication[42, 61]. The Food and Drug Administration (FDA) approved Five NNRTIs drugs from 1996-2011[11] (**Figure 7**).

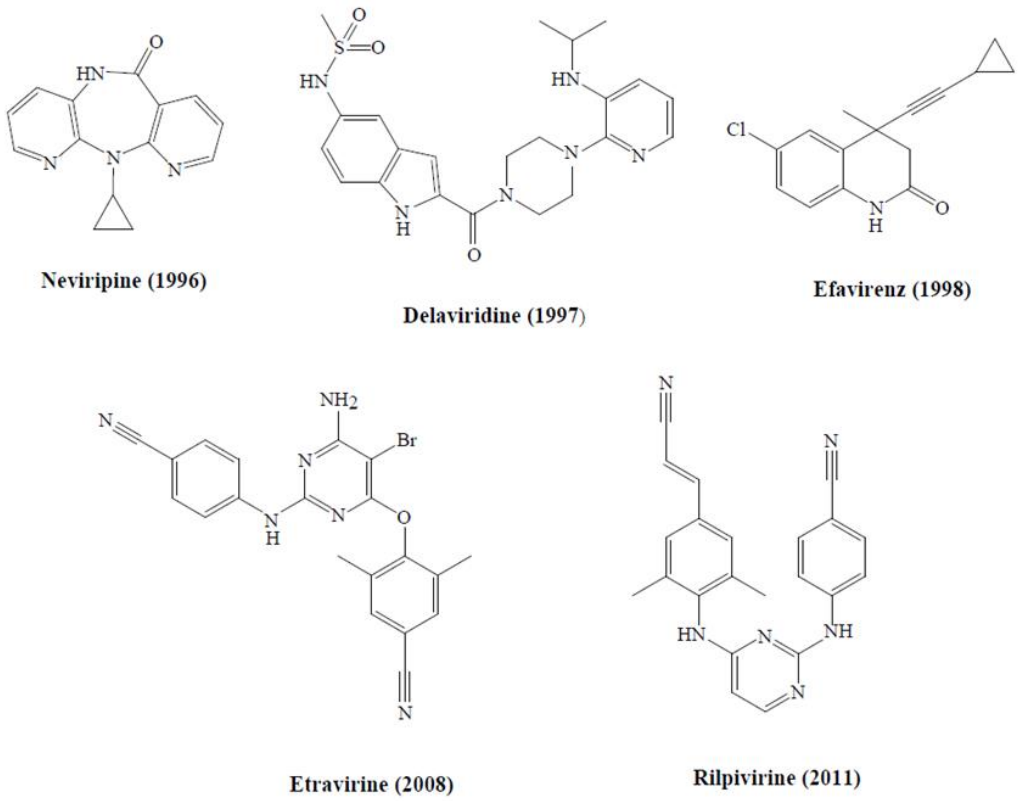
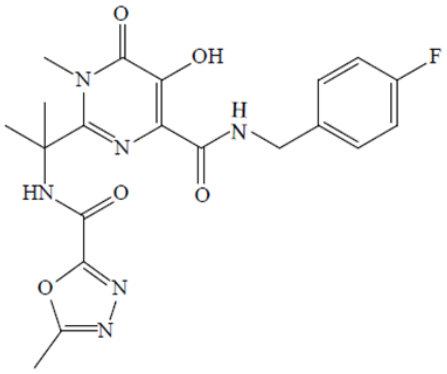


Figure 7. Structure of FDA approved NNRTI drugs[11]

1.4.3 Retroviral Integrase Inhibitors

Integrase inhibitors belong to a class of antiretroviral drug, designed to inhibit the activity of that enzyme. This viral enzyme transmits the viral genome into the DNA of the host cell and since integration is also an important viral phase in retroviral replication, inhibiting it can prevent further spreading of the virus[1, 36]. Development of HIV-1 RIs is lagging behind owing to the fact that the structure of the retroviral integrase has not been fully elucidated. In 2007 the FDA approved Raltegravir (**Figure 8**) as the first integrase inhibitor[62].

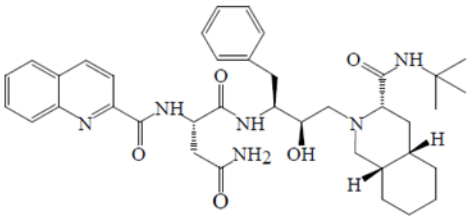


Raltegravir (2007)

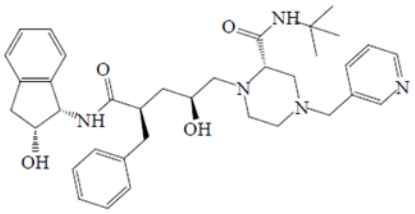
Figure 8. Schematic representation of approved FDA HIV-1 integrase inhibitor[62]

1.4.4 Protease Inhibitors

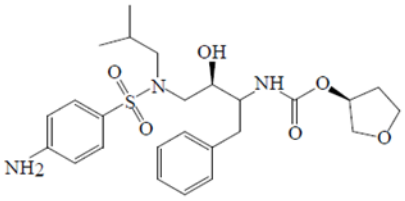
HIV protease inhibitors are peptide-like chemicals that competitively inhibits the cleavage of gag and gag-pol from transmitting into a functional virus[30, 63, 64]. The FDA approved protease inhibitor drugs include; Saquinavir (is the first and most widely administered drug)[65], Indinavir, Tipranavir, Amprenavir, Nelfinavir, Lopinavir, Ritronavir and Atazanavir[18, 66] (**Figure 9)**. Human immune virus protease inhibitors (HIV-1 PRIs) advantages include; better viral suppression, prolonged viral control, reduction of death rates. These PR inhibitors have been reviewed intensively and they show antiviral, pharmacodynamics, pharmacokinetics (which includes; absorption, distribution, metabolism and excretion of drugs) properties[67-69]. Therefore, HIV treatment is worthwhile for patients infected with HIV virus[32, 36]. Although, some setbacks have been drawn from mutations and polymorphisms that occur within the protease and this has led to the development of combined antiretroviral therapy (cART), where HIV-1 PR is the top target[30, 63, 70].



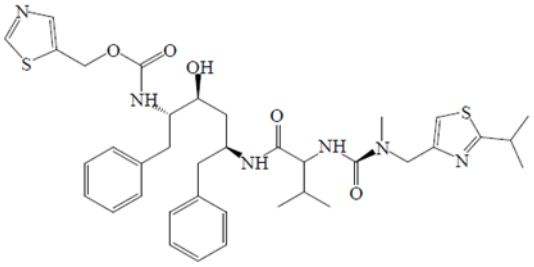
Saquinavir (1995)



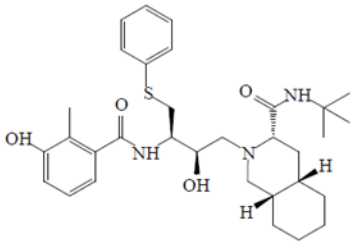
Indinavir (1996)



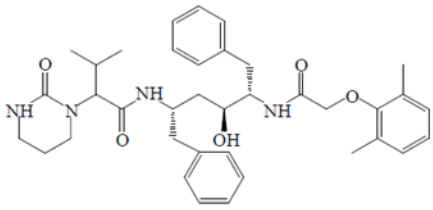
Ampenavir (1999)



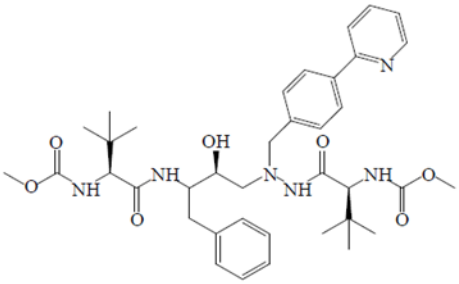
Ritonavir (1996)



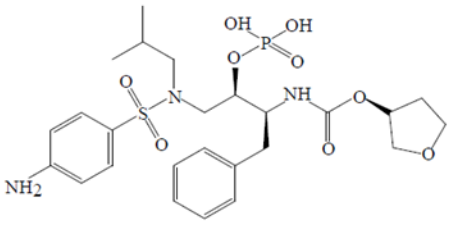
Nelfinavir (1997)



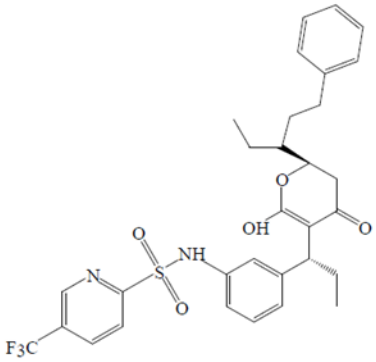
Lopinavir (2000)



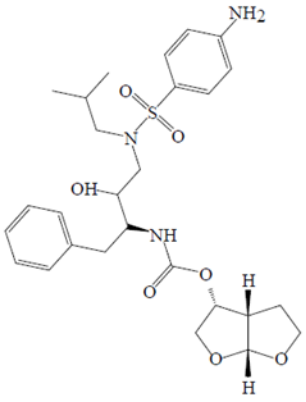
Atazanavir (2003)



Fosamprenavir (2003)



Tipranavir (2005)



Darunavir (2006)

Figure 9. Structures of HIV-1 approved protease inhibitors[18, 71, 72]

1.5 HIV-1 Protease

Human immune virus type 1 protease, belong to the family of aspartyl protease[26, 29], and since it plays a key role in HIV virus mutations, it has become an important target for antiretroviral drugs[64, 73]. The HIV-1 PR exists as an identical active C2-symmentric homodimer with 99 amino acids in each dimer[65, 66]. The active sites consist of Asp25-Thr26-Gly27 triad at each half enclosed by a two beta-sheets flaps[4, 30]. Asp25 and 25' of the active site interact directly with the inhibitors and substrates, while the role of the other amino acids (Thr26, 26', Gly27, 27') is still largely unknown[39]. The nucleotide sequence of HIV-1 PR has been observed to change from generations due to high error rates of HIV reverse transcriptase[8]. However, the catalytic triad residues of HIV PR, Asp25-Thr26-Gly27, is well conserved and it is similar in all HIV subtypes[29, 30, 74].The protease enzyme has three domains as shown in **Figure 10** and will be discussed in the next section.

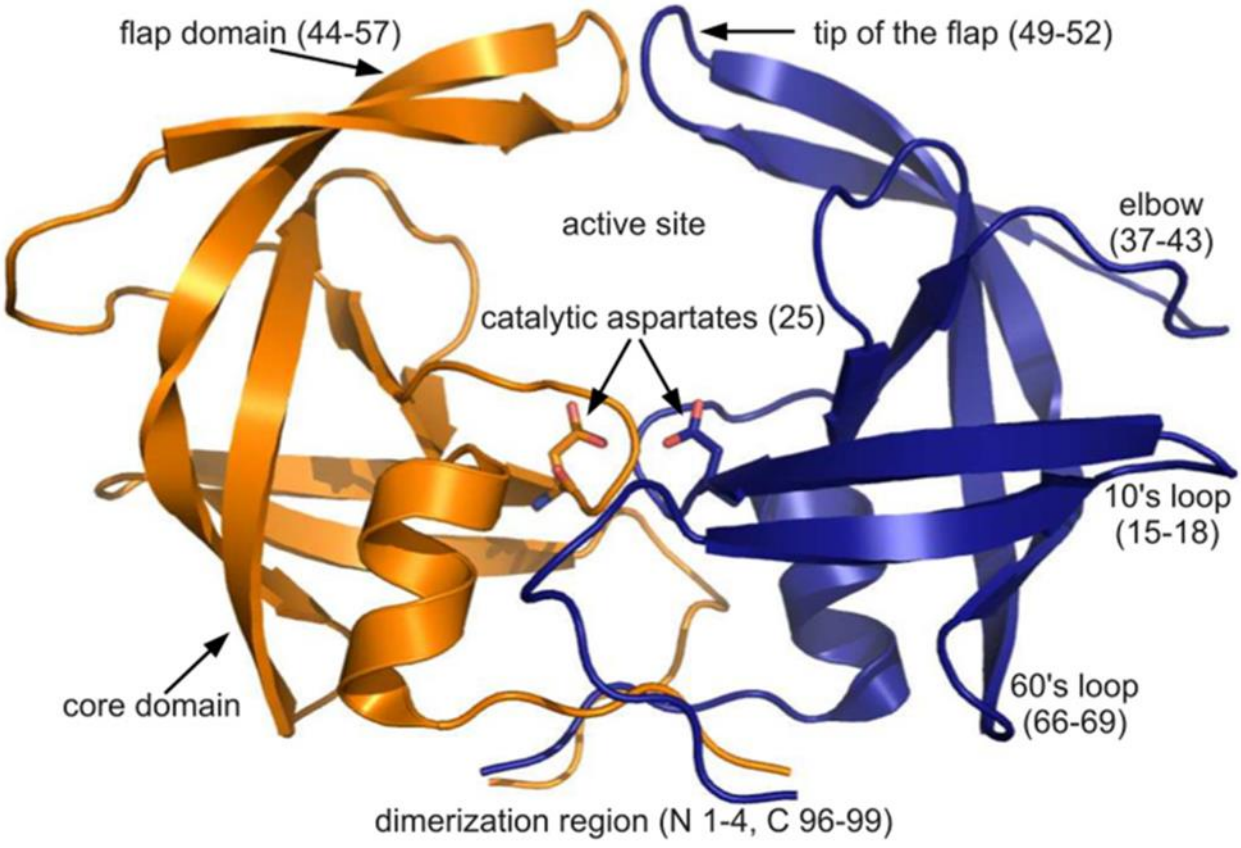


Figure 10. Ribbon representation of HIV-1 protease[75]

1.5.1 Domains in HIV-1 PR Dimer

(i) The core domain or active site cavity

(ii) The terminal domain or dimerization domain and

(iii) The flap domain

1.5.1.1 Core Domain (active site cavity)

The active site cavity is important in the stabilization of the dimer as well as the catalytic site. This domain is composed of four beta-strand structures. The catalytic tripeptide Asp25-Thr26-Gly27 is situated at the core domain interface from the monomers[8, 76, 77].

1.5.1.2 Terminal Domain (dimerization domain)

This domain plays a vital role in the stabilization of active site and dimer formation. The terminal domain is composed of β-antiparallel sheets that comprise of four termini dimer residues 1-4 and 95-99, the helix residues 86-94, and a turn encompassing residues 4-9 of each monomer[78].

1.5.1.3 Flap Domain

The flap domain consists of two β-hairpin structures[30], 43-58 amino acids residues found in both monomers[77, 79]. Studies have revealed that the HIV-1 PR exhibit structural arrangement from closed form, semi-open form and wide-open form, which covers the active site and this plays a vital role in the ligand—substrate binding interaction[80, 81]. Though the HIV protease has been crystallized as a single monomer (PDB code: 1HHP)[82], it functions as a dimeric form. The dimer is established due to intermolecular forces contributing to its stability and it is observed in complexes of the HIV-1 PR with inhibitors[82].

1.5.2 Dimer Stability of HIV-1 PR

HIV-1 PR dimer stabilization is attained by factors which include side chains hydrophobic interactions, non-covalent interactions, and interactions including the catalytic residues[78]. A network of hydrogen bonds referred to as "firearm" is produced by the PR catalytic residue 'Asp25-Thr26-Gly27' triad[80, 83] and it is depicted in (**Figure 11**).

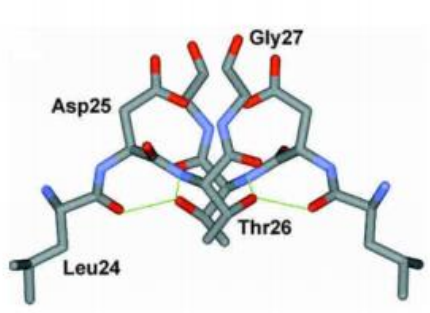


Figure 11. Schematic representation of the catalytic residues triad Asp25-Thr26-Gly27[83]

1.5.3 Substrate Binding

The active site cavity is composed of distinct subsites, where the substrates side chains are accommodated upon binding. The defined subsites are named S_1 to S_n and S_1' to S_n' for each monomer and while the corresponding amino acids side chain of the substrate that is cleaved during hydrolysis are named P_1 to P_n and P_1' to P_n', starting from the scissile bond[84] (**Figure 12**). However, the protease subsites S_1 and S_1' are the same, as a result of the C2 symmetric nature and this holds for all other subsites[8, 64, 85].

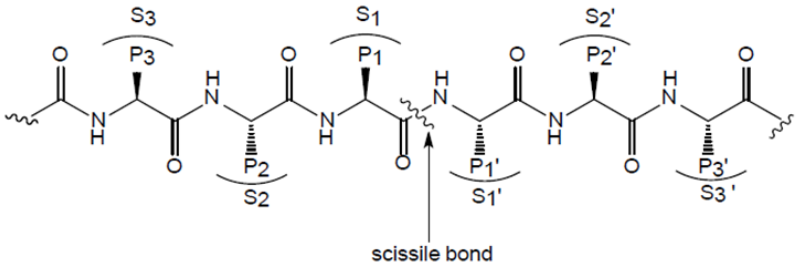


Figure 12. P_1___P_n, $P_{1'}$____$P_{n'}$_is a standard nomenclature used in representing amino acids of peptide residues, while protease binding sites are represented by S_1___S_n, $S_{1'}$____ $S_{n'}$_[8, 86].

1.5.4 HIV-1 PR Catalytic Mechanism

Despite all that was investigated about HIV-1 PR and the importance of the protease family, the mechanism by which it catalysis the proteolysis of the peptide remains a formidable challenge[13, 19, 74]. Experimental studies have shown several factors such as; the catalytic aspartate protonation state and the role of structural water, and also the identification of the Asp25-Thr26-Gly27 catalytic triad [13, 19]. Based on structural observations of different Aspartic protease complexes,

several distinct mechanisms have been proposed[74]. This falls generally into two categories: general Acid-Base (GA-GB) catalytic mechanism and catalysis through a covalently bonded intermediate[13] and it is represented in **Figure 13**. Although there is no definitive proof yet as to which mechanism is right, most studies indicate a general acid-base reaction mechanism for peptide cleavage bond[74, 87]. The peptide carbon hybridization changes from sp^2 to sp^3 caused by a nucleophilic attack from water molecule at the scissile bond[74, 87-90]. The un-protonated ASP-PR which acts as the general base helps in the ionization of the water molecule and becomes protonated, the resulting hydroxyl ion cleaves and bond covalently with the peptide carbon, initiating the carbon-nitrogen (peptide) bond breakage[87, 88].

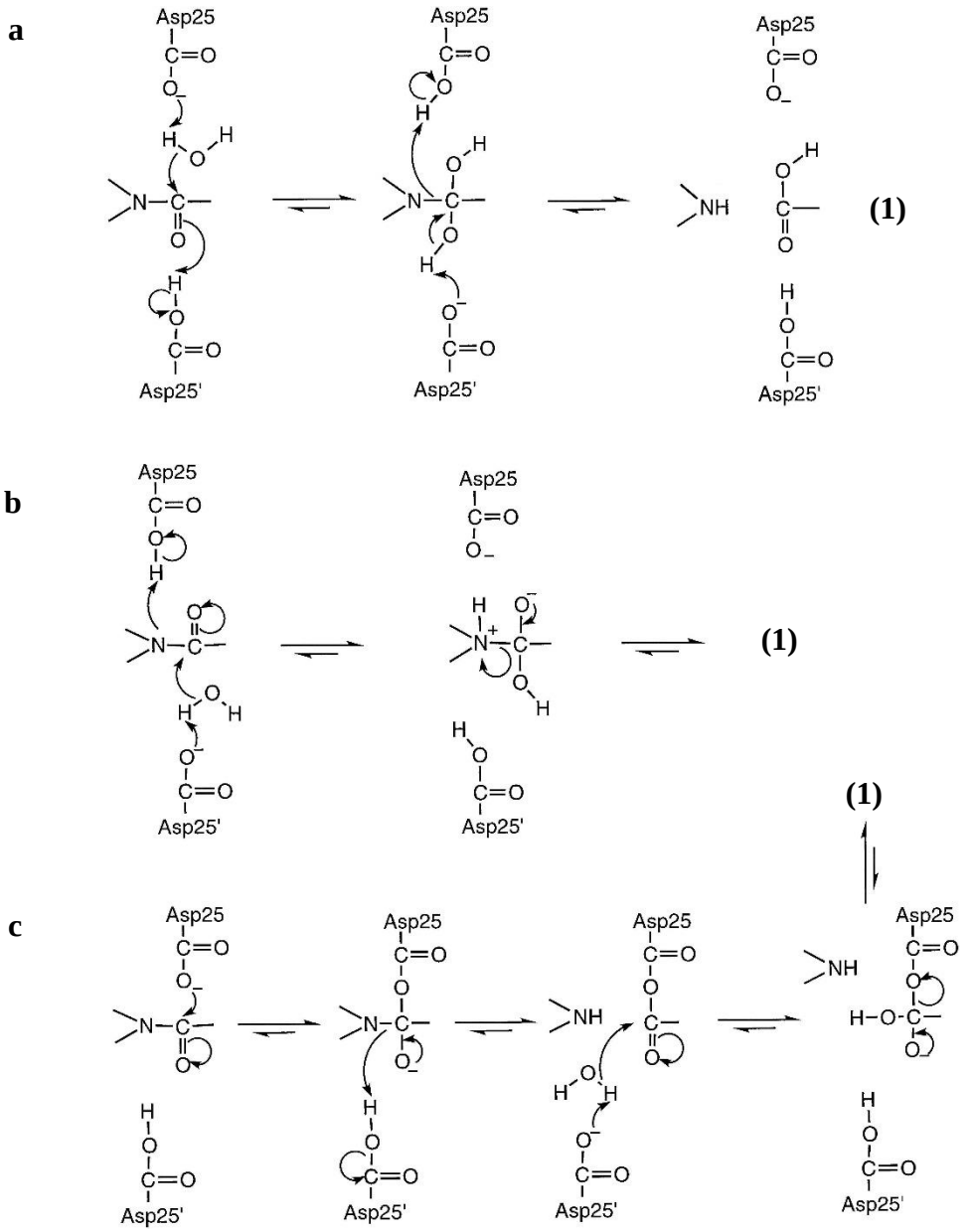


Figure 13. General HIV-1 PR catalytic mechanisms (a) Acid-base general mechanism with water, (b) Acid-base reaction with a zwitterion molecule, (c) Direct nucleophilic reaction with covalently-bonded intermediate[8, 13, 88].

1.6 HIV-1 PR Mutations

Over time, multiple mutations of the HIV-1 PR have developed that show resistance to the available anti-retroviral drugs[91]. The lack of proofreading of the reverse transcriptase and high viral replication rate contributes significantly to a build-up of drug resistance of this protease[92].

The insertion, substitution, and deletion that occur in the amino acids of the virus sequence are attributed to mutations that arise through resistance[31]. The structure of the active site in the wildtype HIV PR is influenced by these mutations, which then diminishes the efficacy of the protease inhibitors[93-97].

The structure of subtype B differs from the subtype C-SA PR as a result of a mutation that occur in eight different positions away from the active pocket of the wild-type subtype B[31]. Likewise, two new mutations have been found to occur from the C-SA HIV PR designated I36T↑T and L38L↑N↑L PRs[98]. The former (I36T↑T) has a 100 amino acids in each monomer and the mutation occur at position 36 from isoleucine to threonine, with an insertion at the second residue of threonine. While, the later (L38L↑N↑L) with an insertion in the asparagine and leucine at position 38 has 101 amino acids in each monomer was reported[98, 99]. The protease has five mutations E35D, N37S, I36G, D60E and M46L and was discovered in a drug-naïve patient (with respect to PIs)[98].

1.7 Objectives of this Study

The primary scope of this study is to test and improve a pre-existing computational model to correctly reflect the experimental data in terms of binding interactions of available HIV PR inhibitors. This is accomplished by the following objectives;

- To create a computational chemistry model using a ONIOM (two-layer) method to calculate theoretical binding free energies. These results will be compared with reported experimental binding energies of FDA approved HIV protease inhibitors against subtype B and C-SA HIV PR.
- To test the performance of the same ONIOM computational model for the C-SA HIV PR mutant designated L38L↑N↑L PR. The theoretical binding free energies will be compared with our recently reported experimental binding energies of the FDA approved PR drugs to this mutant.

1.8 Thesis Outline

This dissertation continues as follows:

A computational study background is presented in **Chapter 2**.

ONIOM theoretical model to calculate the binding interactions and reflect the experimental binding energies of FDA approved HIV-1 PR inhibitors against subtype B and C-SA PR is the focus in **Chapter 3**.

The tested computational chemistry model (ONIOM) in investigating the binding interactions and also reflect experimental binding energies of PR inhibitors against C-SA HIV PR mutant L38L↑N↑L PR is discussed in **Chapter 4**.

Summary of the dissertation is followed in **Chapter 5**.

References

[1] Shehu-Xhilaga, M., and Oelrichs, R. (2009) Basic HIV virology, *HIV Management in Australasia*, 9-18.

[2] Klatzmann, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., Gluckman, J.-C., and Montagnier, L. (1983) T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV, *Nature 312*, 767-768.

[3] Gallo, R. C., Sarin, P. S., Gelmann, E., Robert-Guroff, M., Richardson, E., Kalyanaraman, V., Mann, D., Sidhu, G. D., Stahl, R. E., and Zolla-Pazner, S. (1983) Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS), *Science 220*, 865-867.

[4] Kar, P., and Knecht, V. (2012) Origin of decrease in potency of darunavir and two related antiviral inhibitors against HIV-2 compared to HIV-1 protease, *The Journal of Physical Chemistry B 116*, 2605-2614.

[5] Vangelista, L., Secchi, M., and Lusso, P. (2008) Rational design of novel HIV-1 entry inhibitors by RANTES engineering, *Vaccine 26*, 3008-3015.

[6] Glasier, A., Gülmezoglu, A. M., Schmid, G. P., Moreno, C. G., and Van Look, P. F. (2006) Sexual and reproductive health: a matter of life and death, *The Lancet 368*, 1595-1607.

[7] HIV/AIDS, J. U. N. P. O., and HIV/AIDS, J. U. N. P. O. (2016) Global AIDS update 2016, *Geneva, Switzerland*.

[8] Brik, A., and Wong, C.-H. (2003) HIV-1 protease: mechanism and drug discovery, *Organic & Biomolecular Chemistry 1*, 5-14.

[9] Rodríguez-Barrios, F., and Gago, F. (2004) HIV protease inhibition: limited recent progress and advances in understanding current pitfalls, *Current Topics in Medicinal Chemistry 4*, 991-1007.

[10] Lenhard, J. M., Croom, D. K., Weiel, J. E., and Winegar, D. A. (2000) HIV protease inhibitors stimulate hepatic triglyceride synthesis, *Arteriosclerosis, Thrombosis, and Vascular Biology 20*, 2625-2629.

[11] Pomerantz, R. J., and Horn, D. L. (2003) Twenty years of therapy for HIV-1 infection, *Nature Medicine 9*.

[12] Kempf, D. J., Norbeck, D. W., Codacovi, L., Wang, X. C., Kohlbrenner, W. E., Wideburg, N. E., Paul, D. A., Knigge, M. F., and Vasavanonda, S. (1990) Structure-based, C2 symmetric inhibitors of HIV protease, *Journal of Medicinal Chemistry 33*, 2687-2689.

[13] Chatfield, D. C., Eurenius, K. P., and Brooks, B. R. (1998) HIV-1 protease cleavage mechanism: a theoretical investigation based on classical MD simulation and reaction path

calculations using a hybrid QM/MM potential, *Journal of Molecular Structure: THEOCHEM 423*, 79-92.

[14] Pomerantz, R. J. (2002) HIV: a tough viral nut to crack, *Nature 418*, 594-595.

[15] Tomasselli, A. G., and Heinrikson, R. L. (2000) Targeting the HIV-protease in AIDS therapy: a current clinical perspective, *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology 1477*, 189-214.

[16] Kempf, D. J., Marsh, K. C., Denissen, J. F., McDonald, E., Vasavanonda, S., Flentge, C. A., Green, B. E., Fino, L., Park, C. H., and Kong, X.-P. (1995) ABT-538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans, *Proceedings of the National Academy of Sciences 92*, 2484-2488.

[17] Roberts, N. A., Martin, J. A., Kinchington, D., Broadhurst, A. V., Craig, J. C., Duncan, I. B., Galpin, S. A., Handa, B. K., Kay, J., and Krohn, A. (1990) Rational design of peptide-based HIV proteinase inhibitors, *Science 248*, 358-361.

[18] Ahmed, S. M., Kruger, H. G., Govender, T., Maguire, G. E., Sayed, Y., Ibrahim, M. A., Naicker, P., and Soliman, M. E. (2013) Comparison of the molecular dynamics and calculated binding free energies for nine FDA-approved HIV-1 PR drugs against subtype B and C-SA HIV PR, *Chemical Biology & Drug Design 81*, 208-218.

[19] Mildner, A. M., Rothrock, D. J., Leone, J. W., Bannow, C. A., Lull, J. M., Reardon, I. M., Sarcich, J. L., Howe, W. J., and Tomich, C.-S. C. (1994) The HIV-1 protease as enzyme and substrate: mutagenesis of autolysis sites and generation of a stable mutant with retained kinetic properties, *Biochemistry 33*, 9405-9413.

[20] Meintjes, G., Conradie, J., Cox, V., Dlamini, S., Fabian, J., Maartens, G., Manzini, T., Mathe, M., Menezes, C., and Moorhouse, M. (2014) Adult antiretroviral therapy guidelines 2014, *Southern African Journal of HIV Medicine 15*, 121-143.

[21] Saen-oon, S., Aruksakunwong, O., Wittayanarakul, K., Sompornpisut, P., and Hannongbua, S. (2007) Insight into analysis of interactions of saquinavir with HIV-1 protease in comparison between the wild-type and G48V and G48V/L90M mutants based on QM and QM/MM calculations, *Journal of Molecular Graphics and Modelling 26*, 720-727.

[22] Hensen, C., Hermann, J. C., Nam, K., Ma, S., Gao, J., and Höltje, H.-D. (2004) A combined QM/MM approach to protein-ligand interactions: polarization effects of the HIV-1 protease on selected high affinity inhibitors, *Journal of Medicinal Chemistry 47*, 6673-6680.

[23] Morokuma, K. (2002) New challenges in quantum chemistry: quests for accurate calculations for large molecular systems, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 360*, 1149-1164.

[24] Vreven, T., and Morokuma, K. (2000) On the application of the IMOMO (integrated molecular orbital+molecular orbital) method, *Journal of Computational Chemistry 21*, 1419-1432.

[25] Lundberg, M., Sasakura, Y., Zheng, G., and Morokuma, K. (2010) Case studies of ONIOM (DFT:DFTB) and ONIOM (DFT:DFTB:MM) for enzymes and enzyme mimics, *Journal of Chemical Theory and Computation 6*, 1413-1427.

[26] Promsri, S., Chuichay, P., Sanghiran, V., Parasuk, V., and Hannongbua, S. (2005) Molecular and electronic properties of HIV-1 protease inhibitor C 60 derivatives as studied by the ONIOM method, *Journal of Molecular Structure: THEOCHEM 715*, 47-53.

[27] Ruangpornvisuti, V. (2004) Recognition of carboxylate and dicarboxylates by azophenol–thiourea derivatives: a theoretical host–guest investigation, *Journal of Molecular Structure: THEOCHEM 686*, 47-55.

[28] Dapprich, S., Komáromi, I., Byun, K. S., Morokuma, K., and Frisch, M. J. (1999) A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives, *Journal of Molecular Structure: THEOCHEM 461*, 1-21.

[29] Robbins, A. H., Coman, R. M., Bracho-Sanchez, E., Fernandez, M. A., Gilliland, C. T., Li, M., Agbandje-McKenna, M., Wlodawer, A., Dunn, B. M., and McKenna, R. (2010) Structure of the unbound form of HIV-1 subtype A protease: comparison with unbound forms of proteases from other HIV subtypes, *Acta Crystallographica Section D: Biological Crystallography 66*, 233-242.

[30] Genoni, A., Morra, G., Merz Jr, K. M., and Colombo, G. (2010) Computational study of the resistance shown by the subtype B/HIV-1 protease to currently known inhibitors, *Biochemistry 49*, 4283-4295.

[31] Mosebi, S., Morris, L., Dirr, H. W., and Sayed, Y. (2008) Active-site mutations in the south african human immunodeficiency virus type 1 subtype C protease have a significant impact on clinical inhibitor binding: kinetic and thermodynamic study, *Journal of Virology 82*, 11476-11479.

[32] Deeks, S. G., Smith, M., Holodniy, M., and Kahn, J. O. (1997) HIV-1 protease inhibitors: a review for clinicians, *Jama 277*, 145-153.

[33] Vacca, J. P., and Condra, J. H. (1997) Clinically effective HIV-1 protease inhibitors, *Drug Discovery Today 2*, 261-272.

[34] Turner, B. G., and Summers, M. F. (1999) Structural biology of HIV, *Journal of Molecular Biology 285*, 1-32.

[35] Thomas, S., (2014) Diagram of HIV virion. Available at:https://upload.wikimedia.org/wikipedia/commons/thumb/5/5e/HI-virion-structure_en.svg/2000px-HI-virion-structure_en.svg.png [26 June 2016].

[36] Bon, D., Stephan, C., Keppler, O., and Herrmann, E. (2012) Viral dynamic model of antiretroviral therapy including the integrase inhibitor raltegravir in patients with HIV-1, *BIOMATH 1*, Article ID: 1209251.

[37] Feinberg, M. B. (1996) Changing the natural history of HIV disease, *The Lancet 348*, 239-246.

[38] Seibert, C., Ying, W., Gavrilov, S., Tsamis, F., Kuhmann, S. E., Palani, A., Tagat, J. R., Clader, J. W., McCombie, S. W., and Baroudy, B. M. (2006) Interaction of small molecule inhibitors of HIV-1 entry with CCR5, *Virology 349*, 41-54.

[39] Mager, P. P. (2001) The active site of HIV-1 protease, *Medicinal Research Reviews 21*, 348-353.

[40] Mitsuya, H., Yarchoan, R., and Broder, S. (1990) Molecular targets for AIDS therapy, *Science 249*, 1533-1544.

[41] Katz, R. A., and Skalka, A. M. (1994) The retroviral enzymes, *Annual Review of Biochemistry 63*, 133-173.

[42] De Clercq, E. (1998) The role of non-nucleoside reverse transcriptase inhibitors (NNRTIs) in the therapy of HIV-1 infection, *Antiviral Research 38*, 153-179.

[43] InfoNet, T. A., (2013) HIV Life Cycle [ONLINE]. Available at:http://www.aidsinfonet.org/img/uploaded/image/106-life-cycle2013.jpg [30 october, 2016].

[44] Roberts, M., Reiss, M. J., and Monger, G. (2000) Advanced biology, *Nelson Thornes, UK*.

[45] Rakesh, S. (2012) Enzyme inhibition: mechanism and scope, enzyme inhibition and bioapplications. *DOI: 10.5772/39273*.

[46] Berg, J., Tymoczko, J., Stryer, L., and Stryer, L. (2002) Biochemistry, *5th Edition, WH Freeman, New York*.

[47] Zhang, Z., Gan, J., Zhang, D., Zhu, M., and Humphreys, W. (2007) Protocols for assessment of in vitro and in vivo bioactivation potential of drug candidates, *Drug Metabolism in Drug Design and Development: Basic Concepts and Practice*, 447-476.

[48] Wong, C.-H., and Whitesides, G. M. (1994) Enzymes in synthetic organic chemistry, *Vol. 12, Academic Press, Massachusetts*.

[49] Makatini, M. M. (2011) Design, synthesis and screening of novel PCU-peptide/peptoid derived HIV protease inhibitors, *University of KwaZulu-Natal, Westville*.

[50] Whitcomb, J. M., Huang, W., Fransen, S., Limoli, K., Toma, J., Wrin, T., Chappey, C., Kiss, L. D., Paxinos, E. E., and Petropoulos, C. J. (2007) Development and characterization of a novel single-cycle recombinant-virus assay to determine human immunodeficiency virus type 1 coreceptor tropism, *Antimicrobial Agents and Chemotherapy 51*, 566-575.

[51] Ho, H.-T., Fan, L., Nowicka-Sans, B., McAuliffe, B., Li, C.-B., Yamanaka, G., Zhou, N., Fang, H., Dicker, I., and Dalterio, R. (2006) Envelope conformational changes induced by human immunodeficiency virus type 1 attachment inhibitors prevent CD4 binding and downstream entry events, *Journal of Virology 80*, 4017-4025.

[52] Shaheen, F., and Collman, R. G. (2004) Co-receptor antagonists as HIV-1 entry inhibitors, *Current Opinion in Infectious Diseases 17*, 7-16.

[53] Zhang, C., Hou, T., Feng, Z., and Li, Y. (2013) Structure-based development of antagonists for chemokine receptor CXCR4, *Current Computer-Aided Drug Design 9*, 60-75.

[54] Chibi, B. (2013) Computational studies of pentacycloundecane peptide based HIV-1 protease inhibitors, *University of KwaZulu-Natal,Westville*.

[55] Dau, B., and Holodniy, M. (2009) Novel targets for antiretroviral therapy, *Drugs 69*, 31-50.

[56] Matthews, T., Salgo, M., Greenberg, M., Chung, J., DeMasi, R., and Bolognesi, D. (2004) Enfuvirtide: the first therapy to inhibit the entry of HIV-1 into host CD4 lymphocytes, *Nature Reviews Drug Discovery 3*, 215-225.

[57] Ayouba, A., Cannou, C., Nugeyre, M.-T., Barré-Sinoussi, F., and Menu, E. (2008) Distinct efficacy of HIV-1 entry inhibitors to prevent cell-to-cell transfer of R5 and X4 viruses across a human placental trophoblast barrier in a reconstitution model in vitro, *Retrovirology 5*, 31.

[58] Boso, G., Orvell, C., and Somia, N. V. (2015) The nature of the N-Terminal amino acid residue of HIV-1 RNase H is critical for the stability of reverse transcriptase in viral particles, *Journal of Virology 89*, 1286-1297.

[59] Painter, G. R., Almond, M. R., Mao, S., and Liotta, D. C. (2004) Biochemical and mechanistic basis for the activity of nucleoside analogue inhibitors of HIV reverse transcriptase, *Current Topics in Medicinal Chemistry 4*, 1035-1044.

[60] De Clercq, E. (2000) Novel compounds in preclinical/early clinical development for the treatment of HIV infections, *Reviews in Medical Virology 10*, 255-277.

[61] Drake, S. M. (2000) NNRTIs—a new class of drugs for HIV, *Journal of Antimicrobial Chemotherapy 45*, 417-420.

[62] Pommier, Y., Johnson, A. A., and Marchand, C. (2005) Integrase inhibitors to treat HIV/AIDS, *Nature Reviews Drug Discovery 4*, 236-248.

[63] Hohlfeld, K., Wegner, J. r. K., Kesteleyn, B., Linclau, B., and Unge, J. (2015) Disubstituted bis-THF moieties as new P2 ligands in nonpeptidal HIV-1 protease inhibitors (II), *Journal of Medicinal Chemistry 58*, 4029-4038.

[64] Clavel, F., and Mammano, F. (2010) Role of gag in HIV resistance to protease inhibitors, *Viruses 2*, 1411-1426.

[65] Braz, A. S., Tufanetto, P., Perahia, D., and Scott, L. P. (2012) Relation between flexibility and positively selected HIV-1 protease mutants against inhibitors, *Proteins: Structure, Function, and Bioinformatics 80*, 2680-2691.

[66] Bonini, C., Chiummiento, L., De Bonis, M., Di Blasio, N., Funicello, M., Lupattelli, P., Pandolfo, R., Tramutola, F., and Berti, F. (2010) Synthesis of new thienyl ring containing HIV-1 protease inhibitors: promising preliminary pharmacological evaluation against recombinant HIV-1 proteases, *Journal of Medicinal Chemistry 53*, 1451-1457.

[67] Flexner, C. (1998) HIV-protease inhibitors, *New England Journal of Medicine 338*, 1281-1292.

[68] Abdel-Rahman, H. M., Al-karamany, G. S., El-Koussi, N. A., Youssef, A. F., and Kiso, Y. (2002) HIV protease inhibitors: peptidomimetic drugs and future perspectives, *Current Medicinal Chemistry 9*, 1905-1922.

[69] Tomasselli, A., Howe, W., Sawyer, T., Wlodawer, A., and Heinrikson, R. (1991) The complexities of AIDS: an assessment of the HIV protease as a therapeutic target, *Chimica Oggi 9*, 6-27.

[70] Titanji, B. K., Aasa-Chapman, M., Pillay, D., and Jolly, C. (2013) Protease inhibitors effectively block cell-to-cell spread of HIV-1 between T cells, *Retrovirology 10*, 1-11.

[71] Vacca, J., Dorsey, B., Schleif, W., Levin, R., McDaniel, S., Darke, P., Zugay, J., Quintero, J., Blahy, O., and Roth, E. (1994) L-735,524: An orally bioavailable human immunodeficiency virus type 1 protease inhibitor, *Proceedings of the National Academy of Sciences 91*, 4096-4100.

[72] Kaldor, S. W., Kalish, V. J., Davies, J. F., Shetty, B. V., Fritz, J. E., Appelt, K., Burgess, J. A., Campanale, K. M., Chirgadze, N. Y., and Clawson, D. K. (1997) Viracept (nelfinavir mesylate, AG1343): a potent, orally bioavailable inhibitor of HIV-1 protease, *Journal of Medicinal Chemistry 40*, 3979-3985.

[73] Li, D., Ji, B., Hwang, K.-C., and Huang, Y. (2011) Strength of hydrogen bond network takes crucial roles in the dissociation process of inhibitors from the HIV-1 protease binding pocket, *PloS One 6*, e19268.

[74] Silva, A. M., Cachau, R. E., Sham, H. L., and Erickson, J. W. (1996) Inhibition and catalytic mechanism of HIV-1 aspartic protease, *Journal of Molecular Biology 255*, 321-340.

[75] Venkatakrishnan, B., Palii, M.-L., Agbandje-McKenna, M., and McKenna, R. (2012) Mining the protein data bank to differentiate error from structural variation in clustered static structures: an examination of HIV protease, *Viruses 4*, 348-362.

[76] Wlodawer, A., and Vondrasek, J. (1998) Inhibitors of HIV-1 protease: a major success of structure-assisted drug design 1, *Annual Review of Biophysics and Biomolecular Structure 27*, 249-284.

[77] Castro, H. C., Abreu, P. A., Geraldo, R. B., Martins, R. C., dos Santos, R., Loureiro, N. I., Cabral, L. M., and Rodrigues, C. R. (2011) Looking at the proteases from a simple perspective, *Journal of Molecular Recognition 24*, 165-181.

[78] Briz, V., Poveda, E., and Soriano, V. (2006) HIV entry inhibitors: mechanisms of action and resistance pathways, *Journal of Antimicrobial Chemotherapy 57*, 619-627.

[79] Scott, W. R., and Schiffer, C. A. (2000) Curling of flap tips in HIV-1 protease as a mechanism for substrate entry and tolerance of drug resistance, *Structure 8*, 1259-1265.

[80] Chen, J., Yang, M., Hu, G., Shi, S., Yi, C., and Zhang, Q. (2009) Insights into the functional role of protonation states in the HIV-1 protease-BEA369 complex: molecular dynamics simulations and free energy calculations, *Journal of Molecular Modeling 15*, 1245-1252.

[81] Hornak, V., Okur, A., Rizzo, R. C., and Simmerling, C. (2006) HIV-1 protease flaps spontaneously close to the correct structure in simulations following manual placement of an inhibitor into the open state, *Journal of the American Chemical Society 128*, 2812-2813.

[82] Spinelli, S., Liu, Q., Alzari, P., Hirel, P., and Poljak, R. (1991) The three-dimensional structure of the aspartyl protease from the HIV-1 isolate BRU, *Biochimie 73*, 1391-1396.

[83] Ingr, M., Uhlíková, T. á., Střížovský, K., Majerová, E., and Konvalinka, J. (2003) Kinetics of the dimerization of retroviral proteases: the “fireman's grip” and dimerization, *Protein Science 12*, 2173-2182.

[84] Rao, M. B., Tanksale, A. M., Ghatge, M. S., and Deshpande, V. V. (1998) Molecular and biotechnological aspects of microbial proteases, *Microbiology and Molecular Biology Reviews 62*, 597-635.

[85] Tözsér, J. (2010) Comparative studies on retroviral proteases: substrate specificity, *Viruses 2*, 147-165.

[86] Reetz, M. T., Merk, C., and Mehler, G. (1998) Preparation of novel HIV-protease inhibitors, *Chemical Communications*, 2075-2076.

[87] Trylska, J., Grochowski, P., and McCammon, J. A. (2004) The role of hydrogen bonding in the enzymatic reaction catalyzed by HIV-1 protease, *Protein Science 13*, 513-528.

[88] Calixto, A. R., Brás, N. r. F., Fernandes, P. A., and Ramos, M. J. (2014) Reaction mechanism of human renin studied by quantum mechanics/molecular mechanics (QM/MM) calculations, *ACS Catalysis 4*, 3869-3876.

[89] Lipkowitz, K. B., and Boyd, D. B. (1997) Reviews in computational chemistry:* *Volume 13*, *Wiley Online Library*.

[90] Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., and Weiner, P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins, *Journal of the American Chemical Society 106*, 765-784.

[91] Verheyen, J., Knops, E., Kupfer, B., Hamouda, O., Somogyi, S., Schuldenzucker, U., Hoffmann, D., Kaiser, R., Pfister, H., and Kücherer, C. (2009) Prevalence of C-terminal gag cleavage site mutations in HIV from therapy-naive patients, *Journal of Infection 58*, 61-67.

[92] Freeman, S., and Herron, J. C. (2013) Evolutionary analysis: pearson new international edition, *Pearson Higher Edition, New York*.

[93] Wensing, A. M., van Maarseveen, N. M., and Nijhuis, M. (2010) Fifteen years of HIV protease inhibitors: raising the barrier to resistance, *Antiviral Research 85*, 59-74.

[94] Rhee, S.-Y., Taylor, J., Fessel, W. J., Kaufman, D., Towner, W., Troia, P., Ruane, P., Hellinger, J., Shirvani, V., and Zolopa, A. (2010) HIV-1 protease mutations and protease inhibitor cross-resistance, *Antimicrobial Agents and Chemotherapy 54*, 4253-4261.

[95] Agniswamy, J., Shen, C.-H., Aniana, A., Sayer, J. M., Louis, J. M., and Weber, I. T. (2012) HIV-1 protease with 20 mutations exhibits extreme resistance to clinical inhibitors through coordinated structural rearrangements, *Biochemistry 51*, 2819-2828.

[96] Weber, I. T., and Agniswamy, J. (2009) HIV-1 protease: structural perspectives on drug resistance, *Viruses 1*, 1110-1136.

[97] Weikl, T. R., and Hemmateenejad, B. (2013) How conformational changes can affect catalysis, inhibition and drug resistance of enzymes with induced-fit binding mechanism such as the HIV-1 protease, *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1834*, 867-873.

[98] Maseko, S. B., Natarajan, S., Sharma, V., Bhattacharyya, N., Govender, T., Sayed, Y., Maguire, G. E., Lin, J., and Kruger, H. G. (2016) Purification and characterization of naturally occurring HIV-1 (south african subtype C) protease mutants from inclusion bodies, *Protein Expression and Purification 122*, 90-96.

[99] Lockhat, H. A., Silva, J. R., Alves, C. N., Govender, T., Lameira, J., Maguire, G. E., Sayed, Y., and Kruger, H. G. (2015) Binding free energy calculations of nine FDA-approved protease inhibitors against HIV-1 subtype C I36T↑ T containing 100 amino acids per monomer, *Chemical Biology & Drug Design, 87*, 487-498.

CHAPTER TWO

COMPUTATIONAL BACKGROUND

2.1 Introduction to Computational Techniques

Theoretical structural calculations was first introduced in chemistry by Walter Heitler and Fritz London in 1927[1], and ever since, there has been an outstanding progress and dynamic use of computational techniques in research. An intensive background on molecular modelling is readily available in literature[2-6]. The combination of molecular modelling, spectroscopic techniques and enzymatic drug design was comprehensively reviewed and this also describes an overview of computational methods for this study[7, 8].

2.2 Theoretical Models

The 3D structure of macro-molecules especially proteins can be predicted through comparative modelling since not all proteins readily form crystal structures suitable for X-ray analysis and their determination from Nuclear Magnetic Resonance (NMR) is time consuming[7, 9]. Computational chemistry involves diverse theoretical fields ranging from quantum mechanics (QM), molecular mechanics (MM), dynamics (MD), conformational analysis and energy minimizations. All these are used to determine the chemical behaviour of molecular systems[8, 10].

2.2.1 Quantum Mechanics

Quantum Mechanics (QM) is based on electronic structure theories meant to solve the Erwin Schrödinger equation (1926) usually expressed in terms of wave functions[11, 12], and is used to understand the properties of molecules by studying the electronics of the compound[13]. This theory is more accurate but costly considering the computational time and resources, hence more practical for smaller systems[7, 14, 15]. There are several families of QM methods available to researchers.

2.2.1.1 Semi-empirical Methods

Currently, these methods are basically utilized for larger systems such as proteins and biomolecules. Semi-empirical methods are based on quantum mechanics (QM), but due to quote

severe approximations only the valence electrons are considered. It can be applied to large molecular systems in acceptable time frames[7, 8, 16, 17]. Semi-empirical methods are often used for modelling large molecules such as proteins and the common theoretical methods are PM6, AM1, PM3, ZINDO and INDO[8, 13, 18, 19]. The different theoretical approximations are designed for specific series of compounds; parameterization involves approximation utilizing data from experimental or high level *ab initio* calculation results[20]. Researchers should therefore verify the suitability of the chosen method for the specific chemical system they want to study.

2.2.1.2 Ab initio Methods

Ab initio methods remain the most widely used electronic structure method adopted by both theoretical and computational chemists, despite its demanding requirements in terms of time and resources[13, 21, 22]. These methods are centred on the laws of QM derived from theoretical principles[21]. Examples for different levels of theory are Moller-Plesset (MP*n*) which includes electron correlation, Hartree-Fock (HF) and Configuration Interaction (CI). QM *ab initio* methods approach the absolute experimental value as the level of theory is improved and the basis set increased[20].

2.2.1.3 Density Function Theory Methods (DFT)

Kohn, Hohenberg and Sham proposed an alternative method that uses density functional models in the design of a more effective electronic structural method and this compute energies using electron density instead of the wave function[23, 24]. DFT methods are faster than MP methods and produce similar quality of results[25]. DFT is an extension of Hartree-Fock (HF) theory and Time-Dependent DFT is an extension of DFT, and is broadly used to simulate various properties of both organic and inorganic compounds[26].

The hybrid model Becke3LYP by Becke[27] and Lee *et al*[28]., is the most popularly used DFT[29, 30] functional especially for organic compounds. Although, poor exchange-correlations interactions and low standard in estimating barrier heights have been noted[31, 32]. B3LYP have been shown to be in agreement with high level *ab initio* results and give relatively good geometrics of molecular, organic and organometallic compounds[33-36]. Hence, Becke, Lee, Yang and Parr's (Becke3LYP) DFT with 6-31G(d)[37, 38] basis set was used for certain parts of the system studied

herein. The choice of the basis set is based on its sufficiency for the geometry optimization for the chosen systems.

2.2.2 Molecular Mechanics (MM)

This method is usually referred to as a non-quantum mechanical method since it does not make use of quantum mechanics instead uses the law of classical physics to describe and understand the behavioural pattern of atoms and molecules from a single geometry[11, 39].

Molecular mechanics can be used to model large systems such as proteins and biological molecules, and are characterized by force fields with fixed empirical parameters that gives it an edge over quantum mechanics in terms of the required computational resources and time[40]. Popular force-fields includes; Merk Molecular Force Field (MMFF)[41], AMBER[42] and Universal Force Field (UFF)[43]. It is important to realise that these force fields were designed for families of compounds and researchers should verify the suitability of the force field for the proposed system under investigation. In our case, AMBER[42] force field was used which implements a simple harmonic model that includes torsions, bond stretch and angle bends, and also standard functions for Van der Waal and electrostatic interactions[40, 44, 45]. Although, Gaussian package automatically assigns the atom types for UFF calculations, AMBER calculations require all atom types to be clearly defined within the molecule specification section.

Although MM methods are less expensive, it does not explicitly calculate the electronic effect of chemical systems[13, 39], but rather uses a very crude approximation. One can therefore not study bond formation/breaking reactions with MM methods.

In order to increase the accuracy of this method, there has been a recent increase in the application of combined/hybrid QM and MM methods to predict drug resistance and metabolism and also in the design of drugs and catalyst[7, 46, 47].

In this study, geometry optimization will be carried out using a combined QM and MM methods (a two-layered ONIOM[48-50] approach), in investigating the binding free interaction energies of the HIV-1 PR inhibitors against subtype B and C-SA.

2.3 ONIOM Hybrid Method (QM:MM)

Our own N-layered Integrated molecular Orbital and Molecular mechanics (ONIOM) hybrid (QM:MM) method was first developed by Morokuma *et al*[48, 51, 52], and will be used in this study. ONIOM first appeared in Gaussian 98 and has since undergone several substantial improvements in Gaussian 03[53]. ONIOM can be used to perform optimizations, calculate energies, predict electric and magnetic effects and also vibrational frequency, which makes it applicable to a large molecular systems[48, 54].

ONIOM model is applied to large systems and are treated at diverse level of accuracy by defining two or three regions within the structure known as layers[54]. The high layer constitutes the smallest part of the region usually treated with the most accurate method (QM or DFT), and it is referred to as the Model System, bond formation and breaking occur in this site[52, 55]

A middle region is applied in the case of a three-layer ONIOM model and it comprises a much larger part of the whole system than the high layer. This is usually treated with an intermediate accurate model between the high and low level methods (lower level QM/DFT or semi-empirical). The middle layer is also referred to as the Intermediate Model System, while the whole molecule is also called the Real System[51, 56].

The low layer comprises the rest of the system/molecule and it is usually treated with an inexpensive method such as a MM method[54]. In our ONIOM hybrid approach, a two-layer model was implemented, in which the two catalytic aspartate residues as well as the HIV PR inhibitors was treated at a DFT theory and the rest of the system at a low level MM (AMBER) theory.

Both the two and three-layer ONIOM calculations are performed in a similar manner, and are represented respectively as[54, 56, 57];

$$E_{ONIOM} = E_{real}^{low} + E_{SM}^{high} - E_{SM}^{low} \quad \textbf{2}$$

$$E_{ONIOM} = E_{real}^{low} + E_{IM}^{medium} + E_{SM}^{high} - E_{IM}^{low} - E_{SM}^{medium} \quad \textbf{3}$$

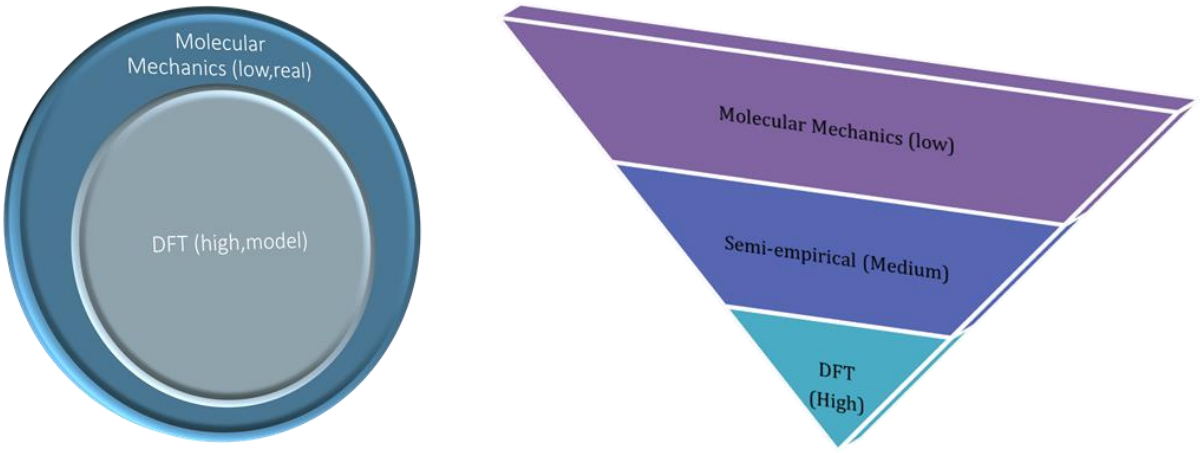


Figure 4. Schematic representation of two and three-layer ONIOM models[54, 56, 57]

2.3.1 Calculating Binding Free Energy with ONIOM

All the inhibitor—enzyme complexes were treated, viewed and modelled using the Gauss view program, which is a pre and post-processor graphical and user interface (GUI)[58]. The latest version of the Gaussian program is the Gaussian 09 package, which started with Gaussian 70, 92, 94, 98 and 03 versions[59].

The binding free energies of the nine FDA HIV-1 protease inhibitors were investigated towards both HIV-1 subtype B and C-SA PR using the Gauss view program[58] and ONIOM model implemented in Gaussian09[60] package; ONIOM(B3LYP/6-31G (d):AMBER).

The binding free ONIOM energy can be calculated as follow[61]:

$$\Delta G_{ONIOM} \approx \Delta G_{bind} = G_{complex} - G_{protein} + G_{ligand} \qquad \mathbf{5}$$

Where, ΔG is the total binding energy of the system, $G_{complex}$, $G_{protein}$ and G_{ligand} is the energies obtained from the complex (enzyme and ligand), enzyme and ligand respectively. Specific computational details are presented in the respective chapters.

References

[1] Heitler, W., and London, F. (1927) Interaction of neutral atoms and homopolar bond to quantum mechanics, in *Magazine for Physics 44*, 455-472.

[2] Barcellos, G. B., Pauli, I., Caceres, R. A., Timmers, M., Saraiva, L. F., Dias, R., de Azevedo, J., and Filgueira, W. (2008) Molecular modeling as a tool for drug discovery, *Current Drug Targets 9*, 1084-1091.

[3] Li, Y., and Harte, W. E. (2002) A review of molecular modeling approaches to pharmacophore models and structure-activity relationships of ion channel modulators in CNS, *Current Pharmaceutical Design 8*, 99-110.

[4] Hinchliffe, A. (2005) Molecular modelling for beginners, *John Wiley & Sons, New Jersey*.

[5] Mukesh, B., and Rakesh, K. (2011) Molecular docking: a review, *International Journal of Research Ayurveda Pharmacy 2*, 746-1751.

[6] Shibata, J. H. (2012) Review of molecular modeling basics, *Journal of Chemical Education 89*, 1489-1490.

[7] Honarparvar, B., Govender, T., Maguire, G. E., Soliman, M. E., and Kruger, H. G. (2013) Integrated approach to structure-based enzymatic drug design: molecular modeling, spectroscopy, and experimental bioactivity, *Chemical Reviews 114*, 493-537.

[8] Schreiner, P. R. (2007) Relative energy computations with approximate density functional theory—a caveat!, *Angewandte Chemie International Edition 46*, 4217-4219.

[9] Cavasotto, C. N., and Phatak, S. S. (2009) Homology modeling in drug discovery: current trends and applications, *Drug Discovery Today 14*, 676-683.

[10] Thishana, S., Krishna, B., and Hendrik, G. K. (2006) A computational study of the mechanism of formation of the penta-cycloundecane (PCU) cage lactam., *Trends and Perspectives in Modern Computational Science, 7*, 511-514.

[11] Atkins, P., and De Paula, J. (2006) Atkins' physical chemistry, *New York*.

[12] Schrödinger, E. (1926) An undulatory theory of the mechanics of atoms and molecules, *Physical Review 28*, 1049.

[13] Foresman, J., and Frish, E. (1996) Exploring chemistry, *Gaussian Inc., Pittsburg, USA*.

[14] Joseph-McCarthy, D. (1999) Computational approaches to structure-based ligand design, *Pharmacology & Therapeutics 84*, 179-191.

[15] Cramer, C. J. (2013) Essentials of computational chemistry: theories and models, *John Wiley & Sons, New Jersey.*

[16] Lewars, E. G. (2010) Computational chemistry: introduction to the theory and applications of molecular and quantum mechanics, *Springer Science & Business Media, Germany*.

[17] Nicolaides, C. A. (2014) Quantum chemistry and its "ages", *International Journal of Quantum Chemistry 114*, 963-982.

[18] Reimers, J. R. (2011) Computational methods for large systems: electronic structure approaches for biotechnology and nanotechnology, *John Wiley & Sons, New Jersey*.

[19] Liu, X.-T., Zhao, Y., Ren, A.-M., and Feng, J.-K. (2011) A comparative study of one-and two-photon absorption properties of pyrene and perylene diimide derivatives, *Journal of Molecular Modeling 17*, 1413-1425.

[20] Frisch, A., and Foresman, J. (1996) Exploring chemistry with electronic structure methods, *Pittsburgh PA: Gaussian Inc 302*.

[21] Ghosh, S. K., and Chattaraj, P. K. (2013) Concepts and methods in modern theoretical chemistry: electronic structure and reactivity, *CRC Press*, *Florida*.

[22] Leszczynski, J. (2012) Handbook of computational chemistry, *Vol. 2, Springer Science & Business Media, Germany.*

[23] Hohenberg, P., and Kohn, W. (1964) Inhomogeneous electron gas, *Physical Review 136*, B864.

[24] Kohn, W., and Sham, L. J. (1965) Self-consistent equations including exchange and correlation effects, *Physical Review 140*, A1133.

[25] Görling, A. (1992) Kohn-Sham potentials and wave functions from electron densities, *Physical Review A 46*, 3753.

[26] Tomberg, A. (2013) Gaussian 09W Tutorial, *An introduction to computational chemistry using G09W and avogadro software*, 1-36.

[27] Becke, A. D. (1993) Density-functional thermochemistry. III. The role of exact exchange, *The Journal of Chemical Physics 98*, 5648-5652.

[28] Lee, C., Yang, W., and Parr, R. G. (1988) Development of the colle-salvetti correlation-energy formula into a functional of the electron density, *Physical Review B 37*, 785.

[29] Kohn, W., Becke, A. D., and Parr, R. G. (1996) Density functional theory of electronic structure, *The Journal of Physical Chemistry 100*, 12974-12980.

[30] Neumann, R., Nobes, R. H., and Handy, N. C. (1996) Exchange functionals and potentials, *Molecular Physics 87*, 1-36.

[31] Zhao, Y., González-García, N., and Truhlar, D. G. (2005) Benchmark database of barrier heights for heavy atom transfer, nucleophilic substitution, association, and unimolecular reactions and its use to test theoretical methods, *The Journal of Physical Chemistry A 109*, 2012-2018.

[32] Zhao, Y., and Truhlar, D. G. (2008) Density functionals with broad applicability in chemistry, *Accounts of Chemical Research 41*, 157-167.

[33] Kruger, H. G., Mdluli, P., Power, T. D., Raasch, T., and Singh, A. (2006) Experimental and computational studies of the regioselective protection of hydantoins using anhydride, *Journal of Molecular Structure: THEOCHEM 771*, 165-170.

[34] Goldstein, E., Beno, B., and Houk, K. (1996) Density functional theory prediction of the relative energies and isotope effects for the concerted and stepwise mechanisms of the diels-alder reaction of butadiene and ethylene, *Journal of the American Chemical Society 118*, 6036-6043.

[35] Osuna, S., Morera, J., Cases, M., Morokuma, K., and Sola, M. (2009) Diels−Alder reaction between cyclopentadiene and C60: an analysis of the performance of the ONIOM method for the study of chemical reactivity in fullerenes and nanotubes, *The Journal of Physical Chemistry A 113*, 9721-9726.

[36] Ansbacher, T., Srivastava, H. K., Martin, J. M., and Shurki, A. (2010) Can DFT methods correctly and efficiently predict the coordination number of copper (I) complexes? a case study, *Journal of Computational Chemistry 31*, 75-83.

[37] Hariharan, P. C., and Pople, J. A. (1973) The influence of polarization functions on molecular orbital hydrogenation energies, *Theoretica Chimica Acta 28*, 213-222.

[38] Rassolov, V. A., Pople, J. A., Ratner, M. A., and Windus, T. L. (1998) 6-31G* basis set for atoms K through Zn, *The Journal of Chemical Physics 109*, 1223-1229.

[39] Bultinck, P., De Winter, H., Langenaeker, W., and Tollenare, J. P. (2003) Computational medicinal chemistry for drug discovery, *CRC Press, Florida*.

[40] Rogers, D. W. (2003) Computational chemistry using the PC, *John Wiley & Sons, New Jersey*.

[41] Halgren, T. A. (1996) Merck molecular force field. I. basis, form, scope, parameterization, and performance of MMFF94, *Journal of Computational Chemistry 17*, 490-519.

[42] Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The AMBER biomolecular simulation programs, *Journal of Computational Chemistry 26*, 1668-1688.

[43] Rappé, A. K., Casewit, C. J., Colwell, K., Goddard Iii, W., and Skiff, W. (1992) UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations, *Journal of the American Chemical Society 114*, 10024-10035.

[44] Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., and Weiner, P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins, *Journal of the American Chemical Society 106*, 765-784.

[45] Sansom, C. E., and Smith, C. A. (1998) Computer applications in the biomolecular sciences. Part 1: molecular modelling, *Biochemical Education 26*, 103-110.

[46] Verma, J., Khedkar, V. M., and Coutinho, E. C. (2010) 3D-QSAR in drug design-a review, *Current Topics in Medicinal Chemistry 10*, 95-115.

[47] De Benedetti, P. G., and Fanelli, F. (2010) Computational quantum chemistry and adaptive ligand modeling in mechanistic QSAR, *Drug Discovery Today 15*, 859-866.

[48] Dapprich, S., Komáromi, I., Byun, K. S., Morokuma, K., and Frisch, M. J. (1999) A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives, *Journal of Molecular Structure: THEOCHEM 461*, 1-21.

[49] Humbel, S., Sieber, S., and Morokuma, K. (1996) The IMOMO method: integration of different levels of molecular orbital approximations for geometry optimization of large systems: test for n-butane conformation and SN_2 reaction: RCl^+ Cl^-, *The Journal of Chemical Physics 105*, 1959-1967.

[50] Svensson, M., Humbel, S., Froese, R. D., Matsubara, T., Sieber, S., and Morokuma, K. (1996) ONIOM: a multilayered integrated MO+MM method for geometry optimizations and single point energy predictions. A test for diels-alder reactions and $Pt(P(t\text{-}Bu)_3)^{2+}$ H2 oxidative addition, *The Journal of Physical Chemistry 100*, 19357-19363.

[51] Vreven, T., and Morokuma, K. (2006) Hybrid methods: ONIOM (QM:MM) and QM/MM, *Annual Reports in Computational Chemistry 2*, 35-51.

[52] Vreven, T., Byun, K. S., Komáromi, I., Dapprich, S., Montgomery, J. A., Morokuma, K., and Frisch, M. J. (2006) Combining quantum mechanics methods with molecular mechanics methods in ONIOM, *Journal of Chemical Theory and Computation 2*, 815-826.

[53] Vreven, T., Morokuma, K., Farkas, Ö., Schlegel, H. B., and Frisch, M. J. (2003) Geometry optimization with QM/MM, ONIOM, and other combined methods. I. Microiterations and constraints, *Journal of Computational Chemistry 24*, 760-769.

[54] Chung, L. W., Sameera, W., Ramozzi, R., Page, A. J., Hatanaka, M., Petrova, G. P., Harris, T. V., Li, X., Ke, Z., and Liu, F. (2015) The ONIOM method and its applications, *Chemical Reviews 115*, 5678-5796.

[55] Software, N. S. f. C. C. (2012) ONIOM case study on the protein alpha-conotoxin G1, *EPSRC, UK*.

[56] Keiji Morokuma, and Vreven, T. (2000) Tutorial on the ONIOM method for inorganic chemistry, in *ACS National Meeting* ACS, *Washington, DC*.

[57] Lundberg, M., Sasakura, Y., Zheng, G., and Morokuma, K. (2010) Case studies of ONIOM (DFT:DFTB) and ONIOM (DFT:DFTB:MM) for enzymes and enzyme mimics, *Journal of Chemical Theory and Computation 6*, 1413-1427.

[58] Dennington, R., Keith, T., and Millam, J. (2009) Semichem Inc, *Shawnee Mission KS, GaussView, Version 5*.

[59] Frisch, M., GWT, H., and Schlegel, G. Scuseria,; MA Robb, JRC, JA Montgomery, Jr, *T. Vreven*, 4-6.

[60] Frisch, M., Trucks, G., Schlegel, H., Scuseria, G., Robb, M., Cheeseman, J., Scalmani, G., Barone, V., Mennucci, B., and Petersson, G. (2009) 09, Revision D. 01, Gaussian, *Inc., Wallingford, CT*.

[61] Poongavanam, V., Steinmann, C., and Kongsted, J. (2014) Inhibitor ranking through QM based chelation calculations for virtual screening of HIV-1 RNase H inhibition, *PloS one 9*, e98659.

CHAPTER THREE

Investigating the binding free energies of FDA approved drugs against subtype B and C-SA HIV PR: ONIOM approach

Zainab K. Sanusi,[a] Thavendran Govender,[a] Glenn E. M. Maguire,[a,b] Sibusiso B Maseko,[a] Johnson Lin,[c] Hendrik G. Kruger[a*] and Bahareh Honarparvar[a*]

[a]Catalysis and Peptide Research Unit, School of Health Sciences, University of KwaZulu-Natal, Durban 4001, South Africa.
[b]School of Chemistry and Physics, University of KwaZulu-Natal, 4001 Durban, South Africa.
[C]School of Life Sciences, University of KwaZulu-Natal, Durban 4001, South Africa.

***Corresponding authors:** Honarparvar@ukzn.ac.za (Dr Bahareh Honarparvar), kruger@ukzn.ac.za (Prof. Hendrik G. Kruger), Telephone: + 27 31 2601845, Fax: +27 31 2603091, Catalysis and Peptide Research Unit, School of Health Sciences, University of KwaZulu-Natal, Durban 4041, South Africa.

Abstract

Human immune virus subtype C is the most widely spread HIV subtype in Sub-Sahara Africa and South Africa. A profound structural insight in terms of finding potential lead compounds is therefore necessary for drug discovery. The focus of this study is to test an ONIOM computational model for the nine Food and Drugs Administration (FDA) HIV antiviral drugs complexed to subtype B and C-SA PR respectively. The experimental binding energies of the PR inhibitors were compared to the ONIOM calculated results. To achieve this, an integrated two-layered ONIOM model was used to optimize the geometrics of the FDA approved HIV-1 PR inhibitors for subtype B. In our hybrid ONIOM model, the HIV-1 PR inhibitors as well as the ASP 25/25' catalytic active residues were treated at high level quantum mechanics (QM) theory using B3LYP/6-31G(d), and the remaining HIV PR residues were considered using the AMBER force field. The theoretical binding free energies (ΔG_{bind}) for subtype B follow a similar trend to the experimental results, with one exemption. The computational model was less suitable for C-SA PR. Analysis of the results provided valuable information about the shortcomings of this approach. We finally conclude that the computational model can be

improved by adding water to the active pocket of the catalytic residues and treat the water at a minimum semi-empirical level of theory (PM6).

Keywords: HIV subtype B/C-SA PR; HIV PR inhibitors; Inhibitor—enzyme interactions; Our Own N-layered Integrated molecular Orbital and molecular Mechanics (ONIOM); Binding free energies.

1.0 Introduction

Acquired immune deficiency syndrome (AIDs) caused by the Human immune virus (HIV) remains a major disease worldwide, mostly in Africa. The illness was first discovered in the 80's by Luc Montagnier[1, 2] and further categorized by Gallo *et al*[3]. The size of infected population and the number of HIV patients increase yearly and this has become a major health concern[4]. The human immune virus is classified into two main types (HIV-type 1 and HIV-type 2); type 1 is further categorize into ten subtypes[5-7], of which subtype C-SA is prevalent in South Africa (95%)[5, 8-10]. The active form of the HIV-1 protease consists of C2-symmentric, identical 99-amino acids homodimer, bonded by hydrophobic and electrostatic forces as well as hydrogen bonds[11, 12].

The HIV-1 aspartic protease active site is composed of Asp25-Thr26-Gly27 catalytic triads enclosed by an extended two glycine rich β-sheets hairpins known as flaps[7, 13]. The triads are located at the edge of substrate binding site, and support the catalysis cleavage (especially the ASP) of the scissile substrate peptide[10, 14, 15].

Antiviral inhibitors were developed against HIV-1 protease as it was considered to be the most eminent target[16, 17]. A number of HIV-1 protease antiretroviral drugs approved by the FDA, were developed for subtype B, which is the more common strain found in Australia, Western Europe and North America. These drugs exhibit weaker activities against subtype C and A found in sub-Sahara Africa and India[6-8].

The single crystal X-ray structure of the South African HIV-1 protease subtype C (C-SA HIV-1 PR)[18] was resolved and a successful modelling of this enzyme based on the peptide sequence has also been reported by our group[19-21]. Experimental binding free energies for C-SA HIV PR have

previously only been determined for four of the FDA approved protease inhibitors[8, 22]. Recently, our laboratory determined the binding free energies for all nine approved HIV PR drugs against C-SA HIV PR[23].

It was demonstrated before [24] that experimental binding free energies (ΔG_{bind}) represent the most suitable comparative index for calculated binding free energies[5, 25-27].

The inhibitory activities of a selected number of commercial inhibitors against subtype B/mutants and subtype A, C, and F HIV-1 PRs were investigated using molecular dynamics simulations[7, 28]. It was suggested that the HIV-1 PR mutations can alter the energetic and dynamic complexation of the enzyme and this can affect the binding properties of the inhibitors[7] leading to drug resistance. The effect of mutations ultimately leads to drug resistance in the HIV-1 PR causing differences in the binding affinities of the protease inhibitors[28]. Our group also carried out a theoretical binding free energies of the FDA approved protease inhibitors against both subtype B and C-SA HIV PRs using molecular dynamics (MM-GBSA)[5, 10]. The observed results show that the absolute experimental values versus theoretical values differ due to the available parameterization implemented in the theoretical models which is an approximate of experimental data[29], but still largely follow a similar trend with the experimental data for subtype B[5, 10].

Our Own N-layered Integrated molecular Orbital and molecular Mechanics (ONIOM) method was developed by Morokuma[30-32]. In this multi-layered approach, the active site is treated with a high level of theory (Density Functional Theory or *Ab Initio*) while the rest of the system/enzyme is treated at a lower level (AMBER). This hybrid approach allows for treatment of large molecular systems in different research areas[33-36]. Several studies have utilized the ONIOM model for calculating interaction energies of wild-type HIV PR-1 with selected commercial inhibitors[37-39] . In all cases, the catalytic aspartate residue was treated at high level of theory, B3LYP/6-31G(d,p) and the rest of the system was modeled at MM level of theory. The binding free energies derived from the ONIOM model were compared with other binding energies derived by different computational approaches. It was concluded from the results obtained that the choice of the protonation state of the Asp dyad has an effect on the dynamic behavior of the enzyme[37-39]. The choice of protonation state for the catalytic HIV PR aspartates

(Asp 25/25') for ONIOM calculations has been studied extensively in literature[20, 21, 40-42]. It can be concluded that the p*K*a of one of the catalytic aspartate increases to 5.2 [15] when bound to the inhibitor; the corresponding value is p*K*a 4.5 [43] when unbound[44]. This implies that one of the two catalytic Asp groups should be protonated for binding studies, while the other Asp is unprotonated [24].

In other enzymatic studies, the ONIOM approach was extensively used to calculate the binding free energies of ligand—protein interactions at the B3LYP/6-31G(d):MM level of theory[34, 45, 46]. Comparison of the computed ONIOM results of the high and low potent inhibitors (-111.7 -112.3 kcal/mol) with the reported K_d experimental dissociation constants (1.1 – 29 nM) indicated that the two-layer ONIOM binding energies are not always supported by experimental findings[35]. The inconsistency in the calculated ONIOM energies and experimental dissociation constants was attributed to the selected layer for the QM region in the two-layer ONIOM calculations[35].

In this study, a QM:MM two-layer ONIOM hybrid method was adopted to investigate the binding affinity of the FDA approved protease drugs against subtype B and C-SA HIV-1 protease. For further analysis of the obtained ONIOM binding free energies, we investigated the thermodynamic parameters, electrostatic and hydrogen bonding interactions for these inhibitor—enzyme complexes.

2.0 Materials and Methods

The X-ray crystal structures for the nine commercial protease drugs complexed with subtype B HIV PR have been reported: 4YOA (DRV)[47], 4L1A (LPV)[48], 4EYR (RTV)[49], 3WSJ (IDV)[50], 3S56 (SQV)[51], 3S45 (APV)[51], 2PYM (NFV)[52], 4NJU (TPV)[53], 3EM4 (ATV)[54] (**Figure 2**). C-SA HIV PR (PDB code: 3U71)[18] is different at eight point amino acids residue mutations with respect to subtype B (PDB code:1HXW)[55]: R41K, L19I, T12S, H69K, I93L, I15V, L89M and M36I[8] (**Figure 1)**. Structural preparation of subtype C-SA HIV PR complexed with the nine drugs were performed as reported before[5, 10], since the X-ray structures for C-SA PR complexed with the FDA approved drugs have not yet been recorded in the Protein Data Bank (PDB). Further structural analysis were performed on these complexes to evaluate the mode of

interactions between the ligands and the corresponding subtype B and C-SA HIV PRs using Ligplot[56] and Accelrys (Discovery) Visualizer[57] software.

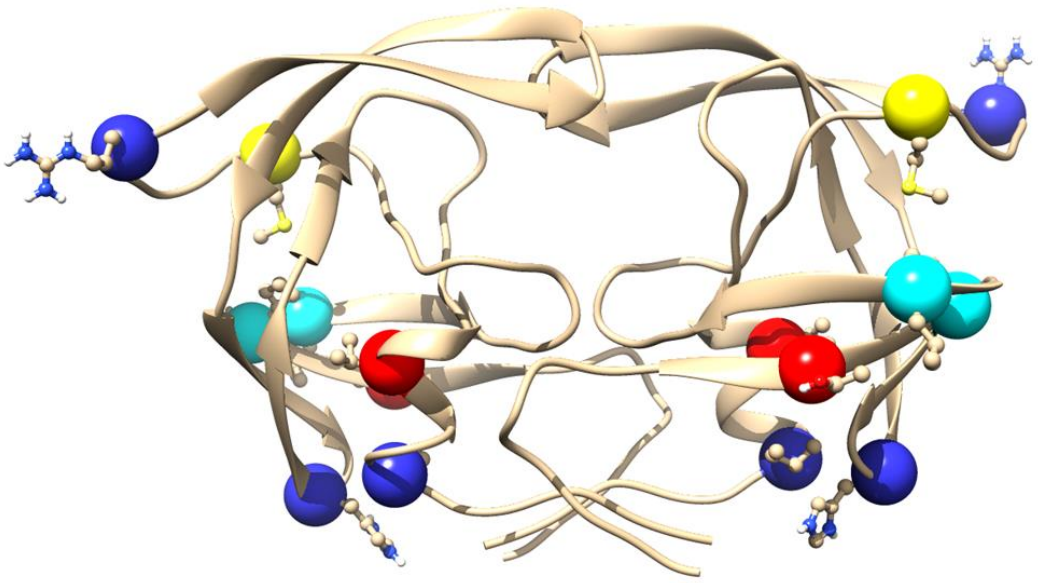


Figure 5. Homo dimeric X-ray structure of subtype C-SA protease (PDB code:3U71) [18] adopted in this study, showing the positions occupied by the eight amino acids polymorphism R41K, L19I, T12S, H69K, I93L, I15V, L89M and M36I[8] located outside the active site. This structure was created using chimera[58].

Based on the reported protonation state[20, 21, 40-42] of the HIV PR, a mono-protonated Asp 25/25' state was induced in the binding site at physiological pH 7. The protonation state of both HIV enzymes adopted in this study were assigned using PROPKA[59, 60] based on standard p*K*a values at pH 7. It is notable that all the Asp, Glu, Lys, Arg amino acids with the C and N-terminal groups were also charged, whereas His was kept in its neutral form[39].

To ensure that the selected inhibitors maintains the same pose inside the binding site of the subtype C-SA as the subtype B protease, the C-SA PR were superimposed with the subtype B HIV PR—inhibitor crystal complexes using PyMOL[61]. The PyMOL evaluates the root mean square (RMS), which is a helpful measure of how well the inhibitor—enzyme complexes were superimposed. An optimal superimposition is recognized if the RMS is less than 2 Å[62-64].

The structures of all inhibitor—enzyme complexes were refined afterwards by removing the ions and crystallographic water, that are present from the protein manually from the PDB file using a

text editor. Thereafter, protons were added to the required catalytic aspartate using GaussView[65]. The active Asp25/25' residues and the inhibitors were considered at a high level (QM/DFT[66, 67] - B3LYP[68, 69]/6-31G(d)[70, 71]) and the remaining part of the system at low layer (MM - AMBER[72]) for subsequent ONIOM[32, 73] calculations.

Superimposed structures of all the subtype C-SA HIV PR—inhibitor complexes are provided with the supplementary material **Figure S1**. The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.

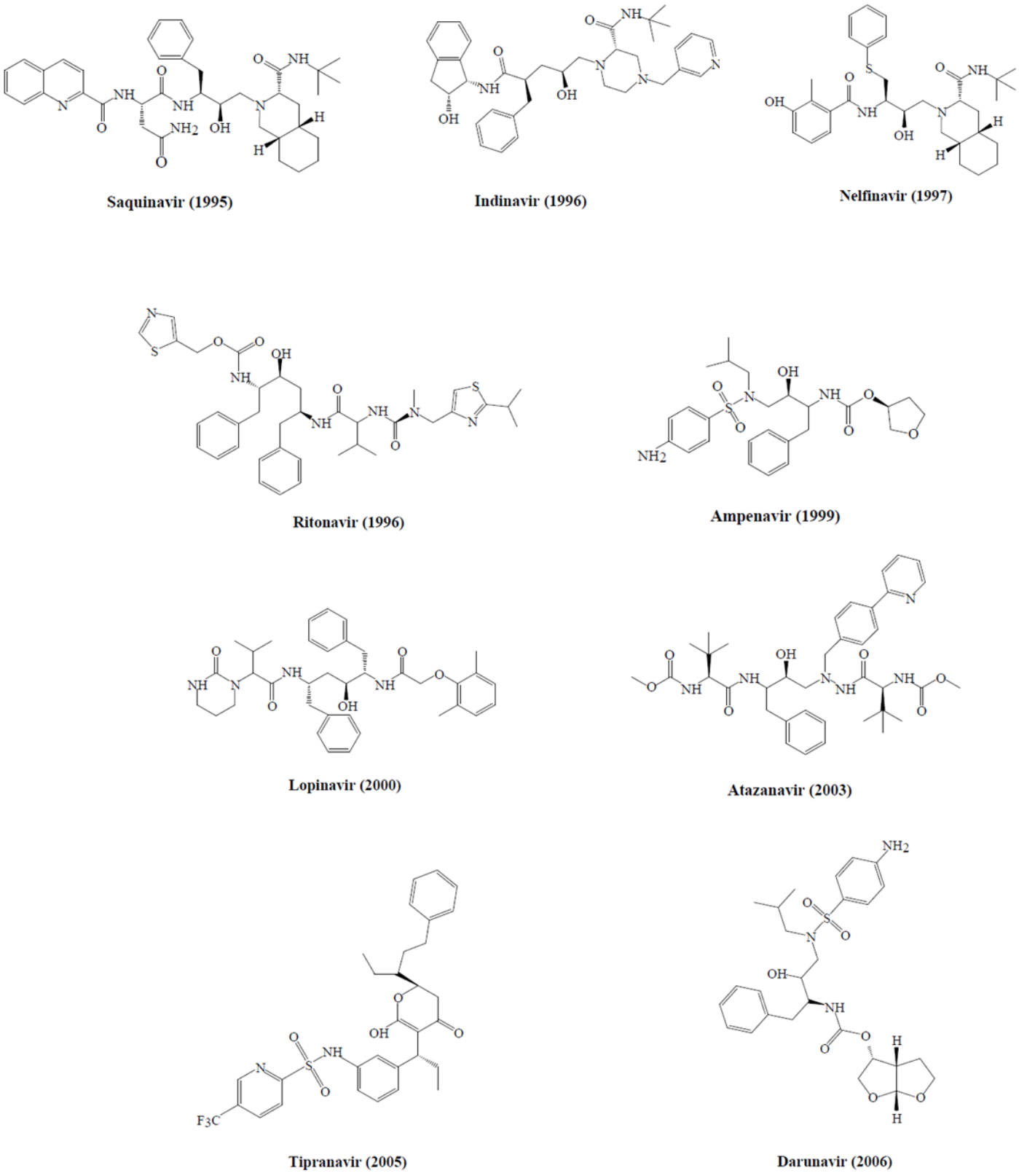


Figure 2. The structures of HIV-1 FDA approved protease inhibitors[5, 10, 74, 75]

2.1 ONIOM Binding Free Energies

A two-layer ONIOM[31, 76, 77] approach was applied to calculate the binding free energies of the different PR drugs with the subtype B and C-SA HIV PR. Preceding studies showed that the Becke3LYP method gives better relative energies and are in agreement with high level *ab-initio* methods[45, 78, 79]. Hence, full optimization calculations were carried out on both the ligands and inhibitor—enzyme complexes with Gaussian09 package[80], using ONIOM (B3LYP/6-31G(d):AMBER) for the QM:MM level of theory. A schematic illustration of the ONIOM2 model is depicted in **Figure 3**.

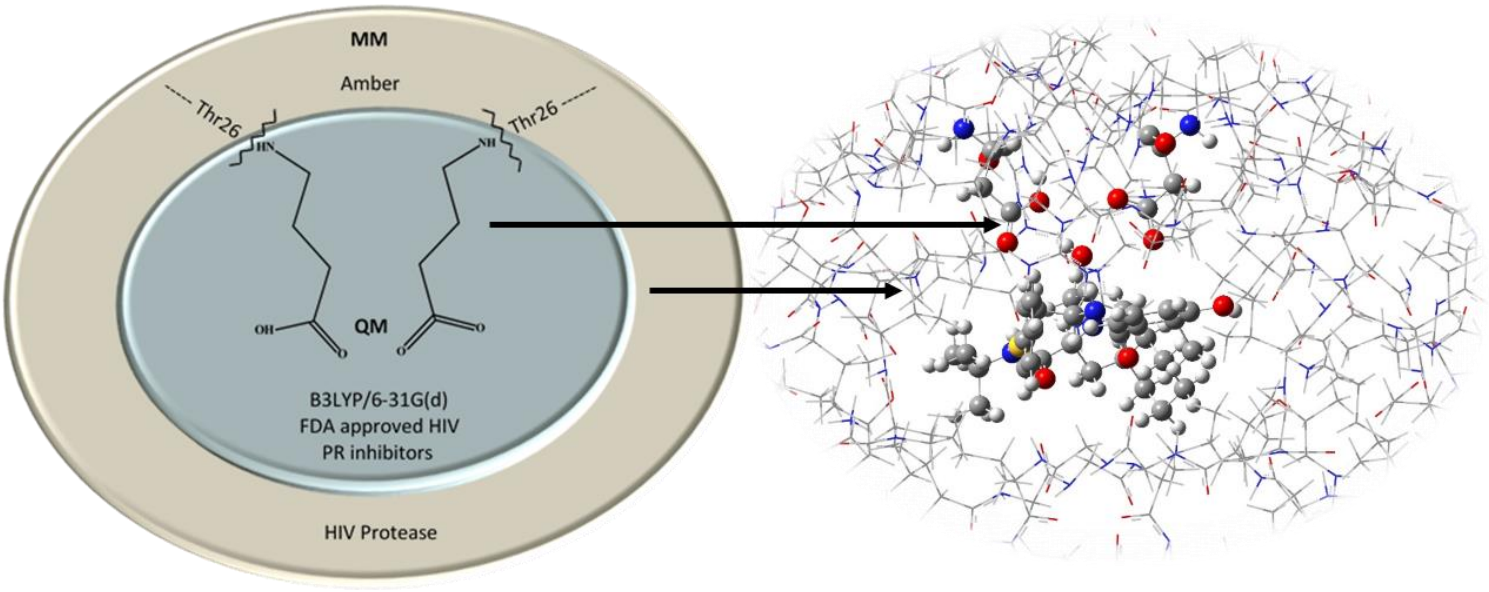


Figure 3. Schematic diagram of the two-layered ONIOM model (B3LYP/6-31G(d):AMBER) of subtype C-SA—NFV HIV PR complex.

The total interaction energy attained from the ONIOM2 calculations[81, 82] is defined as:

$$\Delta E_{ONIOM} = \Delta E_{model,high} + \Delta E_{real,low} - \Delta E_{model,low} \qquad \mathbf{1}$$

Where, ΔE_{model} is the energies of the model system calculated at the high and low level respectively and ΔE_{real} presents the energy of the entire (real) system.

The relative standard Gibbs free energies (ΔG) of all reactions, were derived from the frequency calculations at ONIOM (B3LYP/6-31G(d):AMBER) QM:MM level of theory. Thus, the corresponding binding free ONIOM energies of the complex systems are calculated from:

$$\Delta G_{ONIOM} \approx \Delta G_{bind} = G_{complex} - G_{protein} + G_{ligand} \quad \mathbf{2}$$

The thermodynamics quantities (enthalpy and entropy) changes were also obtained from the ONIOM calculation within the enzyme system for C-SA HIV PR. The entropy (ΔS) of all the reactions is estimated by a thermodynamic equation[83].

$$\Delta S_{total} = \Delta S_{surroundings} + \Delta S_{system} \quad \mathbf{3}$$

$$\Delta S_{total} = \Delta S_{translational} + \Delta S_{vibrational} + \Delta S_{rotational} \quad \mathbf{4}$$

3.0 Results and Discussion

The RMS values of the superimposed inhibitor—enzyme complexes observed were in the range of 0.5-0.8 Å for all the 18 systems set up for this study (**Figure 4**), which shows a reasonably good superimposed prediction[62-64]. Visual inspection was also performed to compare the initial inhibitor—enzyme complex conformation before and after optimization. In each optimization process, it was observed that the selected inhibitors remained inside the active pocket of the individual subtype B and C-SA HIV PRs as the starting X-ray structures.

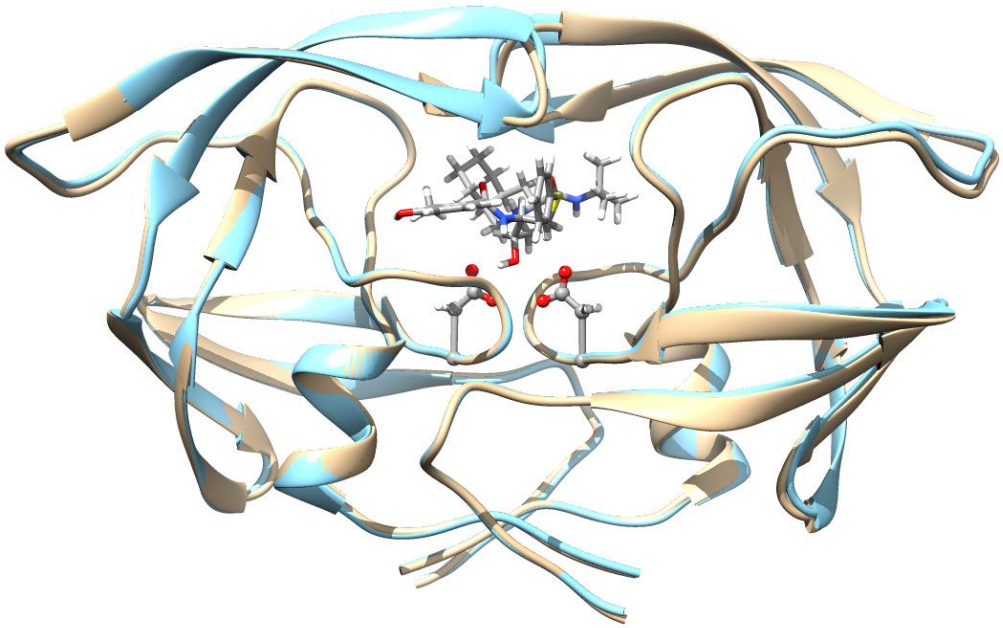


Figure 4. Super-imposed inhibitor—enzyme complex of subtype B—NFV (blue) with C-SA PR (brown). This structure was created using PyMOL[61]

The ONIOM binding free energies for subtype B and C-SA HIV PRs with the various FDA approved drugs are reported in **Table 1**. The experimental Gibbs energies for the aforementioned complexes[22, 25-27] were used to evaluate the accuracy of our theoretical model.

It is evident from the compiled experimental data that the second generation protease inhibitors (DRV, TPV, ATV, and LPV) show better binding affinities compared to the first generation (APV, SQV, RTV, NFV, IDV) against subtype B protease. However, for subtype C-SA PR the only second generation inhibitor that demonstrates an improved binding energy is ATV.

Likewise, our calculated ONIOM (B3LYP/6-31G(d):AMBER) binding free energies for subtype B reveal better binding affinities for the second generation inhibitors. The calculated energy for ATV is considerably more negative (-85.3 kcal/mol) than other inhibitors complexed to the enzyme and the reasons are discussed afterwards. The rest of the theoretical binding free energies for subtype B follow the same trend as the experimental energies (**Table 1**).

For C-SA HIV PR there are two exemptions (outliers): TPV (-78.9 kcal/mol) and NFV (-38.6 kcal/mol), unlike for subtype B, there appears to be no clear correlation between the trend of the calculated binding free energies and the reported experimental energies[23]. In the previous report for a two-layer ONIOM model, it was realized that the general trend for theoretical binding energies does not always correlate with experimental data[35]. Despite the uncertain correlation, the following more general observation was made for both experimental and computed binding energies for C-SA PR. ATV, APV, IDV, RTV, and DRV exhibit comparable better binding affinities, while SQV, LPV and NFV revealed weaker binding affinities.

Table 1. The binding free energies (kcal mol^{-1}) for HIV PIs against Subtype B and C-SA obtained by ONIOM (B3LYP/6-31G(d):AMBER).

Inhibitors	ΔG_{bind} values for Subtype B kcal/mol		ΔG_{bind} values for Subtype C-SA kcal/mol	
First generation PIs				
	Exp[a]	Theory[b]	Exp[c]	Theory[b]
RTV	-13.7	-62.8	-13.9	-62.9
APV	-13.2	-56.4	-13.9	-69.0
SQV	-13.0	-54.0	-13.4	-57.1
NFV	-12.8	-46.2	-13.5	-38.6
IDV	-12.4	-45.8	-14.0	-64.0
Second generation PIs				
DRV	-15.2	-65.9	-13.8	-62.8
LPV	-15.1	-65.7	-13.2	-56.5
TPV	-14.6	-63.1	-13.2	-78.9
ATV	-14.3	-85.3	-14.4	-66.9

Note: HIV protease inhibitors are ranked in terms of their binding free energies.
[a]Experimental binding Gibbs free energies taken from literature against subtype B PR[5, 22, 25-27]
[b]Calculated binding free energies using ONIOM for subtype B and C-SA PR respectively.
[c]Experimental binding Gibbs free energies taken from literature against subtype C-SA PR[23].
The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.

Several factors can potentially contribute to the different results for the two enzymes adopted in this study. First, there are eight point mutations in C-SA compared with subtype B HIV PR[8]. The experimental study for subtype B and the first resolved crystal structure for subtype C-SA protease divulge that polymorphisms at residue 36 of the C-SA HIV PR have a significant influence on the stability of the enzyme hinge region. Also, the lack of the E35-R57 salt bridge results in reduced stability of the hinge region; the latter contributes to increased flaps flexibility[18] and the flap movement plays a major role in the complexation event and thus the binding free energies[84-86].

Subsequently, an attempt was made to achieve further insight into the detailed hydrogen bond interactions of the inhibitor—enzyme complexes. The hydrogen bond distances were measured between both catalytic aspartates in the active pocket and hydroxyl group of the selected FDA approved inhibitors. All the inhibitor—enzyme interactions were plotted using Accelrys (Discovery) Visualizer[57] before and after optimization.

Comparison of the average hydrogen bond (HB) distances between active residues of subtype B and C-SA PR with the selected ligands, shows that in all cases, the hydroxyl group of the inhibitors form hydrogen bonding with Asp25/25' of HIV proteases (**Figure 5**).

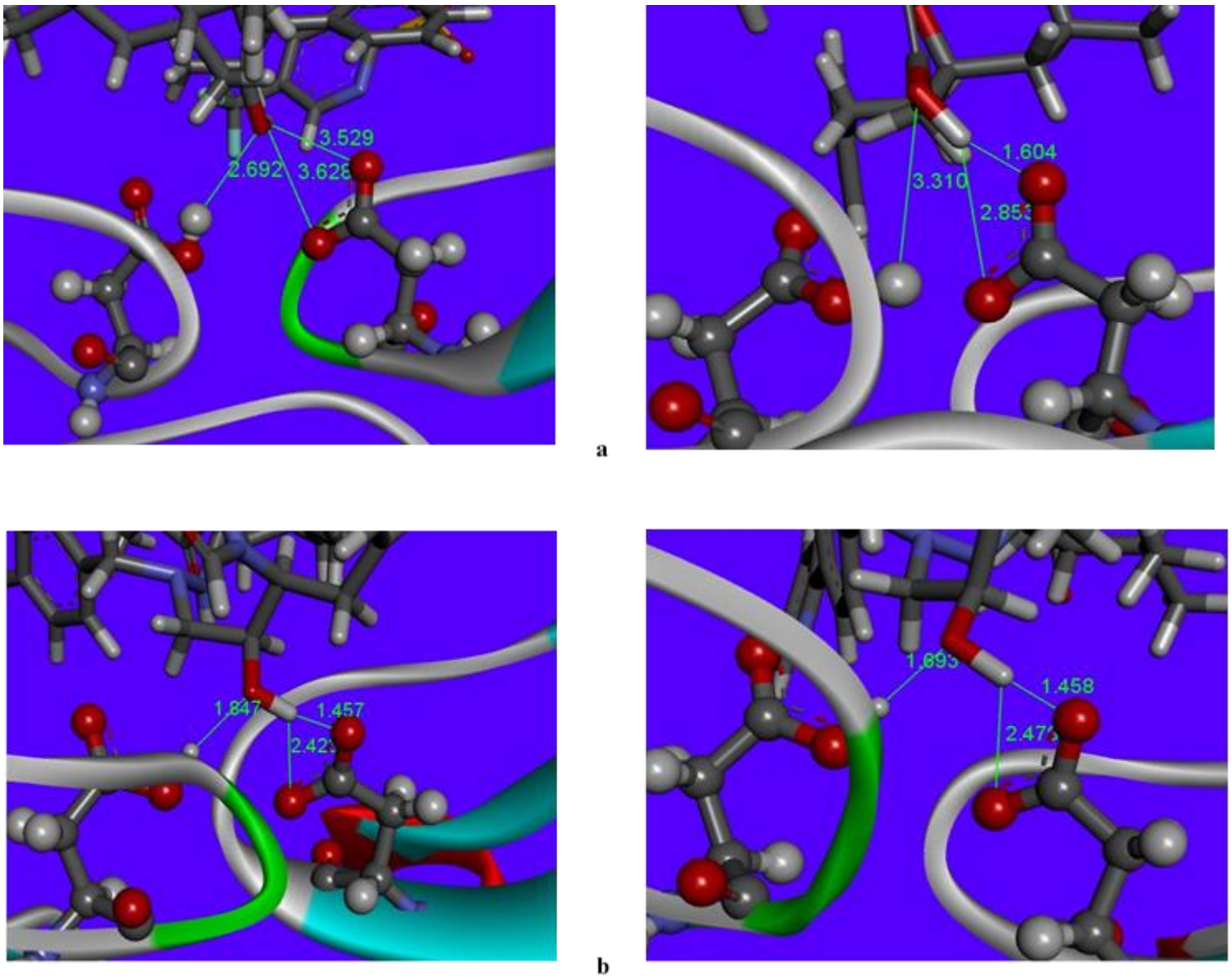


Figure 5. Hydrogen bond distances between the hydroxyl groups of the TPV and ATV drugs with the catalytic ASP25 and ASP25' residues of **a:** subtype B—TPV PR, **b:** subtype B—ATV PR before and after optimization using Accelrys (Discovery) visualizer. Detailed comparative plots for all inhibitor—enzyme complexes are provided with the supplementary material **Figure S3**. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.)

The change in hydrogen bond distances between the inhibitor -OH group and Asp 25/25' were calculated (**Figure S3,** supplementary material) for all the inhibitor—enzyme complex structures before and after optimization. For subtype B, the first generation inhibitors (with weaker binding energies) experience a slight reduction in the average hydrogen bond distance of 0.1 Å. While, the second generation inhibitors (with better binding energies) revealed a larger reduction with an average hydrogen bond distance of 0.5 Å upon optimization of the inhibitor—enzyme complex. When the optimized inhibitor -OH and Asp25/25' hydrogen bond distances for subtype B/TPV (3.3 Å; -63.1 kcal/mol) are compared to that of the outlier subtype B/ATV complex (1.7 Å; -85.3 kcal/mol). It is clear that ATV experiences a much stronger hydrogen bond interaction, most likely leading to the increased theoretical binding free energy.

In the case of the subtype C-SA PR outliers: For TPV, comparison of the hydrogen bond distances with ATV reveal that TPV (3.1 Å; -78.9 kcal/mol) has a stronger hydrogen bond interaction than ATV (4.4 Å; -66.9 kcal/mol). This also explains the difference in calculated binding free energies of these two complexes. For the other C-SA outlier: NFV, a comparison with SQV that has a similar experimental binding free energy, the corresponding hydrogen bond distance for NFV (4.4 Å; -38.6 kcal/mol) is longer than that of SQV (4.1 Å; -57.1 kcal/mol). This greater distance for NFV corresponds to a weaker theoretical binding free energy.

To further probe the nature of these differences in calculated binding free energies, the electrostatic and hydrogen bond interactions for both subtype B and C-SA HIV-1 PRs complexed with the various FDA approved inhibitors were plotted using Ligplot[56] and depicted in **Figure 6** (for subtype B—ATV) and in the supplementary material (**Figure S2**). The plots show hydrogen bonding and electrostatic interactions occurring between the inhibitors, catalytic aspartate and other residues in both proteases. However, the mode of interaction differs from subtype B and C-SA PR, this may be due to the polymorphism occurring in C-SA HIV PR which causes the enzyme to be more flexible[18].

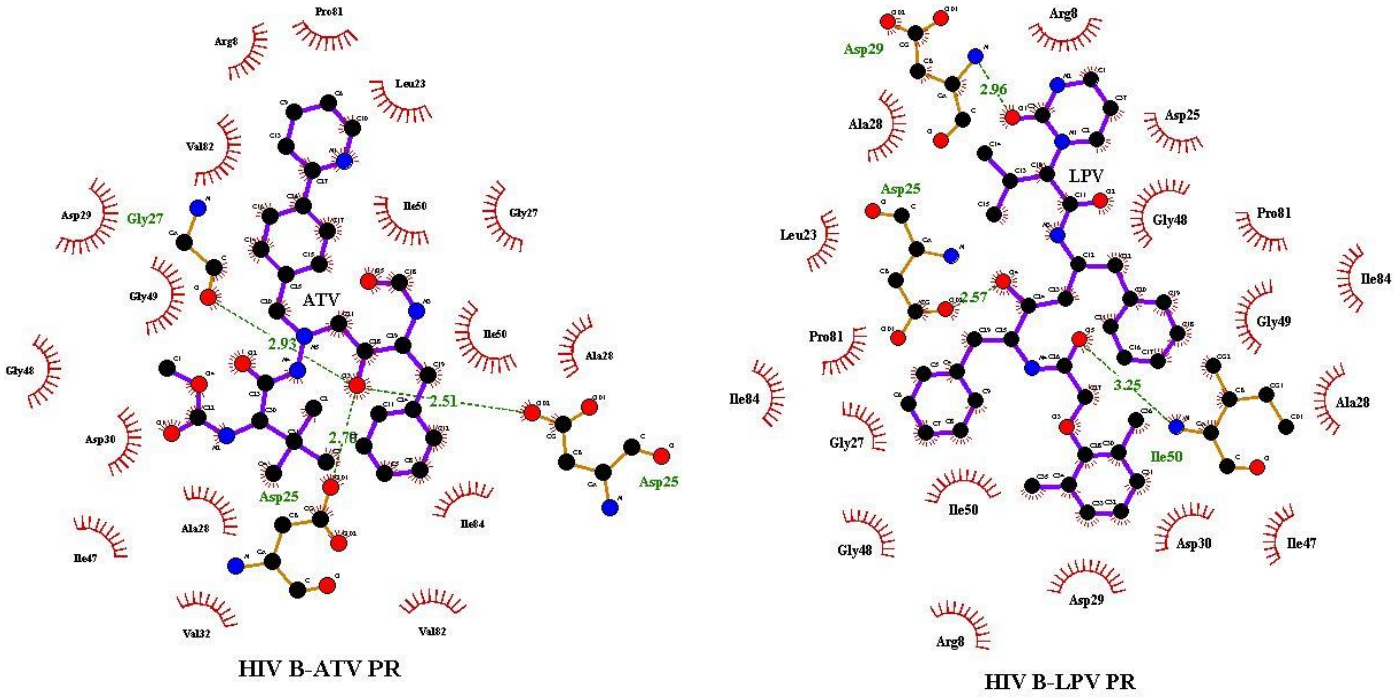


Figure 6. Electrostatic and hydrogen bond interactions plot of HIV subtype B PR complexed with ATV and LPV. These plots were created after ONIOM optimization of each complex system using Ligplot[56]. Detailed plots showing the electrostatic and hydrogen bond interactions for the remaining complexes are provided with the supplementary material **Figure S2**. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.)

The calculated thermochemical properties [binding free interaction energy (ΔG), enthalpy (ΔH) and entropy (ΔS) contributions] of the various drugs complexes with subtype B (**Table 2**) and C-SA HIV PR are provided in supplementary material (**Table S1**).

The calculated binding free energies (ΔG) for the various FDA approved inhibitors complexes with subtype B follow the same trend as the experimental data (**Table 1**), except for ATV, while the C-SA HIV PR has two outliers (TPV and NFV) as discussed earlier. The thermochemical properties can be used to explain these outliers since ΔG is a function of both enthalpy (ΔH) and entropy (ΔS).

For the subtype B outlier, ATV: the ΔH value (-104.2 kcal/mol) (**Table 2)** suggests a better binder than the rest of the drugs complexed to the enzyme. As demonstrated before, this may arise from stronger hydrogen bond interactions between the ATV hydroxyl group and the

catalytic aspartate residue, which translates into a larger binding free energy ΔG (-85.3 kcal/mol).

For the two C-SA, outliers TPV and NFV (**Table S1**, supplementary material), it is noticeable that the ΔH (-129.3 kcal/mol) for TPV also suggests better binding energy than the rest of the drugs in complex with the enzyme. While ΔH (-59.3 kcal/mol) for NFV indicates a weaker contact. These enthalpy values are attributed to the nature of the hydrogen bond interaction between the inhibitor-OH group and the catalytic aspartate (**Figure 5** and subsequent discussion). Thus, it reflects a stronger binding free energy ΔG (-78.9 kcal/mol) for TPV and a weaker binding free energy ΔG (-38.6 kcal/mol) for NFV.

Table 2. The binding free energies, enthalpies and entropy of the various FDA approved HIV PIs against subtype B PR.

Inhibitors	ΔG^a kcal mol^{-1}	ΔH kcal mol^{-1}	ΔS_{total} cal mol^{-1} K^{-1}	ΔS_{trans} cal mol^{-1} K^{-1}	ΔS_{vib} cal mol^{-1} K^{-1}	ΔS_{rot} cal mol^{-1} K^{-1}
First generation PIs						
RTV	-62.8	-70.5	-25.1	-43.9	56.8	-38.1
APV	-56.4	-81.8	-85.4	-44.5	-3.3	-37.6
SQV	-54.0	-70.1	-53.3	-45.3	31.3	-39.3
NFV	-46.2	-58.5	-41.4	-44.8	41.6	-38.2
IDV	-45.8	-67.8	-71.4	-45.0	12.9	-39.3
Second generation PIs						
DRV	-65.9	-76.8	-37.2	-43.1	42.1	-36.3
LPV	-65.7	-99.7	-113.2	-43.5	-32.5	-37.2
TPV	-63.1	-75.1	-39.9	-43.4	41.2	-37.6
ATV	-85.3	-104.2	-63.3	-43.2	16.3	-36.5

HIV protease inhibitors (HIV PIs) are ranked in terms of their binding free energies (ΔG), the thermodynamics values for subtype C-SA are provided with supplementary material **(Table S1**).

[a]Calculated binding free energies using ONIOM for subtype B PR (taken from **Table 1**).

The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.

For the various FDA approved drugs, it should be noted that more negative entropies (higher entropy penalty) is indicative of greater restrictions of movement for the ligand in the active site, due to stronger non-covalent inhibitor—enzyme interactions for certain parts of the inhibitor[87]. The translational and rotational entropy contributions are in close range for both subtype B and C-SA HIV PRs (**Table S1**), it can be seen that the major entropy contribution is from the vibrational energy component. LPV demonstrates a larger entropy penalty ΔS_{total} (-113.2 cal/mol K) for the subtype B than any of the other inhibitors. An entropy penalty is normally paid upon restriction of the inhibitor[87] and in this case most possibly due to stabilization of the ligand side chains. Hydrogen bond interactions occur for two side chain residues of LPV with amino acids in the active site of the enzyme (**Figure 5**), which explain the observed increase in entropy value for LPV (**Table 2**). These restrictions were not observed to the same extend for the other inhibitors.

The calculated binding free energies of the FDA approved drugs against subtype B in comparison to C-SA HIV PR, reveals that the model works better for subtype B. The outlier results for subtype B and C-SA PR appears to be due to the approximations applied in our ONIOM model. Movement of the inhibitors closer to the Asp25/25' residues during optimization, suggest that omission of water molecules in the model is an over-simplification and should be addressed in future studies.

4.0 Conclusion

In this study, the binding free energies of the nine selected HIV-1 PR inhibitors were investigated using a two-layered ONIOM (B3LYP/6-31G(d):AMBER) model. The difference in binding affinities of the PIs with the two proteases seems to relate to the insertion and mutants experienced by the mutant and the extent of binding interactions between the catalytic aspartates, Asp25/25', and the inhibitors.

The calculated binding free energies for subtype B HIV PR show a satisfactory trend with the experimental data with one exception. For subtype C-SA HIV PR, some discrepancies occur in terms of the trend with the experimental data, which means that the existing model requires further optimization before it can be used for C-SA PR.

It is known that water molecules facilitate the catalytic interactions of the substrate with the Asp 25/25' residues, therefore the level of accuracy of the two layer ONIOM model can be improved by adding explicit water molecules to the active site of the HIV PR and in the least treated at a semi-empirical level (PM6).

The information obtained from this research is helpful to improve the computational model for the potential design of more potent C-SA HIV PR inhibitors.

Competing interests

The authors declare that they have no competing interests.

Acknowledgment

We thank the College of Health Sciences (CHS), Aspen Pharmacare, MRC and the NRF for financial support. We are also grateful to the CHPC (www.chpc.ac.za) and UKZN HPC cluster as our computational resources.

References

[1] Shehu-Xhilaga, M., and Oelrichs, R. (2009) Basic HIV virology, *HIV Management in Australasia*, 9-18.

[2] Klatzmann, D., Champagne, E., Chamaret, S., Gruest, J., Guetard, D., Hercend, T., Gluckman, J.-C., and Montagnier, L. (1983) T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV, *Nature 312*, 767-768.

[3] Gallo, R. C., Sarin, P. S., Gelmann, E., Robert-Guroff, M., Richardson, E., Kalyanaraman, V., Mann, D., Sidhu, G. D., Stahl, R. E., and Zolla-Pazner, S. (1983) Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS), *Science 220*, 865-867.

[4] HIV/AIDS, J. U. N. P. O., and HIV/AIDS, J. U. N. P. O. (2016) Global AIDS update 2016, *Geneva, Switzerland*.

[5] Ahmed, S. M., Kruger, H. G., Govender, T., Maguire, G. E., Sayed, Y., Ibrahim, M. A., Naicker, P., and Soliman, M. E. (2013) Comparison of the molecular dynamics and calculated binding free energies for nine FDA-approved HIV-1 PR drugs against subtype B and C-SA HIV PR, *Chemical Biology & Drug Design 81*, 208-218.

[6] Robbins, A. H., Coman, R. M., Bracho-Sanchez, E., Fernandez, M. A., Gilliland, C. T., Li, M., Agbandje-McKenna, M., Wlodawer, A., Dunn, B. M., and McKenna, R. (2010) Structure of the unbound form of HIV-1 subtype A protease: comparison with unbound forms of proteases from other HIV subtypes, *Acta Crystallographica Section D: Biological Crystallography 66*, 233-242.

[7] Genoni, A., Morra, G., Merz Jr, K. M., and Colombo, G. (2010) Computational study of the resistance shown by the subtype B/HIV-1 protease to currently known inhibitors, *Biochemistry 49*, 4283-4295.

[8] Mosebi, S., Morris, L., Dirr, H. W., and Sayed, Y. (2008) Active-site mutations in the south african human immunodeficiency virus type 1 subtype C protease have a significant impact on clinical inhibitor binding: kinetic and thermodynamic study, *Journal of Virology 82*, 11476-11479.

[9] Ahmed, S. M., Maguire, G. E., Kruger, H. G., and Govender, T. (2014) The Impact of active site mutations of south african HIV PR on drug resistance: insight from molecular dynamics simulations, binding free energy and per-residue footprints, *Chemical Biology & Drug Design 83*, 472-481.

[10] Lockhat, H. A., Silva, J. R., Alves, C. N., Govender, T., Lameira, J., Maguire, G. E., Sayed, Y., and Kruger, H. G. (2015) Binding free energy calculations of nine FDA-approved protease inhibitors against HIV-1 subtype C I36T↑ T containing 100 amino acids per monomer, *Chemical Biology & Drug Design 87*, 487-498.

[11] Braz, A. S., Tufanetto, P., Perahia, D., and Scott, L. P. (2012) Relation between flexibility and positively selected HIV-1 protease mutants against inhibitors, *Proteins: Structure, Function, and Bioinformatics 80*, 2680-2691.

[12] Bonini, C., Chiummiento, L., De Bonis, M., Di Blasio, N., Funicello, M., Lupattelli, P., Pandolfo, R., Tramutola, F., and Berti, F. (2010) Synthesis of new thienyl ring containing HIV-1 protease inhibitors: promising preliminary pharmacological evaluation against recombinant HIV-1 proteases, *Journal of Medicinal Chemistry 53*, 1451-1457.

[13] Vangelista, L., Secchi, M., and Lusso, P. (2008) Rational design of novel HIV-1 entry inhibitors by RANTES engineering, *Vaccine 26*, 3008-3015.

[14] Kurth, R., and Bannert, N. (2010) Retroviruses: molecular biology, genomics and pathogenesis, *Horizon Scientific Press, UK*.

[15] Brik, A., and Wong, C.-H. (2003) HIV-1 protease: mechanism and drug discovery, *Organic & Biomolecular Chemistry 1*, 5-14.

[16] Abdel-Rahman, H. M., Al-karamany, G. S., El-Koussi, N. A., Youssef, A. F., and Kiso, Y. (2002) HIV protease inhibitors: peptidomimetic drugs and future perspectives, *Current Medicinal Chemistry 9*, 1905-1922.

[17] Wlodawer, A. (2002) Rational approach to AIDS drug design through structural biology*, *Annual Review of Medicine 53*, 595-614.

[18] Naicker, P., Achilonu, I., Fanucchi, S., Fernandes, M., Ibrahim, M. A., Dirr, H. W., Soliman, M. E., and Sayed, Y. (2013) Structural insights into the south african HIV-1 subtype C protease: impact of hinge region dynamics and flap flexibility in drug resistance, *Journal of Biomolecular Structure and Dynamics 31*, 1370-1380.

[19] Makatini, M. M., Petzold, K., Alves, C. N., Arvidsson, P. I., Honarparvar, B., Govender, P., Govender, T., Kruger, H. G., Sayed, Y., and JerônimoLameira. (2013) Synthesis, 2D-NMR and molecular modelling studies of pentacycloundecane lactam-peptides and peptoids as potential HIV-1 wild type C-SA protease inhibitors, *Journal of Enzyme Inhibition and Medicinal Chemistry 28*, 78-88.

[20] Karpoormath, R., Sayed, Y., Govender, P., Govender, T., Kruger, H. G., Soliman, M. E., and Maguire, G. E. (2012) Pentacycloundecane derived hydroxy acid peptides: a new class of irreversible non-scissile ether bridged type isoster as potential HIV-1 wild type C-SA protease inhibitors, *Bioorganic Chemistry 40*, 19-29.

[21] Honarparvar, B., Makatini, M. M., Pawar, S. A., Petzold, K., Soliman, M. E., Arvidsson, P. I., Sayed, Y., Govender, T., Maguire, G. E., and Kruger, H. G. (2012) Pentacycloundecane-diol-based HIV-1 protease inhibitors: biological screening, 2D NMR, and molecular simulation studies, *ChemMedChem 7*, 1009-1019.

[22] Velazquez-Campoy, A., Vega, S., and Freire, E. (2002) Amplification of the effects of drug resistance mutations by background polymorphisms in HIV-1 protease from african subtypes, *Biochemistry 41*, 8613-8619.

[23] Sibusiso, B. M. E., Padayachee; Thavendran, Govendera; Yasien, Sayed; Gert, Kruger; Glenn, E.M Maguire; Johnson, Linb. (2016) I36T↑T IN C-SA HIV-1 Protease significantly alters protease drug interactions, *Antiviral Research, Submitted for Publication*.

[24] Honarparvar, B., Govender, T., Maguire, G. E., Soliman, M. E., and Kruger, H. G. (2013) Integrated approach to structure-based enzymatic drug design: molecular modeling, spectroscopy, and experimental bioactivity, *Chemical Reviews 114*, 493-537.

[25] Ohtaka, H., Velázquez-Campoy, A., Xie, D., and Freire, E. (2002) Overcoming drug resistance in HIV-1 chemotherapy: the binding thermodynamics of amprenavir and TMC-126 to wild-type and drug-resistant mutants of the HIV-1 protease, *Protein Science 11*, 1908-1916.

[26] King, N. M., Prabu-Jeyabalan, M., Nalivaika, E. A., Wigerinck, P., de Béthune, M.-P., and Schiffer, C. A. (2004) Structural and thermodynamic basis for the binding of TMC114, a next-generation human immunodeficiency virus type 1 protease inhibitor, *Journal of Virology 78*, 12012-12021.

[27] Muzammil, S., Armstrong, A., Kang, L., Jakalian, A., Bonneau, P., Schmelmer, V., Amzel, L., and Freire, E. (2007) Unique thermodynamic response of tipranavir to human immunodeficiency virus type 1 protease drug resistance mutations, *Journal of Virology 81*, 5144-5154.

[28] Batista, P. R., Wilter, A., Durham, E. H., and Pascutti, P. G. (2006) Molecular dynamics simulations applied to the study of subtypes of HIV-1 protease common to brazil, africa, and asia, *Cell Biochemistry and Biophysics 44*, 395-404.

[29] Frisch, A., and Foresman, J. (1996) Exploring chemistry with electronic structure methods, *Pittsburgh PA: Gaussian Inc 302*.

[30] Morokuma, K. (2002) New challenges in quantum chemistry: quests for accurate calculations for large molecular systems, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 360*, 1149-1164.

[31] Dapprich, S., Komáromi, I., Byun, K. S., Morokuma, K., and Frisch, M. J. (1999) A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives, *Journal of Molecular Structure: THEOCHEM 461*, 1-21.

[32] Vreven, T., and Morokuma, K. (2000) On the application of the IMOMO (integrated molecular orbital+molecular orbital) method, *Journal of Computational Chemistry 21*, 1419-1432.

[33] Zheng, F., and Zhan, C.-G. (2008) Rational design of an enzyme mutant for anti-cocaine therapeutics, *Journal of Computer-Aided Molecular Design 22*, 661-671.

[34] Ruangpornvisuti, V. (2004) Recognition of carboxylate and dicarboxylates by azophenol–thiourea derivatives: a theoretical host–guest investigation, *Journal of Molecular Structure: THEOCHEM 686*, 47-55.

[35] Samanta, P. N., and Das, K. K. (2016) Prediction of binding modes and affinities of 4-substituted-2, 3, 5, 6-tetrafluorobenzenesulfonamide inhibitors to the carbonic anhydrase receptor by docking and ONIOM calculations, *Journal of Molecular Graphics and Modelling 63*, 38-48.

[36] Promsri, S., Chuichay, P., Sanghiran, V., Parasuk, V., and Hannongbua, S. (2005) Molecular and electronic properties of HIV-1 protease inhibitor C 60 derivatives as studied by the ONIOM method, *Journal of Molecular Structure: THEOCHEM 715*, 47-53.

[37] Fong, P., McNamara, J. P., Hillier, I. H., and Bryce, R. A. (2009) Assessment of QM/MM scoring functions for molecular docking to HIV-1 protease, *Journal of Chemical Information and Modeling 49*, 913-924.

[38] Saen-oon, S., Aruksakunwong, O., Wittayanarakul, K., Sompornpisut, P., and Hannongbua, S. (2007) Insight into analysis of interactions of saquinavir with HIV-1 protease in comparison between the wild-type and G48V and G48V/L90M mutants based on QM and QM/MM calculations, *Journal of Molecular Graphics and Modelling 26*, 720-727.

[39] Wittayanarakul, K., Aruksakunwong, O., Saen-oon, S., Chantratita, W., Parasuk, V., Sompornpisut, P., and Hannongbua, S. (2005) Insights into saquinavir resistance in the G48V HIV-1 protease: quantum calculations and molecular dynamic simulations, *Biophysical Journal 88*, 867-879.

[40] Shi, S., Hu, G., Chen, J., Zhang, S., and Zhang, Q. (2009) Molecular dynamics simulations on the role of protonation states in HIV-1 protease-indinavir complex, *Acta Chim. Sin 67*, 2791-2797.

[41] Tong, Y., Mei, Y., Zhang, J. Z., Duan, L. L., and Zhang, Q.-g. (2009) Quantum calculation of protein solvation and protein–ligand binding free energy for HIV-1 protease/water complex, *Journal of Theoretical and Computational Chemistry 8*, 1265-1279.

[42] Makatini, M. M. (2011) Design, synthesis and screening of novel PCU-peptide/peptoid derived HIV protease inhibitors, *University of KwaZulu-Natal, Westville*.

[43] Smith, R., Brereton, I. M., Chai, R. Y., and Kent, S. B. (1996) Ionization states of the catalytic residues in HIV-1 protease, *Nature Structural & Molecular Biology 3*, 946-950.

[44] Honarparvar, B., Pawar, S. A., Alves, C. N., Lameira, J., Maguire, G. E., Silva, J. R. A., Govender, T., and Kruger, H. G. (2015) Pentacycloundecane lactam vs lactone norstatine type

protease HIV inhibitors: binding energy calculations and DFT study, *Journal of Biomedical Science 22*, 1-15.

[45] Remko, M., Walsh, O. A., and Richards, W. G. (2001) Theoretical study of molecular structure, tautomerism, and geometrical isomerism of moxonidine: two-layered ONIOM calculations, *The Journal of Physical Chemistry A 105*, 6926-6931.

[46] Li, W., Qin, S., Su, Z., Hu, C., and Feng, X. (2012) Theoretical study on the mechanism and stereochemistry of salicylaldehyde–Al (III)-catalyzed hydrophosphonylation of benzaldehyde, *Computational and Theoretical Chemistry 989*, 44-50.

[47] Kuiper, B. D., Keusch, B. J., Dewdney, T. G., Chordia, P., Ross, K., Brunzelle, J. S., Kovari, I. A., MacArthur, R., Salimnia, H., and Kovari, L. C. (2015) The L33F darunavir resistance mutation acts as a molecular anchor reducing the flexibility of the HIV-1 protease 30s and 80s loops, *Biochemistry and Biophysics Reports 2*, 160-165.

[48] Liu, Z., Yedidi, R. S., Wang, Y., Dewdney, T. G., Reiter, S. J., Brunzelle, J. S., Kovari, I. A., and Kovari, L. C. (2013) Crystallographic study of multi-drug resistant HIV-1 protease lopinavir complex: mechanism of drug recognition and resistance, *Biochemical and Biophysical Research Communications 437*, 199-204.

[49] Liu, Z., Yedidi, R. S., Wang, Y., Dewdney, T. G., Reiter, S. J., Brunzelle, J. S., Kovari, I. A., and Kovari, L. C. (2013) Insights into the mechanism of drug resistance: X-ray structure analysis of multi-drug resistant HIV-1 protease ritonavir complex, *Biochemical and Biophysical Research Communications 431*, 232-238.

[50] Kuhnert, M., Steuber, H., and Diederich, W. E. (2014) Structural basis for HTLV-1 protease inhibition by the HIV-1 protease inhibitor indinavir, *Journal of Medicinal Chemistry 57*, 6266-6272.

[51] Tie, Y., Wang, Y. F., Boross, P. I., Chiu, T. Y., Ghosh, A. K., Tozser, J., Louis, J. M., Harrison, R. W., and Weber, I. T. (2012) Critical differences in HIV-1 and HIV-2 protease specificity for clinical inhibitors, *Protein Science 21*, 339-350.

[52] Kožíšek, M., Bray, J., Řezáčová, P., Šašková, K., Brynda, J., Pokorná, J., Mammano, F., Rulíšek, L., and Konvalinka, J. (2007) Molecular analysis of the HIV-1 resistance development: enzymatic activities, crystal structures, and thermodynamics of nelfinavir-resistant HIV protease mutants, *Journal of Molecular Biology 374*, 1005-1016.

[53] Yedidi, R. S., Garimella, H., Aoki, M., Aoki-Ogata, H., Desai, D. V., Chang, S. B., Davis, D. A., Fyvie, W. S., Kaufman, J. D., and Smith, D. W. (2014) A conserved hydrogen-bonding network of P2 bis-tetrahydrofuran-containing HIV-1 protease inhibitors (PIs) with a protease active-site amino acid backbone aids in their activity against PI-resistant HIV, *Antimicrobial Agents and Chemotherapy 58*, 3679-3688.

[54] King, N. M., Prabu-Jeyabalan, M., Bandaranayake, R. M., Nalam, M. N., Nalivaika, E. A., Özen, A. e. l., Haliloğlu, T. r., Yılmaz, N. e. K., and Schiffer, C. A. (2012) Extreme entropy–enthalpy compensation in a drug-resistant variant of HIV-1 protease, *ACS Chemical Biology 7*, 1536-1546.

[55] Kempf, D. J., Norbeck, D. W., Codacovi, L., Wang, X. C., Kohlbrenner, W. E., Wideburg, N. E., Paul, D. A., Knigge, M. F., and Vasavanonda, S. (1990) Structure-based, C2 symmetric inhibitors of HIV protease, *Journal of Medicinal Chemistry 33*, 2687-2689.

[56] Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand–protein interaction diagrams for drug discovery, *Journal of Chemical Information and Modeling 51*, 2778-2786.

[57] Studio, D. (2013) 4.0 Tutorials, receptor-ligand interaction. *Accelrys Inc, San Diego, CA, USA*.

[58] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF chimera—a visualization system for exploratory research and analysis, *Journal of Computational Chemistry 25*, 1605-1612.

[59] Li, H., Robertson, A. D., and Jensen, J. H. (2005) Very fast empirical prediction and rationalization of protein pKa values, *Proteins: Structure, Function, and Bioinformatics 61*, 704-721.

[60] http://nbcr-222.ucsd.edu/pdb2pqr_2.0.0/

[61] DeLano, W. L. (2002) The PyMOL molecular graphics system.

[62] Cole, J. C., Murray, C. W., Nissink, J. W. M., Taylor, R. D., and Taylor, R. (2005) Comparing protein–ligand docking programs is difficult, *Proteins: Structure, Function, and Bioinformatics 60*, 325-332.

[63] Gohlke, H., Hendlich, M., and Klebe, G. (2000) Knowledge-based scoring function to predict protein-ligand interactions, *Journal of Molecular Biology 295*, 337-356.

[64] Kontoyianni, M., McClellan, L. M., and Sokol, G. S. (2004) Evaluation of docking performance: comparative data on docking algorithms, *Journal of Medicinal Chemistry 47*, 558-565.

[65] Dennington, R., Keith, T., and Millam, J. (2009) Semichem Inc, *Shawnee Mission KS, GaussView, Version 5*.

[66] Kohn, W., Becke, A. D., and Parr, R. G. (1996) Density functional theory of electronic structure, *The Journal of Physical Chemistry 100*, 12974-12980.

[67] Neumann, R., Nobes, R. H., and Handy, N. C. (1996) Exchange functionals and potentials, *Molecular Physics 87*, 1-36.

[68] Becke, A. D. (1993) Density-functional thermochemistry. III. The role of exact exchange, *The Journal of Chemical Physics 98*, 5648-5652.

[69] Lee, C., Yang, W., and Parr, R. G. (1988) Development of the colle-salvetti correlation-energy formula into a functional of the electron density, *Physical Review B 37*, 785-789.

[70] Hariharan, P. C., and Pople, J. A. (1973) The influence of polarization functions on molecular orbital hydrogenation energies, *Theoretica Chimica Acta 28*, 213-222.

[71] Rassolov, V. A., Pople, J. A., Ratner, M. A., and Windus, T. L. (1998) 6-31G* basis set for atoms K through Zn, *The Journal of Chemical Physics 109*, 1223-1229.

[72] Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The AMBER biomolecular simulation programs, *Journal of Computational Chemistry 26*, 1668-1688.

[73] Vreven, T., and Morokuma, K. (2006) Hybrid methods: ONIOM (QM: MM) and QM/MM, *Annual Reports in Computational Chemistry 2*, 35-51.

[74] Vacca, J., Dorsey, B., Schleif, W., Levin, R., McDaniel, S., Darke, P., Zugay, J., Quintero, J., Blahy, O., and Roth, E. (1994) L-735,524: An orally bioavailable human immunodeficiency virus type 1 protease inhibitor, *Proceedings of the National Academy of Sciences 91*, 4096-4100.

[75] Kaldor, S. W., Kalish, V. J., Davies, J. F., Shetty, B. V., Fritz, J. E., Appelt, K., Burgess, J. A., Campanale, K. M., Chirgadze, N. Y., and Clawson, D. K. (1997) Viracept (nelfinavir mesylate, AG1343): a potent, orally bioavailable inhibitor of HIV-1 protease, *Journal of Medicinal Chemistry 40*, 3979-3985.

[76] Humbel, S., Sieber, S., and Morokuma, K. (1996) The IMOMO method: Integration of different levels of molecular orbital approximations for geometry optimization of large systems: Test for n-butane conformation and SN_2 reaction: RCl^+ Cl^-, *The Journal of ChemicalPphysics 105*, 1959-1967.

[77] Svensson, M., Humbel, S., Froese, R. D., Matsubara, T., Sieber, S., and Morokuma, K. (1996) ONIOM: a multilayered integrated MO+MM method for geometry optimizations and single point energy predictions. A test for diels-alder reactions and $Pt(P(t\text{-}Bu)_3)^{2+}$ H_2 oxidative addition, *The Journal of Physical Chemistry 100*, 19357-19363.

[78] Kapp, J., Remko, M., and Schleyer, P. v. R. (1996) H_2XO and $(CH_3)_2XO$ Compounds (X= C, Si, Ge, Sn, Pb): Double bonds vs carbene-like structures can the metal compounds exist at all?, *Journal of the American Chemical Society 118*, 5745-5751.

[79] Johnson, B. G., Gill, P. M., and Pople, J. A. (1993) The performance of a family of density functional methods, *The Journal of Chemical Physics 98*, 5612-5626.

[80] Frisch, M., Trucks, G., Schlegel, H., Scuseria, G., Robb, M., Cheeseman, J., Scalmani, G., Barone, V., Mennucci, B., and Petersson, G. (2009) 09, Revision D. 01, Gaussian, *Inc., Wallingford, CT*.

[81] Chung, L. W., Sameera, W., Ramozzi, R., Page, A. J., Hatanaka, M., Petrova, G. P., Harris, T. V., Li, X., Ke, Z., and Liu, F. (2015) The ONIOM method and its applications, *Chemical Reviews 115*, 5678-5796.

[82] Lundberg, M., Sasakura, Y., Zheng, G., and Morokuma, K. (2010) Case studies of ONIOM (DFT:DFTB) and ONIOM (DFT:DFTB:MM) for enzymes and enzyme mimics, *Journal of Chemical Theory and Computation 6*, 1413-1427.

[83] Ochterski, J. W. (2000) Thermochemistry in gaussian, *Gaussian Inc*, 1-19.

[84] Jenwitheesuk, E., and Samudrala, R. (2003) Improved prediction of HIV-1 protease-inhibitor binding energies by molecular dynamics simulations, *BMC Structural Biology 3*, 1.

[85] Perryman, A. L., Lin, J. H., and McCammon, J. A. (2004) HIV-1 protease molecular dynamics of a wild-type and of the V82F/I84V mutant: possible contributions to drug resistance and a potential new target site for drugs, *Protein Science 13*, 1108-1123.

[86] Freedberg, D. I., Ishima, R., Jacob, J., Wang, Y. X., Kustanovich, I., Louis, J. M., and Torchia, D. A. (2002) Rapid structural fluctuations of the free HIV protease flaps in solution: relationship to crystal structures and comparison with predictions of dynamics calculations, *Protein Science 11*, 221-232.

[87] Ryde, U. (2014) A fundamental view of enthalpy–entropy compensation, *MedChemComm 5*, 1324-1336.

CHAPTER FOUR

Exploring the Binding Free energies of FDA approved PR inhibitors against novel C-SA HIV-1 PR mutant L38L↑N↑L PR: ONIOM method

Zainab K. Sanusi,[a] Thavendran Govender,[a] Glenn E. M. Maguire,[a,b] Sibusiso B Maseko,[a] Johnson Lin,[c] Hendrik G. Kruger[a*] and Bahareh Honarparvar[a*]

[a]Catalysis and Peptide Research Unit, School of Health Sciences, University of KwaZulu-Natal, Durban 4001, South Africa.
[b]School of Chemistry and Physics, University of KwaZulu-Natal, 4001 Durban, South Africa.
[C]School of Life Sciences, University of KwaZulu-Natal, Durban 4001, South Africa.

***Corresponding authors:** Honarparvar@ukzn.ac.za (Dr Bahareh Honarparvar), kruger@ukzn.ac.za (Prof. Hendrik G. Kruger), Telephone: + 27 31 2601845, Fax: +27 31 2603091, Catalysis and Peptide Research Unit, School of Health Sciences, University of KwaZulu-Natal, Durban 4041, South Africa.

Abstract

The aspartate protease of the human immune deficiency type-1 virus (HIV-1) has become a crucial anti-viral target in which many useful antiretroviral inhibitors have been developed. However, it seems the emergence of new HIV-1 PR mutations leads to the increase in drug resistance, hence, the available FDA approved drugs show less activity towards the protease. A mutation and insertion designated L38L↑N↑L PR was recently reported from a novel subtype C-SA HIV-1 PR. An integrated two-layered ONIOM (QM:MM) method was employed in this study to examine the binding affinities of the nine FDA approved HIV PR inhibitors against this mutant. The catalytic Asp 25/25' and HIV PR inhibitors (HIV PIs) were treated at a high level QM theory using B3LYP/6-31G(d), while the MM level was assigned for the remaining part of the system using AMBER force field. The computed results as well as experimental data in general revealed a reduced inhibitory activity for all the various FDA approved drugs towards the L38L↑N↑L PR in comparison with subtype C-SA HIV-1 PR. This suggests that the insertion and mutations significantly affect the binding affinities or characteristics of the HIV PIs and/or parent PR. TPV (-61.2 kcal/mol) and NFV (-24.8 kcal/mol) were outliers with respect to

calculated binding free energies. The same trend for the computational binding free interaction energies was observed for eight of the nine inhibitors (TPV gave a better energy than experiment) with respect to the experimental binding free energies. The outcome of this study shows that ONIOM method can be used as a reliable computational approach to rationalize lead compounds against specific targets. The investigation carried out in this work can help to improve the efficiency of future HIV-1 PR antiviral drugs.

Keywords: L38L↑N↑L PR; HIV PR inhibitors; Inhibitor—enzyme interactions; Our Own N-layered Integrated molecular Orbital and molecular Mechanics (ONIOM); Binding free energies.

1.0 Introduction

The human immune deficiency virus (subtype B) protease is one of the most recognized and researched enzymes regarding its function and structure and these efforts have paved the way for the development of effective inhibitors[1, 2]. HIV-1 protease is a di-symmetric aspartyl protease that belongs to the group of retroviruses known as lentiviruses[2-5].

The protease (PR) is normally composed of two identical 99-amino acids in each monomer, in which the amino and carboxyl termini combine at the β-sheet dimer interface and form a hydrophobic binding site with two aspartic acids (ASP25/25') at the active site[6-8].

The HIV-1 PR functions by cleaving the gag and pol viral polypeptides into building blocks for protein synthesis to create a mature functional HIV virus[9-12]. HIV-1 protease is one of the principal target for HIV/AIDS drug inhibitors since the PR is essential for the development of viable progeny.[3, 13-15]. The presence of these inhibitors significantly retracts the function of the PR, preventing the virus from maturing in the affected individual[10, 16]. The nine HIV-1 protease antiretroviral drugs approved by the FDA were developed for subtype B which is the more common strain found in Australia, Western Europe and North America. These drugs exhibit weaker activities against subtype C and A found in sub-Sahara Africa and India[17-19].

The development of HIV PR resistance due to mutations causes enormous hindrance in attaining long-term suppression of HIV replication in patients receiving anti-retroviral drugs[20]. The alterations in HIV PR was investigated by Otto et al[21]. They argued the actual importance of the

observed alterations is unclear since the mutations repeatedly occur with extra mutations in the protease and also in the viral genome[21]. Accessing the growth properties of the mutations and effect of the alteration in the viral genome will help in understanding the function of the protease gene[22, 23]. Hence, improvements can only be made on the new designed drugs since much cannot be done on the genetic source of drug resistance[24-27].

Our group has reported inhibitor synthesis for C-SA HIV protease and also computational methods to rationalize observed experimental data[28-35]. A computational model to calculate the binding free energies of HIV PR inhibitors against subtype B and C-SA PR utilizing MM/MD methods was developed by our group[36-38]. These computational results revealed that the absolute experimental values for the binding free energies versus theoretical values differ due to the available parameterization implemented in the theoretical model, which is an approximation of experimental data[39]. However, the calculated results follow a similar trend with the experimental data for subtype B[36, 37].

We have also recently investigated an ONIOM[40, 41] model to calculate the binding free interaction energies of the nine FDA HIV-1 protease inhibitors against subtype B and C-SA PR;[42] these were compared with the experimental binding free energies data[43]. It was observed that the obtained theoretical data follow a satisfactory trend with the reported experimental data for the subtype B PR. However, the applied model for the C-SA HIV PR revealed limitations[42].

Several other studies have utilized the ONIOM model for calculating interaction energies of wild-type HIV PR-1 with selected commercial inhibitors[44-46]. In all cases, the catalytic aspartate residues were treated at a high-level theory (DFT), and the rest of the system was modeled with MM. The choice of protonation state for the catalytic HIV PR aspartates (Asp 25/25') for ONIOM calculations have been studied extensively in literature[28, 33, 47-49]. The p*K*a of one of the catalytic aspartate increases to 5.2[2] when bound to the inhibitor and the corresponding value is p*K*a 4.5[50] when unbound[51]. This implies that one of the two catalytic Asp groups should be protonated for binding studies, while the other Asp is unprotonated[30].

Herein, we report a hybrid ONIOM model to compute the free binding energies of the nine FDA approved HIV-1 protease inhibitors towards a new HIV subtype C-SA PR mutant that also

experience two insertions. The C-SA HIV-1 protease mutant consists of two extra amino acids, resulting in a C-SA PR mutant (L38L↑N↑L)[52] with 101-amino acid residues in each monomer. The arrow (↑) before the amino acids shows that asparagine and leucine are inserted at position 38 respectively. In addition, the following five mutations have also occurred in the protease E35D, I36G, N37S, M46L and D60E[52].

This variant was found in a patient that was drug-naïve to commercially accessible HIV protease drugs but reacted positively to the following reverse transcriptase inhibitors (RTIs): d4t (stavudine), 3TC (lamivudine), and efavirenz[52].

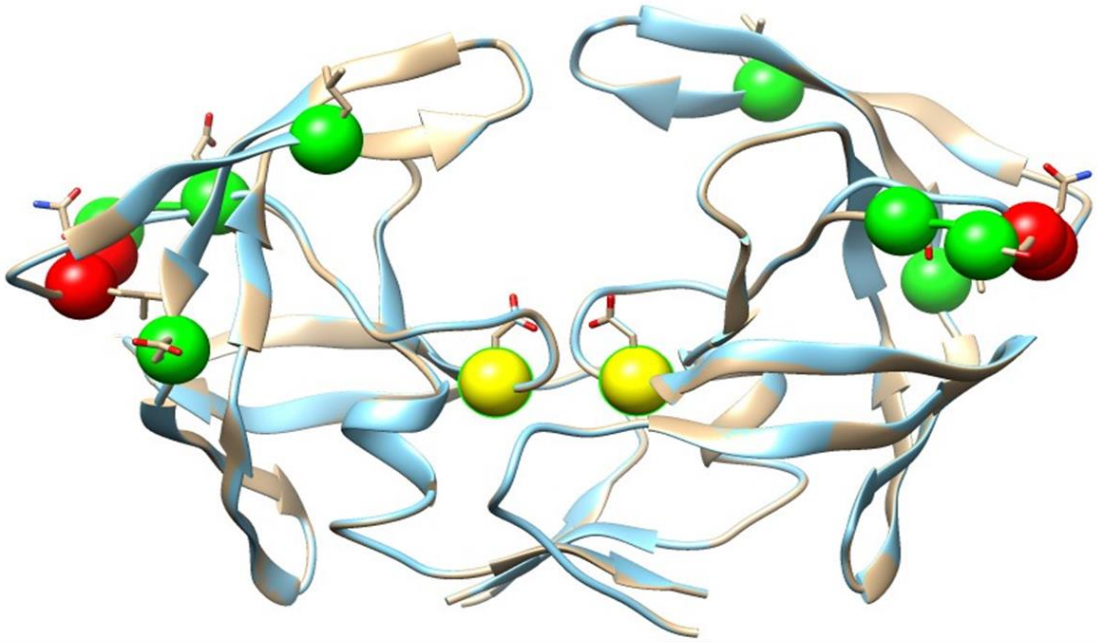


Figure 6. Schematic representation of super-imposed subtype C-SA(blue) and mutant L38L↑N↑L PR(brown), showing the position of the mutations (E35D, I36G, N37S, M46L and D60E) in green, insertions (asparagine and leucine) are depicted in red, and Asp25/25' in yellow.

2.0 Computational Methods

Homology modelling[30, 53-56] was used to generate the three-dimensional structure for L38L↑N↑L protease that is necessary to start the computational process, as there is no single crystal X-ray structure available presently. The same computational model as our recent ONIOM paper on Subtype B and C-SA was utilized[42]. In this multi-layered approach, the active site was treated

with a high level Density Functional theory[57, 58] using B3LYP[59, 60]/6-31G(d)[61, 62] basis set, while the rest of the protease was treated at a low MM level using AMBER[63] force field. Further structural analysis to estimate the mode of interactions between the ligand and L38L↑N↑L protease will be performed using the Accelrys (Discovery) Visualizer[64] and Ligplot[65] software.

2.1 Structural Preparation of Inhibitor—Enzyme Complexes

The structures of clinically available FDA inhibitors complexed with the HIV-1 subtype B protease were taken from the Protein Data Bank (PDB). The PDB codes are as follows: 4YOA (DRV)[66], 4L1A (LPV)[67], 4EYR (RTV)[27], 3WSJ (IDV)[68], 3S56 (SQV)[69], 3S45 (APV)[69], 2PYM (NFV)[70], 4NJU (TPV)[71], 3EM4 (ATV)[72] (**Figure S1**).

Since the X-ray structures for L38L↑N↑L PR complexed with the FDA approved drugs have not yet been reported, structural preparation input structures for the nine drugs complexed to L38L↑N↑L PR, were performed using the same overlay method reported previously[36, 37, 42]. The 3D structures for all the inhibitor—L38L↑N↑L mutant PR complexes were generated by superimposing the corresponding subtype B HIV PR—inhibitor crystal complexes with the L38L↑N↑L PR using PyMOL[73], in order for the inhibitors to maintain the same position as in the subtype B PR. PyMOL evaluates the root mean square (RMS), which is a helpful measure of how well the inhibitor—enzyme complexes were superimposed. An optimal superimposition is considered acceptable if the RMS is less than 2 Å[74-76].

The protonation state of the L38L↑N↑L PR structure was assigned using PROPKA[77, 78] based on the P*K*a values at pH 7 as presented before[42]. This depicts that the catalytic aspartate at the active binding site in both chains Asp25 and Asp25' are deprotonated and protonated respectively. It is notable that the carbonyl and amino terminus as well as the Asp, Gly, Arg, Lys amino acids were charged, while His is kept in its neutral form.

The structures of all inhibitor—enzyme complexes were refined afterwards by manually removing the ions and crystallographic water, that are present from the protein from the PDB file using a text editor. Thereafter, protons were added to the required catalytic aspartate using GaussView[79]. The catalytic Asp25/25' residues and the inhibitors that constitute the smallest part of the system were considered at a high level (QM/DFT[57, 58] - Beck3LYP[59, 60]/6-31G(d)[61, 62])

and the remaining larger part of the system at low layer (MM - AMBER[63]) for subsequent ONIOM[40, 41] calculations. Details about the preparation of the ONIOM starting structures were presented before[42].

All the 3D inhibitor—protease complex structures are created using PyMOL[73] software and are provided in supplementary material **Figure S2**. The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided with the supplementary material.

2.2 The ONIOM Evaluation Method

An ONIOM[40, 80-82] computational model was used to calculate the binding free energies of the various PR inhibitors with the L38L↑N↑L PR. Preceding studies showed that the DFT level theory at the high level is the most popular approach due to its popular balance of accuracy and efficiency[59, 60, 83], and B3LYP method provides good energies and are in excellent agreement with *ab initio* high level results[84-86]. Therefore, the geometry of the various nine FDA approved drugs complexed with the L38L↑N↑L PR were optimized using the two-layer ONIOM approach (B3LYP/6-31G(d):AMBER) QM:MM level of theory in Gaussian09[87] developed version. The extrapolated energy E_{ONIOM2}[88, 89] is defined as:

$$\Delta E_{ONIOM2} = \Delta E_{real,low} + \Delta E_{model,high} - \Delta E_{model,low} \qquad \mathbf{1}$$

Where, ΔE_{real} is the energy of the entire (real) system and ΔE_{model} is the energies of the model system calculated at the high and low level respectively. A schematic representation of the ONIOM2 model is presented in **Figure 2**.

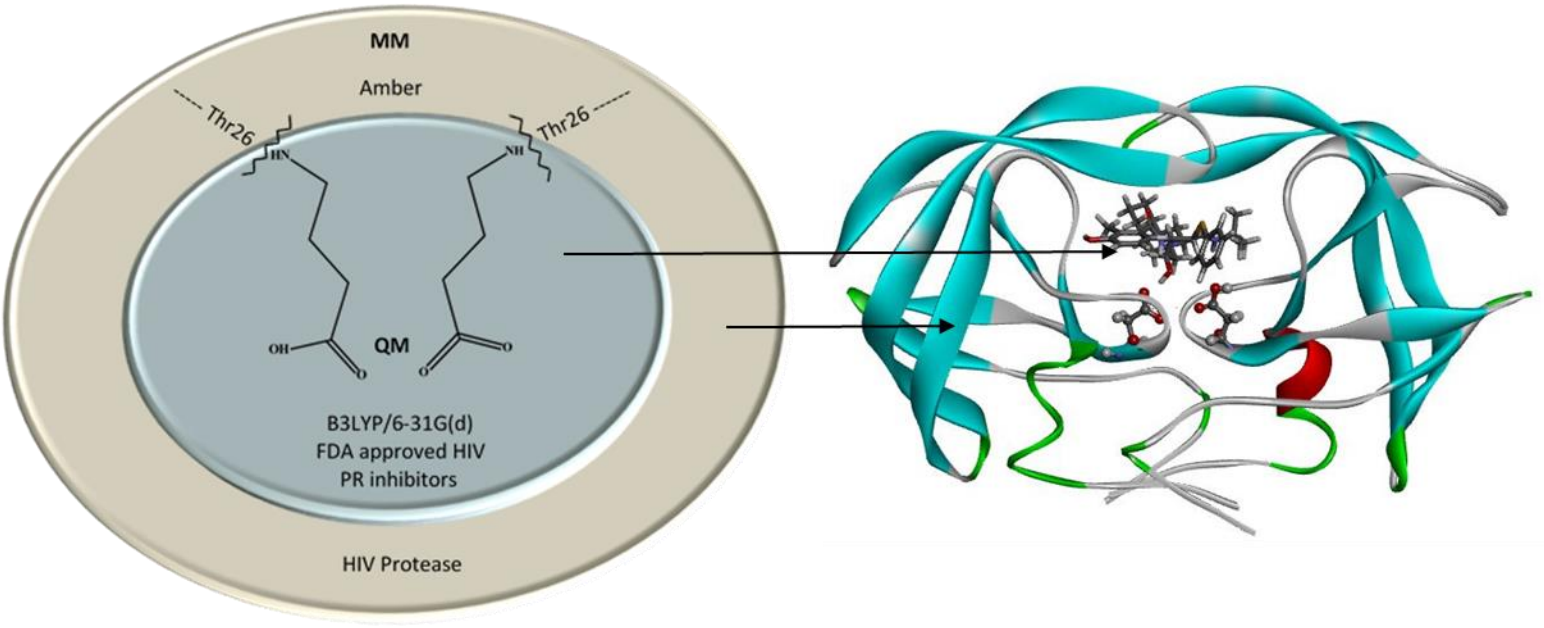


Figure 2. Schematic diagram of the calculated two-layered ONIOM theoretical levels (B3LYP/6-31G(d):AMBER) of L38L↑N↑L PR complex with APV.

The Gibbs binding free energies (ΔG) for the systems was achieved from frequency calculations of the optimized complexed structures using the 2-layer ONIOM algorithm as discussed above. Thus, the change in free energies of the corresponding complex systems is expressed as:

$$\Delta G_{ONIOM} \approx \Delta G_{bind} = G_{complex} - G_{ligand} + G_{protein} \qquad \mathbf{2}$$

3.0 Results and Discussion

In order to ascertain the efficiency of the FDA approved drugs against this subtype, binding free interaction energy calculations were performed on all the inhibitor—protease complexes with the ONIOM computational model. The average RMS values of the aligned inhibitor—L38L↑N↑L complex is 0.7 Å, which indicates that the complexes were correctly superimposed[74-76]. The calculated results for L38L↑N↑L PR were ranked with the experimental data[43], and also with our previously reported computational data for C-SA HIV PR[42] (**Table 1**).

The experimental results of the FDA approved drugs exhibit reduced binding free energies for the L38L↑N↑L PR mutant compared to the subtype C-SA HIV PR. Only APV possesses a reasonable binding affinity for the mutant protease in comparison to the rest of the inhibitors.

The trends of the calculated binding free energies for the FDA drugs towards C-SA HIV PR has been discussed in our previous work[42]. Here, the ONIOM (B3LYP/6-31G(d):AMBER) binding free energies of L38L↑N↑L PR demonstrate the same general trend as the experimental Gibb's free energy data but exhibit reduced theoretical binding affinities compared to subtype C-SA PR (**Table 1)**. However, there were two exemptions (outliers); TPV shows much better binding free interaction energy (-61.2 kcal/mol) than other drugs complexed to the mutant enzyme. Although NFV exhibits the weakest binding affinity in both experimental and theoretical results for L38L↑N↑, the computed result appears significantly smaller than the other inhibitors (**Table 1**).

Table 1. The binding free energies[a] (kcal/mol) for FDA HIV PIs against C-SA and L38L↑N↑L PRs [reported experimental results[90] and calculated ONIOM (B3LYP/6-31G(d):AMBER) results].

Inhibitors	ΔG_{bind} (Exp)[b]	ΔG_{bind} (Exp)[c]	ΔG_{bind} (Calc)[d]	ΔG_{bind} (Calc)[e]
	C-SA	L38L↑N↑L	C-SA	L38L↑N↑L
APV	-13.9	-13.1	-69.0	-56.9
RTV	-13.9	-12.9	-62.9	-56.4
IDV	-14.0	-12.1	-64.0	-52.1
LPV	-13.2	-11.6	-56.5	-51.8
SQV	-13.4	-10.1	-57.1	-45.8
DRV	-13.8	-9.88	-62.8	-43.9
ATV	-14.4	-9.69	-66.9	-43.3
TPV	-13.2	-9.47	-78.9	-61.2
NFV	-13.5	-9.38	-38.6	-24.8

[a]HIV PR inhibitors with respect to their binding energies.
[b,c]Experimental data for wildtype C-SA HIV PR and L38L↑N↑L PR carried out by our group respectively[43, 90].
[d]Calculated binding energies of wild type C-SA HIV PR previously reported by our group[42].
[e]Calculated binding energies for L38L↑N↑L PR.
The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided with the supplementary material.

As it was argued in our previous study[42], the reason for these can potentially be attributed to the simplified computational model (the omission of water in the active site of the protease), rendering the model less accurate and this possibility will be interrogated further.

The insertions and mutations of amino acids residues is expected to change the structure of the mutant PR having an impact on the binding interactions between the inhibitors and the protease. Hence, detailed hydrogen bond interactions of the inhibitor—enzyme complexes were measured using Accelrys (Discovery) Visualizer[64]. The change in hydrogen bond distances were measured between both catalytic aspartates Asp25/25' in the binding site and the hydroxyl group of the FDA approved inhibitors before and after optimization (**Figure 3**). In all cases, the hydroxyl group of the inhibitors form similar hydrogen bond interactions with the Asp25/25' of L38L↑N↑L PR (**Figure S3**, supplementary material). For the inhibitors exhibiting slightly better binding affinities with the L38L↑N↑L PR (APV, RTV, IDV and LPV), a large reduction in the average hydrogen bond distance 0.7 Å was observed after optimization. While inhibitors with weaker binding affinities (SQV, DRV, and ATV) revealed a smaller reduction of the average hydrogen bond distance 0.1 Å.

For the outliers; TPV and NFV, a comparison of the hydrogen bond distances (**Figure 3**) reveals that TPV (1.6 Å; -61.2 kcal/mol) has a stronger HB interaction than NFV (1.9 Å; -24.8 kcal/mol), explaining the difference in theoretical binding free energies for these two outliers.

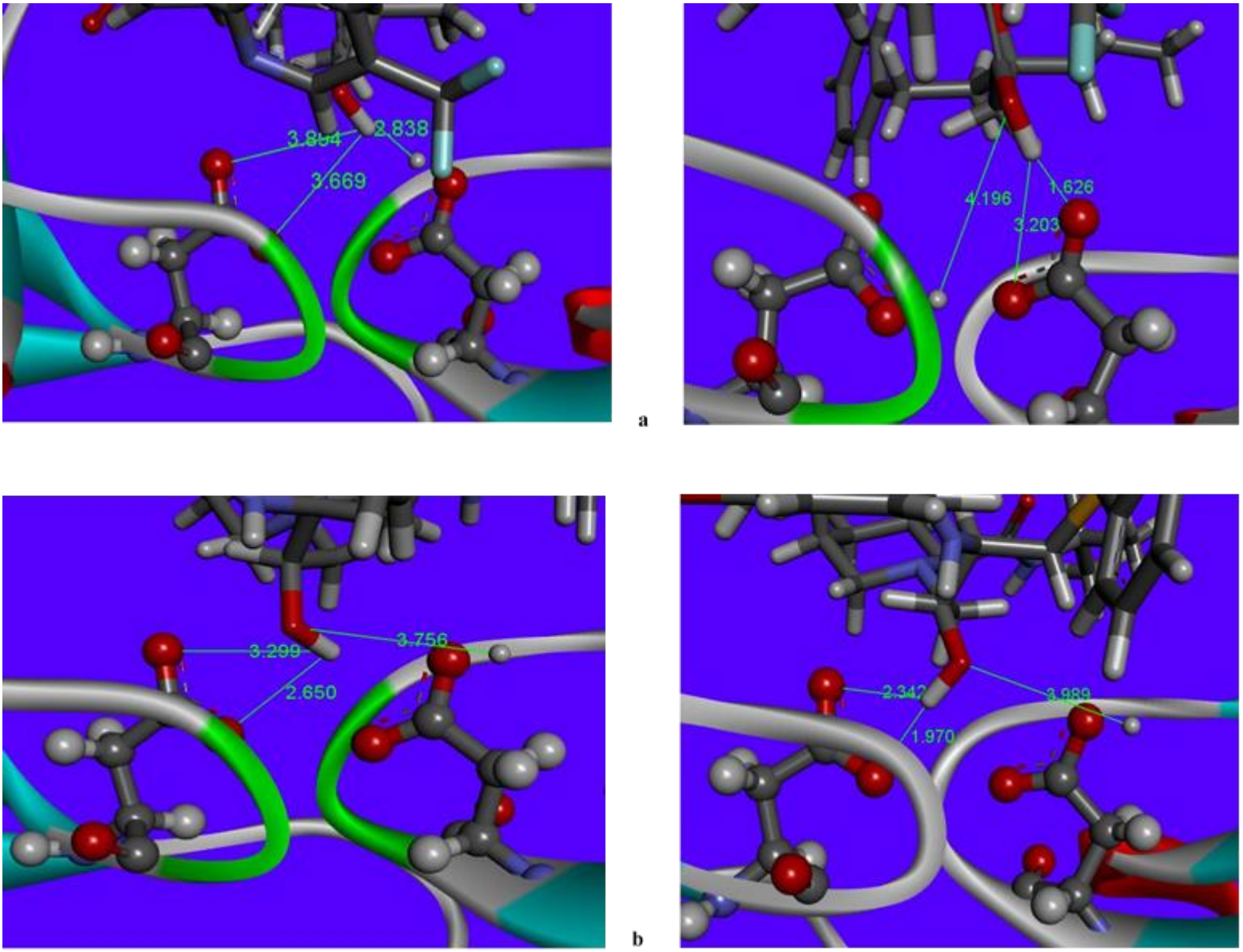


Figure 3. Hydrogen bond distances between the hydroxyl groups of TPV and NFV drugs with the catalytic ASP25 and ASP25' residues of **a:** L38L↑N↑L—TPV PR, **b:** L38L↑N↑L—NFV PR before and after optimization. Detailed comparative plots for all inhibitor—enzyme complexes are provided in supporting information **Figure S3**. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided with the supplementary material.)

In an attempt to further understand the binding environment within the L38L↑N↑L PR, electrostatic and hydrogen bond interactions were plotted for all inhibitor—enzyme complexes (**Figure S4** in the supplementary material) with Ligplot[65]. The two exemptions; TPV and NFV are shown in **Figure 4**. The plot for TPV reveals hydrogen bond and electrostatic interactions between the inhibitor and other side chain residues (NH-OH, NH-NH) in the protease which

most likely contribute to the improved binding calculation result. For NFV hydrogen bond interactions occur with the catalytic aspartates and one side chain (mainly OH-OH) resulting in weaker binding affinity.

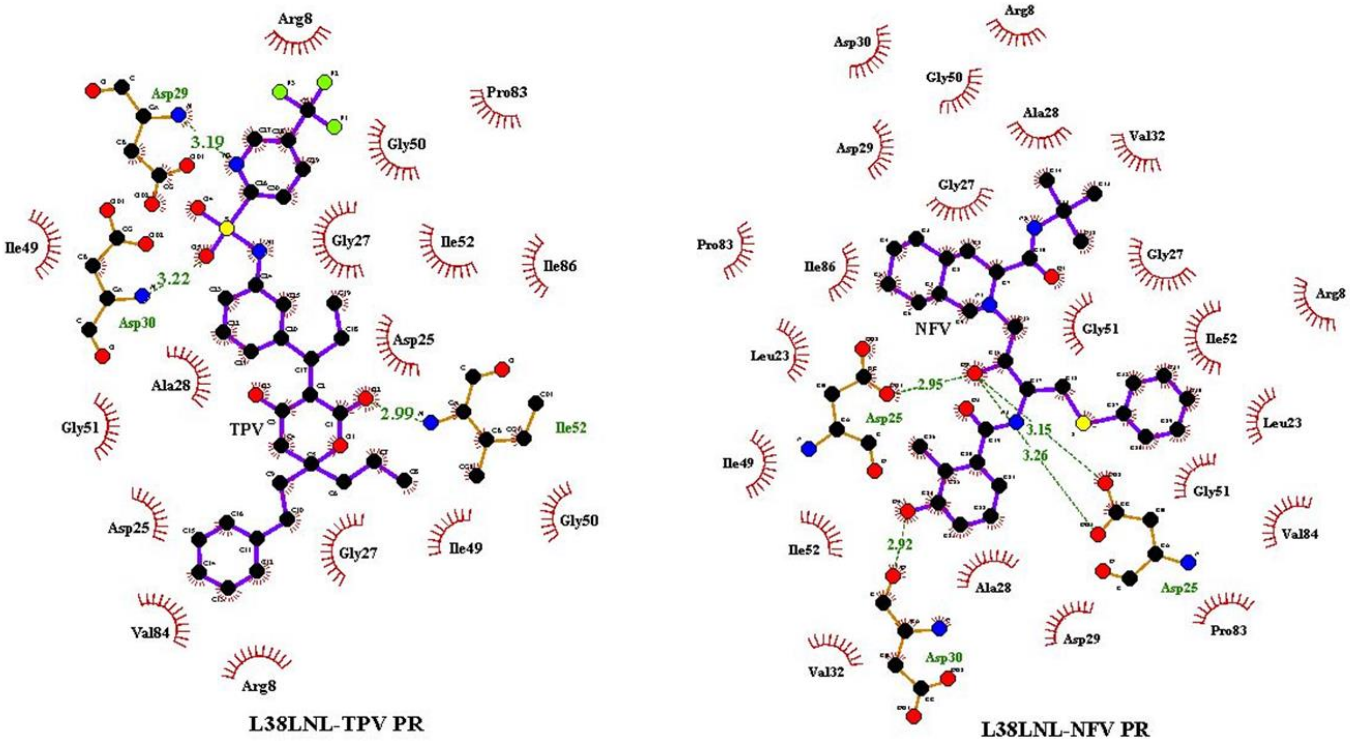


Figure 4. Electrostatic and hydrogen bond interactions plots of L38L↑N↑L PR complexed with TPV and NPV. The plot were created after optimization of each complex system using Ligplot[65]. Detailed plots showing the electrostatic and hydrogen bond interactions are provided in supporting information **Figure S4**. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided with the supplementary material.)

As previously discussed, the calculated binding free energies (ΔG) for the various FDA approved inhibitors complex with L38L↑N↑L PR follow the same trend with the corresponding experimental data (**Table 1**), except for TPV and NFV. The thermochemical properties can be used to rationalize these outliers, since ΔG is a function of both enthalpy (ΔH) and entropy (ΔS). The calculated energies of the various drugs complexed with L38L↑N↑L PR are depicted in **Table 2**.

The enthalpy contribution of TPV (ΔH -82.3 kcal/mol) is the highest of all cases, explaining why this drug binds better in theory than the other inhibitors (the enthalpy contribution is far greater than entropy to the Gibbs free binding energies). As argued before (see discussion around **Figure 3**), this is most likely the result of our simplified computational model, where the interaction between the TPV hydroxyl group and the Asp25/25' residues are over emphasized. On the other hand, NFV reveals a much reduced enthalpy contribution ΔH = -42.1 kcal/mol in comparison to the other inhibitor—enzyme complexes, explaining the weak calculated binding free interaction energy (ΔG -24.8 kcal/mol).

As for the entropy results (ΔS) of the various drug-complexes, it should be noted that a more negative entropy indicates greater restrictions of movement for the ligand in the active site, due to steric restrictions as well as stronger non-covalent inhibitor-enzyme interactions for certain parts of the inhibitor[91]. Such cases result in higher entropy penalties. The translational and rotational entropy contributions are in a close approximate range (-40 cal/mol K and -30 cal/mol K respectively) for all the inhibitors and only vary in the residual vibrational entropy ΔS_{vib} (**Table 2**).

Table 2. The binding free energies, enthalpies and entropy of the various FDA approved HIV PIs against L38L↑N↑L PR.

Inhibitors	ΔG^a kcal mol^{-1}	ΔH kcal mol^{-1}	ΔS_{total} cal mol^{-1} K^{-1}	ΔS_{trans} cal mol^{-1} K^{-1}	ΔS_{vib} cal mol^{-1} K^{-1}	ΔS_{rot} cal mol^{-1} K^{-1}
APV	-56.9	-80.6	-79.4	-44.5	2.8	-37.7
RTV	-56.4	-80.5	-80.7	-45.5	4.7	-39.9
IDV	-52.1	-80.6	-95.7	-45.0	-11.3	-39.4
LPV	-51.8	-79.9	-94.1	-45.0	-10.1	-39.0
SQV	-45.8	-72.5	-89.4	-44.8	-5.2	-39.4
DRV	-43.9	-68.8	-83.4	-44.5	0.5	-39.4
ATV	-43.3	-72.0	-94.1	-44.8	-11.1	-38.2
TPV	-61.2	-82.3	-70.7	-45.0	13.6	-39.3
NFV	-24.8	-42.1	-57.9	-44.8	25.3	-38.4

HIV protease inhibitors (HIV PIs) are ranked in terms of their binding free energies (ΔG).
[a]Calculated binding free energies using ONIOM for L38L↑N↑L PR from **Table 1**.
The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided with the supplementary material.

The level of consistency attained by our theoretical method with the experimental data (**Table 1)** in this study suggests the computational model can be used to rationalize the different drugs—mutant PR binding interactions. However, the movement of the inhibitors closer to the Asp25/25' residues during optimization, imply that omission of water molecules in the model is an over-simplification[42] and should be addressed in future studies to further improve the computational model.

4.0 Conclusion

This study involves the testing of a computational two-layered ONIOM computational model (QM:MM) to calculate the binding free energies of the nine FDA approved HIV-1 PIs. The calculated binding free energies for L38L↑N↑L HIV PR show a satisfactory trend with the experimental data with two exceptions. Two outliers were observed, and analysis of hydrogen bond interactions and enthalpy contributions explained the observed anomalies. It was concluded that our ONIOM (B3LYP/6-31G(d):AMBER) model can be further improved by the

addition of water in the active pocket of the protease. Explicit water molecules should at least be treated at a semi-empirical level (PM6). These results will assist to systematically improve our computational model for the potential design of more potent HIV protease inhibitors.

Competing interests

The authors declare that they have no competing interests.

Acknowledgment

We thank the College of Health Sciences (CHS), Aspen Pharmacare, MRC and the NRF for financial support. We are also grateful to the CHPC (www.chpc.ac.za) and UKZN HPC cluster as our computational resources.

References

[1] Mildner, A. M., Rothrock, D. J., Leone, J. W., Bannow, C. A., Lull, J. M., Reardon, I. M., Sarcich, J. L., Howe, W. J., and Tomich, C.-S. C. (1994) The HIV-1 protease as enzyme and substrate: mutagenesis of autolysis sites and generation of a stable mutant with retained kinetic properties, *Biochemistry 33*, 9405-9413.

[2] Brik, A., and Wong, C.-H. (2003) HIV-1 protease: mechanism and drug discovery, *Organic & Biomolecular Chemistry 1*, 5-14.

[3] Hensen, C., Hermann, J. C., Nam, K., Ma, S., Gao, J., and Höltje, H.-D. (2004) A combined QM/MM approach to protein-ligand interactions: polarization effects of the HIV-1 protease on selected high affinity inhibitors, *Journal of Medicinal Chemistry 47*, 6673-6680.

[4] Shehu-Xhilaga, M., and Oelrichs, R. (2009) Basic HIV virology, *HIV Management in Australasia*, 9-18.

[5] Pearl, L. H., and Taylor, W. R. (1987) A structural model for the retroviral proteases, *Nature 329*, 351-354.

[6] Braz, A. S., Tufanetto, P., Perahia, D., and Scott, L. P. (2012) Relation between flexibility and positively selected HIV-1 protease mutants against inhibitors, *Proteins: Structure, Function, and Bioinformatics 80*, 2680-2691.

[7] Bonini, C., Chiummiento, L., De Bonis, M., Di Blasio, N., Funicello, M., Lupattelli, P., Pandolfo, R., Tramutola, F., and Berti, F. (2010) Synthesis of new thienyl ring containing HIV-1 protease inhibitors: promising preliminary pharmacological evaluation against recombinant HIV-1 proteases, *Journal of Medicinal Chemistry 53*, 1451-1457.

[8] Mager, P. P. (2001) The active site of HIV-1 protease, *Medicinal Research Reviews 21*, 348-353.

[9] Clavel, F., and Mammano, F. (2010) Role of Gag in HIV resistance to protease inhibitors, *Viruses 2*, 1411-1426.

[10] Deeks, S. G., Smith, M., Holodniy, M., and Kahn, J. O. (1997) HIV-1 protease inhibitors: a review for clinicians, *Jama 277*, 145-153.

[11] Mitsuya, H., Yarchoan, R., and Broder, S. (1990) Molecular targets for AIDS therapy, *Science 249*, 1533-1544.

[12] Katz, R. A., and Skalka, A. M. (1994) The retroviral enzymes, *Annual Review of Biochemistry 63*, 133-173.

[13] Li, D., Ji, B., Hwang, K.-C., and Huang, Y. (2011) Strength of hydrogen bond network takes crucial roles in the dissociation process of inhibitors from the HIV-1 protease binding pocket, *PLoS One 6*, e19268.

[14] Kramer, R., Schaber, M., Skalka, A., Ganguly, K., Wong-Staal, F., and Reddy, E. (1986) HTLV-III gag protein is processed in yeast cells by the virus pol-protease, *Science 231*, 1580-1584.

[15] Huff, J. R. (1991) HIV protease: a novel chemotherapeutic target for AIDS, *Journal of Medicinal Chemistry 34*, 2305-2314.

[16] Hohlfeld, K., Wegner, J. r. K., Kesteleyn, B., Linclau, B., and Unge, J. (2015) Disubstituted bis-THF moieties as new P2 ligands in nonpeptidal HIV-1 protease inhibitors (II), *Journal of Medicinal Chemistry 58*, 4029-4038.

[17] Genoni, A., Morra, G., Merz Jr, K. M., and Colombo, G. (2010) Computational study of the resistance shown by the subtype B/HIV-1 protease to currently known inhibitors, *Biochemistry 49*, 4283-4295.

[18] Robbins, A. H., Coman, R. M., Bracho-Sanchez, E., Fernandez, M. A., Gilliland, C. T., Li, M., Agbandje-McKenna, M., Wlodawer, A., Dunn, B. M., and McKenna, R. (2010) Structure of the unbound form of HIV-1 subtype A protease: comparison with unbound forms of proteases from other HIV subtypes, *Acta Crystallographica Section D: Biological Crystallography 66*, 233-242.

[19] Mosebi, S., Morris, L., Dirr, H. W., and Sayed, Y. (2008) Active-site mutations in the south african human immunodeficiency virus type 1 subtype C protease have a significant impact on clinical inhibitor binding: kinetic and thermodynamic study, *Journal of Virology 82*, 11476-11479.

[20] Verheyen, J., Knops, E., Kupfer, B., Hamouda, O., Somogyi, S., Schuldenzucker, U., Hoffmann, D., Kaiser, R., Pfister, H., and Kücherer, C. (2009) Prevalence of C-terminal gag cleavage site mutations in HIV from therapy-naive patients, *Journal of Infection 58*, 61-67.

[21] Hodge, C. N., Aldrich, P. E., Bacheler, L. T., Chang, C.-H., Eyermann, C. J., Garber, S., Grubb, M., Jackson, D. A., Jadhav, P. K., and Korant, B. (1996) Improved cyclic urea inhibitors of the HIV-1 protease: synthesis, potency, resistance profile, human pharmacokinetics and X-ray crystal structure of DMP 450, *Chemistry & Biology 3*, 301-314.

[22] Brown, D., Arzumanov, A. A., Turner, J. J., Stetsenko, D. A., Lever, A. M., and Gait, M. J. (2005) Antiviral activity of steric-block oligonucleotides targeting the HIV-1 trans-activation response and packaging signal stem-loop RNAs, *Nucleosides, Nucleotides and Nucleic Acids 24*, 393-396.

[23] Nijhuis, M., van Maarseveen, N. M., Lastere, S., Schipper, P., Coakley, E., Glass, B., Rovenska, M., de Jong, D., Chappey, C., and Goedegebuure, I. W. (2007) A novel substrate-based HIV-1 protease inhibitor drug resistance mechanism, *PLoS Med 4*, e36.

[24] Lampejo, T., and Pillay, D. (2013) HIV virology, testing and monitoring, *Medicine 41*, 420-424.

[25] Mao, Y. (2011) Dynamical basis for drug resistance of HIV-1 protease, *BMC Structural Biology 11*, 1-9.

[26] Palmisano, L., and Vella, S. (2011) A brief history of antiretroviral therapy of HIV infection: success and challenges, *Annali Dell'Istituto Superiore Di Sanitã 47*, 44-48.

[27] Liu, Z., Yedidi, R. S., Wang, Y., Dewdney, T. G., Reiter, S. J., Brunzelle, J. S., Kovari, I. A., and Kovari, L. C. (2013) Insights into the mechanism of drug resistance: X-ray structure analysis of multi-drug resistant HIV-1 protease ritonavir complex, *Biochemical and Biophysical Research Communications 431*, 232-238.

[28] Honarparvar, B., Makatini, M. M., Pawar, S. A., Petzold, K., Soliman, M. E., Arvidsson, P. I., Sayed, Y., Govender, T., Maguire, G. E., and Kruger, H. G. (2012) Pentacycloundecane-diol-based HIV-1 protease inhibitors: biological screening, 2D NMR, and molecular simulation studies, *ChemMedChem 7*, 1009-1019.

[29] Makatini, M. M., Petzold, K., Alves, C. N., Arvidsson, P. I., Honarparvar, B., Govender, P., Govender, T., Kruger, H. G., Sayed, Y., and JerônimoLameira. (2013) Synthesis, 2D-NMR and molecular modelling studies of pentacycloundecane lactam-peptides and peptoids as potential HIV-1 wild type C-SA protease inhibitors, *Journal of Enzyme Inhibition and Medicinal Chemistry 28*, 78-88.

[30] Honarparvar, B., Govender, T., Maguire, G. E., Soliman, M. E., and Kruger, H. G. (2013) Integrated approach to structure-based enzymatic drug design: molecular modeling, spectroscopy, and experimental bioactivity, *Chemical Reviews 114*, 493-537.

[31] Pawar, S. A., Jabgunde, A. M., Govender, P., Maguire, G. E., Kruger, H. G., Parboosing, R., Soliman, M. E., Sayed, Y., Dhavale, D. D., and Govender, T. (2012) Synthesis and molecular modelling studies of novel carbapeptide analogs for inhibition of HIV-1 protease, *European Journal of Medicinal Chemistry 53*, 13-21.

[32] Pawar, S. A., Jabgunde, A. M., Maguire, G. E., Kruger, H. G., Sayed, Y., Soliman, M. E., Dhavale, D. D., and Govender, T. (2013) Linear and cyclic glycopeptide as HIV protease inhibitors, *European Journal of Medicinal Chemistry 60*, 144-154.

[33] Karpoormath, R., Sayed, Y., Govender, P., Govender, T., Kruger, H. G., Soliman, M. E., and Maguire, G. E. (2012) Pentacycloundecane derived hydroxy acid peptides: a new class of irreversible non-scissile ether bridged type isoster as potential HIV-1 wild type C-SA protease inhibitors, *Bioorganic Chemistry 40*, 19-29.

[34] Makatini, M. M., Petzold, K., Arvidsson, P. I., Honarparvar, B., Govender, T., Maguire, G. E., Parboosing, R., Sayed, Y., Soliman, M. E., and Kruger, H. G. (2012) Synthesis, screening and computational investigation of pentacycloundecane-peptoids as potent CSA-HIV PR inhibitors, *European Journal of Medicinal Chemistry 57*, 459-467.

[35] Karpoormath, R., Sayed, Y., Govender, T., Kruger, H. G., Soliman, M. E., and Maguire, G. E. (2013) Novel PCU cage diol peptides as potential targets against wild-type C-SA HIV-1 protease: synthesis, biological screening and molecular modelling studies, *Medicinal Chemistry Research 22*, 3918-3933.

[36] Ahmed, S. M., Kruger, H. G., Govender, T., Maguire, G. E., Sayed, Y., Ibrahim, M. A., Naicker, P., and Soliman, M. E. (2013) Comparison of the molecular dynamics and calculated binding free energies for nine FDA-approved HIV-1 PR drugs against subtype B and C-SA HIV PR, *Chemical Biology & Drug Design 81*, 208-218.

[37] Lockhat, H. A., Silva, J. R., Alves, C. N., Govender, T., Lameira, J., Maguire, G. E., Sayed, Y., and Kruger, H. G. (2015) Binding free energy calculations of nine FDA-approved protease inhibitors against HIV-1 subtype C I36T↑ T containing 100 amino acids per monomer, *Chemical Biology & Drug design 87*, 487-498.

[38] Ahmed, S. M., Maguire, G. E., Kruger, H. G., and Govender, T. (2014) The impact of active site mutations of south african HIV PR on drug resistance: insight from molecular dynamics simulations, binding free energy and per-residue footprints, *Chemical Biology & Drug Design 83*, 472-481.

[39] Frisch, A., and Foresman, J. (1996) Exploring chemistry with electronic structure methods, *Pittsburgh PA: Gaussian Inc 302*.

[40] Vreven, T., and Morokuma, K. (2000) On the application of the IMOMO (integrated molecular orbital+molecular orbital) method, *Journal of Computational Chemistry 21*, 1419-1432.

[41] Vreven, T., and Morokuma, K. (2006) Hybrid methods: ONIOM (QM:MM) and QM/MM, *Annual Reports in Computational Chemistry 2*, 35-51.

[42] Zainab, K. S. T., Govender; Glenn, E. M. Maguire; Hendrik, G. Kruger; Bahareh, Honarparvar;. (2016) Investigation of the binding free interaction energies of FDA approved HIV-1 PR inhibitors against subtype B and C-SA HIV PR: ONIOM approach, *Biophysical Journal, Submitted for publication*.

[43] Sibusiso, B. M. E., Padayachee; Thavendran, Govendera; Yasien, Sayed; Gert, Kruger; Glenn, E.M Maguire; Johnson, Linb;. (2016) I36T↑T in C-SA HIV-1 protease significantly alters protease drug interactions, *Antiviral Research, Submitted for Pubication*.

[44] Fong, P., McNamara, J. P., Hillier, I. H., and Bryce, R. A. (2009) Assessment of QM/MM scoring functions for molecular docking to HIV-1 protease, *Journal of Chemical Information and Modeling 49*, 913-924.

[45] Saen-oon, S., Aruksakunwong, O., Wittayanarakul, K., Sompornpisut, P., and Hannongbua, S. (2007) Insight into analysis of interactions of saquinavir with HIV-1 protease in comparison between the wild-type and G48V and G48V/L90M mutants based on QM and QM/MM calculations, *Journal of Molecular Graphics and Modelling 26*, 720-727.

[46] Wittayanarakul, K., Aruksakunwong, O., Saen-oon, S., Chantratita, W., Parasuk, V., Sompornpisut, P., and Hannongbua, S. (2005) Insights into saquinavir resistance in the G48V HIV-1 protease: quantum calculations and molecular dynamic simulations, *Biophysical Journal 88*, 867-879.

[47] Shi, S., Hu, G., Chen, J., Zhang, S., and Zhang, Q. (2009) Molecular dynamics simulations on the role of protonation states in HIV-1 protease-indinavir complex, *Acta Chim. Sin 67*, 2791-2797.

[48] Tong, Y., Mei, Y., Zhang, J. Z., Duan, L. L., and Zhang, Q.-g. (2009) Quantum calculation of protein solvation and protein–ligand binding free energy for HIV-1 protease/water complex, *Journal of Theoretical and Computational Chemistry 8*, 1265-1279.

[49] Makatini, M. M. (2011) Design, synthesis and screening of novel PCU-peptide/peptoid derived HIV protease inhibitors, *University of KwaZulu-Natal, Westville*.

[50] Smith, R., Brereton, I. M., Chai, R. Y., and Kent, S. B. (1996) Ionization states of the catalytic residues in HIV-1 protease, *Nature Structural & Molecular Biology 3*, 946-950.

[51] Honarparvar, B., Pawar, S. A., Alves, C. N., Lameira, J., Maguire, G. E., Silva, J. R. A., Govender, T., and Kruger, H. G. (2015) Pentacycloundecane lactam vs lactone norstatine type protease HIV inhibitors: binding energy calculations and DFT study, *Journal of Biomedical Science 22*, 1-15.

[52] Maseko, S. B., Natarajan, S., Sharma, V., Bhattacharyya, N., Govender, T., Sayed, Y., Maguire, G. E., Lin, J., and Kruger, H. G. (2016) Purification and characterization of naturally occurring HIV-1 (south african subtype C) protease mutants from inclusion bodies, *Protein Expression and Purification 122*, 90-96.

[53] Dahiya, S., Gahlaut, A., and Kulharia, M. (2013) Comparative evaluation of commercially available homology modelling tools: a structural bioinformatics perspective, *Drug Invention Today 5*, 207-211.

[54] Weber, I. T. (1990) Evaluation of homology modeling of HIV protease, *Proteins: Structure, Function, and Bioinformatics 7*, 172-184.

[55] Maghsoudi, A. H., Khodagholi, F., Hadi-Alijanvand, H., Esfandiarei, M., Sabbaghian, M., Zakeri, Z., Shaerzadeh, F., Abtahi, S., and Maghsoudi, N. (2011) Homology modeling, docking, molecular dynamics simulation, and structural analyses of coxsakievirus B3 2A protease: an enzyme involved in the pathogenesis of inflammatory myocarditis, *International Journal of Biological Macromolecules 49*, 487-492.

[56] Dhanavade, M. J., Jalkute, C. B., Barage, S. H., and Sonawane, K. D. (2013) Homology modeling, molecular docking and MD simulation studies to investigate role of cysteine protease from xanthomonas campestris in degradation of Aβ peptide, *Computers in Biology and Medicine 43*, 2063-2070.

[57] Kohn, W., Becke, A. D., and Parr, R. G. (1996) Density functional theory of electronic structure, *The Journal of Physical Chemistry 100*, 12974-12980.

[58] Neumann, R., Nobes, R. H., and Handy, N. C. (1996) Exchange functionals and potentials, *Molecular Physics 87*, 1-36.

[59] Becke, A. D. (1993) Density-functional thermochemistry. III. The role of exact exchange, *The Journal of Chemical Physics 98*, 5648-5652.

[60] Lee, C., Yang, W., and Parr, R. G. (1988) Development of the colle-salvetti correlation-energy formula into a functional of the electron density, *Physical Review B 37*, 785-789.

[61] Hariharan, P. C., and Pople, J. A. (1973) The influence of polarization functions on molecular orbital hydrogenation energies, *Theoretica Chimica Acta 28*, 213-222.

[62] Rassolov, V. A., Pople, J. A., Ratner, M. A., and Windus, T. L. (1998) 6-31G* basis set for atoms K through Zn, *The Journal of Chemical Physics 109*, 1223-1229.

[63] Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The AMBER biomolecular simulation programs, *Journal of Computational Chemistry 26*, 1668-1688.

[64] Studio, D. (2013) 4.0 Tutorials, receptor-ligand interaction. *Accelrys Inc, San Diego, CA, USA*.

[65] Laskowski, R. A., and Swindells, M. B. (2011) LigPlot+: multiple ligand–protein interaction diagrams for drug discovery, *Journal of Chemical Information and Modeling 51*, 2778-2786.

[66] Kuiper, B. D., Keusch, B. J., Dewdney, T. G., Chordia, P., Ross, K., Brunzelle, J. S., Kovari, I. A., MacArthur, R., Salimnia, H., and Kovari, L. C. (2015) The L33F darunavir resistance mutation acts as a molecular anchor reducing the flexibility of the HIV-1 protease 30s and 80s loops, *Biochemistry and Biophysics Reports 2*, 160-165.

[67] Liu, Z., Yedidi, R. S., Wang, Y., Dewdney, T. G., Reiter, S. J., Brunzelle, J. S., Kovari, I. A., and Kovari, L. C. (2013) Crystallographic study of multi-drug resistant HIV-1 protease lopinavir complex: mechanism of drug recognition and resistance, *Biochemical and Biophysical Research Communications 437*, 199-204.

[68] Kuhnert, M., Steuber, H., and Diederich, W. E. (2014) Structural basis for HTLV-1 protease inhibition by the HIV-1 protease inhibitor indinavir, *Journal of Medicinal Chemistry 57*, 6266-6272.

[69] Tie, Y., Wang, Y. F., Boross, P. I., Chiu, T. Y., Ghosh, A. K., Tozser, J., Louis, J. M., Harrison, R. W., and Weber, I. T. (2012) Critical differences in HIV-1 and HIV-2 protease specificity for clinical inhibitors, *Protein Science 21*, 339-350.

[70] Kožíšek, M., Bray, J., Řezáčová, P., Šašková, K., Brynda, J., Pokorná, J., Mammano, F., Rulíšek, L., and Konvalinka, J. (2007) Molecular analysis of the HIV-1 resistance development: enzymatic activities, crystal structures, and thermodynamics of nelfinavir-resistant HIV protease mutants, *Journal of Molecular Biology 374*, 1005-1016.

[71] Yedidi, R. S., Garimella, H., Aoki, M., Aoki-Ogata, H., Desai, D. V., Chang, S. B., Davis, D. A., Fyvie, W. S., Kaufman, J. D., and Smith, D. W. (2014) A conserved hydrogen-bonding network of P2 bis-tetrahydrofuran-containing HIV-1 protease inhibitors (PIs) with a protease active-site amino acid backbone aids in their activity against PI-resistant HIV, *Antimicrobial Agents and Chemotherapy 58*, 3679-3688.

[72] King, N. M., Prabu-Jeyabalan, M., Bandaranayake, R. M., Nalam, M. N., Nalivaika, E. A., Özen, A. e. l., Haliloğlu, T. r., Yılmaz, N. e. K., and Schiffer, C. A. (2012) Extreme entropy–enthalpy compensation in a drug-resistant variant of HIV-1 protease, *ACS Chemical Biology 7*, 1536-1546.

[73] DeLano, W. L. (2002) The PyMOL molecular graphics system.

[74] Cole, J. C., Murray, C. W., Nissink, J. W. M., Taylor, R. D., and Taylor, R. (2005) Comparing protein–ligand docking programs is difficult, *Proteins: Structure, Function, and Bioinformatics 60*, 325-332.

[75] Gohlke, H., Hendlich, M., and Klebe, G. (2000) Knowledge-based scoring function to predict protein-ligand interactions, *Journal of Molecular Biology 295*, 337-356.

[76] Kontoyianni, M., McClellan, L. M., and Sokol, G. S. (2004) Evaluation of docking performance: comparative data on docking algorithms, *Journal of Medicinal Chemistry 47*, 558-565.

[77] Li, H., Robertson, A. D., and Jensen, J. H. (2005) Very fast empirical prediction and rationalization of protein pKa values, *Proteins: Structure, Function, and Bioinformatics 61*, 704-721.

[78] Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., and Baker, N. A. (2004) PDB2PQR: an automated pipeline for the setup of poisson–boltzmann electrostatics calculations, *Nucleic Acids Research 32*, W665-W667.

[79] Dennington, R., Keith, T., and Millam, J. (2009) Semichem Inc, *Shawnee Mission KS, GaussView, Version 5*.

[80] Morokuma, K. (2002) New challenges in quantum chemistry: quests for accurate calculations for large molecular systems, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 360*, 1149-1164.

[81] Vreven, T., Morokuma, K., Farkas, Ö., Schlegel, H. B., and Frisch, M. J. (2003) Geometry optimization with QM/MM, ONIOM, and other combined methods. I. Microiterations and constraints, *Journal of Computational Chemistry 24*, 760-769.

[82] Dapprich, S., Komáromi, I., Byun, K. S., Morokuma, K., and Frisch, M. J. (1999) A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives, *Journal of Molecular Structure: THEOCHEM 461*, 1-21.

[83] Parr, R. G., and Yang, W. (1984) Density functional approach to the frontier-electron theory of chemical reactivity, *Journal of the American Chemical Society 106*, 4049-4050.

[84] Johnson, B. G., Gill, P. M., and Pople, J. A. (1993) The performance of a family of density functional methods, *The Journal of Chemical Physics 98*, 5612-5626.

[85] Kapp, J., Remko, M., and Schleyer, P. v. R. (1996) H_2XO and $(CH_3)_2XO$ Compounds (X= C, Si, Ge, Sn, Pb): Double bonds vs carbene-like structures can the metal compounds exist at all?, *Journal of the American Chemical Society 118*, 5745-5751.

[86] Remko, M., Walsh, O. A., and Richards, W. G. (2001) Theoretical study of molecular structure, tautomerism, and geometrical isomerism of moxonidine: two-layered ONIOM calculations, *The Journal of Physical Chemistry A 105*, 6926-6931.

[87] Frisch, M., Trucks, G., Schlegel, H., Scuseria, G., Robb, M., Cheeseman, J., Scalmani, G., Barone, V., Mennucci, B., and Petersson, G. (2009) 09, Revision D. 01, Gaussian, *Inc., Wallingford, CT*.

[88] Lundberg, M., Sasakura, Y., Zheng, G., and Morokuma, K. (2010) Case studies of ONIOM (DFT:DFTB) and ONIOM (DFT:DFTB:MM) for enzymes and enzyme mimics, *Journal of Chemical Theory and Computation 6*, 1413-1427.

[89] Chung, L. W., Sameera, W., Ramozzi, R., Page, A. J., Hatanaka, M., Petrova, G. P., Harris, T. V., Li, X., Ke, Z., and Liu, F. (2015) The ONIOM method and its applications, *Chemical Reviews 115*, 5678-5796.

[90] Sibusiso, B. M. E., Padayachee; Thavendran, Govendera; Yasien, Sayed; Gert, Kruger; Glenn, E.M Maguire; Johnson, Linb;. (2016) Kinetic and thermodynamic characterization of HIV-protease inhibitors against L38L↑N↑L mutant in the south africa HIV-1 subtype C protease, *The Lancet HIV, Submitted for publication*.

[91] Ryde, U. (2014) A fundamental view of enthalpy–entropy compensation, *MedChemComm 5*, 1324-1336.

CHAPTER FIVE

CONCLUSION

A geometry optimization was performed on the nine FDA approved HIV PR inhibitors using density functional theory [B3LYP/6-31G(d)] for the QM level of theory and Amber force field for the MM level. A comparative analysis in ranking the potency of the FDA approved drugs in terms of binding affinities against subtype B, C-SA HIV PR was demonstrated in chapter three of this thesis. It was observed that the calculated binding free interaction energies for the FDA HIV PR drugs complexed with the subtype B PR followed a reasonably trend with respect to the experimental binding free energy results. Some discrepancies occur in the calculated binding free energies of the FDA HIV PR inhibitors complexed with the C-SA HIV PR with respect to the experimental binding free energies. A rational explanation for the observed discrepancies was provided, suggesting that our computational model was over simplified. We finally concluded that the computational model can be improved by the addition of water to the active pocket of the catalytic residues and to treat the water at a minimum semi-empirical level of theory (PM6).

A comparative analysis in ranking the potency of the FDA approved drugs in terms of binding affinities against the L38L↑N↑L mutant was also demonstrated in chapter four of this thesis. It was discovered that the calculated binding free interaction energies for the FDA HIV PR drugs complexed with the L38L↑N↑L PR followed a satisfactory trend with respect to the experimental binding free energy results. A rational explanation for the observed discrepancies was provided, again suggesting that our computational model was over simplified. The outcome of this study revealed that ONIOM method can be used as a reliable computational approach to rationalize lead compounds against specific targets. The investigation carried out in this work can help to improve the efficiency of future HIV-1 PR antiviral drugs.

Altogether, the calculated binding free interaction energies are in reasonable agreement for both subtype B and L38L↑N↑L HIV PR with the reported experimental data, while some discrepancies occur mainly with the subtype C-SA HIV PR. An important conclusion can be drawn that the ONIOM model can potentially be used to compute the binding free energies of available inhibitors and new PR inhibitors against the corresponding HIV enzymes.

The catalytic Asp25 and Asp25' residues with the selected HIV inhibitors were considered at a high level QM; DFT theory. In future, water will be added to the active site and treated at a semi-empirical middle layer, also the catalytic triads Asp25-Thr26-Gly27 will be considered at a high level since they also play an important catalytic role in the enzyme. The results will then be compare to our previous study; also, the basis set used for the high level will also be modified to improve the efficiency of the computational model. This approach will assist in modifying/improving our model, to obtain better insight of the interaction of the inhibitors with the other catalytic residues.

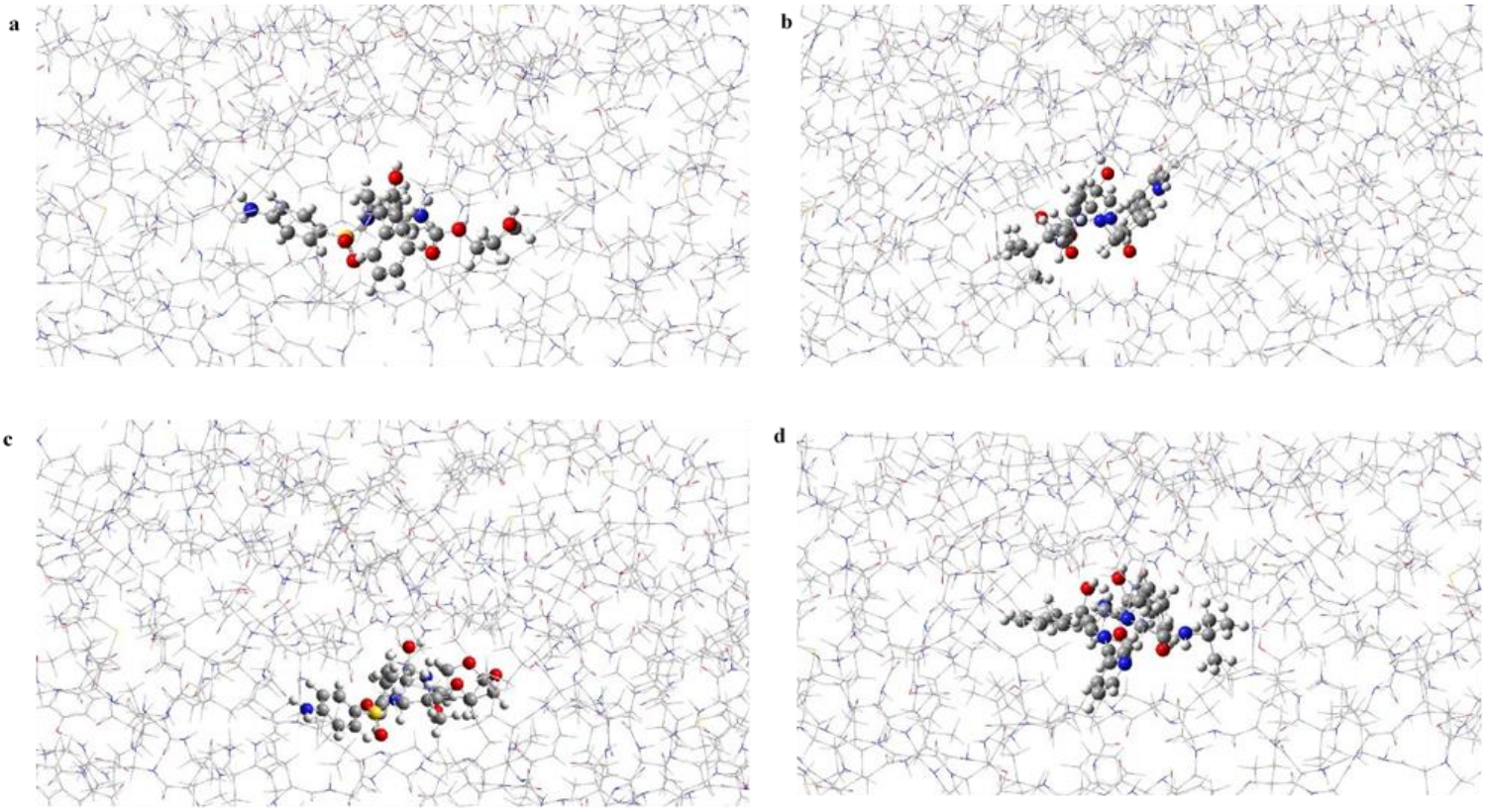
Hence, information gathered in this study will facilitate the design of future potent HIV antiretroviral drugs by optimizing the interactions with the catalytic residues and or other residues of the HIV enzyme.

APPENDIX A

Supplementary Material for Chapter Three

Investigation of the binding free interaction energies of FDA approved HIV-1 PR Inhibitors against Subtype B and C-SA HIV PR: ONIOM Approach

ZKS, TG, GEM, SBM, JL, HG* and BH*



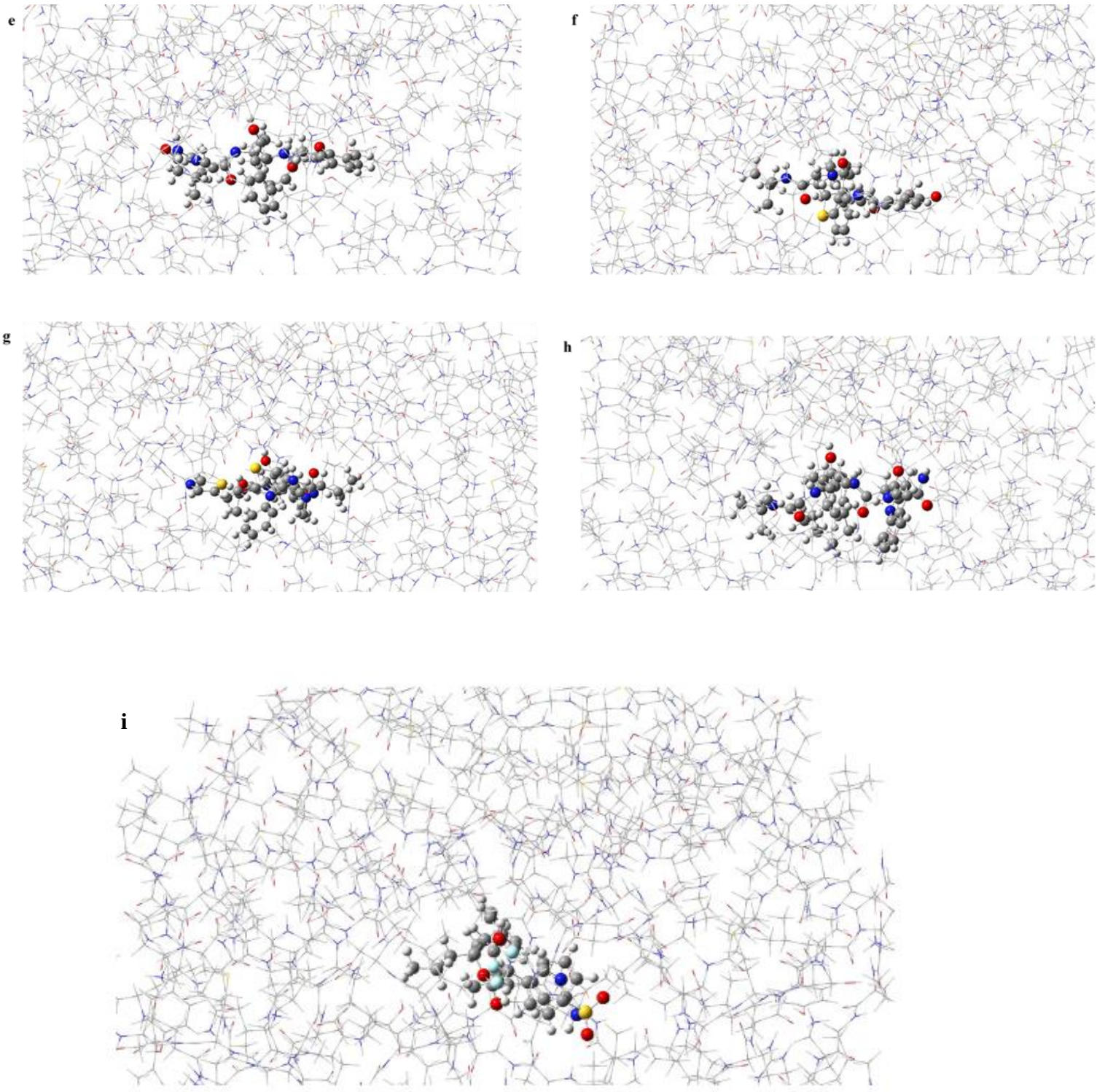
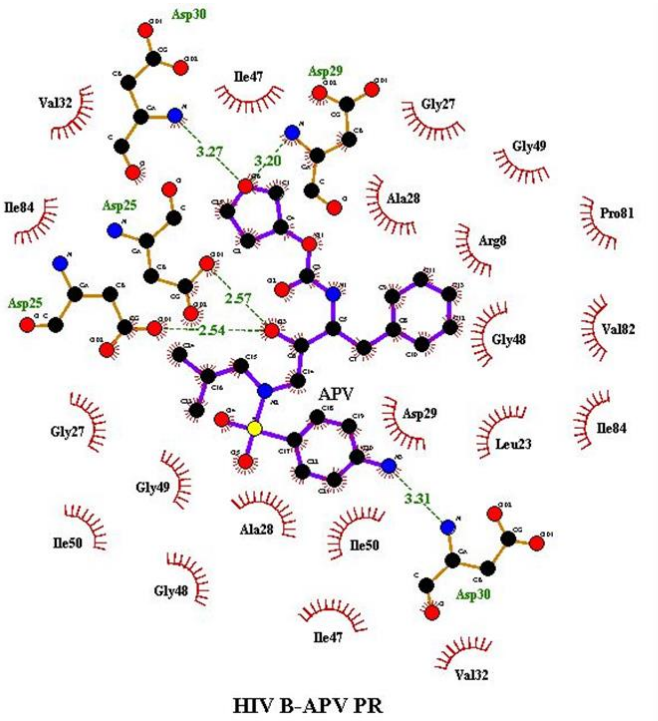
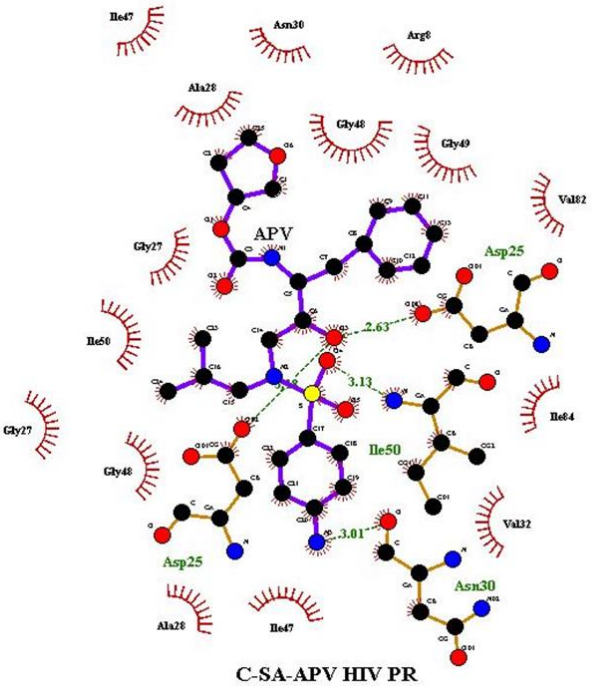


Figure S7. 3D Schematic representation of C-SA HIV PR complexed with **a:** APV, **b:** ATV, **c:** DRV, **d:** IDV, **e:** LPV, **f:** NFV, **g:** RTV, **h:** SQV, and **i:** TPV PIs. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.)

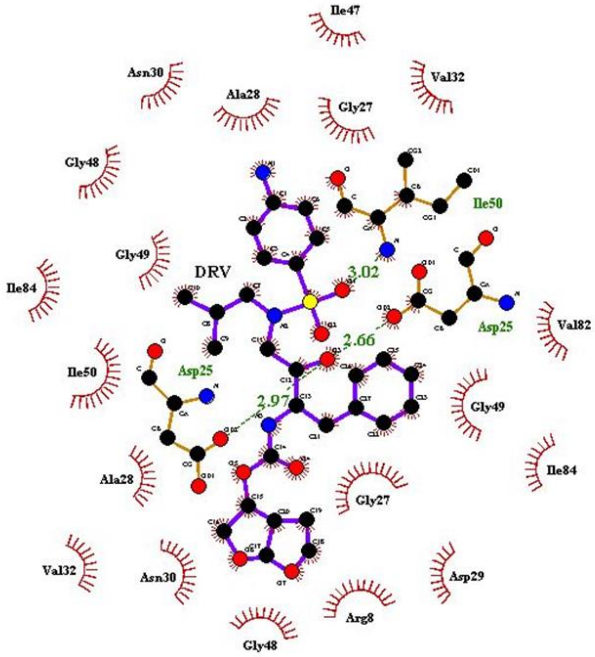
Detailed plots showing the electrostatic interactions and hydrogen bonding of subtype B and C-SA HIV PR respectively with the selected FDA approved HIV PIs.



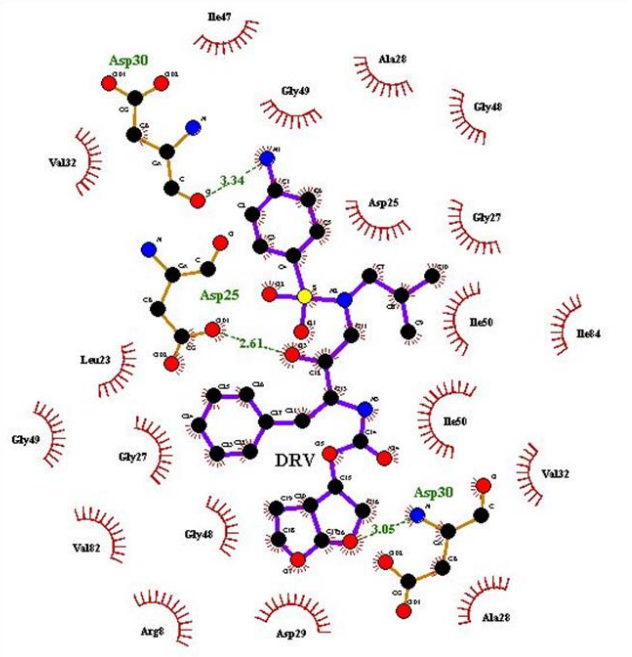
HIV B-APV PR



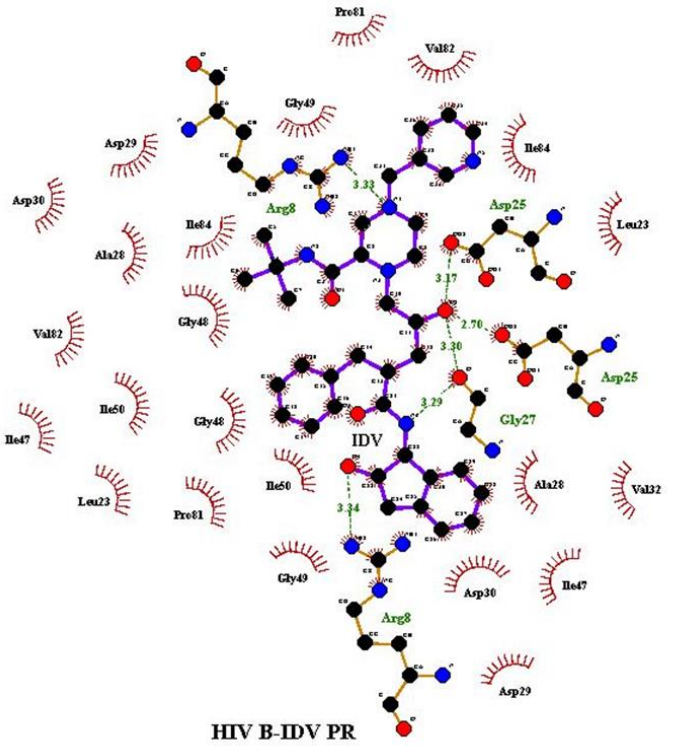
C-SA-APV HIV PR



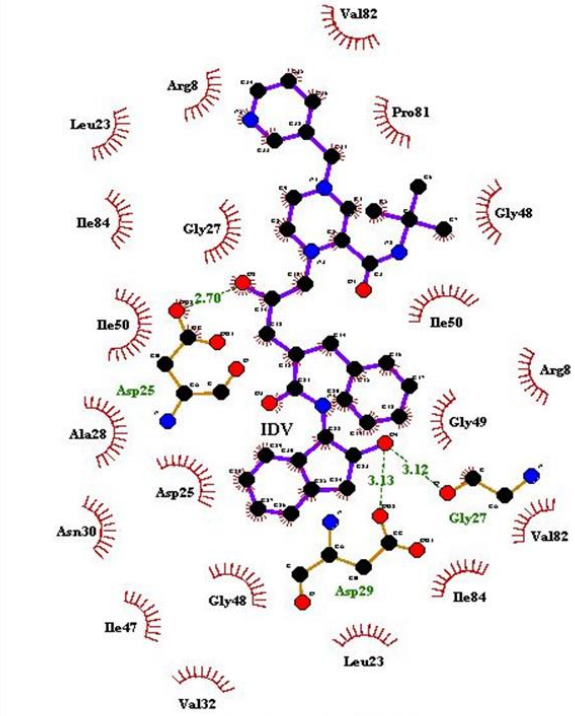
C-SA-DRV HIV PR



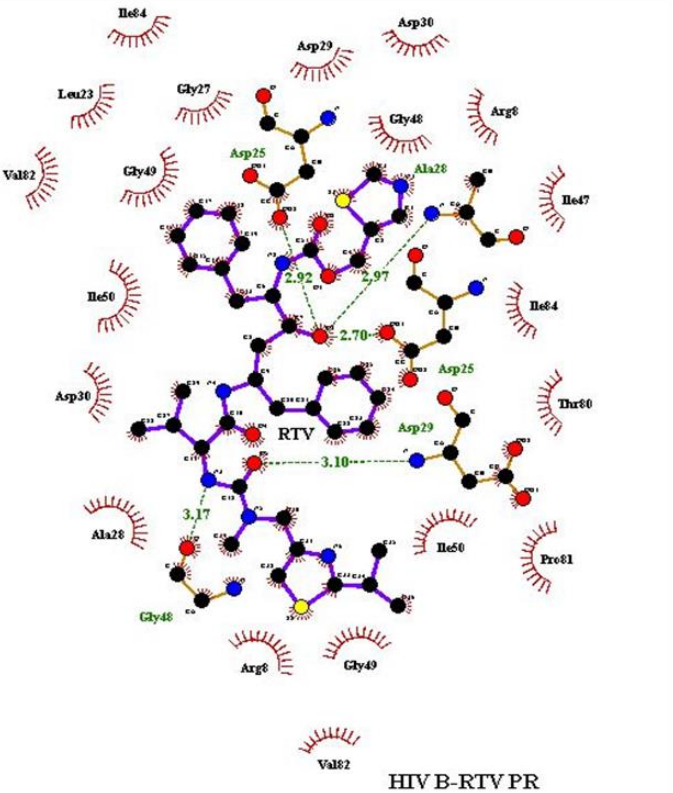
HIV B-DRV PR



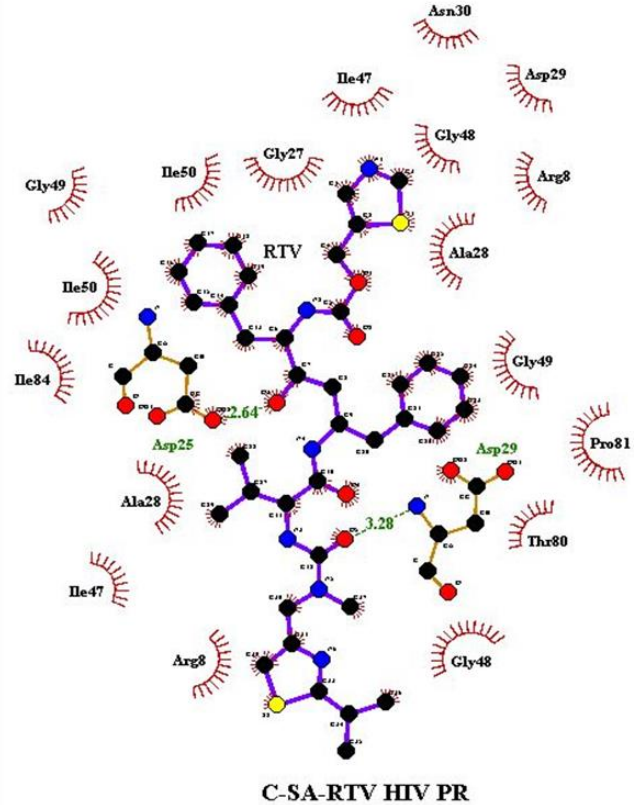
HIV B-IDV PR



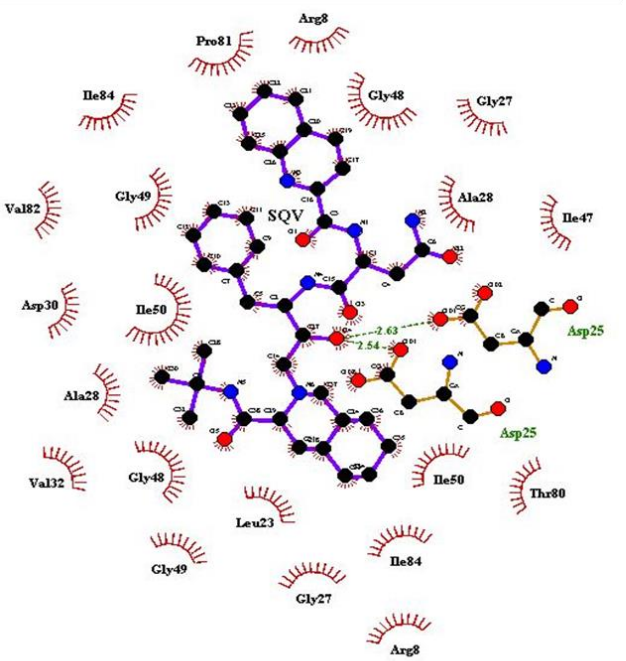
C-SA-IDV HIV PR



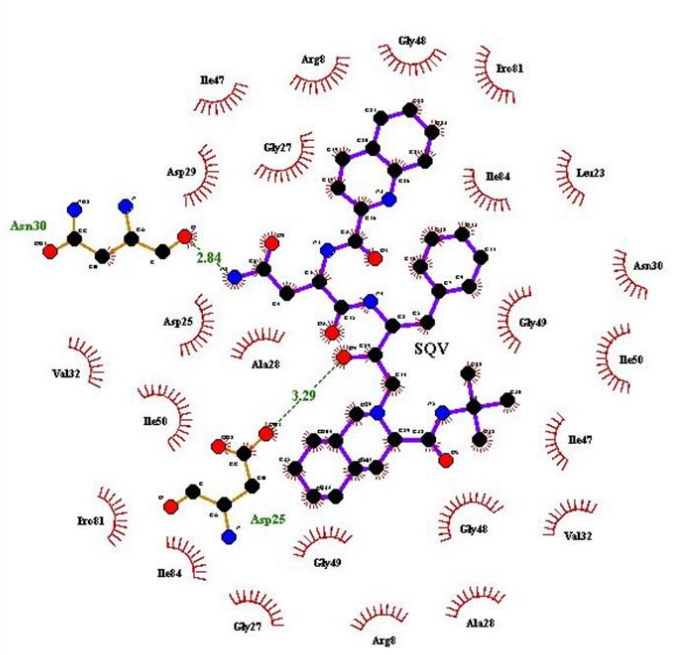
HIV B-RTV PR



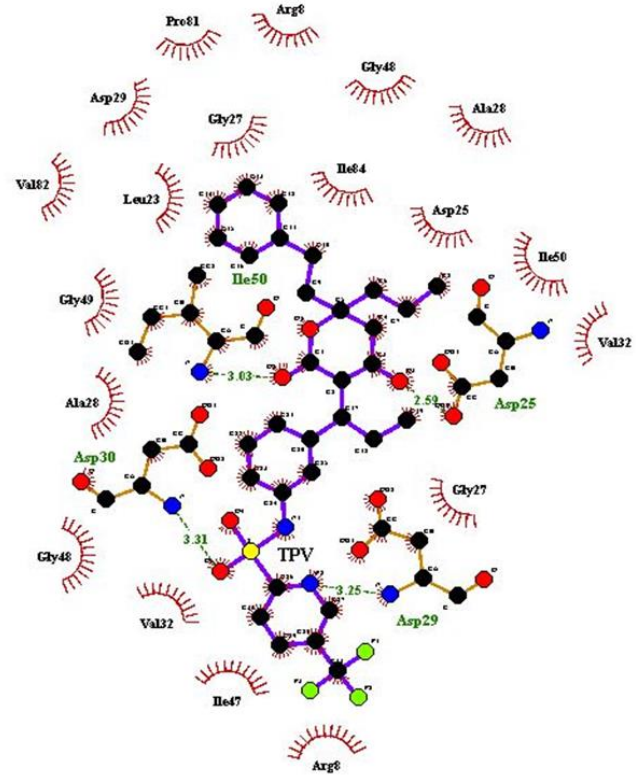
C-SA-RTV HIV PR



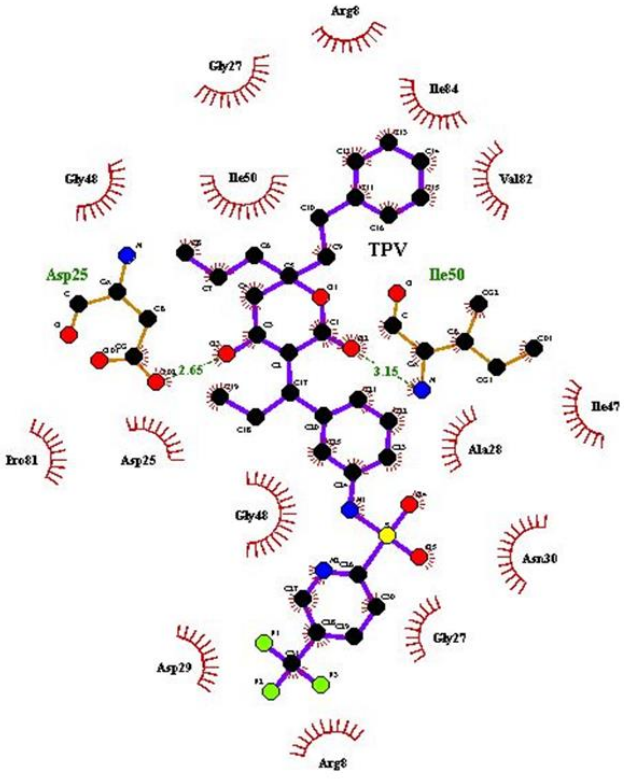
HIV B-SQV PR



C-SA-SQV HIV PR



HIV B-TPV PR



C-SA-TPV HIV PR

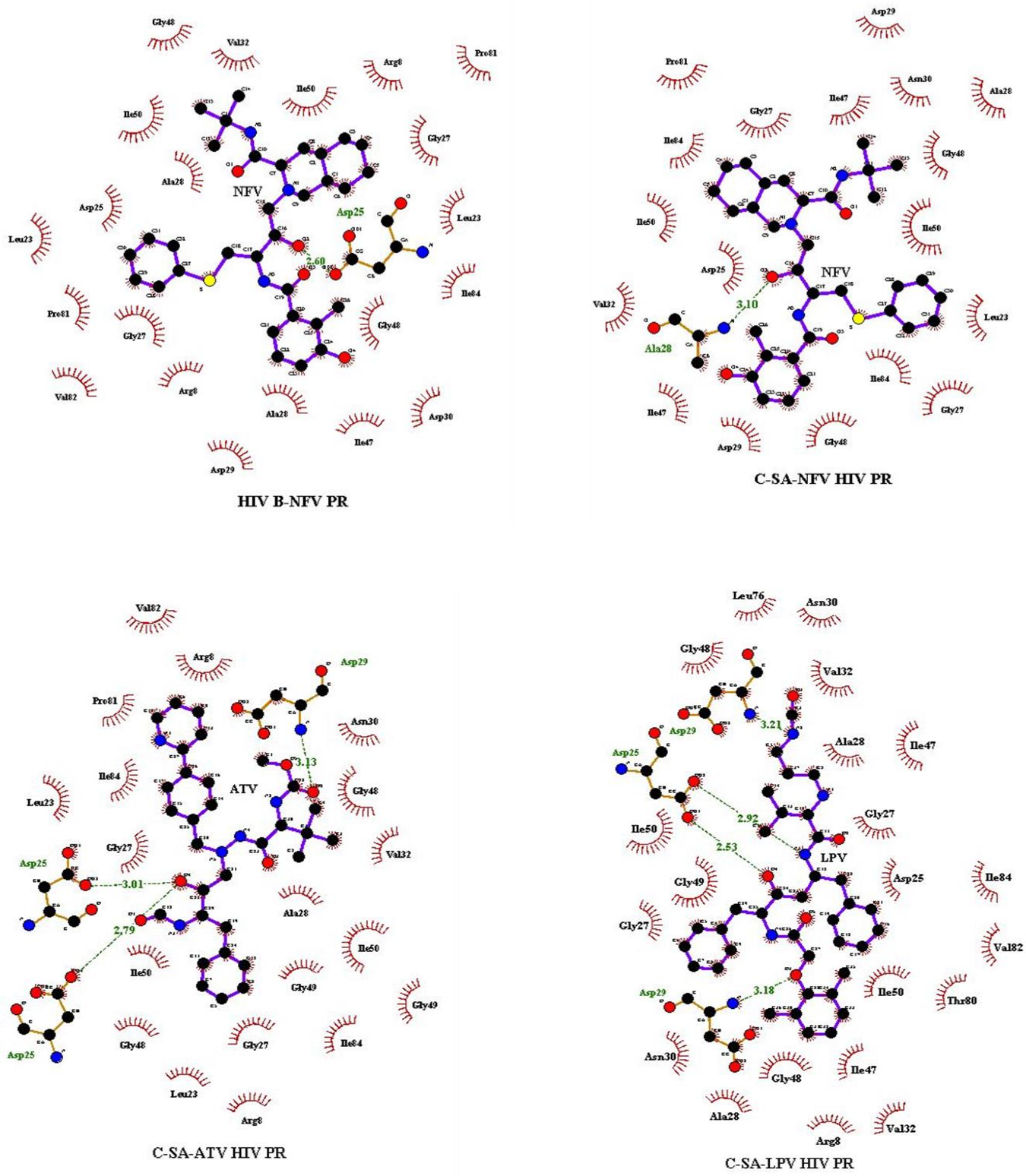


Figure S8. Electrostatic and hydrogen bon interactions plots of subtype B and C-SA HIV PR respectively with the various FDA approved HIV PIs. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.)

The Figure below illustrates before and after geometry optimization was performed on all the 18 enzyme—inhibitor complexes in determining the hydrogen bond (HB) distance between the catalytic Asp 25/25' and hydroxyl group of the selected FDA approved inhibitors.

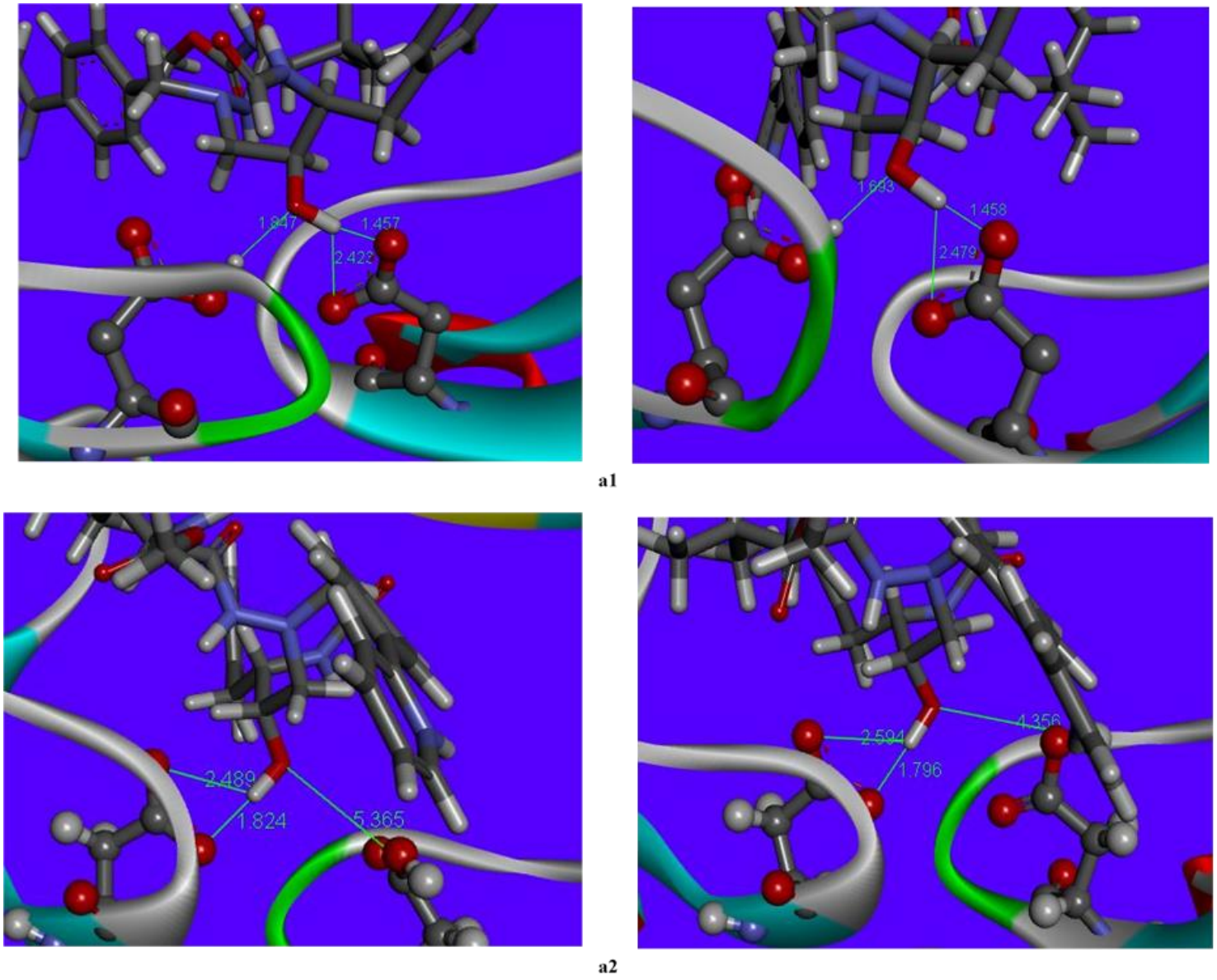


Figure S9. Hydrogen bond distances between the hydroxyl groups of the selected FDA approved ATV drug with the catalytic ASP25 and ASP25' residues of **a1:** subtype B HIV PR, **a2:** subtype C-SA PR before and after optimization. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.)

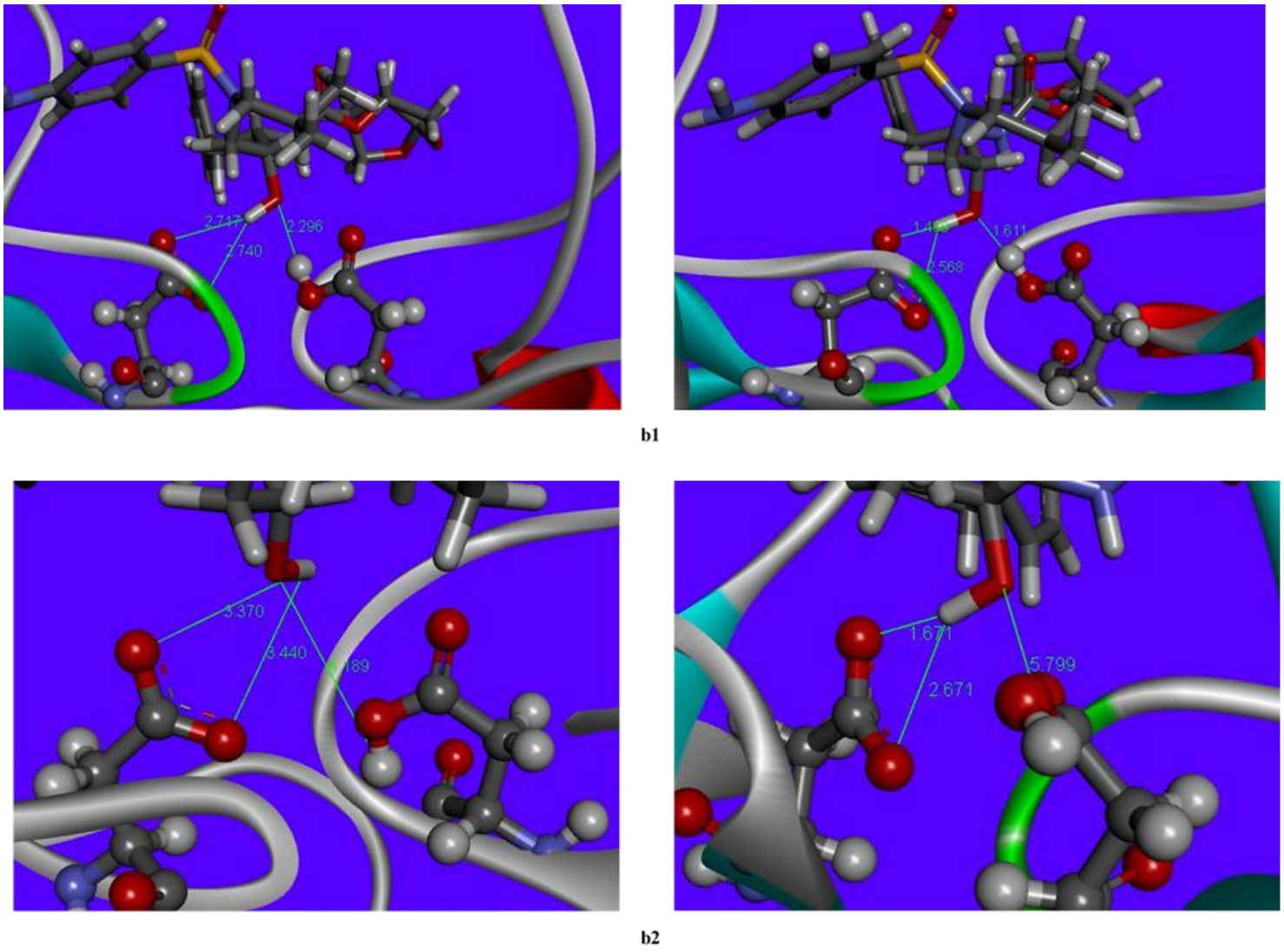


Figure S3. Hydrogen bond distances between the hydroxyl groups of the selected FDA approved DRV drug with the catalytic ASP25 and ASP25' residues of **b1:** subtype B HIV PR, **b2:** subtype C-SA PR before and after optimization. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.)

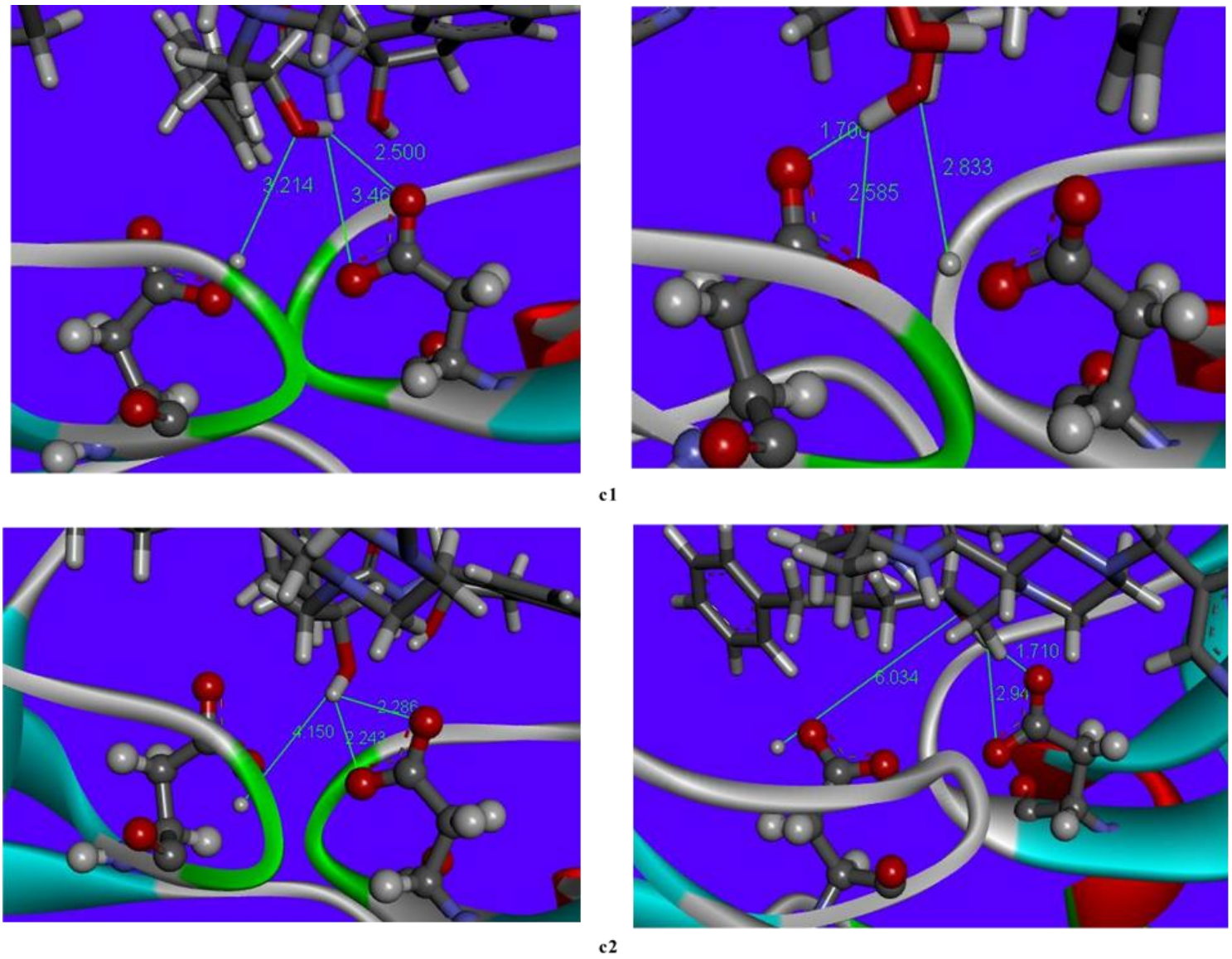
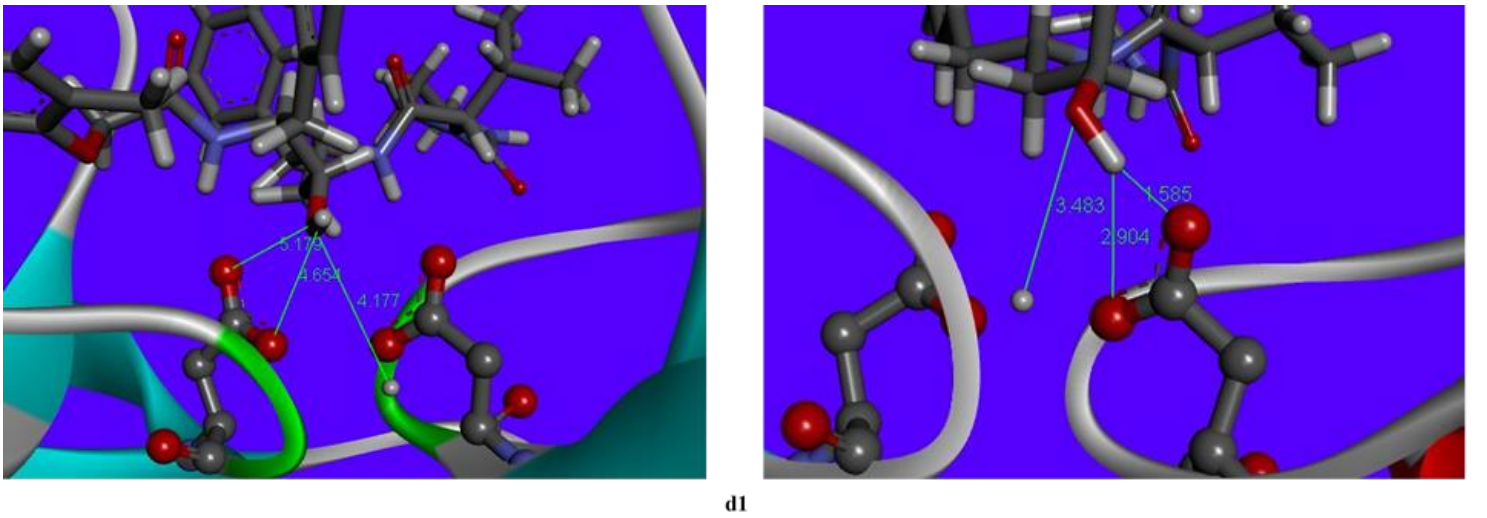
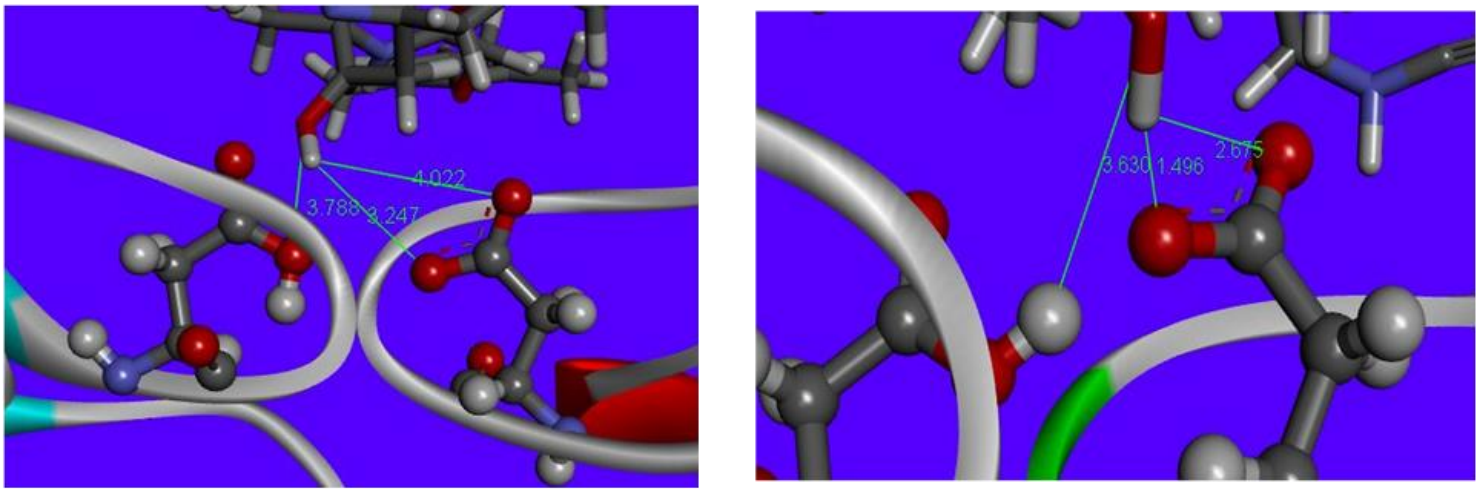


Figure S3. Hydrogen bond distances between the hydroxyl groups of the selected FDA approved IDV drug with the catalytic ASP25 and ASP25' residues of **c1:** subtype B HIV PR, **c2:** subtype C-SA PR before and after optimization. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.)



d1



d2

Figure S3. Hydrogen bond distances between the hydroxyl groups of the selected FDA approved LPV drug with the catalytic ASP25 and ASP25' residues of **d1:** subtype B HIV PR, **d2:** subtype C-SA PR before and after optimization. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.)

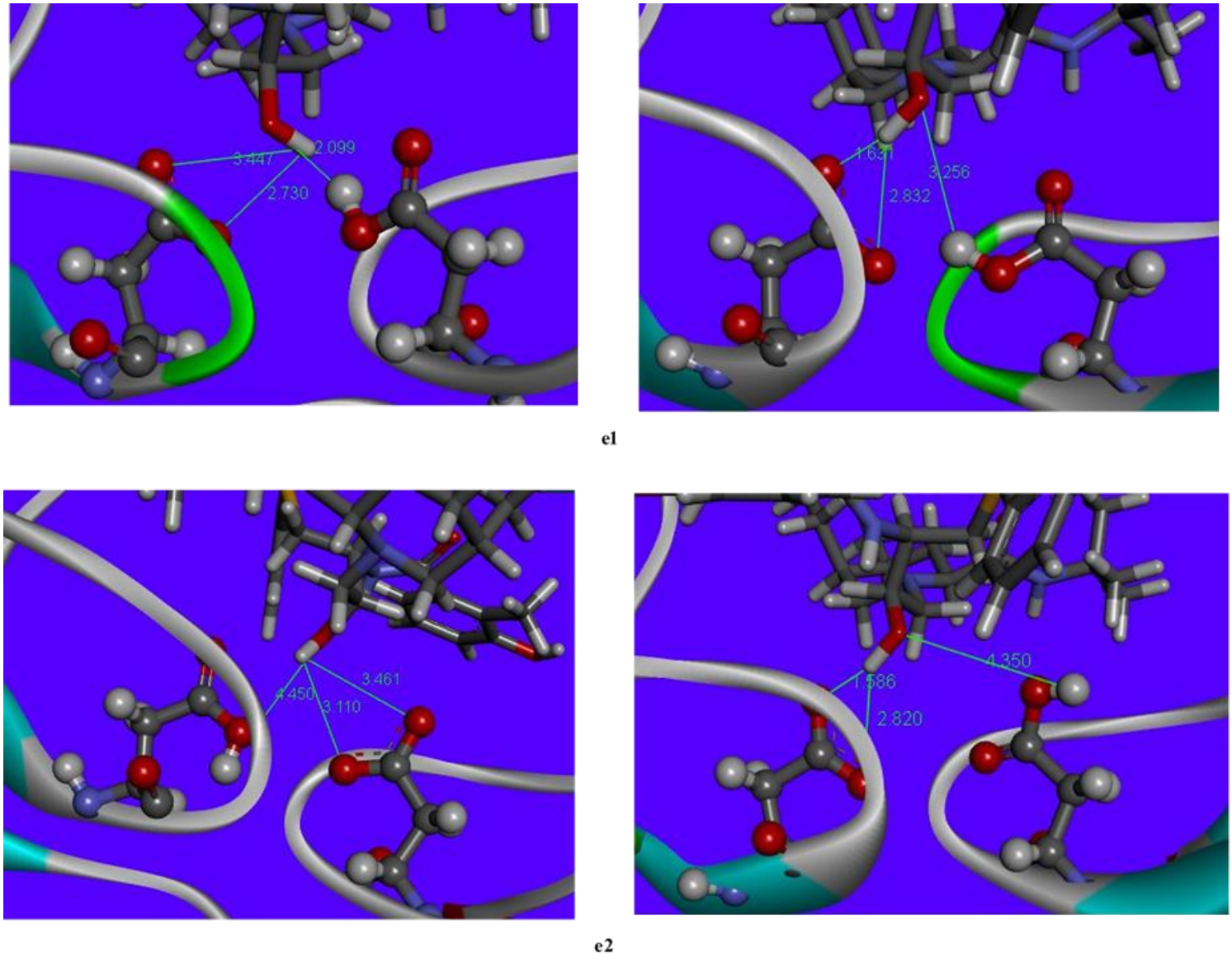
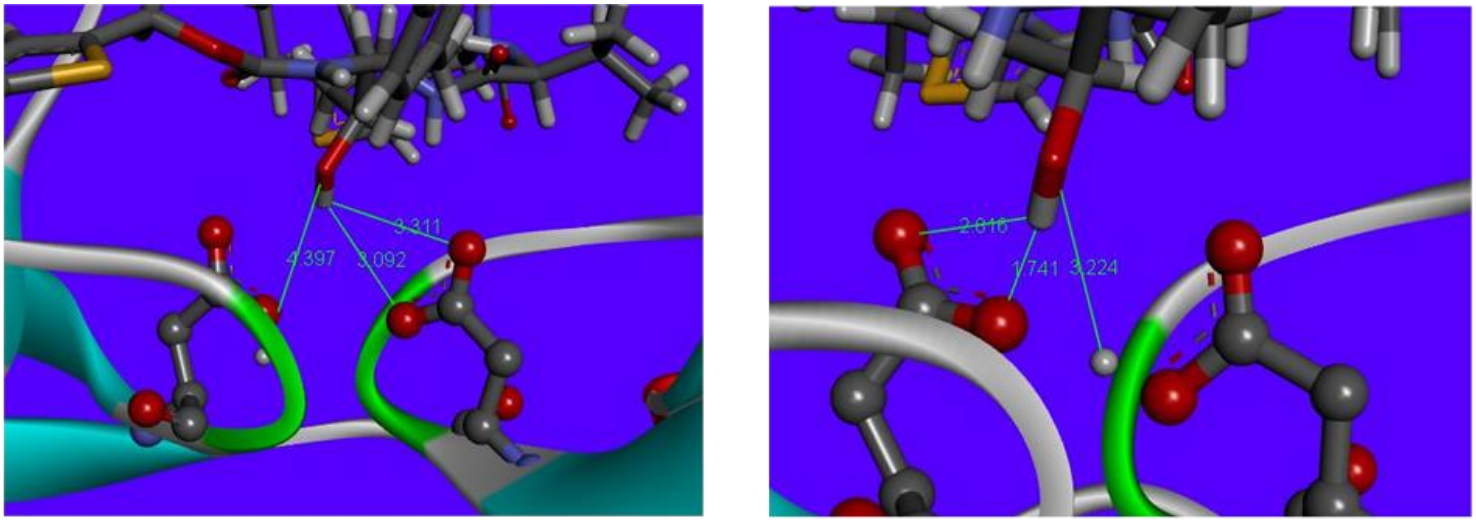
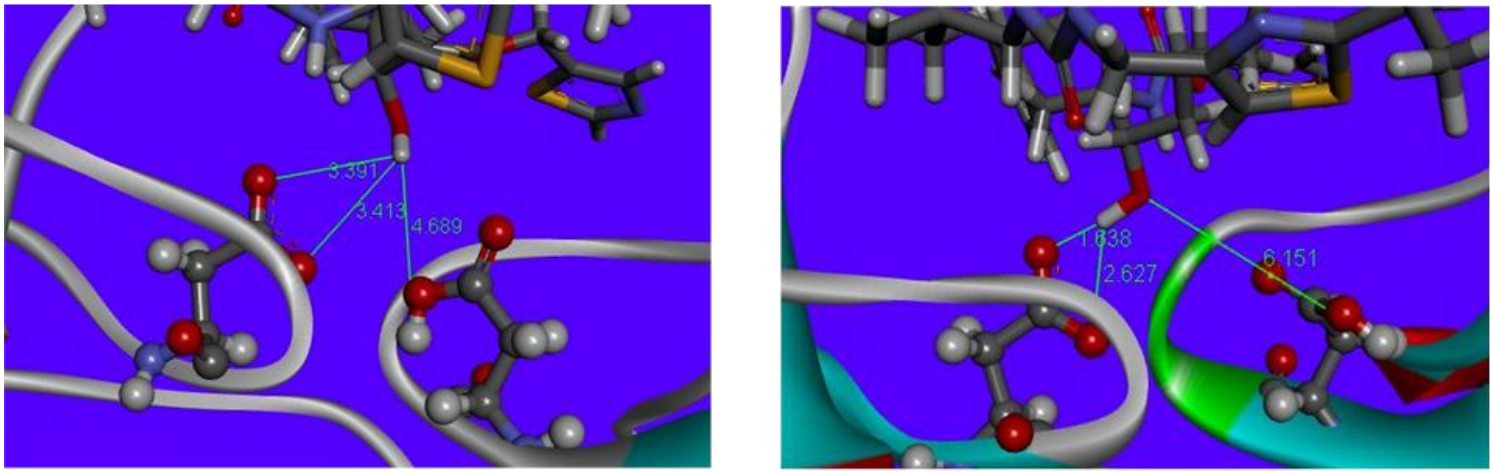


Figure S3. Hydrogen bond distances between the hydroxyl groups of the selected FDA approved NFV drug with the catalytic ASP25 and ASP25' residues of **e1:** subtype B HIV PR, **e2:** subtype C-SA PR before and after optimization. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.)

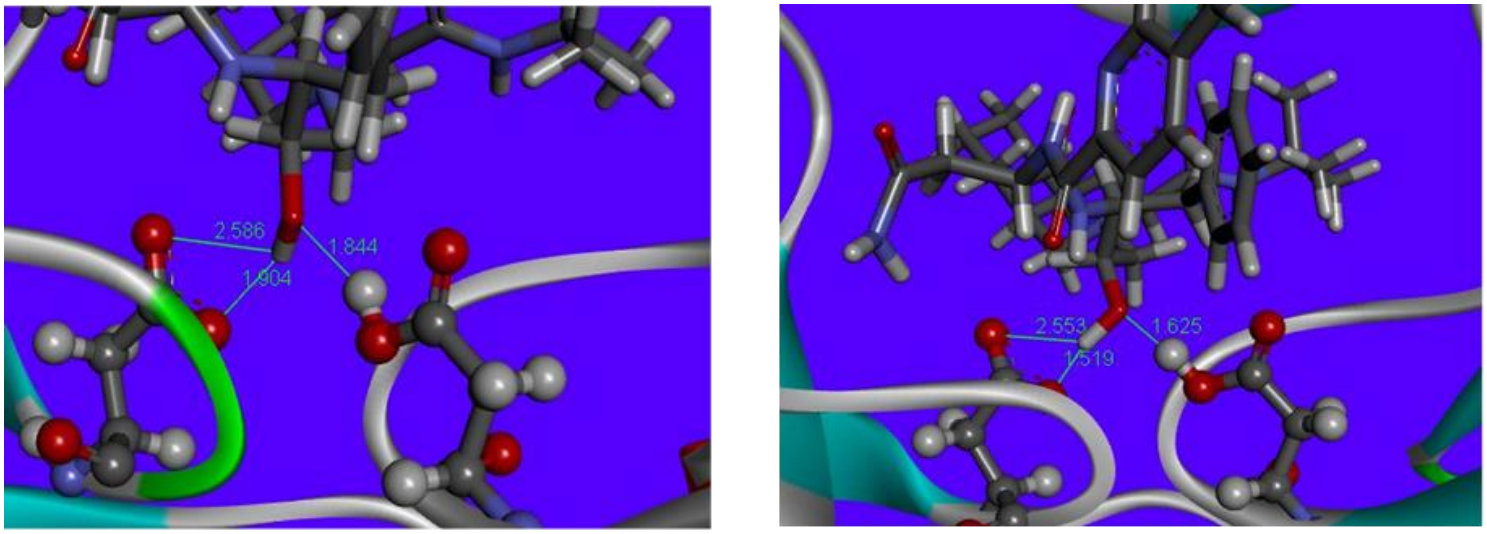


f1

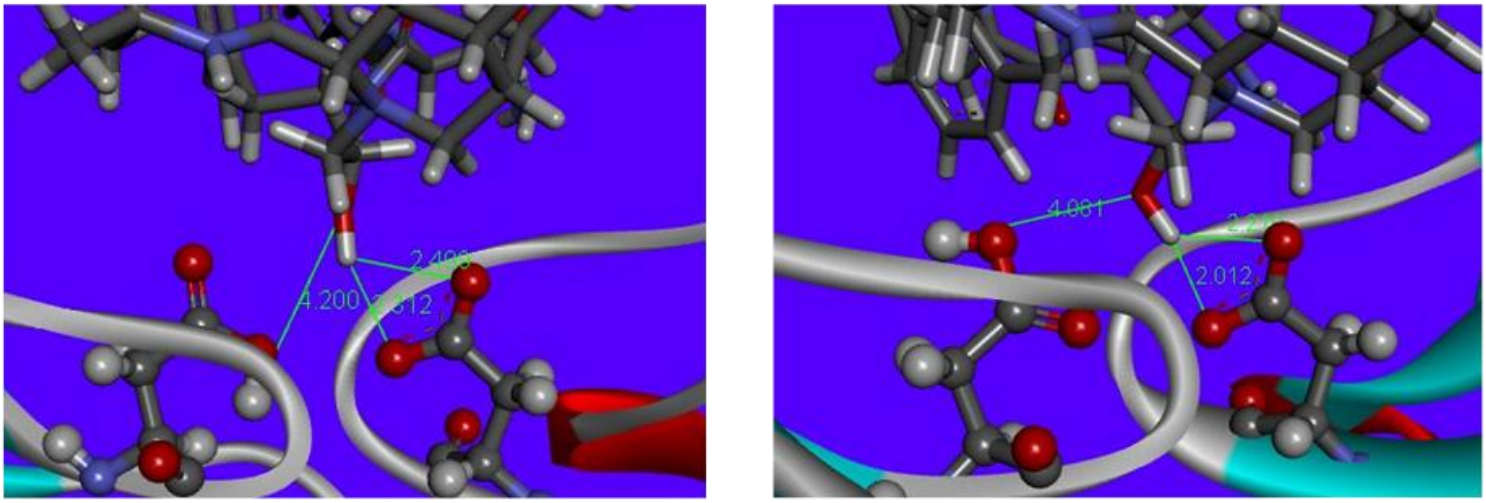


f2

Figure S3. Hydrogen bond distances between the hydroxyl groups of the selected FDA approved RTV drug with the catalytic ASP25 and ASP25' residues of **f1:** subtype B HIV PR, **f2:** subtype C-SA PR before and after optimization. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.)

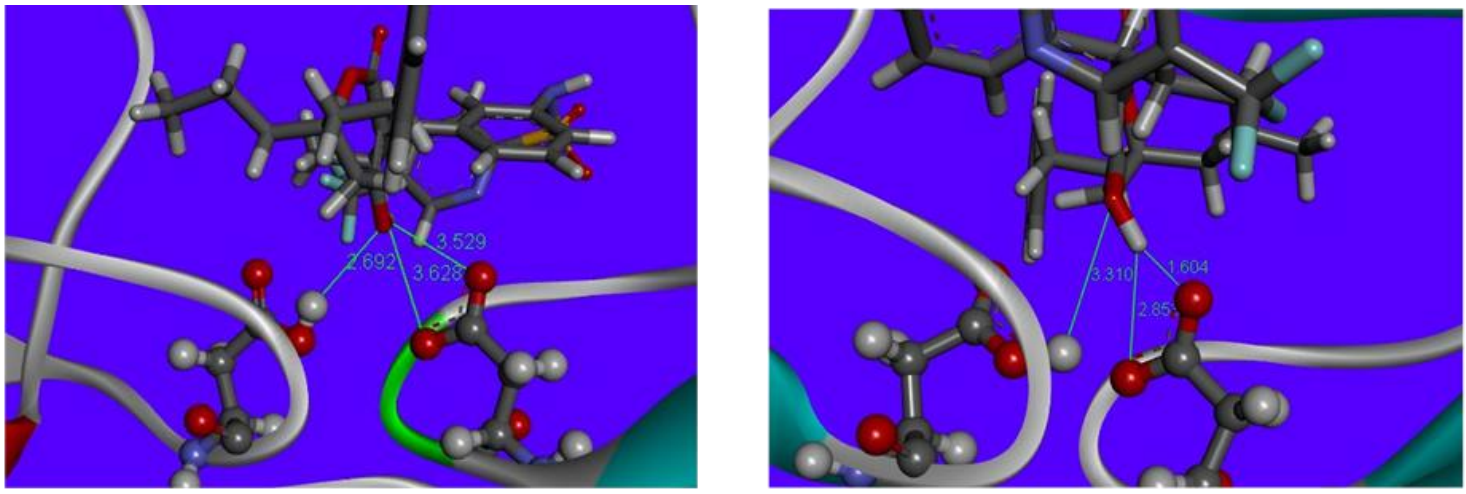


g1

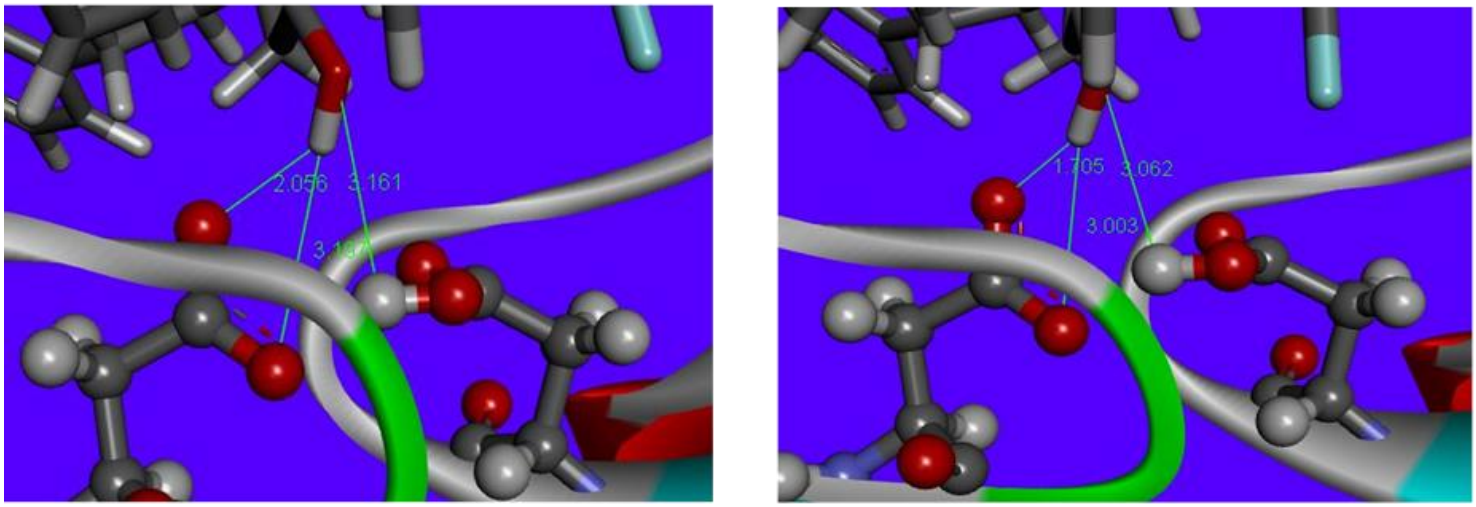


g2

Figure S3. Hydrogen bond distances between the hydroxyl groups of the selected FDA approved SQV drug with the catalytic ASP25 and ASP25' residues of **g1:** subtype B HIV PR, **g2:** subtype C-SA PR before and after optimization. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.)



h1



h2

Figure S3. Hydrogen bond distances between the hydroxyl groups of the selected FDA approved TPV drug with the catalytic ASP25 and ASP25' residues of **h1:** subtype B HIV PR, **h2:** subtype C-SA PR before and after optimization. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.)

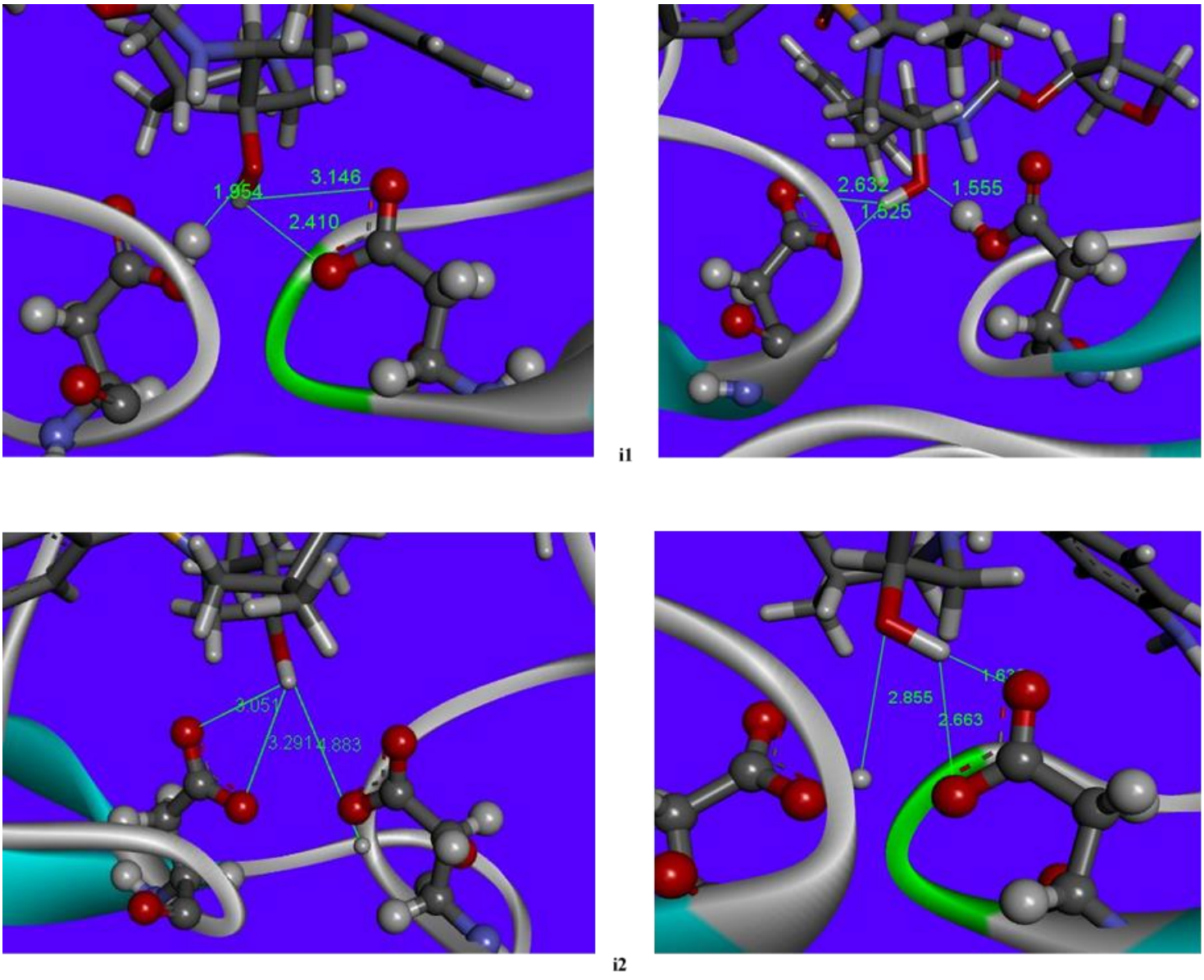


Figure S3. Hydrogen bond distances between the hydroxyl groups of the selected FDA approved ATV drug with the catalytic ASP25 and ASP25' residues of **i1:** subtype B HIV PR, **i2:** subtype C-SA PR before and after optimization. (The ONIOM (Gaussian) input files as well as the optimized output files of all inhibitor—enzyme complexes are also provided in PDB format with the supplementary material.)

Table S3. The Free energies, Entropy and Enthalpies contributions of FDA approved HIV PIs against subtype C-SA PR.

Inhibitors	ΔG^a kcal mol^{-1}	ΔH kcal mol^{-1}	ΔS_{total} cal mol^{-1} K^{-1}	ΔS_{trans} cal mol^{-1} K^{-1}	ΔS_{vib} cal mol^{-1} K^{-1}	ΔS_{rot} cal mol^{-1} K^{-1}
First generation PIs						
RTV	-62.9	-82.6	-67.0	-44.5	15.4	-37.9
APV	-69.0	-102.3	-110.9	-43.5	-31.7	-35.7
SQV	-57.1	-79.3	-74.0	-44.3	7.6	-37.4
NFV	-38.6	-59.4	-69.7	-43.8	10.6	-36.4
IDV	-64.0	-91.3	-91.6	-45.0	-7.2	-39.4
Second generation PIs						
ATV	-66.9	-84.8	-60.3	-43.8	19.9	-36.2
DRV	-62.8	-87.5	-82.4	-43.3	-3.0	-36.1
LPV	-56.5	-81.3	-83.1	-44.1	-2.1	-36.9
TPV	-78.9	-129.3	-169.2	-44.0	-88.0	-37.3

HIV protease inhibitors (HIV PIs) are ranked in terms of their binding free interaction energies (ΔG).

[a]Calculated binding free interaction energies using ONIOM for subtype C-SA PR (taken from **Table 1)**

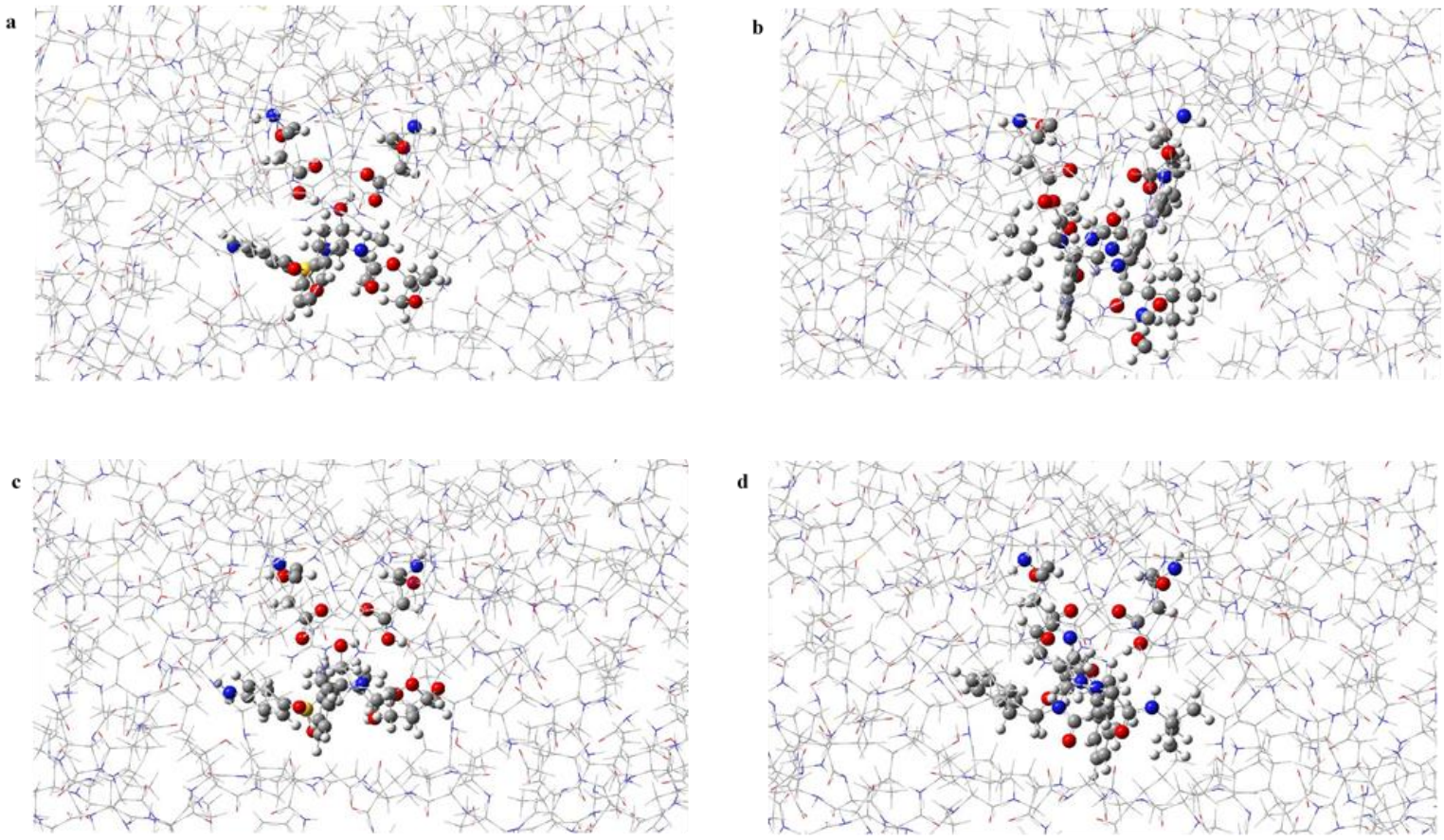
APPENDIX B

Supplementary Material for Chapter Four

Exploring the Binding Free energies of FDA approved PR Inhibitors against novel C-SA HIV-1 PR mutant L38L↑N↑L PR: ONIOM method

ZKS, TG, GEM, SBM, JL, HG* and BH*

The figure below shows the 3D structures of the novel C-SA HIV PR mutants designated L38L↑N↑L PR—inhibitor complexes.



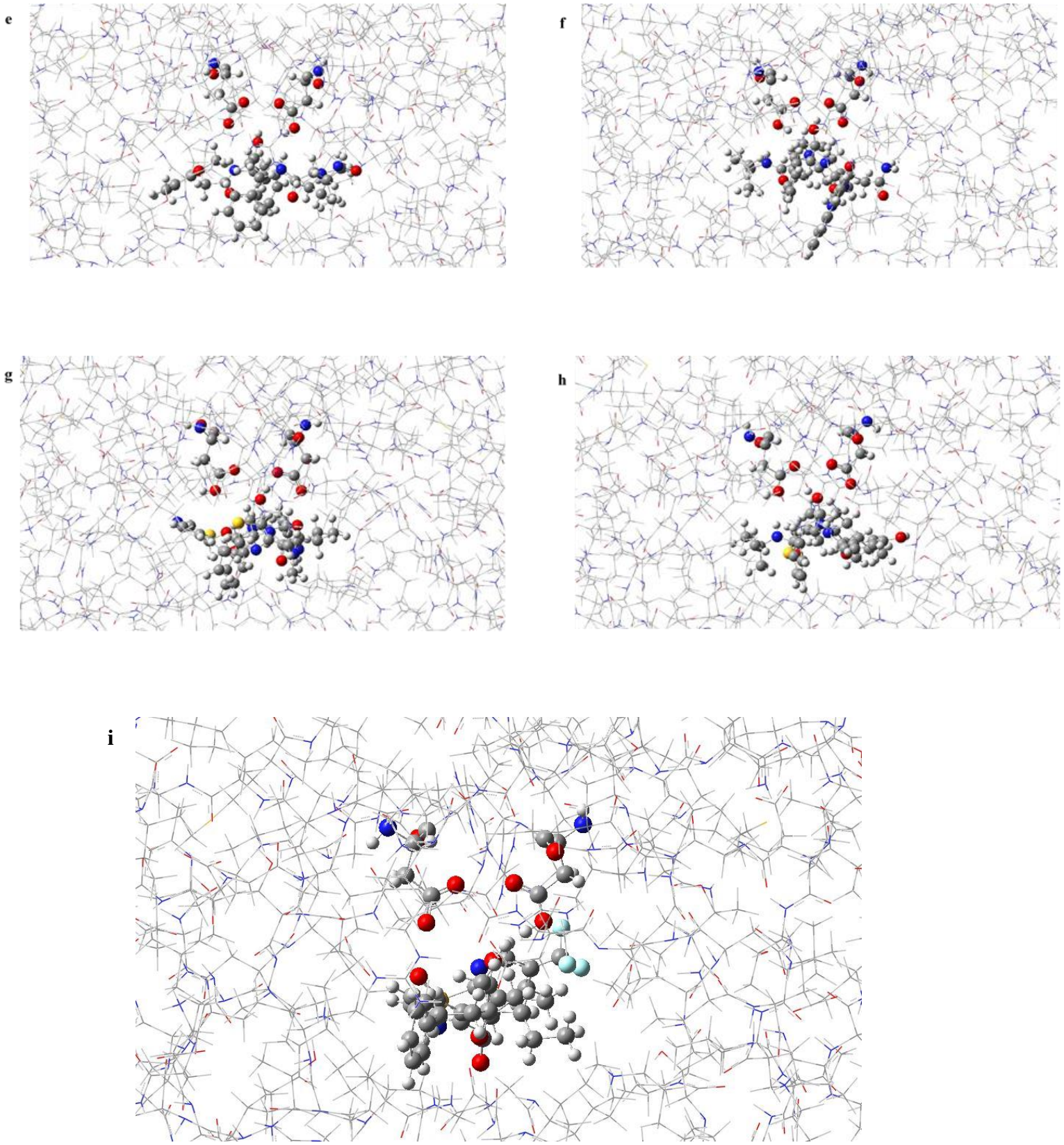
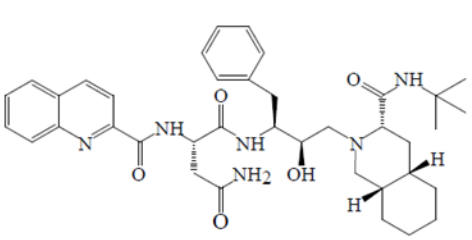
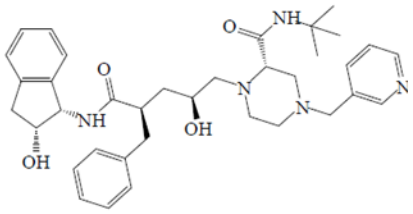


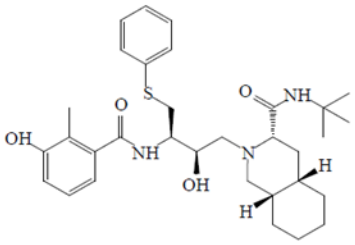
Figure S10. 3D Schematic representation of L38L↑N↑L PR complexed with **a:** APV, **b:** ATV, **c:** DRV, **d:** IDV, **e:** LPV, **f:** SQV, **g:** RTV, **h:** NFV, and **i:** TPV PIs.



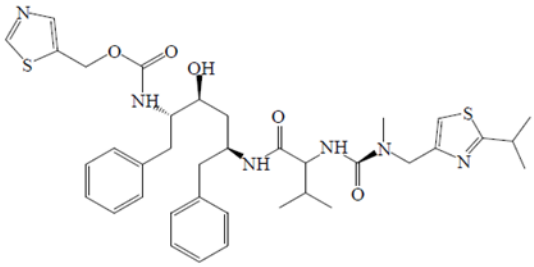
Saquinavir (1995)



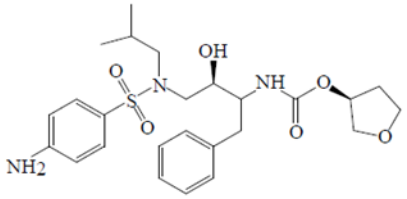
Indinavir (1996)



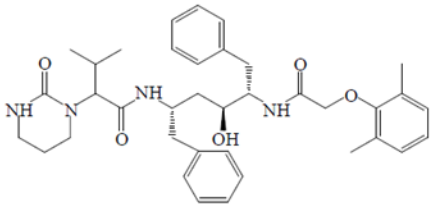
Nelfinavir (1997)



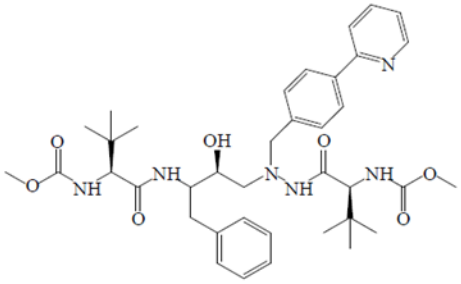
Ritonavir (1996)



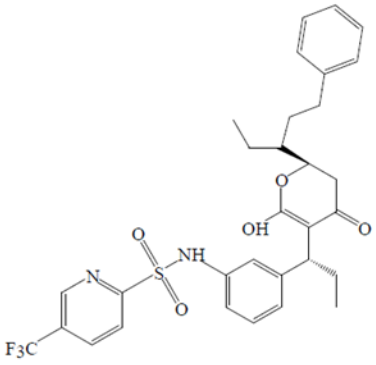
Ampenavir (1999)



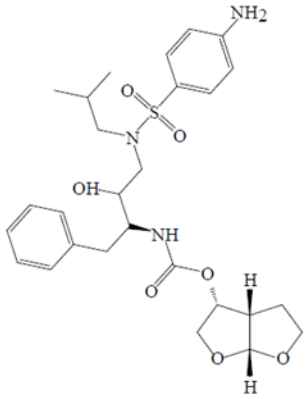
Lopinavir (2000)



Atazanavir (2003)



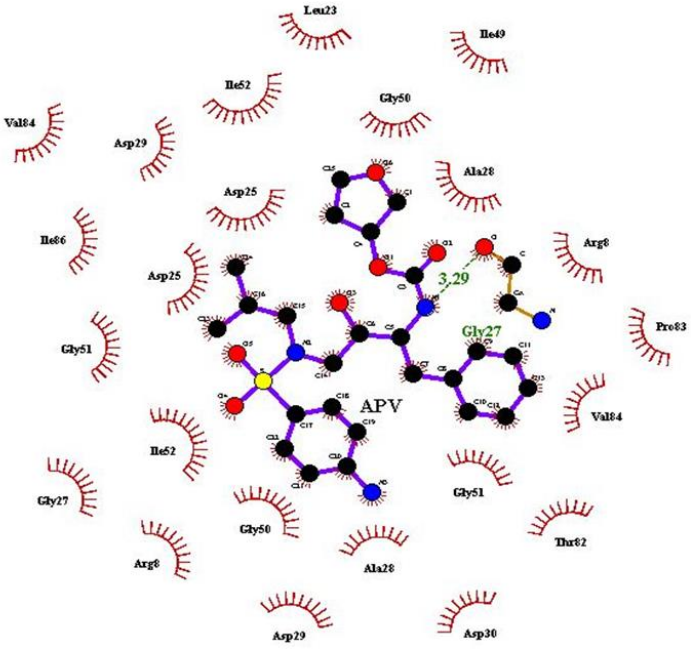
Tipranavir (2005)



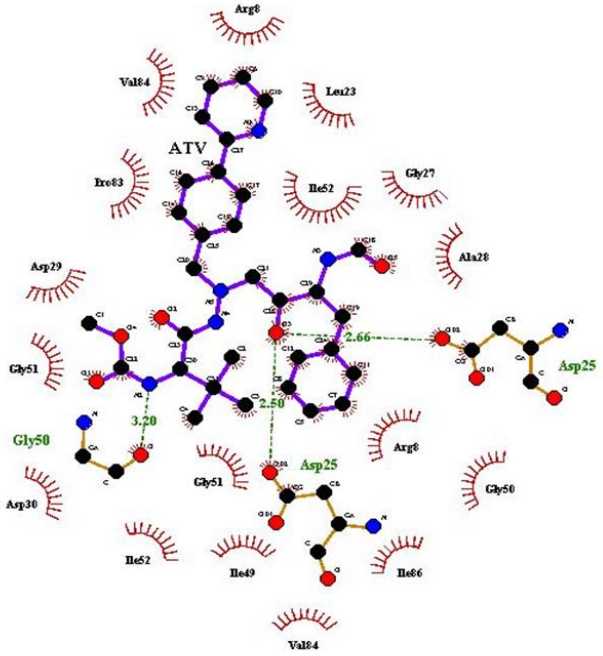
Darunavir (2006)

Figure S11. The structures of the FDA approved HIV PR inhibitors.

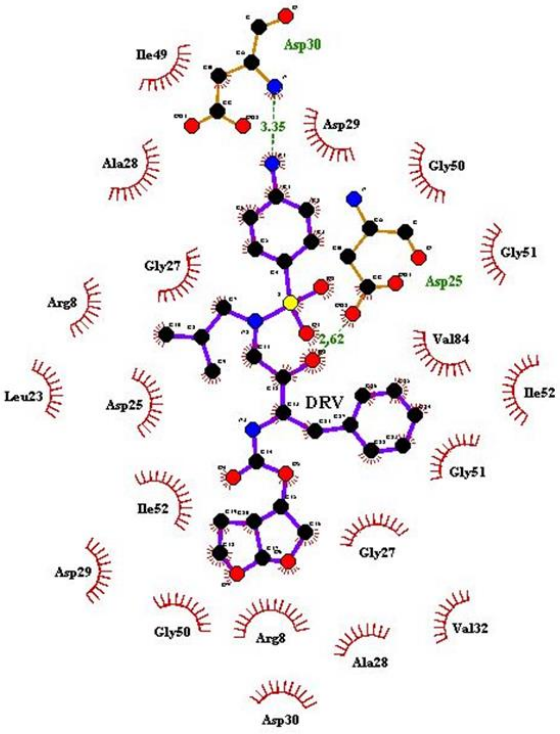
Detailed plots showing the electrostatic interactions and hydrogen bonding of L38L↑N↑L PR with the selected FDA approved HIV PIs.



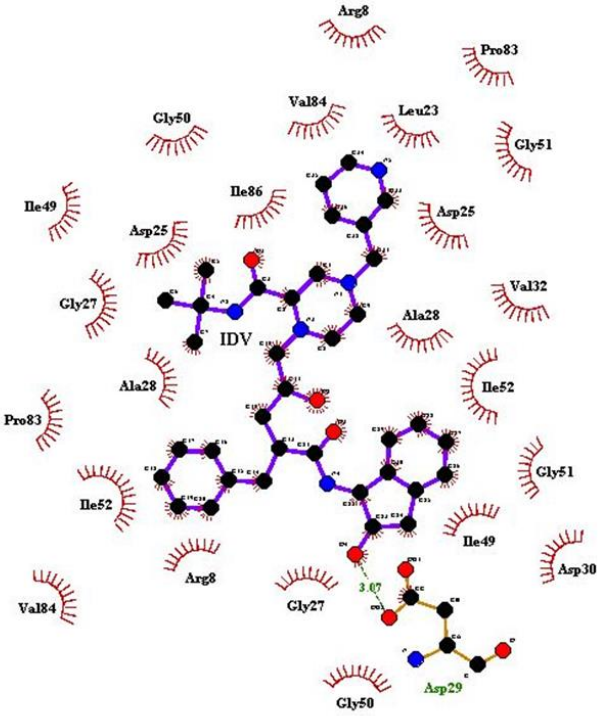
L38LNL-APV PR



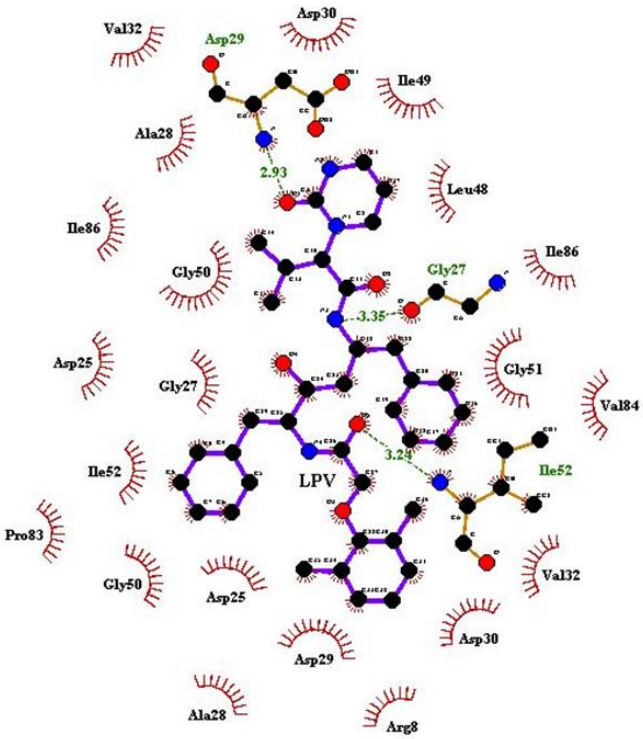
L38LNL-ATV PR



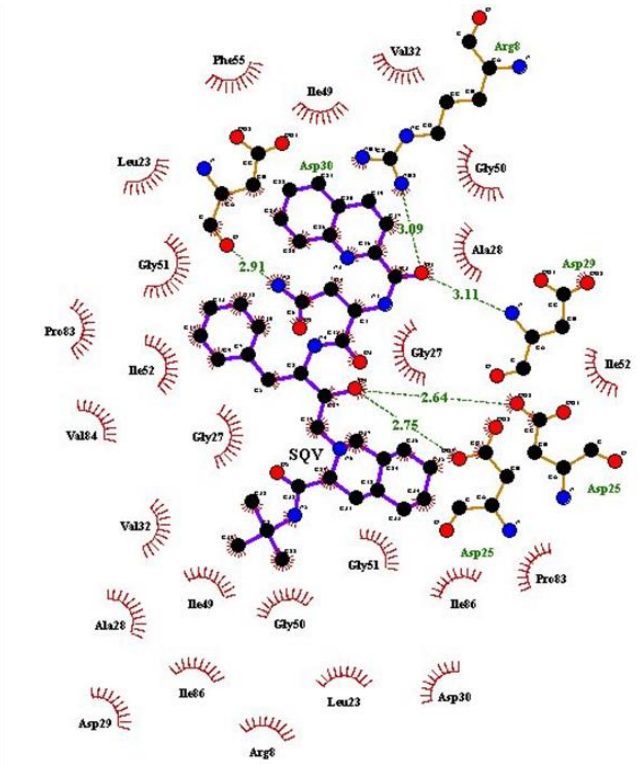
L38LNL-DRV PR



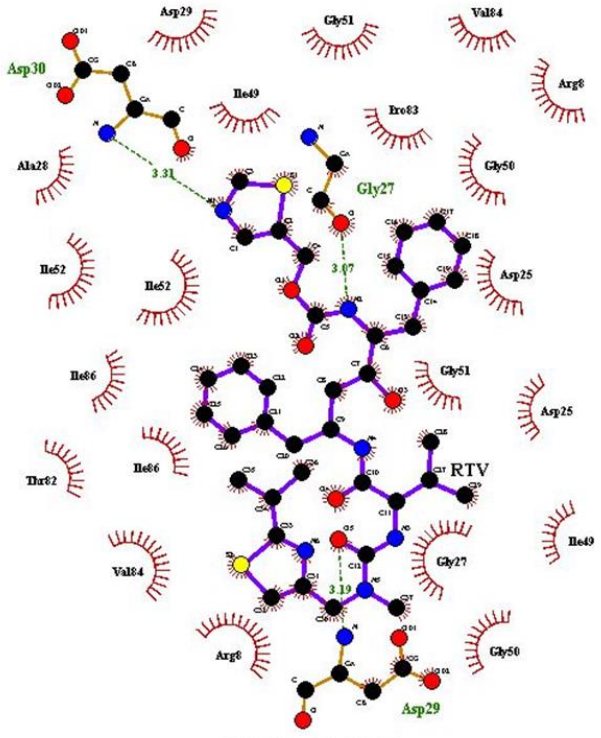
L38LNL-IDV PR



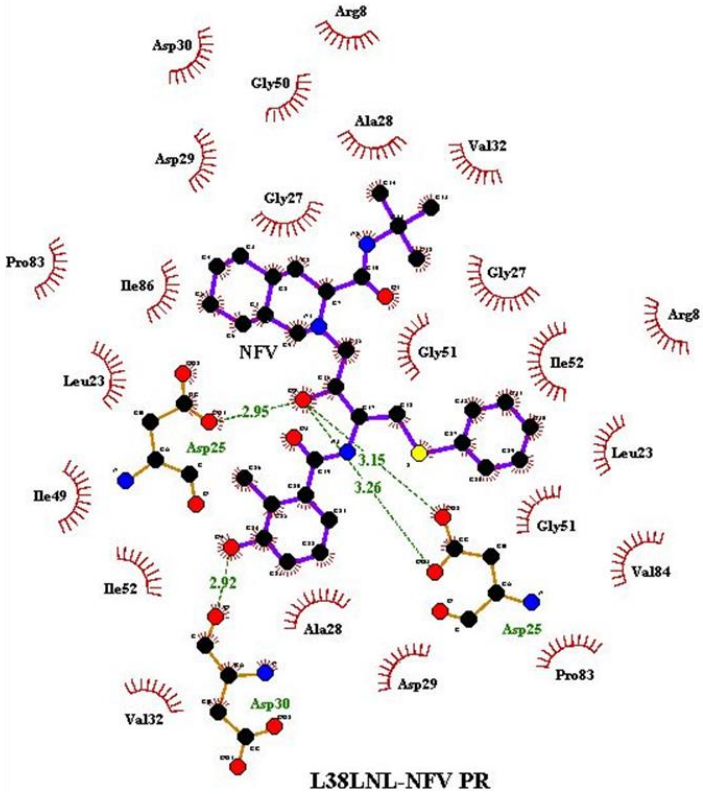
L38LNL-LPV PR



L38LNL-SQV PR



L38LNL-RTV PR



L38LNL-NFV PR

Figure S3. Electrostatic and hydrogen bond interactions plots of L38L↑N↑L PR with the selected FDA approved HIV PIs.

The Figure below illustrates before and after geometry optimization was performed on all L38L↑N↑L PR—inhibitor complexes in determining the hydrogen bond (HB) distance between the catalytic Asp 25/25' and hydroxyl group of the selected FDA approved inhibitors.

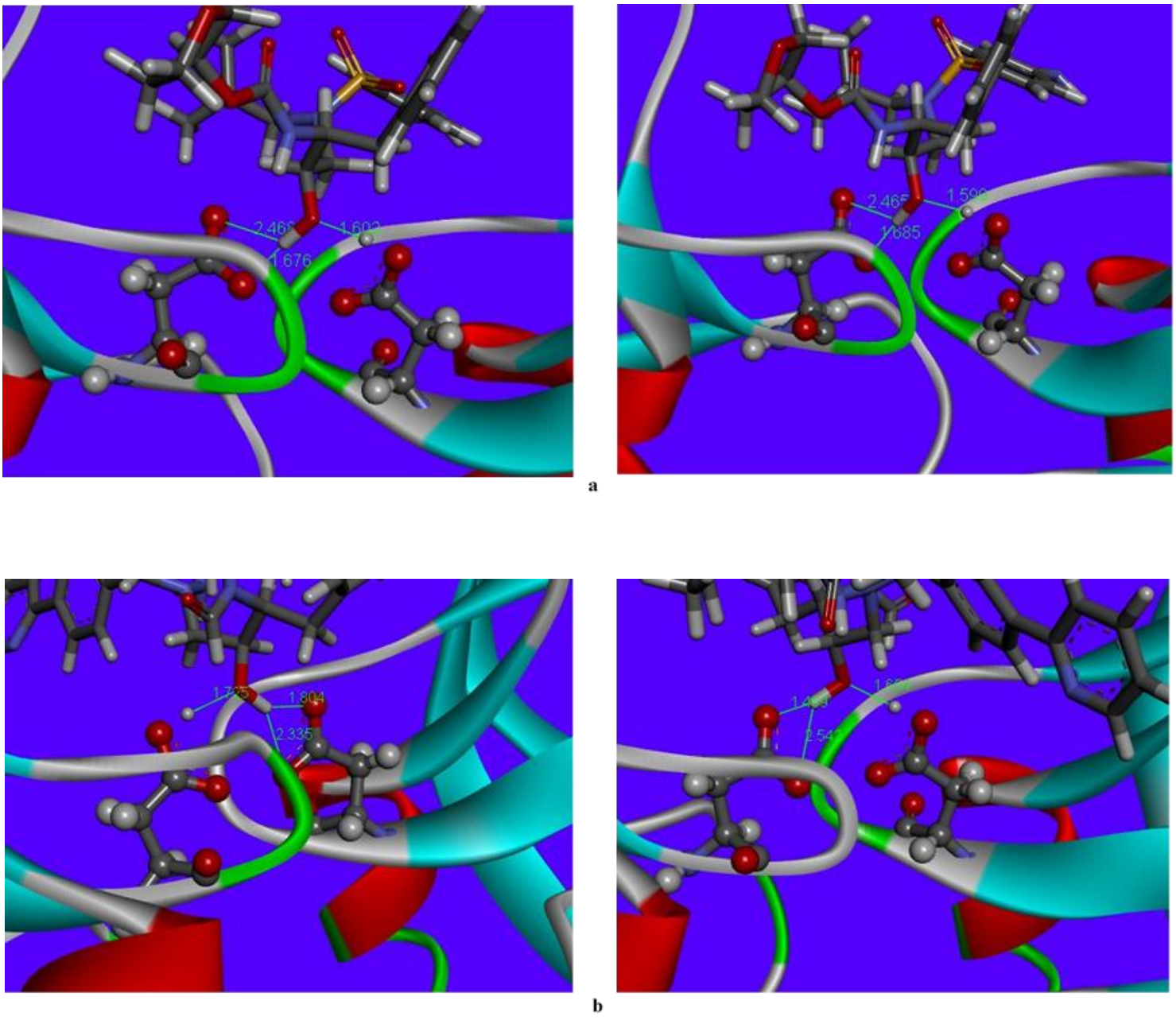
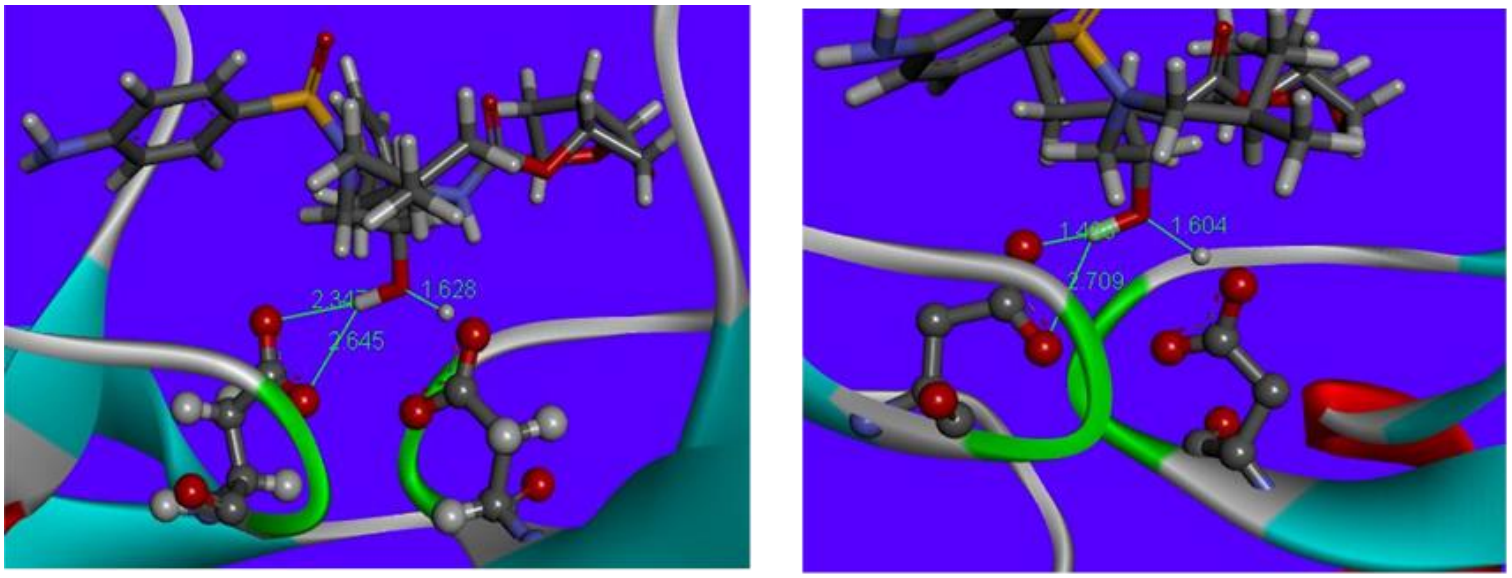
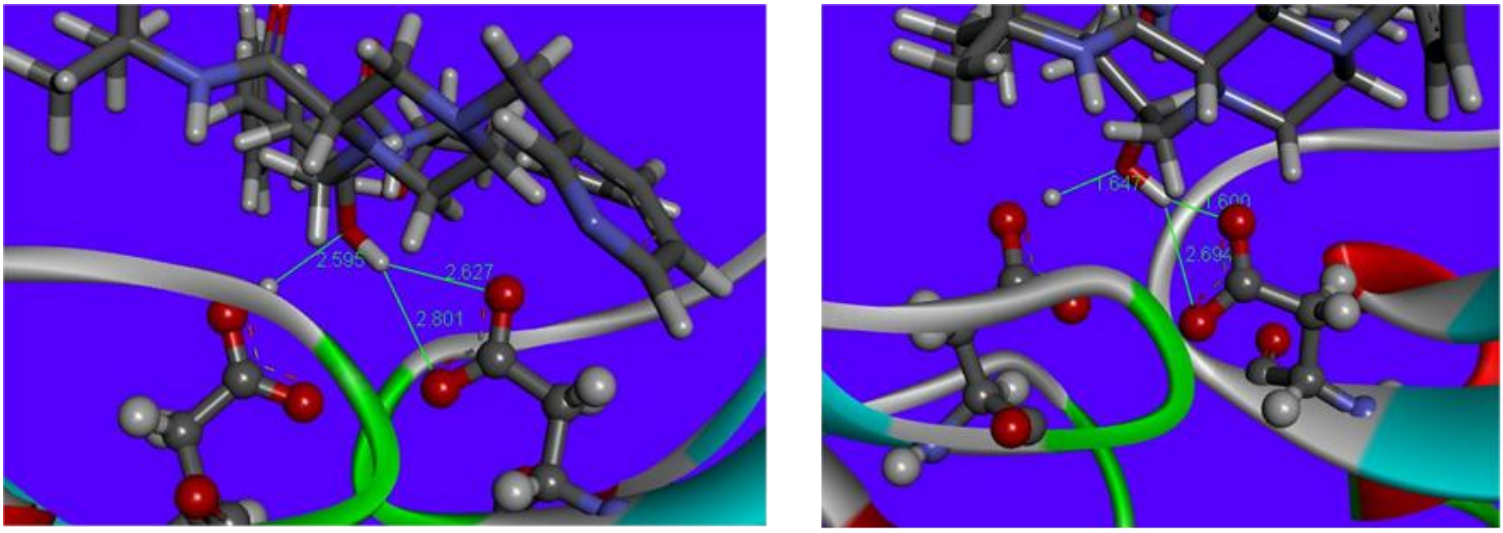


Figure S4. Hydrogen bond distances between the OH group of the FDA approved drugs **a:** APV, **b:** ATV with the catalytic Asp 25/25' residues of L38L↑N↑L before and after optimization.

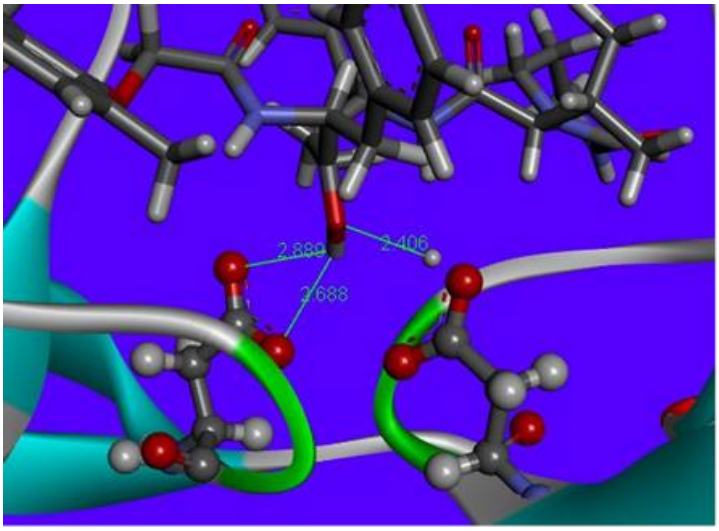


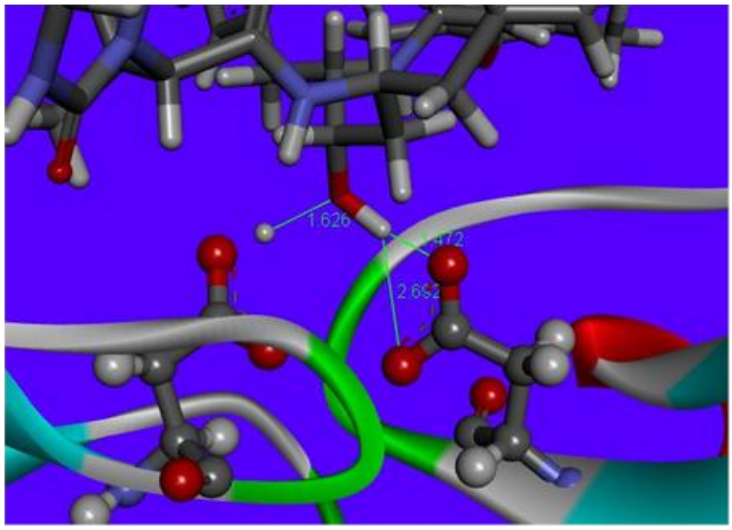
c



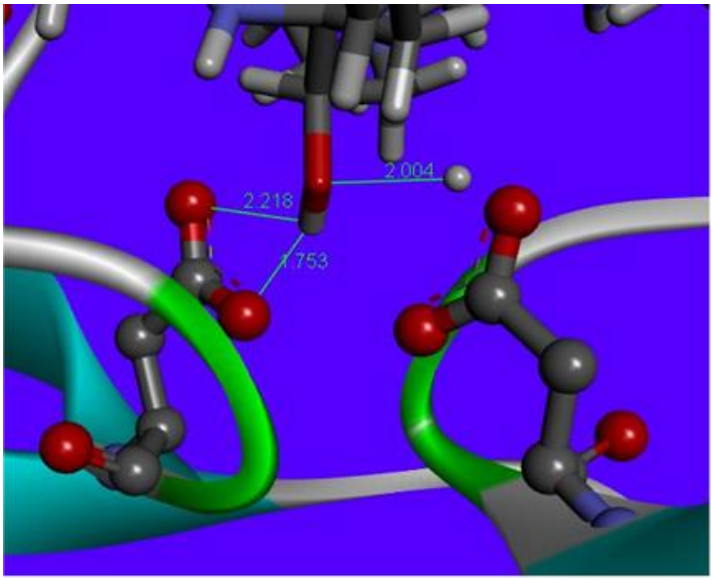
d

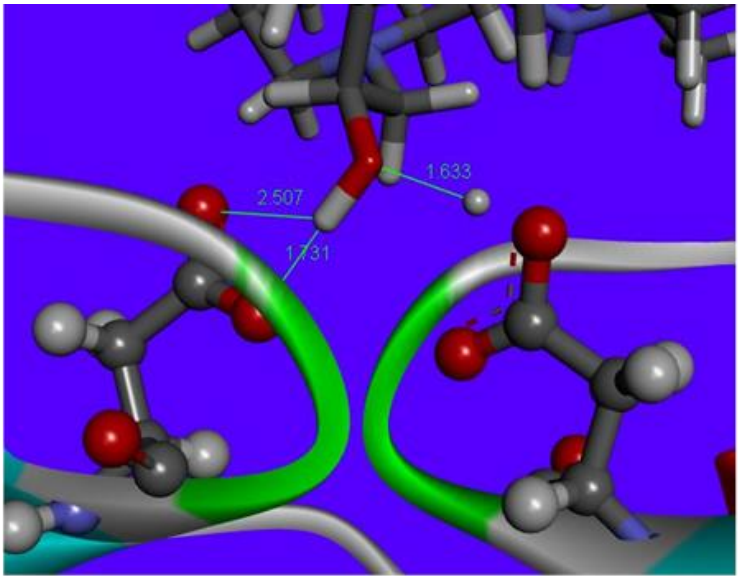
Figure S4. Hydrogen bond distances between the OH group of the FDA approved drugs **c:** DRV, **d:** IDV with the catalytic Asp 25/25' residues of L38L↑N↑L before and after optimization.





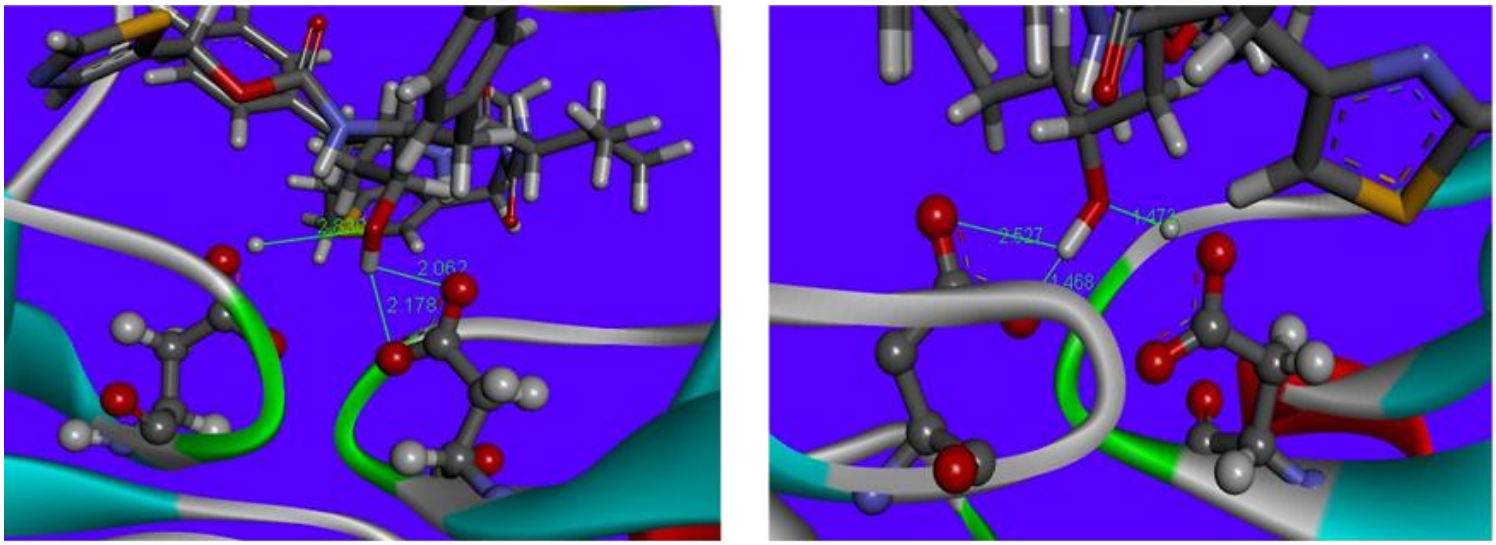
e



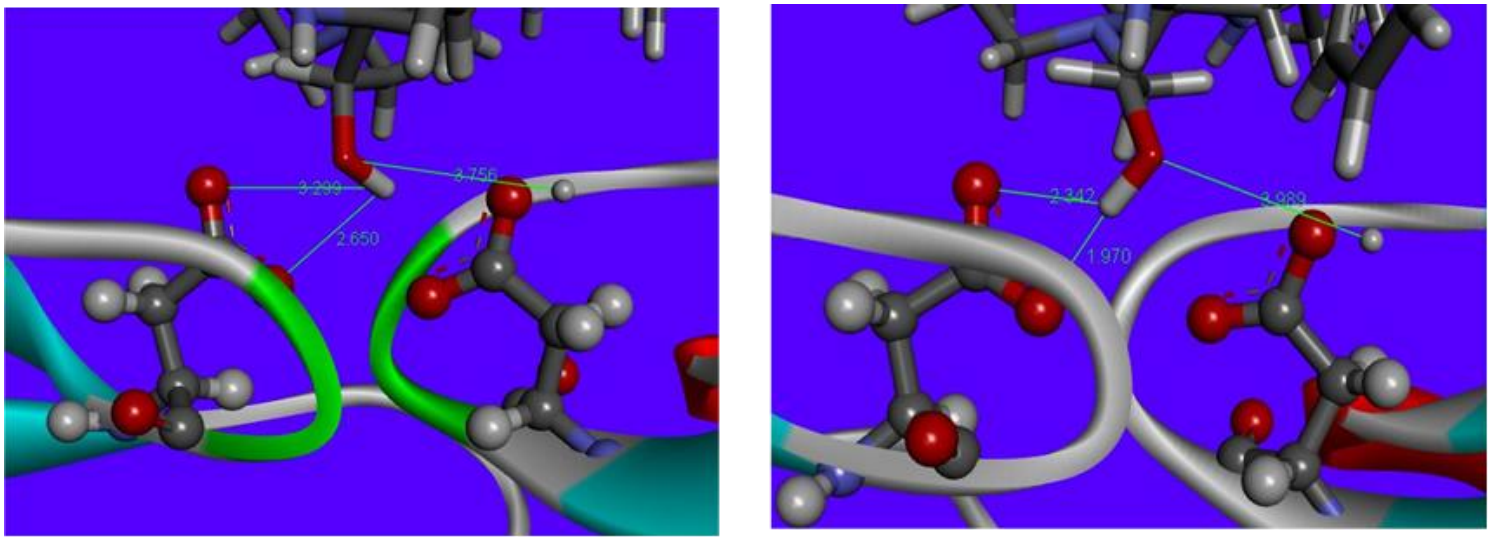


f

Figure S4. Hydrogen bond distances between the OH group of the FDA approved drugs **e:** LPV, **f:** SQV with the catalytic Asp 25/25' residues of L38L↑N↑L before and after optimization.



g



h

Figure S4. Hydrogen bond distances between the hydroxyl groups of the selected FDA approved drugs **g:** RTV, **h:** NFV, with the catalytic Asp 25 and Asp25' residues of L38L↑N↑L PR before and after optimization.