Biochemical Characterization of Highly Mutated South African HIV-1 Subtype C Protease

By

Simeon Eche

(Registration number: 215000273)

Submitted in fulfilment of the academic requirements for the degree of

Doctor of Philosophy in Virology

School of Laboratory Medicine and Medical Sciences, Nelson R Mandel School of Medicine,

University of KwaZulu-Natal

Supervisor: Professor Michelle L. Gordon

2021

Preface

The work described in this thesis was carried out at the HIV Pathogenesis Programme and KwaZulu-Natal, Nelson R Mandela School of Medicine, University of KwaZulu-Natal, Durban, South Africa, from January 2017 to January 2021, under the supervision of Prof. Michelle Gordon.

This work has not been submitted in any form for any degree or diploma to any tertiary institution. Where use has been made of the work of others, it is duly acknowledged in the text.

Signed

……………………………….   10th May 2021

Simeon Eche

Signed

                                                                              10th May 2021

Professor Michelle Lucille Gordon

Declaration

I, Mr. Simeon Eche, declare that:

1. The research reported in this dissertation, except where otherwise indicated, is my original research.

2. This dissertation has not been submitted for any degree or examination at any other university.

3. This dissertation does not contain other scientists' data, pictures, graphs or other information unless specifically acknowledged as being sourced from other scientists.

4. This dissertation does not contain other scientists' writing unless specifically acknowledged as being sourced from other scientists. Where other written sources have been quoted, then their words have been re-written, but the general information attributed to them has been referenced.

5. This dissertation does not contain text, graphics, or tables taken from the internet unless specifically acknowledged, and the source being detailed in the dissertation and in the reference section.

Signed:

…………………………..

Simeon Eche

Dedication

This thesis is dedicated to the Almighty God.

I also dedicate this thesis to the memory of my late mother – Mrs. Patricia Adah, my late friend Gloria Oteh, my dear wife – Mrs. Miracle T. Eche, and to my beloved daughter – Zanita Eche.

Abstract

Understanding the underlying molecular mechanism of HIV-1 protease (PR) inhibition by HIV-1 protease inhibitors (PIs) is essential to gain mechanistic insight into the evolution of resistance to HIV-1 PIs. HIV-1 PIs have improved patient care management, but the accumulation of drug resistance mutations in the HIV-1 PR gene diminishes their inhibitory capacity. The current study investigated the kinetic and structural characteristics of highly mutated South African HIV-1 subtype C PR from clinical isolates obtained from individuals failing a lopinavir (LPV) inclusive regimen at the point of switch to darunavir (DRV) based therapy. In this study, enzyme activity and inhibition assays were used to determine the biochemical fitness of HIV-1 PR variants and the inhibitory constants of HIV-1 PIs for drug-resistant HIV-1 subtype C proteases. The mechanistic insight into the impact of the accumulated drug resistance mutations on the HIV-1 PR structure and its interaction with LPV and DRV was obtained using fluorescence spectroscopy and molecular dynamic simulation.

The study showed that the unfavorable binding landscape caused by the accumulation of drug-resistance mutations resulting from LPV associated drug pressure would shape the outcome of DRV-based therapy after a switch in the treatment regimen. This is related to the distortion of the HIV-1 PR structure associated with increased solvent exposure and instability of the HIV-1 PR dimer caused by these mutations leading to a shorter lifetime of the enzyme-inhibitor complex. Analysis of the binding kinetics of LPV and DRV with the HIV-1 PR variants showed that the drug resistance mutations caused an imbalance between the association and dissociation rate constants favoring a fast dissociation rate. The latter resulted in a reduced inhibitor residence time. Our findings showed that LPV had a longer residence time than DRV when bound to the HIV-1 PR variants; this shows LPV can be a suitable platform for developing newer HIV-1 PIs with a longer residence time. However, the enzyme inhibition mechanism shows both LPV and DRV act via a two-step tight-binding mixed inhibition mechanism, suggesting the existence of a second binding site on HIV-1 PR for these inhibitors. The information provided in this thesis adds to existing knowledge about HIV-1 PI drug resistance and for the design of novel HIV-1 PIs with the potential to evade drug resistance mutations.

Table of content

List of Figures

List of Tables

Abbreviations

3D	Three-dimensional
AIDS	Acquired immune deficiency syndrome
AMBER	Assisted model building with energy refinement
APV	Amprenavir
ARV	Antiretroviral
BSA	Bovine serum albumin
CA	Capsid; p24
CDC	Centre for Disease Control and Prevention
cpx	Complexes
CRF	Circulating recombinant form
EDTA	Ethylenediaminetetraacetic acid
EI	Enzyme-inhibitor complex
ES	Enzyme-substrate complex
ESI	Enzyme-substrate-inhibitor complex
DRV	Darunavir
DRV/r	Darunavir boosted with ritonavir
DTT	Dithiothreitol
FDA	Food and Drug Administration
GST	Glutathione-S-transferase
HAART	Highly active antiretroviral therapy
His-tag	Poly histidine-tag
HIV	Human immunodeficiency virus

HIV-1	Human immunodeficiency virus type I
HIV-2	Human immunodeficiency virus type II
IN	Integrase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KZN	KwaZulu-Natal
LPV	Lopinavir
LPV/r	Lopinavir boosted with ritonavir
LTR	Long terminal repeat
MA	Matrix; p17
MBP	Maltose binding protein
MD	Molecular dynamics
MDR	Multi-drug resistant
MM-GBSA	Molecular mechanics-generalized born surface area
MM-PBSA	Molecular Mechanics-Poisson Boltzmann surface area
MUT	Mutant
NC	Nucleocapsid, P7
NFV	Nelfinavir
NMR	Nuclear magnetic resonance
NNRTI	Non-nucleoside reverse transcriptase inhibitors
NPT	Constant pressure and normal temperature
NRTI	Nucleoside reverse transcriptase inhibitors
nm	nanometres
ns	Nanoseconds

NVT	Constant volume and normal temperature
ORF	Open reading frame
PCR	Polymerase chain reaction
PDB	Protein Databank
PI	Protease inhibitor
ps	Picoseconds
PR	Protease
rpm	Revolutions per minute
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
ROG	Radius of gyration
RT	Reverse transcriptase
RTV	Ritonavir
SASA	Solvent accessible surface area
SDS-PAGE	Sodium dodecyl sulphate–polyacrylamide gel electrophoresis
SIV	Simian immunodeficiency virus
TBS	Tris-Buffered Saline
Trp	Tryptophan
Trx	Thioredoxin
URFs	Unique recombinant forms
vdW	van der Waals
WT	Wild type

Symbols

α	Alpha
β	Beta
Å	Angstroms
Cα	Alpha carbon
K_{cat}	Catalytic constant
K_i	Inhibitory constant
K_m	Michaelis constant
K_{sv}	Stern-Volmer constants
k_5	Association rate constant
k_6	Dissociation rate constant
K_{obs}	Observed rate constant
K	Kelvin
kDa	Kilo Dalton
Kcal	Kilocalories
λ_{max}	Maximal fluorescence emission spectra
π -	pi
t_R -	Inhibitor residence time
v_{max}	Maximal velocity
v_s	Steady state velocity
v_0	Initial velocity

CHAPTER ONE

Introduction and literature review

1.1   Introduction

The world has continued to experience one of its worst public health problems in the human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) epidemic. HIV causes dysregulation of the immune system and chronic infection leading to AIDS if not checked (WHO, 2020). HIV-1 can be transmitted from an infected person to an uninfected person through the exchange of certain body fluids like blood, vaginal secretions, semen, and breast milk. It could also be transmitted during pregnancy and delivery from an infected mother to her child (Shaw and Hunter, 2012). Globally, heterosexual transmission of HIV-1 accounts for close to 70% of infections, and the remaining 30% is ascribed to men who have sex with men (MSM), mother-to-child transmission, and injection drug use (Shaw and Hunter, 2012). While the cure for HIV remains elusive, research has helped to broaden the understanding and knowledge of HIV biology and pathogenesis. This has led to the development of highly active antiretroviral therapy (HAART) to treat HIV-infected individuals (Menéndez-Arias, 2013).

Highly active antiretroviral therapy has improved the treatment of HIV-1 infection. Different HAART regimens have been used to achieve virological suppression, leading to improved clinical outcomes and decreased morbidity and deaths associated with HIV-1 infection (Tseng *et al*., 2015). Notwithstanding the recorded success of HAART in the management of HIV-infected individuals, HAART is inhibitory and not eradicative (Lu *et al*., 2018), and therefore cannot cure the HIV-1 infection. This is due to a stable latent proviral reservoir of HIV-1 in the memory CD4+ T cells. Therefore, people living with HIV need to be placed on lifelong HAART regimens (Lu *et al*., 2018, Kimata *et al*., 2016). HAART involves the use of inhibitors designed to target enzymes crucial to the HIV-1 life cycle (Henes *et al*., 2019). These target enzymes include reverse transcriptase, integrase, and protease (Henes *et al*., 2019).

The introduction of HAART may have helped to improve the management of HIV infection (Maartens *et al*., 2014). Nevertheless, the efficacy is affected by the appearance of mutations in the HIV-1 proteins targeted by the treatment regimen leading to HIV-1 infected individuals experiencing virological failure (Henderson *et al*., 2012). The focus of this study is on the evolution of resistance by highly mutated South African HIV-1 subtype C protease (PR) to lopinavir (LPV) used in the formulation of second-line regimens in low and middle-income countries, and darunavir (DRV) used in the formulation of salvage regimens. Research has shown that the number of HIV-1 infected individuals failing their 2nd-line HIV-1 treatment regimen in sub-Saharan Africa is increasing, with failure to these regimens occurring at 12–18-month of initiation (Edessa *et al*., 2019). A recent study in South Africa has shown

that greater than one major HIV-1 PI resistance mutations were observed in 18% of the study participants (Obasa *et al*., 2020). The clinical impact of HIV-1 PI resistance mutations on virologic failure in people infected with HIV-1 subtype C in South African may be known; however, there is a lack of information on the structural and biochemical characterization of highly mutated HIV-1 subtype C PR and its interaction with LPV and DRV. Therefore, elucidation of the biochemical and structural characteristics of highly mutated South African HIV-1 subtype C PR will provide insight into its molecular interaction with HIV-1 PIs, its structural dynamics, and how it impacts drug binding.

1.2 Epidemiology of HIV

The global statistics of the HIV/AIDS epidemic give a picture of how serious a threat this epidemic is to humanity. By the end of 2019, approximately 38 million people globally live with HIV (UNAIDS, 2020). About 32.7 million people have died from AIDS-associated complications dating back to the time the epidemic started (UNAIDS, 2020). The weight of the global HIV burden is felt more by third-world countries in Sub-Saharan Africa, where about 70% of people living with HIV in the world reside (UNAIDS, 2020). Irrespective of the increase in the use of ART in the management of HIV-infected individuals, about 34% of people living with HIV in Southern Africa as well as East Africa, and 60% of people living with HIV in West and Central Africa do not have access HIV care. This has continually been one of the most common causes of death in sub-Saharan Africa (Roth *et al*., 2018, UNAIDS, 2018).

In Sub-Saharan Africa, South Africa is the most affected country. At the end of 2018, around 7.7 million South Africans were living with HIV making it the country with the highest HIV epidemic globally. There were 240 000 new infections and 71 000 deaths associated with HIV in South Africa in 2018 (UNAIDS, 2018). These figures have led to the South African government putting measures to halt the spread and improve the management of infected individuals. The measures put in place by the government are yielding good results as the deaths associated with HIV have dropped from 140 000 deaths to 71 000 deaths in 2018. There is also a significant decline in the number of new infections within the same period, from 390 000 to 240 000 new infections (UNAIDS, 2018).

1.3 Classification of HIV

The human immunodeficiency virus belongs to the genus Lentivirus, in the family Retroviridae, and the subfamily of Orthoretrovirinae. HIV is classified into types 1 and 2 (HIV-1, HIV-2) based on differences in genetic characteristics and viral antigens. Genetically, HIV-1 and HIV-2 are from the Simian immunodeficiency virus (SIV) and have evolved through cross-species transmission (Kirchner, 2019). Available epidemiologic and phylogenetic data shows HIV-1 was introduced into the human population through the zoonotic transmission of SIV from chimpanzees (*Pan troglodytes troglodytes*)

around 1920 in the central part of Africa (Mourez *et al*., 2013, Locatelli and Peeters, 2012). On the other hand, HIV-2 was zoonotically transmitted from the *Sooty mangabey* (*Cercocebus atys*) to humans around 1940 in the Western part of Africa (Lemey *et al*., 2003). Approximately 97% of global HIV infection is caused by HIV-1, while HIV-2 is found mostly in the Western part of Africa, responsible for about 3% of HIV infections globally (Kirchner, 2019). The focus of this thesis is HIV-1.

1.4   HIV-1 diversity and variability

The presence of genetic variants of HIV-1 has accelerated the global HIV pandemic, having a strong influence on diagnosis, treatment, treatment monitoring, and the development of an effective vaccine against the virus (Hemelaar, 2013). The genetic variation of HIV-1 is a product of numerous factors like the rapid rate of replication of the virus, pressures on the virus from the host immune system, drugs used for treatment, and the recombination events that occur during the process of replication (Ramirez *et al*., 2008, Taylor *et al*., 2008). HIV-1 is classified into groups M, N, O, and P (figure 1.1). Three of these groups: N, O, and P, are limited to the west and central part of Africa (Maartens *et al*., 2014). Group M drives the global HIV-1 epidemic, and it diversifies into different subtypes A, B, C, D, F, G, H, J, K, circulating recombinant forms (CRFs) or unique recombinant forms (URFs) and the newly described HIV-1 subtype L (Hemelaar *et al*., 2019, Yamaguchi *et al*., 2019). HIV-1 subtypes A and F are further classified into sub-sub types A1 to A4 and F1 and F2, respectively (Peeters *et al*., 2013).

The CRFs and URFs are a product of recombination between subtypes. While the CRFs are strains that spread in the population, the URFs are unique recombinant sequences with limited transmission (Robertson *et al*., 2000). Unlike the subtypes, which are assigned letters, the CRFs are annotated with numbers and letters. These numbers reflect their order of description, and the two-letter code reflects the subtypes present in the mosaic structure of the CRF. Examples of CRFs are CRF02_AG and CRF02_AE (Peeters *et al*., 2013). HIV-1 subtype E and I were previously described based on envelope sequences; further analysis of their full-length genomes showed they had a mosaic structure. They were subsequently designated as CRF01_AE (subtype E) and CRF04_cpx (subtype I). Viruses with a complex mosaic structure that have arisen from the recombination of three or more subtypes are called complexes, and the term cpx is used to refer to them (Peeters *et al*., 2013).
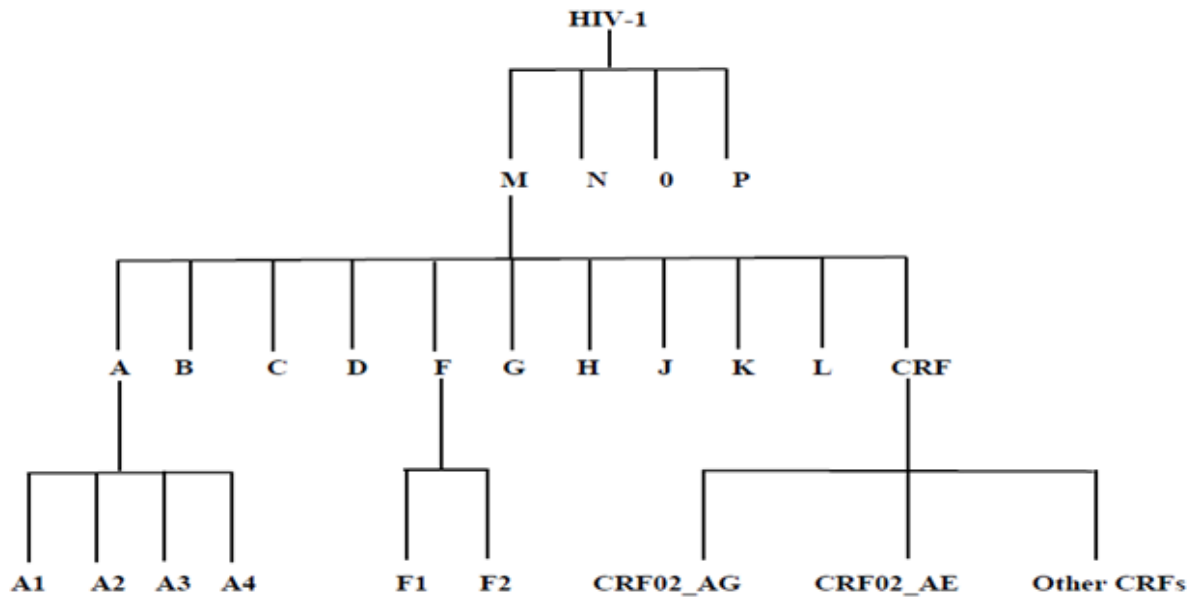
Figure 1. 1: Phylogeny of HIV-1.

1.5 Molecular epidemiology of and global distribution of HIV-1

The distribution of HIV-1 subtypes, CRFs, and URFs differs across different regions globally and may change over time (Figure 1.2). HIV-1 diversity in the central part of Africa is the greatest, with all the HIV-1 subtypes and recombinant forms present. The highest percentage of URFs is also found in this region compared to any other region (Hemelaar *et al.*, 2019). In east Africa, eastern Europe, and central Asia, HIV-1 subtype A is responsible for more than 50% of the HIV-1 infections and accounts for about 10% of global HIV-1 infections. The HIV-1 epidemic in the Americas, the western and central parts of Europe, and the Oceania region are dominated by HIV-1 subtype B. It is accounts for about 12% of global HIV-1 infections. In the southern part of Africa, Ethiopia, and South Asia, the HIV epidemic is dominated by HIV-1 subtype C, which alone accounts for approximately 46% of infections globally (Hemelaar *et al.*, 2019). In South Africa, the HIV-1 epidemic is dominated by the South African HIV-1 subtype C (Morris *et al.*, 2000).

A significant proportion of HIV-1 subtype G is found in the western part of Africa, and it is responsible for about 4·6% of global HIV-1 infection. HIV-1 subtype D found in the East African region causes 2·7% of global HIV-1 infection. The subtypes F, H, J, and K only contribute a combined 0·9% global HIV-1 infection (Hemelaar *et al.*, 2019). The CRF02_AG is responsible for most of the HIV-1 infection in West Africa. This region has the highest percentage of CRF02_AG compared to any other region. The global contribution of CRF02_AG to HIV-1 infection is 7·7%. In southeast Asia as well as east Asia, the HIV-1 epidemic is dominated by CRF01_AE. This variant of HIV-1 contributes 5·3% of global HIV-1 infection (Hemelaar *et al.*, 2019).
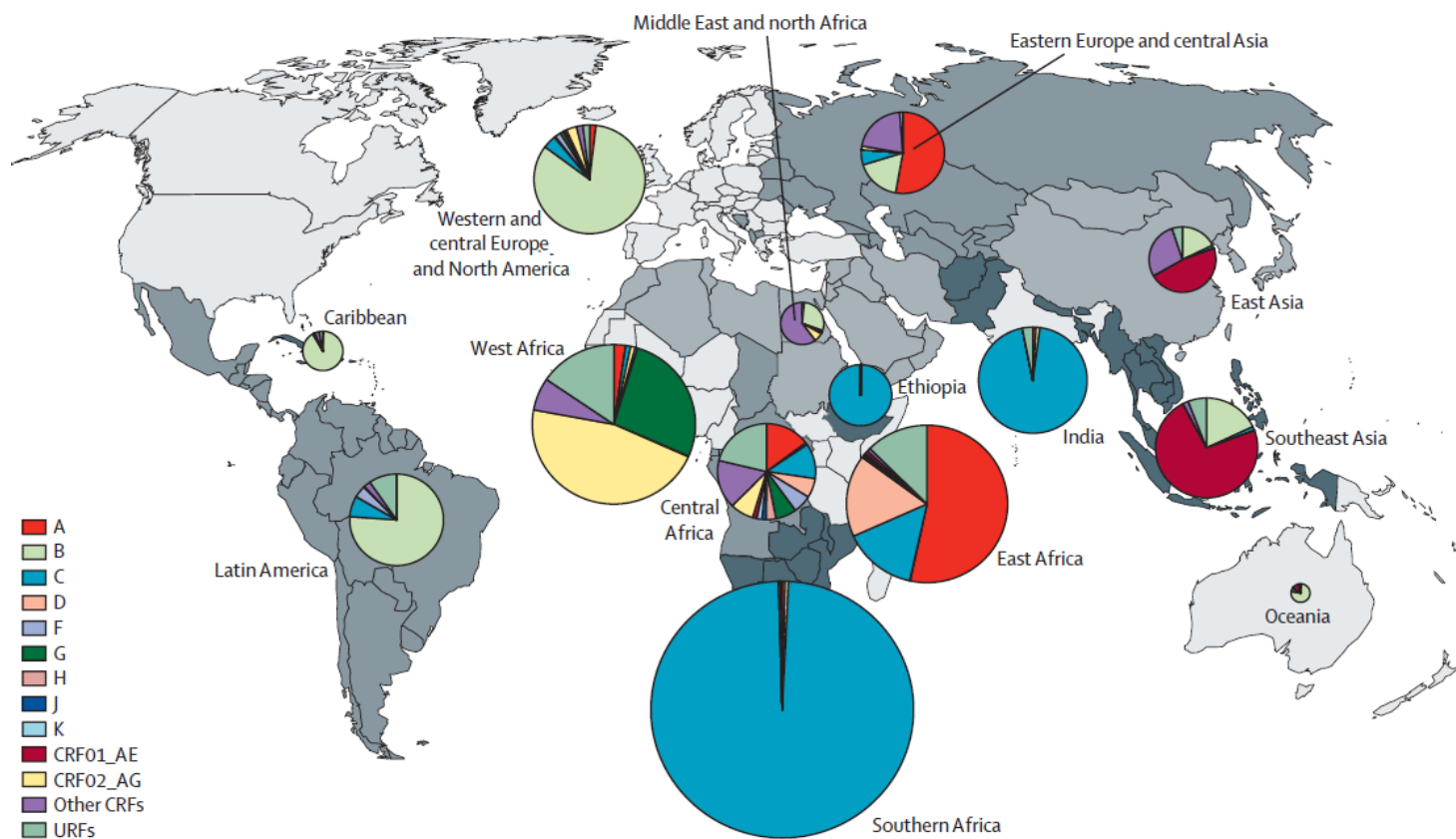
Figure 1. 2: Schematic diagrammatic of HIV-1 variants regional spread from 2010–15. Countries that form a region are shaded in the same color. The pie chart represents the number of HIV-1 infected individuals living in different areas (Taken from Hemelaar *et al.*, (2019).

1.6 HIV-1 Genome organization

The genome of HIV-1 comprises two identical single-stranded RNA molecules found within the core of the virus. It is around 9.8 kb with numerous open reading frames (ORF) transcribed into several viral proteins (Levy, 2007). The HIV-1 genomic DNA is bordered at the 5' and 3' ends by long terminal repeat (LTR) sequences. The LTR at the 5' region "codes for the promotor for transcription of the viral genes" (Blood and Hemotherapy, 2016, Shah *et al.*, 2014). The gag gene reading frame follows the 5' LTR region in the 5' to 3' direction, and this gene codes for proteins that form the outer core membrane like the matrix (MA, p17), capsid (CA, p24), and nucleocapsid (NC, p7). The gag gene also codes for a small nucleic acid-stabilizing protein (Freed, 1998).

The pol reading frame follows the gag reading frame, and it codes for viral enzymes like "protease, (PR, p12), reverse transcriptase which is composed of two subunits p51 and p66 proteins and integrase (IN, p32)" (Blood and Hemotherapy, 2016). Next to the "pol reading frame is the env reading frame, and the two glycoproteins, gp120 (surface protein, SU) and gp41 (transmembrane protein, TM)" (Blood

and Hemotherapy, 2016), arise from the viral envelope (Blood and Hemotherapy, 2016). Besides the viral structural proteins, other regulatory proteins are coded for by the HIV genome. These proteins are the trans-activator protein (Tat) and RNA splicing-regulator (Rev), essential for HIV replication initiation. In addition, regulatory proteins such as "negative regulating factor (Nef), viral infectivity factor (Vif), virus protein r (Vpr), and virus protein unique (Vpu)" (Levy, 2007) are essential in viral replication, viral budding as well as in viral pathogenesis are also coded for the HIV genome (Levy, 2007, Sauter *et al*., 2012).
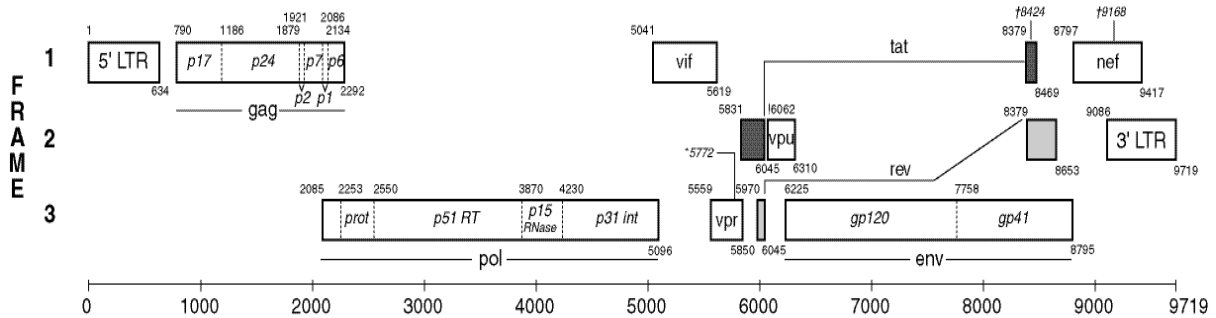
Figure 1. 3: HIV1 Gene map (Taken from Los Alamos HIV database: http://www.hiv.lanl.gov/. [Accessed 15th May 2020]).

1.7   The structure of matured HIV-1

The matured form of HIV measures about100nm in diameter and is round. It has an outer lipid membrane (envelope) and an inner core. The envelope has 72 knobs, made up of trimers of the envelope proteins. The trimers of the transmembrane protein gp41 (TM) anchors the trimers of gp120 surface protein (SU) to the envelope membrane (Gelderblom, 1991). The matrix protein (MA, p17) forms the symmetrical outer capsid membrane, and this outer capsid membrane is in turn covered by the viral envelope. The viral capsid is cone-shaped and formed by the inner capsid protein p24 (CA) (Niedrig *et al*., 1994). The HIV capsid surrounds the nucleocapsid, which encloses the two copies of viral genomic RNA as well as other viral molecules like RT, RNase H, PR, and IN, which are bound to the viral nucleic acid (Zhao *et al*., 2013a). Viral particles also contain oligopeptides, and these are produced as a result of precursor proteins (p55, p160) proteolytic processing (Blood and Hemotherapy, 2016).
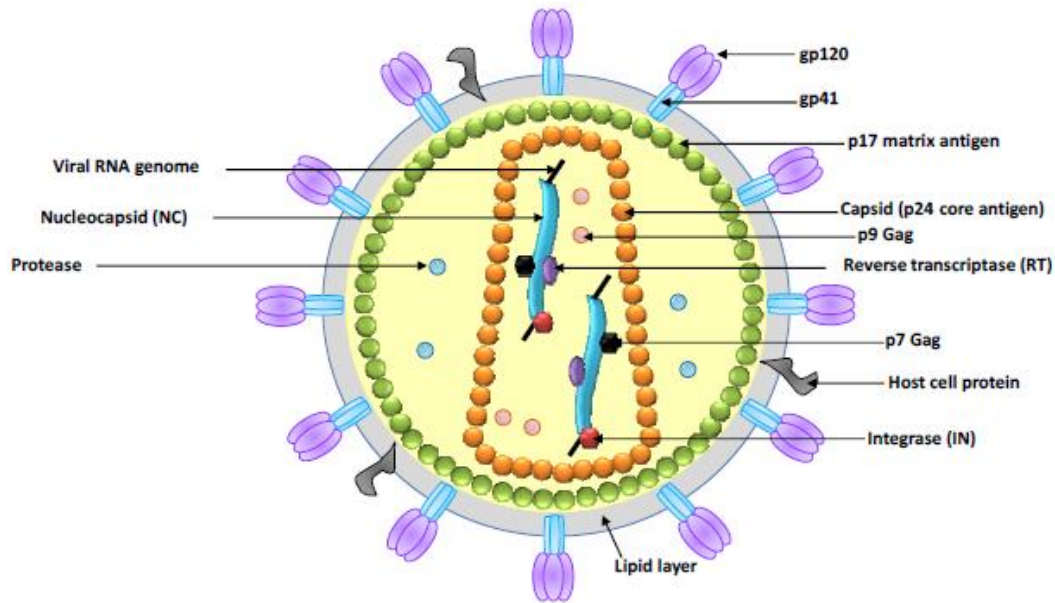
Figure 1. 4: The structure of matured HIV-1. (Taken from Shum *et al*. (2013)).

1.8 Replication and lifecycle of HIV-1

The HIV-1 replication cycle occurs in steps, as shown in the diagram below (figure 1.4). These different steps have become therapeutic targets for commercially available HIV-1 inhibitors (Deeks *et al*., 2015). The first step in the HIV replication cycle is the attachment of HIV-1 to the host cell through the viral glycoprotein (gp120) found in the HIV-1 envelope. These viral glycoproteins bind to CD4, which is their primary receptor (Holec *et al*., 2017). The binding of viral gp120 to a CD4 receptor brings about conformational changes in viral envelop trimer, allowing gp120 to interact with any of the co-receptor, CXCR4 (X4) or CCR5 (R5). Conformational changes associated with the co-receptor interaction enable the insertion of gp41 transmembrane protein fusion peptide into the host cell membrane. This establishes contact between the viral particle and its target cell, facilitating the viral particle fusion and the subsequent release of its content into the host cell (Kirchhoff, 2013).

Contrary to previous knowledge that the entry of the viral core into the host cell precedes uncoating to release the viral genome and viral enzymes (Ambrose and Aiken, 2014), recent studies have shown that uncoating is coupled to reverse transcription and this may protect the viral RNA during conversion to DNA (Müller et al., 2021, Cosnefroy et al., 2016). Viral RNA is transcribed by the reverse transcriptase (RT) enzyme into viral DNA (Kirchhoff, 2013), followed by its entry into the nucleus. Integration of the viral genome with the host genome is facilitated by the viral integrase (IN) enzyme. Once the proviral DNA is integrated into the host cell genome, it serves as a template for transcription of viral RNAs. Some of the newly transcribed RNA coding for HIV proteins (Craigie and Bushman, 2012). After the synthesis of new viral proteins and other components, they are assembled into viral particles. Viral assembly is a well-organized process that involves the multimerization of Gag and Gag-Pol

precursors through interactions amongst Gag proteins (Ganser-Pornillos *et al*., 2008). The matrix protein interacts with the envelop glycoproteins resulting in the recruitment of the former into the newly forming viral particle. Two copies of viral genomic viral RNA are also conscripted into this complex. The buildup of viral proteins and RNA at the host cell plasma membrane stimulates the formation of a rounded membrane-covered particle (Kirchhoff, 2013). The young virions bud off from the infected host cell. At this point, HIV-1 PR is activated and cleaves Gag and Gag-Pol precursors to form the various essential structural proteins needed for maturation into an infectious virus. This viral life cycle stage is an important therapeutic target for drugs that inhibit HIV-1 PR (Briggs and Kräusslich, 2011).
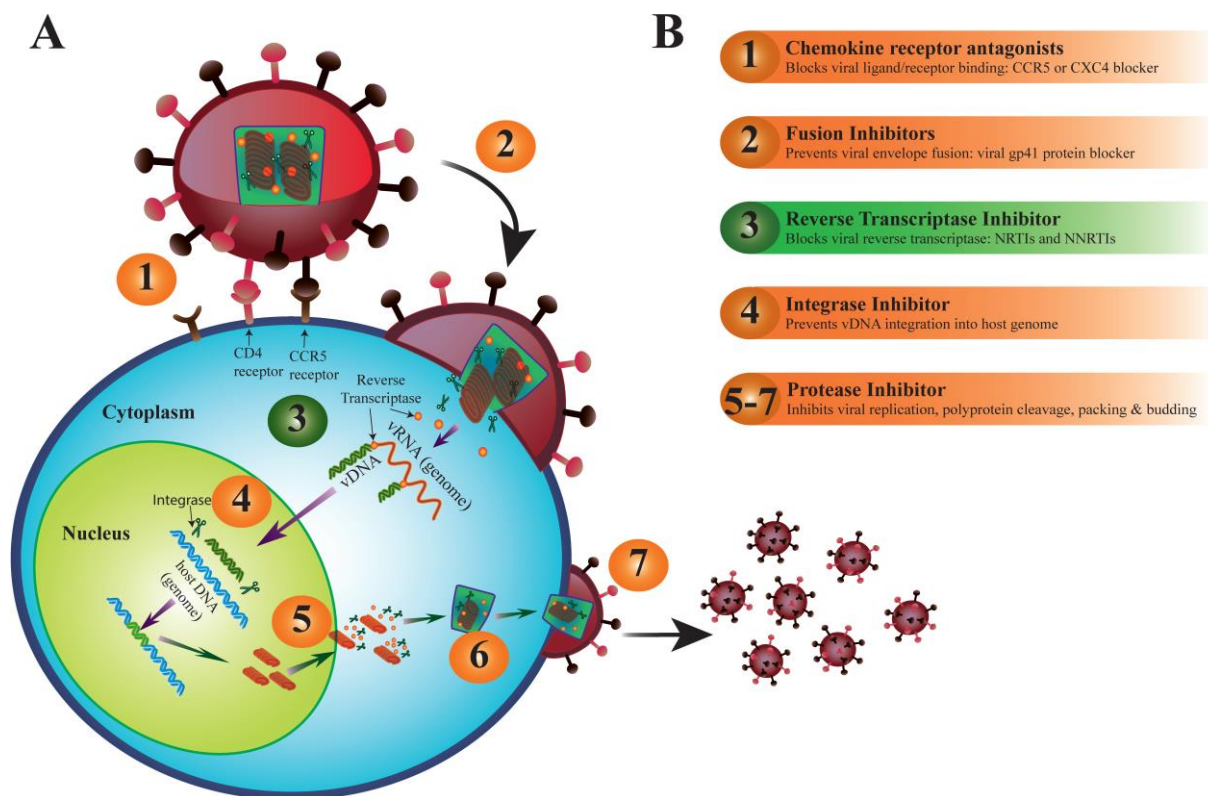
Figure 1. 5: (A) HIV-1 life cycle (B) Antiretroviral drug target. The numbers denote the Antiretroviral drug action site (Taken from Holec *et al*. (2017)).

1.9   HIV-1 Antiretroviral therapy (ART)

Antiretroviral therapy (ART) has evolved from the use of drugs given as monotherapy to standard patient management that involves the use of a combination of different antiretroviral agents (ARVs). Combination therapy, also known as highly active antiretroviral therapy (HAART), has been instrumental in the treatment of HIV-1 infection (Arts and Hazuda, 2012). HAART has helped change HIV infection from an illness that is progressive with a fatal outcome to a chronic disease that can be effectively managed (Maartens *et al*., 2014). The main targets of antiretroviral therapy are the enzymes

involved in the HIV-1 life cycle. Antiretroviral treatment drugs are currently classified into six (Table 1.1) based on their therapeutic mechanisms (Shukla and Chauhan, 2019).

The standard HAART regimens consist of nucleoside reverse transcriptase inhibitors (NRTIs), with a non-nucleoside reverse transcriptase inhibitor (NNRTI), a protease inhibitor (PI), or an integrase inhibitor (INI) (Maartens *et al*., 2014). In high-income countries, the management of HIV-infected individuals includes a starting treatment regimen of dual NRTIs, combined with either an NNRTI, ritonavir-boosted PI, or an INI, due to the similar efficacy as well as tolerability of these regimens (Thompson *et al*., 2012). The WHO recommends that in low and middle-income countries, the choice and use of antiretroviral therapy should be guided by a public health approach, and treatment should include a standardized first-line regimen that combines NNRTI plus dual NRTIs and then a second-line regimen of ritonavir-boosted PIs plus dual NRTIs (WHO, 2016).

Although the results of resistance testing should guide subsequent switches in the antiretroviral therapy regimen in virological failure cases. In low and middle-income countries, resistance testing and viral load monitoring are not often available, thus clinical and CD4 count monitoring is utilized to switch antiretroviral regimens. This may result in needless switches from a first-line to the second-line regimen, and the HIV-1 infected individuals continue to fail first-line treatment, leading to an increased chance of resistance mutations emerging (Sigaloff *et al*., 2011). The different classes of antiviral drugs and their mode of action are presented in Table 1.1 below:

Table 1. 1: Types of ARV drugs and their mode of action (Table modified from Shukla and Chauhan (2019))

Drug Classes	Examples	Mode of action
Receptor/co-receptor antagonist	• Aplaviroc • Maraviroc	Prevents HIV-1 from binding to the host cells receptors (gp120/CCR5/CXCR4)
Fusion Inhibitors (FIs)	• Enfuvirtide (Fuzeon)	Fusion inhibitors block the penetration of HIV-1 through the cell membrane of the host cell
Non-nucleoside reverse transcriptase inhibitor (NNRTIs)	• Delavirdine • Efavirenz • Etravirine • Nevirapine • Rilpivirine • Doravirine	Non-nucleoside reverse transcriptase inhibitors bind in a non-competitive way to HIV-1 reverse transcriptase enzyme and inhibit the conversion of viral RNA into DNA
Nucleoside reverse transcriptase inhibitor (NRTIs)	• Abacavir • Didanosine • Emtricitabine • Lamivudine • Stavudine • Tenofovir • Zidovudine	Nucleoside reverse transcriptase inhibitors directly block HIV-1 reverse transcriptase enzyme from converting viral RNA into DNA
Protease inhibitor (PIs)	• Atazanavir • Darunavir • Fosamprenavir • Indinavir • Lopinavir • Nelfinavir • Ritonavir • Saquinavir • Tipranavir	Protease inhibitors inhibit HIV-1 PR, thus preventing the development of nascent virions into matured ones.
Integrase strand transfer inhibitor (INSTIs)	• Raltegravir • Dolutegravir • Elvitegravir • Bictegravir • Cabotegravir	Integrase strand transfer inhibitors block the integration of viral DNA into the host cell genome

1.10 HIV-1 drug resistance

Resistance to ART is a threat to controlling the HIV-1 epidemic. HIV-1 drugs lead to the inability of antiviral agents to terminate viral replication, and it is associated with changes in the HIV genetic structure. HIV-1 drug resistance can be acquired or transmitted. Acquired drug resistance mutations emerge due to viral replication and the effect of selective pressure associated with treatment using ART (Yan *et al*., 2020). These drug resistance mutations can then be transmitted to ARV naïve people (Yan *et al*., 2020). The development of HIV drug resistance stems from nonadherence, poor tolerability of ART, and drug interactions between ART and other drugs. In addition, the continuous use of ART during virological failure makes the virus select mutations that confers on it the ability to resist the antiviral effect of ART. The rate at which drug resistance develops depends significantly on the selective advantage that the emerging mutations confer on the virus (Deeks *et al*., 2015).

Resistance to ART is associated with the process of viral RNA reverse transcription into viral DNA, facilitated by the reverse transcriptase, which is notoriously error-prone and introducing an average of one mutation per viral genome transcribed (Roberts *et al*., 1988). The high rate of mutation in HIV-1 occurring at each cycle of reverse transcription is associated with a diverse and complex mixture of viral quasispecies that differs by one or several mutations in HIV-1 infected individuals (Clavel and Hance, 2004). Resistance to HIV-1 drugs starts a vicious circle that results in increased cases of treatment failure. This makes the suppression of viral replication with currently available drugs impossible as the mutant HIV-1 variants continuously replicate and become predominant (Clavel and Hance, 2004). Monitoring of HIV-1 infected individuals on ART will help to lower the chances of drug resistance emerging and identify those at risk of treatment failure. Early identification of virological failure will help prevent the sustained use of a failing treatment, thus avoiding the accumulation of drug resistance mutations that may compromise the efficiency of subsequent ART after treatment switch (Godfrey *et al*., 2017). This study focuses on the impact of multidrug-resistant HIV-1 PR mutations on the switch from LPV and DRV and how well DRV will fare in the presence of these drug resistance mutations.

1.11 HIV-1 Protease

HIV-1 PR plays a crucial role in viral replication and maturation. HIV-1 PR cleaves Gag, Gag-Pol, as well as Nef polyproteins resulting in the development of active viral enzymes (reverse transcriptase, protease, and integrase), viral structural proteins (capsid and nucleocapsid) and other viral factors that aid the virus in replicating (Perez *et al*., 2010). The GagPol cleavage step is an ideal drug target, and its inhibition will terminate viral maturation (Fun *et al*., 2012). HIV-1 PR belongs to the family of aspartic proteases. It is a homodimeric molecule consisting of two subunits, each made up of 99 amino acid residues (Yang *et al*., 2012a). The HIV-1 PR is made up of mainly β-sheet, and the aspartic acid residues

(Asp-25) from both monomers form the central active site (Weber and Agniswamy, 2009). There are six structural segments that form HIV-1 PR (Figure 1.7); these segments are; the HIV-1 PR flap region, which is formed by amino acid residues 43–58/43'–58', the flap elbow made up of amino acid residues 35-42/35'-42', the fulcrum (residues 11–22/11'–22'), the cantilever (residues 59–75/59'–75'), the dimer interface (residues 1-5/1'-5', 95-99/95'-99'), and the catalytic site comprising residues 23–30/23'–30' (Perryman *et al*., 2004, Harte *et al*., 1990).
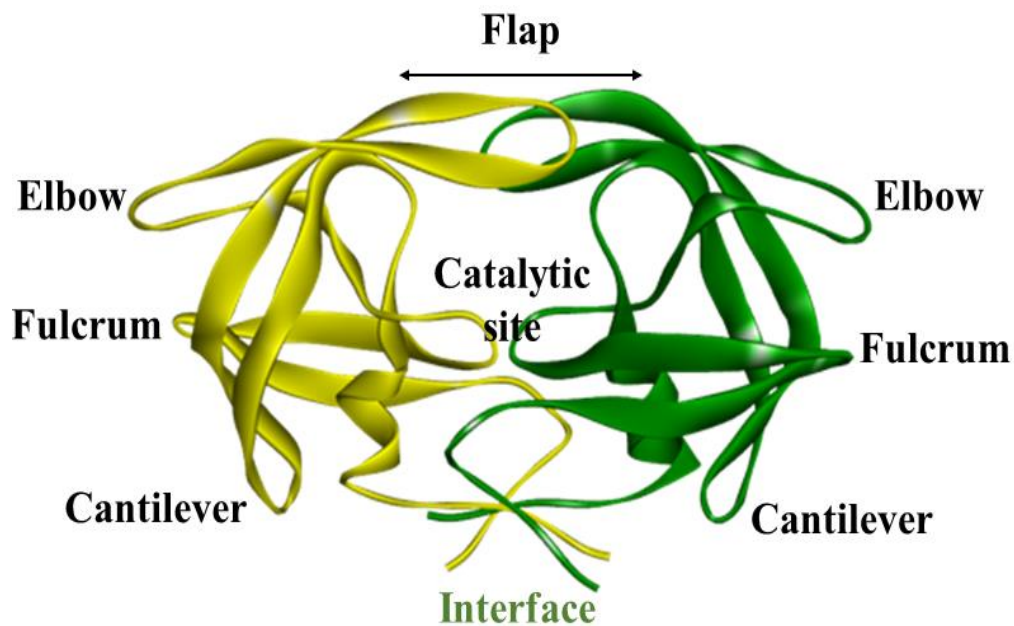
Figure 1. 6: HIV-1 PR structure showing the different structural segments

The HIV-1 PR dimer interface is highly stable, and it is formed by four anti-parallel β-strands (Yang *et al*., 2012a). The active site core is hydrophobic. The two aspartic acid residues in the active site help stabilize water addition across susceptible polypeptide amide links to create an intermediate tetrahedral transition state. This intermediate form of HIV-1 PR substrate is cleaved to produce a C-terminal carboxylic acid and N-terminal amine, followed by cleavage of the HIV-1 PR substrate (Martinez-Cajas and Wainberg, 2007). The active site is covered by the two flaps formed by flexible anti-parallel β-sheets from both monomers. Its flexibility is crucial in regulating the entry of both HIV-1 PR substrates and HIV-1 PIs into the catalytic site (Yu *et al*., 2017). The flaps of HIV-1 PR assume an open conformation in a free enzyme state but a closed conformation when a ligand is in the active site (Yang *et al*., 2012a).

1.11.1 HIV-1 Protease cleavage sites

HIV-1 PR is involved in several proteolytic reactions needed to develop a nascent virion into a fully matured and viable viral particle. This proteolytic cleave reactions are: five reactions in Gag (p17/p24, p24/p2, p2/NC, p7/p1 and p1/p6gag), six reactions in Gag-Pol (NC/TFP, TFP/p6pol, p6pol/PR, PR/RT,

RT/p66, and p66/IN) and one reaction in Nef. As described in figure 1.8, the different reactions occur at unique cleavage sites with a distinct amino acid composition (De Oliveira *et al*., 2003). The cleavage of Pr55gag polyprotein results in the development of the main HIV-1 structural proteins: the "matrix (MA, p17), capsid (CA, p24), nucleocapsid (NC, p7), p2, p1, and p6gag" (De Oliveira *et al*., 2003). HIV-1 viral enzymes are formed from the cleavage of Pr160gag-pol, and this Pr160gag-pol polyprotein contains several proteins like p17, p24, and p2. The cleavage of the C-terminal Pr160gag-pol results in the following products: NC, the trans frame protein (TFP), p6pol, PR, RT (RTp51), RT-Rnase H (RTp66), and IN (Ikuta *et al*., 2000, De Oliveira *et al*., 2003). The HIV-1 PR enzyme is encoded as part of the Gag-Pol polyprotein formed primarily by the embedded HIV-1 PR monomers. The embedded HIV-1 PR undertakes a series of self-cleavage, which results in the release of the matured HIV-1 PR (Davis *et al*., 2012).
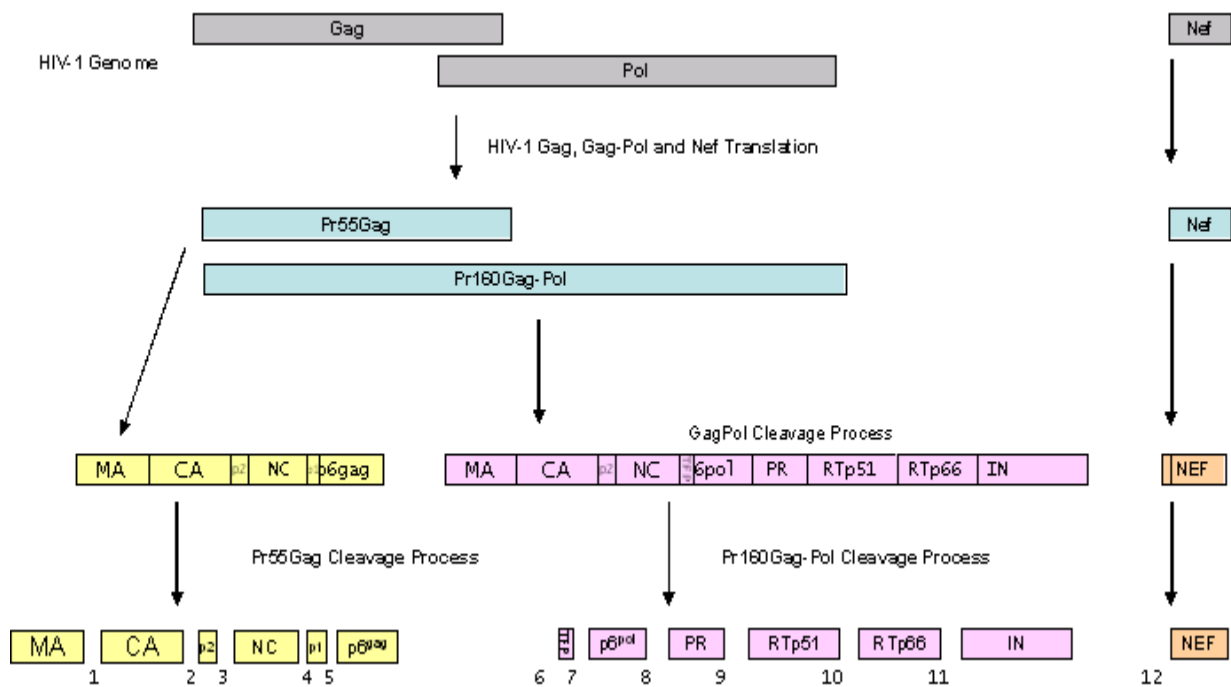
Figure 1. 7: Diagrammatic representation of the HIV-1 PR cleavage sites (Taken from De Oliveira *et al*. (2003))

1.11.2 Mechanism of HIV-1 substrate cleavage

The HIV-1 PR binds to its substrate in an extended conformation needed for adequate interaction with the active site residues (Trylska *et al*., 2007). The peptide bonds within the substrate are hydrolyzed by HIV-1 PR with high catalytic efficiency and sequence selectivity. HIV-1 PR peptide-substrate amino acid residues are labeled by standard nomenclature as P1 to Pn, P1' to Pn,' and the complementary binding sites on the protease enzyme are S1 to Sn, S1' to Sn' subsites (figure 1.9) (Brik *et al*., 2003). Using this standard nomenclature, S1 and S1' subsites are structurally identical (as well as S2 and S2').

The two S1 and S2 subsites are hydrophobic except for the active site residues; Asp-29, Asp-29', Asp-30, and Asp-30'. The S3 subsites next to the S1 subsites are also mostly hydrophobic (Brik *et al*., 2003).
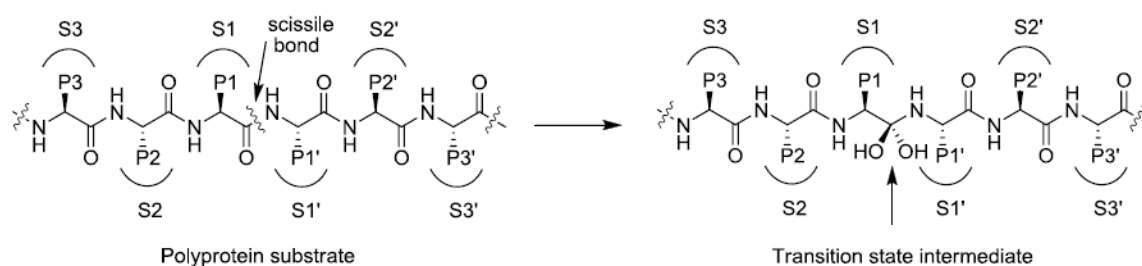
Figure 1. 8: Diagram showing HIV-1 PR substrate and complementary binding sites using standard nomenclature (adapted from Ali *et al*. (2010)).

Though the exact mechanism of HIV-1 PR substrate cleavage is not well understood, it is believed to involve the coordination of water molecules by the aspartic acid residues in the HIV-1 PR catalytic site (Anderson *et al*., 2009). In the catalytic site, the Asp25' residue exists in a deprotonated state and functions as a base where it receives a proton from a water molecule, forming a hydroxide ion that serves as a reactive nucleophile. The Asp25 residue functions as an acid; it donates a proton to form a tetrahedral intermediate (Guarna and Trabocchi, 2014). The termination of the catalytic cycle involves the donation of a proton to the peptide substrate amide group by Asp25, thereby attaining its primary configuration and resulting in the cleavage of a peptide bond. Consequently, the Asp25' residue receives a proton, returning to the original configuration (Guarna and Trabocchi, 2014). This reaction results in the formation of a transition state intermediate, which is used as the basis for the development of transition-state analog peptidomimetic HIV-1 PIs (Adamson, 2012).

1.11.3 HIV-1 Protease Inhibitors (PIs)

The United States Food and drug administration (FDA) has so far approved ten HIV PIs. These HIV-1 PIs include saquinavir (SQV), indinavir (IDV), ritonavir (RTV), nelfinavir (NFV), amprenavir (APV), fosamprenavir (FPV), lopinavir (LPV), atazanavir (ATV), tipranavir (TPV), and darunavir (DRV) (Lv *et al*., 2015). HIV-1 PIs can inhibit the function of HIV-1 PR to the point where the enzyme can no longer cleave Gag and Gag-Pol, leading to the production of immature viruses. The HIV-1 PIs provide a second class of drugs used in antiretroviral therapy, and this has made possible HAART combination therapy (Adamson, 2012). The co-administration of HIV-1 PIs with ritonavir has significantly enhanced the bioavailability and the half-life of the former, leading to a higher concentration of HIV-1 PI in the plasma (increased bioavailability). Ritonavir inhibits the cytochrome P450 3A4 isoenzyme, which plays a role in the metabolism of all HIV-1 PIs (Eagling *et al*., 1997).

HIV-1 PIs are tight-binding competitive inhibitors; they bind to the active site of HIV-1 PR. HIV-1 PIs mimic its substrate and are designed based on the mechanism of HIV-1 PR substrate cleavage (Ali *et al*., 2010). A characteristic feature of HIV-1 PIs is the secondary hydroxyl group, a replacement for the P1 carbonyl moiety found in the HIV-1 PR substrates. This P1 carbonyl moiety makes significant interactions with residues of the active site (Asp 25 and Asp 25′) through hydrogen bonds and ensures the tight binding of the inhibitor with HIV-1 PR (Lv *et al*., 2015, Ali *et al*., 2010). Notwithstanding the success recorded with the introduction and use of HIV-1 PIs, the development of drug-resistance mutations affects the efficacy of HIV-1 PIs in managing HIV-1 infected individuals (Lv *et al*., 2015).

1.11.4 HIV-1 Protease Inhibitor resistance and HIV-1 Protease Mutations

The genetic basis of drug resistance to available HIV-1 PIs is associated with the production of genetically diverse viral quasispecies subject to evolution. Their distinct genetic flexibility allows the population to respond differently to selection pressures (Wensing *et al*., 2019). HIV-1 PI drug selection pressure results in several mutations being accumulated within the viral PR conferring resistance to HIV-1 PIs by diminishing the capacity of the latter to bind to the active site of HIV-1 PR. During this process, a greater percentage of residues within HIV-1 PR mutate in different combinations resulting in HIV-1 PI drug resistance (Ali *et al*., 2010). The complex and interdependent combination of several in the HIV-1 PR gene results in highly mutated HIV-1 PR variants. These mutations act synergistically to confer resistance to HIV-1 PIs while preserving viral fitness (Adamson, 2012). HIV-1 PR mutations can occur either at the enzyme active site or at distant sites outside the active site (Figure 1.10 and Figure 1.11). These mutations, whether in the active site or outside the active side, impact the recognition and binding of HIV-1 PIs (Ali *et al*., 2010).

1.11.4.1 HIV-1 PR active site mutations

Mutations in the HIV-1 PR active site directly decrease its affinity for HIV-1 PIs. These mutations are also called primary or major mutations, Eg: D30N, V32I, L33F, M46I/L, I47A/V, G48V, I50L/V, V82A/F/L/S/T, and I84A/V (Johnson *et al*., 2011). Several primary HIV-1 PI resistance mutations have been characterized. Some of these mutations confer resistance to one HIV-1 PI, while others cause resistance to two or more HIV-1 PIs (figure 1.11). For instance, the mutation D30N confers resistance to HIV-1 PR against NFV therapy, while mutation occurring at amino acid residue 82 confers resistance on HIV-1 PR to RTV and SQV, and the mutation G48V leads to resistance against SQV and ATV. Some primary mutations may cause severe and cross-resistance to most HIV-1 PIs, like the I84V mutation (Figure 1.11) (Ali *et al*., 2010, Wensing *et al*., 2019).

The mutation of just a single residue in the HIV-1 PR active site is sufficient to cause loss of interactions between HIV-1 PR with the inhibitor, and this could be through the loss of some van der Waal contacts,

as seen in the case of V82A mutation (Weber and Agniswamy, 2009). In some cases, the combination of mutations may result in total loss of all van der Waal contacts between the HIV-1 PIs and the active site residues of HIV-1 PR. An example is the combination of D25N and V82A (Prabu-Jeyabalan *et al*., 2003). Primary HIV-1 PR mutations also alter HIV-1 PI recognition through the loss of hydrogen bonds (Weber and Agniswamy, 2009). Asides from loss of interaction with residues in the active site, primary mutations may also affect HIV-1 PR flap flexibility or cause rearrangement of the HIV-1 PR backbone, resulting in decreased affinity for HIV-1 PIs (Ali *et al*., 2010).
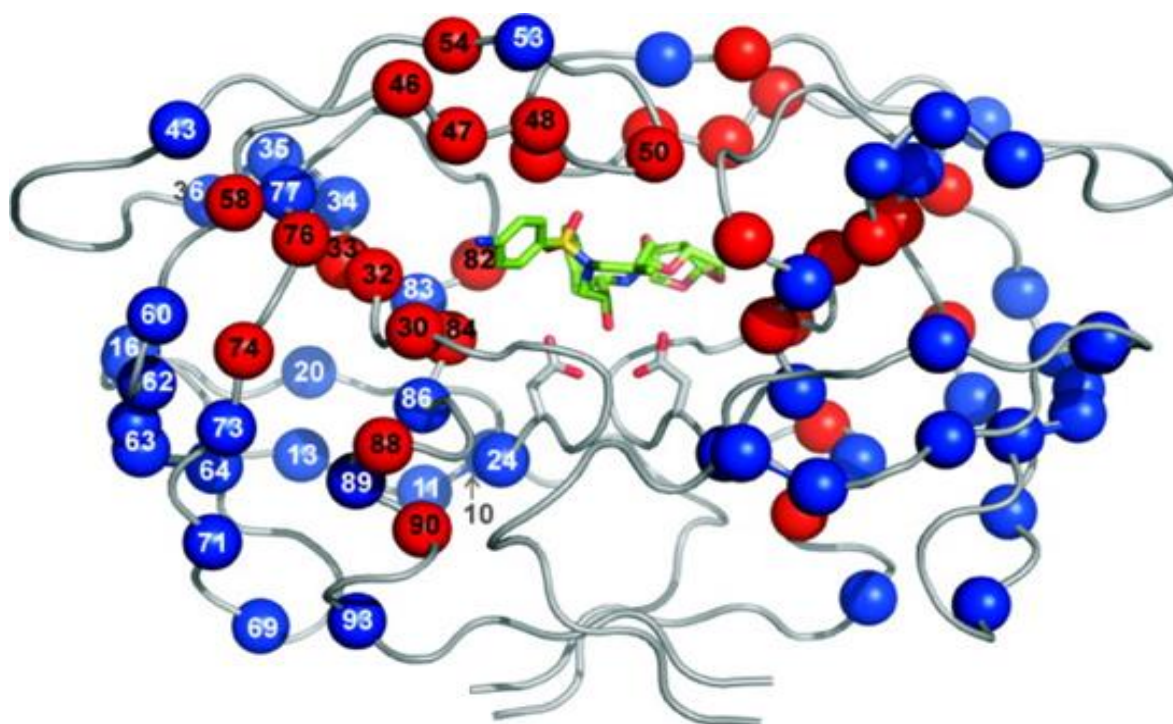
Figure 1. 9: Diagrammatic representation of HIV-1 PR dimer three-dimensional structure. The major (primary) mutations are represented in red balls, while minor (secondary) mutations are represented in the blue balls (Taken from Wensing *et al*. (2010).

1.11.4.2 HIV-1 PR non-active site mutations

Non-active site mutations do not make contact with the inhibitors; they can still cause resistance to HIV-1 PIs (Weber and Agniswamy, 2009). Though the mechanism of how mutations outside the active site cause resistance to HIV-1 PIs is unknown, these mutations may contribute to the development of resistance by inducing conformational changes in the HIV-1 PR structure, thus altering the binding landscape of HIV-1 PIs (Ali *et al*., 2010). Some of the non-active site mutations are minor mutations but still have a negative impact on the efficacy of HIV-1 PIs, like the L24I mutation, which has been shown to induce resistance to several HIV-1 PIs by causing an alteration in HIV-1 PR dimer stability (Weber and Agniswamy, 2009). Similarly, it has been shown that non-active site mutations may

indirectly cause resistance by impacting the hydrogen bond network between HIV-1 PIs and the active site (Bastys *et al*., 2020).

Non-active site mutations such as the N88S and L76V have been shown to induce resistance to currently used HIV-1 PIs and initiate sensitivity to other HIV-1 PIs (Bastys *et al*., 2020). The Flap mutation I54M that is selected as a major resistant mutation in HIV-1 infected individuals undergoing treatment with darunavir, though it makes no contact with HIV-1 PIs (Liu *et al*., 2008), causes changes in the amino acid residues 80-82 (the 80's loop), which makes contact with inhibitors. Another example is the mutation of residue 90 of HIV-1 PR found in the hydrophobic pocket next to the catalytic residues (Weber and Agniswamy, 2009).
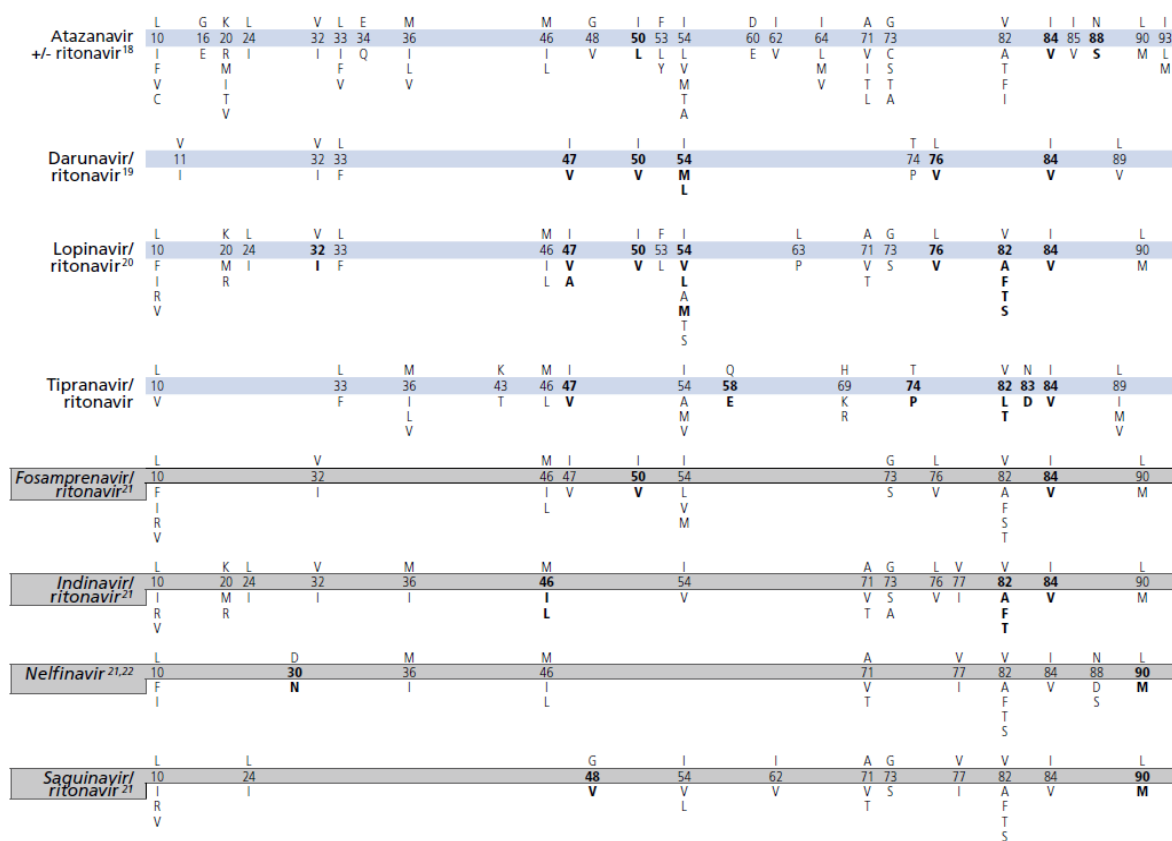
Figure 1. 10: Diagram showing the mutations in the HIV-1 PR gene associated with resistance to HIV-1 PIs (Taken from Wensing *et al*., (2019)).

1.11.5 Polymorphisms in non-B HIV-1 PR

Polymorphisms in HIV-1 proteins, including HIV-1 PR, emerges as a result of genomic variability. Naturally occurring amino acid polymorphisms are found in the HIV-1 PR from non-B subtypes (A, C, and G) that dominate the African HIV epidemic (Table 1.2) (Velázquez-Campoy *et al*., 2003). These polymorphisms are located outside the active site, where they do not influence catalytic activity or substrate binding. However, in the G subtype, some of these amino acid polymorphisms occur at

positions which cause resistance to HIV-1 PIs in the HIV-1 subtype B (Descamps *et al.*, 1998). The isoleucine occurring naturally at position 36 (M36I) in non-B African subtypes HIV-1 PR is a minor mutation in HIV-1 subtype B PR (Gulnik *et al.*, 1995). This polymorphism in the presence of major mutations can cause resistance to HIV-1 PIs. Likewise, the leucine polymorphism at position 93 (I93L) found in the South African HIV-1 subtype C PR and the isoleucine polymorphism at position 20 (K20I) in the HIV-1 subtype G PR is associated with HIV-1 PI resistance and classified as minor drug resistance mutations (Velázquez-Campoy *et al.*, 2003).

Table 1. 2: Naturally occurring amino acid polymorphisms found in the HIV-1 PR from subtypes that dominate the African HIV epidemic

Subtype	Polymorphisms
Subtype A	I13V, E35D, M36I, S37N, R41K, H69K, L89M
Subtype C	M36I, S37N, R41K, H69K, L89M
C-SA PR	T12S, I15V, L19I, M36I, S37N, R41K, H69K, L89M, I93L
Subtype G	I13V, K20I, E35D, M36I, S37N, R41K, C67S, H69K, V82I, L89M

Polymorphisms in non-B HIV-1 PR have been shown to affect its structural stability, and this variation in structural stability impacts the HIV-1 PI binding landscape (Velazquez-Campoy *et al.*, 2002). The South African HIV-1 subtype C PR has four of the polymorphisms (I15V, M36I, and L89M I93L) located in its hydrophobic core which have been shown to impact the interaction of HIV-1 PIs (Naicker *et al.*, 2014). HIV-1 PR polymorphisms have been shown to act synergistically with multidrug-resistant mutations to impact the kinetics of interaction of HIV-1 PIs (Kantor and Katzenstein, 2003, Velázquez-Campoy *et al.*, 2003). Therefore, this study is vital as it biochemically characterized highly mutated South African HIV-1 subtype C PR and its interaction with LPV and DRV, which has not been shown in any study.

1.11.6 Characterization of HIV-1 PR using enzyme kinetics and inhibition

The enzyme kinetic characterization of wild-type and mutated HIV-1 PR will help to determine at the molecular level the degree of interactions with both substrate and HIV-1 PIs invitro. The determined kinetic constants provide an invaluable source of information to assess the fitness of the HIV-1 PR variants (Shuman *et al.*, 2004). Enzyme kinetic characterization helps to gain insight into how mutated HIV-1 PR maintains its affinity for its substrate with moderate catalytic activity even when the mutation seems to influence its biochemical fitness (Gulnik *et al.*, 1995, Ohtaka *et al.*, 2003). How multidrug resistance mutations affect the inhibitory effect and interaction of LPV and DRV with HIV-1 PR was probed using enzyme Kinetics in this current study (measuring the rate of hydrolysis of the chromogenic

substrate) and intrinsic tryptophan fluorescence of HIV-1 PR. These techniques are briefly explained below, but a more detailed elaboration can be found in the subsequent chapters.

1.11.6.1 Enzyme Kinetics

The use of enzyme kinetics to understand enzyme activity and function is invaluable in drug development and understanding the evolution of resistance to a drug. It helps to understand the specific mechanism of action of an enzyme and quantitatively assesses an inhibitor's performance against its target enzyme. (Samuele *et al*., 2013). Enzyme kinetics experiments involve using a chromogenic substrates that produce a colored product, radiometric assays that involve the incorporation or release of radioactivity, mass spectrometry, and others (Rogers and Gibon, 2009, Eisenthal and Danson, 2002). During the kinetic interaction, the enzyme forms an intermediate complex with the substrate (ES), which is subsequently followed by the release of the colored product (P) (Rogers and Gibon, 2009). This process can be illustrated using the Michaelis-Menten kinetics equation below:

$$E + S \rightleftarrows ES \rightarrow E + P$$

In enzyme kinetics, the number of moles of a product formed per second is referred to as the rate of enzyme catalysis (V). An enzyme rate of catalysis (V) increases linearly as the concentration of its substrate ([S]) increases (Figure 1.12) until it reaches a point where it starts leveling off and approaches a maximum at a high concentration of substrate (Rogers and Gibon, 2009). The equation (Michaelis and Menten equation) relating the velocity of an enzyme-catalyzed reaction and the concentration of its substrate is shown below:

$$V_0 = \frac{V_{max}[S]}{K_m + [S]}$$

In the equation above, V_0 is the rate at which an enzyme-catalyzed reaction progresses at a given substrate concentration; it is also known as the initial velocity. The V_{max} is the maximum speed an enzyme-catalyzed reaction attains at saturating substrate concentration, and the Michaelis–Menten constant is denoted as K_m. The K_m value is a reflection of the affinity of an enzyme for a substrate. The higher an enzyme's affinity for a substrate, the lower the K_m for that substrate (figure 1.12 a). The K_m is equal to the substrate concentration when the reaction rate is half of the maximum reaction rate. The Michaelis–Menten equation above has been transformed to obtain the Lineweaver- Burk equation. Graphical representation of this equation gives a straight line (plotted as $1/V_0$ versus $1/[S]$), and the intercept on the x-axis is $-1/K_m$, the intercept on the y-axis is $1/V_{max}$, and the slope is K_m/V_{max} (figure 1.12 b). The Lineweaver–Burk plot has become a useful tool for calculating enzyme kinetic parameters (K_m and V_{max}) (Talens-Perales *et al*., 2016).
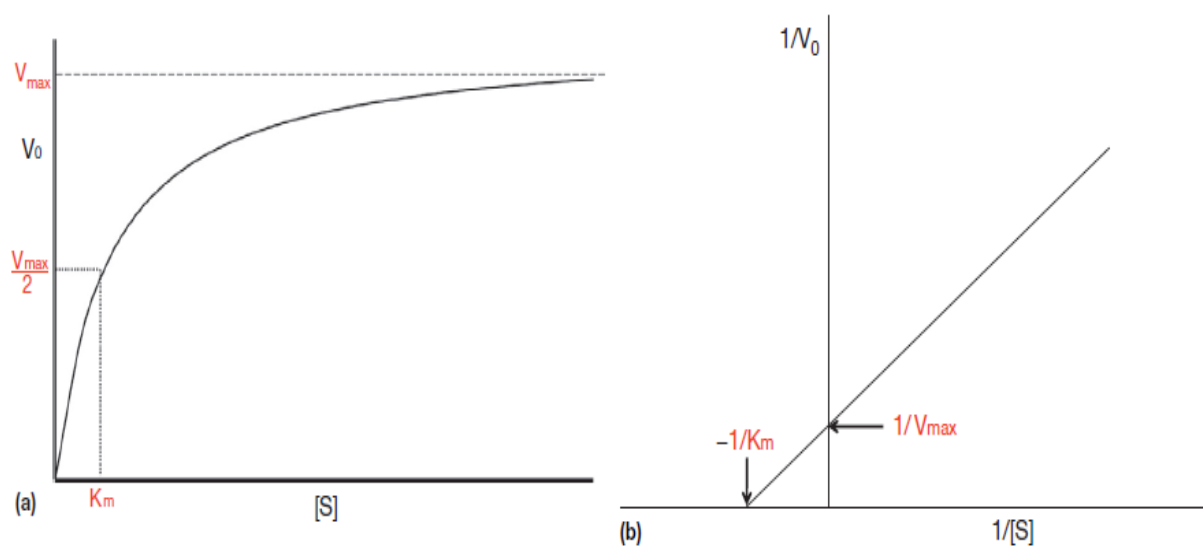
Figure 1. 11: (a) Graphical representation of Michaelis–Menten kinetics showing the velocity (V) of enzyme-catalyzed reaction with a varying substrate concentration ([S]). (b) Lineweaver–Burk/ double reciprocal plot of Michaelis–Menten kinetics (Taken from Talens-Perales *et al*. (2016)).

1.11.6.2 Enzyme inhibition

Enzyme inhibition reduces the activity of an enzyme through the binding of an inhibitor to the active site, a regulatory site on the target enzyme, or the enzyme-substrate complex in the case of uncompetitive inhibition (Rogers and Gibon, 2009). Enzyme inhibitors are classified into two classes: irreversible and reversible inhibitors. The irreversible inhibitors act by binding covalently to the target enzymes, inhibiting the enzyme's activity permanently. On the other hand, reversible inhibitors bind noncovalently to an enzyme resulting in a temporary decrease in the enzyme activity (Roskoski, 2007). This study focuses on reversible FDA-approved inhibitors: LPV and DRV, which are commonly used in the formulation of salvage regimens. Reversible inhibitors act through three basic mechanisms and are classified based on these mechanisms of action. They could either be competitive, non-competitive, and uncompetitive inhibitors. In some cases, the reactions are more complex and may be a combination of all these mechanisms resulting in mixed inhibition (Roskoski, 2015).

The type of reversible inhibition can be identified from the Lineweaver-Burk plot from Michaelis-Menten equations. Plotting I/V against the varying inhibitor concentration ([I]), when the substrate concentration ([S]) is kept constant, gives a straight line. When this is done at two different concentrations of substrate concentrations [S1] and [S2], the two straight lines intersect at a point, as shown in figure 1.13. Where these lines intersect can be read directly as the K_i. The straight lines each represents the reciprocal form of the equation below (Roskoski, 2007):

$$1/V = 1/[S1]\ (1 + [I]/K_i) = 1/[S2]\ (1 + [I]/K_i)$$

1.11.6.2.1 Competitive Inhibition

A competitive inhibitor is generally a close analog of the original substrate of the enzyme. It competes with the substrate for binding at the enzyme's active site though it does not undergo catalysis. This causes a decrease in the enzyme's catalytic efficiency as the substrate cannot bind to the enzyme-inhibitor (EI) complex. The reaction scheme of competitive inhibitor is shown below (Roskoski, 2007);

$$\text{E + S} \underset{K_{-1}}{\overset{K_1}{\rightleftharpoons}} \text{ES} \xrightarrow{K_{cat}} \text{E + P}$$

$$+$$

$$I$$

$$K_{-2} \updownarrow k_2$$

$$\text{EI + S}$$
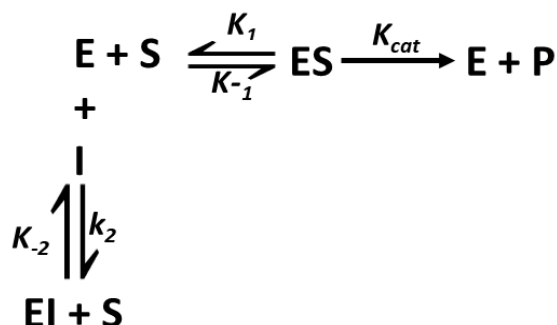
The rate equation of competitive inhibition is given in Eq. (A) below (Roskoski, 2007).

$$V = \frac{V_{max}\,[S]}{K_m\left(1 + \dfrac{[I]}{K_i}\right) + [S]}$$

Eq. (A)

As can be observed, as the concentration of the competitive inhibitor increases, the apparent K_m of the enzyme increases alongside but has no impact on the V_{max}. The Lineweaver–Burke plot of competitive inhibition is shown in figure 1.13A.

1.11.6.2.2 Non-competitive inhibition

Non-competitive inhibitors do not compete with the substrate for binding to the enzyme's active site. These inhibitors bind to the enzyme at different binding pockets leading to a decrease in the catalytic efficiency of the enzyme. Both the substrate and the inhibitor bind to the enzyme independently. The non-competitive inhibition reaction scheme and the equation is presented below (Roskoski, 2007)

$$E + S \underset{K_{-1}}{\overset{K_1}{\rightleftharpoons}} ES \xrightarrow{K_{cat}} E + P$$

(Scheme with non-competitive inhibition showing E+S, EI+S, ES, EIS with rate constants K_1, K_{-1}, K_{cat}, K_{-2}, k_2, K_3, K_{-3}, k_4, K_{-4})
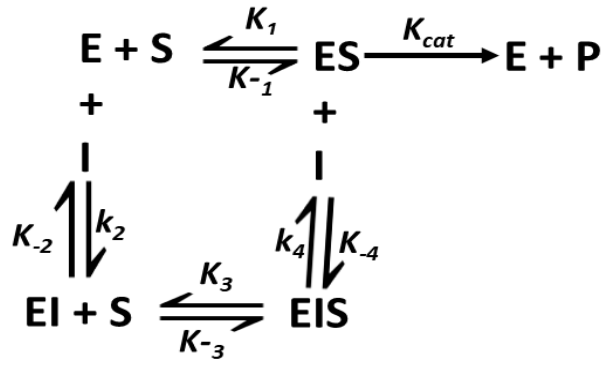
The rate equation of non-competitive inhibition is given in Eq. (B). As the inhibitor binds to the enzyme, the K_m is unaffected, but the K_{cat} decreases. The Lineweaver–Burke plot of non-competitive inhibition is shown in figure 1.13B.

$$V = \frac{K'_{cat}[E]_t\,[S]}{K_m + [S]} \qquad \text{Eq. (B)}$$

1.11.6.2.3 Uncompetitive Inhibition

Uncompetitive inhibitors do not bind directly to the enzyme; they bind to only the enzyme-substrate complex (figure 1.13C). The separation of the interaction of the enzyme-substrate complex by the inhibitor will lead to a decrease in the catalytic efficiency of the enzyme. The scheme is presented below:

$$E + S \underset{K_{-1}}{\overset{K_1}{\rightleftharpoons}} ES \xrightarrow{K_{cat}} E + P$$

(Scheme showing ESI formed from ES with rate constants k_2, K_{-2})
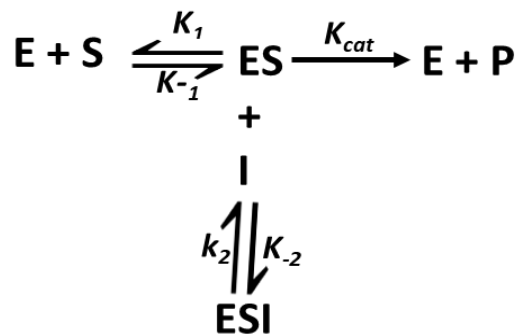
The binding of uncompetitive inhibitors cannot be affected by increasing the substrate concentration; as a result, the apparent V_{max} is reduced, and the apparent Km increases (Roskoski, 2007). The rate equation of uncompetitive inhibition (Eq. (C)) is presented below:

$$V = \frac{V_{max}\,[S]}{K_m + [S]\left(1 + \dfrac{[I]}{K_i}\right)} \qquad \text{Eq. (C)}$$
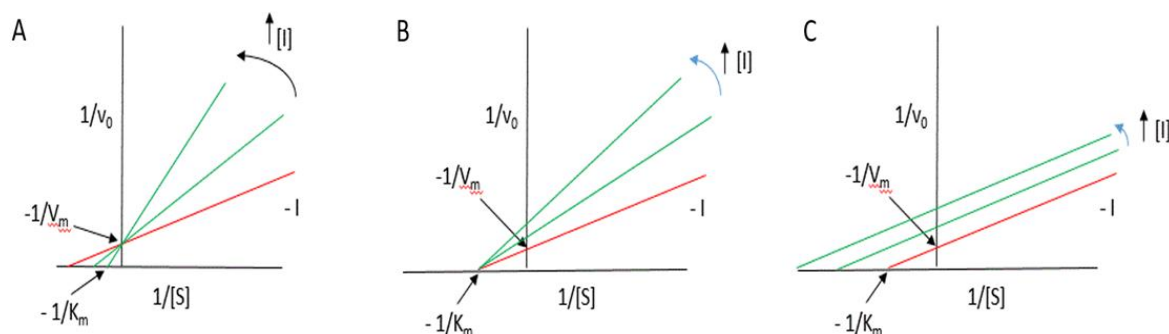
Figure 1. 12: (A) Graphical representation of competitive inhibition. (B) Graphical representation of non-competitive inhibition. (C) Graphical representation of uncompetitive inhibition (Taken from Jakubowski (2015)).

1.11.7 Fluoresce spectroscopy and fluoresce quenching

The elucidation of changes in the three-dimensional structure of proteins and determination of ligand binding affinity utilizing intrinsic tryptophan (Trp) fluorescence of proteins is based on the sensitivity Trp to changes in its local environment. The utilization of intrinsic Trp fluorescence capacity as a fluorescent probe helps to avoid the laborious process of labeling proteins with fluorescent tags (Ghisaidoobe and Chung, 2014). Tryptophan can be excited at around 280nm wavelengths, and the emission spectra recorded across a wide range of wavelengths to obtain the emission maxima. Analysis of the obtained fluorescence spectra gives a clue about changes in the three-dimensional structures of proteins. The inhibitor-induced fluorescence quenching can be used to determine the binding affinity of proteins for ligands (Ghisaidoobe and Chung, 2014).

 HIV-1 PR conformational dynamics have been accessed using the intrinsic tryptophan fluorescence. Two tryptophan residues are found at positions 6 and 42 in each HIV-1 PR monomer. Their position in the enzyme makes them suitable probes to monitor the HIV-1 PR tertiary structure (Szeltner and Polgár, 1996, Dash and Rao, 2001, Ullrich *et al*., 2000). Fluorescence quenching of HIV-1 PR utilizing intrinsic tryptophan fluorescence is based on the assumption that when inhibitors bind to the active site, they make contact with the HIV-1 PR flaps inner surface. Thus the flap movement will influence the intrinsic fluorescence capacity of Trp-42, which is close to the flap (Dash and Rao, 2001). Also, the binding of inhibitors to the active site constraints HIV-1 PR conformational dynamics around the Trp-42 residue, making it an excellent probe to investigate HIV-1 conformational dynamics (Fidy *et al*., 2001, Ullrich *et al*., 2000). In this study, fluorescence spectroscopy was used to explore the conformational changes associated with HIV-1 PR interaction with LPV and DRV.

1.11.8 Structural characterization using Molecular Dynamic (MD) Simulations

Understanding the relationship between the physical properties and functions of biological macromolecules has been made possible with Molecular dynamics simulations (MD) (Karplus and McCammon, 2002). It has helped to predict and monitor the movement of atoms over time in molecular systems. Molecular dynamic simulation shows that protein function is a product of the internal motions and conformational changes in proteins (Karplus and McCammon, 2002). In addition to highlighting the impact of conformational changes on protein function, MD can also be used to study critical biomolecular processes like ligand binding, protein folding, protein stability, and examination of the actual dynamics of the molecule (Hollingsworth and Dror, 2018, Karplus and McCammon, 2002). The MD setup consists of the input molecular structures, the force fields, and topologies. Several online, as well as offline tools, are readily available for MD simulations. This includes Groningen Machine for Chemical Simulations (GROMACS), Assisted Model Building with Energy Refinement (AMBER), Chemistry of Harvard Macromolecular Mechanics (CHARMM), amongst others (Gajula *et al*., 2016).

The combination of MD with experimental laboratory techniques is often employed to study protein structures. These techniques include X-ray crystallography, nuclear magnetic resonance (NMR), and fluorescence spectroscopy (Hollingsworth and Dror, 2018). Laboratory-based protein structure determination using the techniques mentioned above may present the natural form of proteins and their interaction with ligands. However, as the binding of ligands to their target is accompanied by macromolecular motions occurring very fast in millionths of a second, standard laboratory-based techniques may not give a good understanding of these processes of drug binding (Durrant and McCammon, 2011). Thus, MD effectively fills this gap where laboratory-based means of protein structure analysis are deficient (Durrant and McCammon, 2011). In addition to measuring the dynamic motions of ligand-receptor interaction, MD helps to fully understand the overall flexibility of the ligand−protein interaction system as this is vital for accurately predicting ligand binding, kinetic and thermodynamic properties of proteins (De Vivo *et al*., 2016). In this study, MD was used to gain insight into the impact of multidrug-resistant HIV-1 PR mutations on the structural dynamics of South African HIV-1 subtype C PR and to determine the influence of these mutations on LPV and DRV binding landscape.

1.12 Study significance

Research has shown that the number of HIV-1 infected individuals who have developed resistance to and failing second-line HIV-1 PI-based regimen in South Africa is increasing (Steegen *et al*., 2016). Studies assessing drug resistance profiles in South Africans infected with HIV have found at least one major PI mutation in about 16.4 – 18% of the study participants (Steegen *et al*., 2016, Obasa *et al*., 2020). The emergence of these mutations compromises the clinical management of HIV-1 infected

individuals in the face of resource constraints in Sub-Saharan Africa and limited ARV options. These mutations may also result in cross-resistance to other ARVs (Lopes *et al*., 2015). Though the impact of HIV-1 PI resistance mutations on virologic failure in people in South Africa infected with HIV-1 subtype C has been established, there is a lack of information on the biochemical characterization of highly mutated South African HIV-1 subtype C PR and its inhibition with LPV and DRV. It is essential to characterize highly mutated South African HIV-1 subtype C PR as the signature differences between it and the HIV-1 subtype B may impact the inhibitory effects of HIV-1 PIs (Coman *et al*., 2007).

For this reason, the wild-type and mutant South African HIV-1 subtype C PRs were cloned and expressed to characterize their enzymatic behavior and response to two HIV-1 PIs: LPV, commonly used in formulating second-line regimen in South Africa, and DRV, which is used as salvage therapy in HIV-1 infected individuals failing an LPV inclusive regimen. The changes in the HIV-1 PR structure associated with mutations that emerged from drug pressure due to the exposure to an LPV-inclusive regimen were assessed and how these may impact the LPV to DRV binding landscapes. This study also shows how the changes in the conformation of HIV-1 PR affect the microkinetic constants that determine the inhibitory capacity of LPV and DRV.

1.13   Project aims and objectives

1.13.1   Aims

I.   Biochemical characterization of highly mutated HIV-1 subtype C protease
II.   Evaluate the changes in the conformation of HIV-1 protease on inhibitor binding and the evolution of drug resistance

1.13.2   Study objectives

I.   Determine the biochemical fitness and inhibitory constants of HIV protease inhibitors for drug-resistant HIV-1 subtype C proteases.
II.   Evaluate the mechanistic interaction of HIV-1 protease inhibitors with HIV-1 subtype C protease.
III.   Determine the impact of mutations on the conformation and stability of HIV-1 subtype C protease, and how these changes affect the evolution of drug resistance using fluorescence assay and structural dynamics.

1.14   Thesis outline
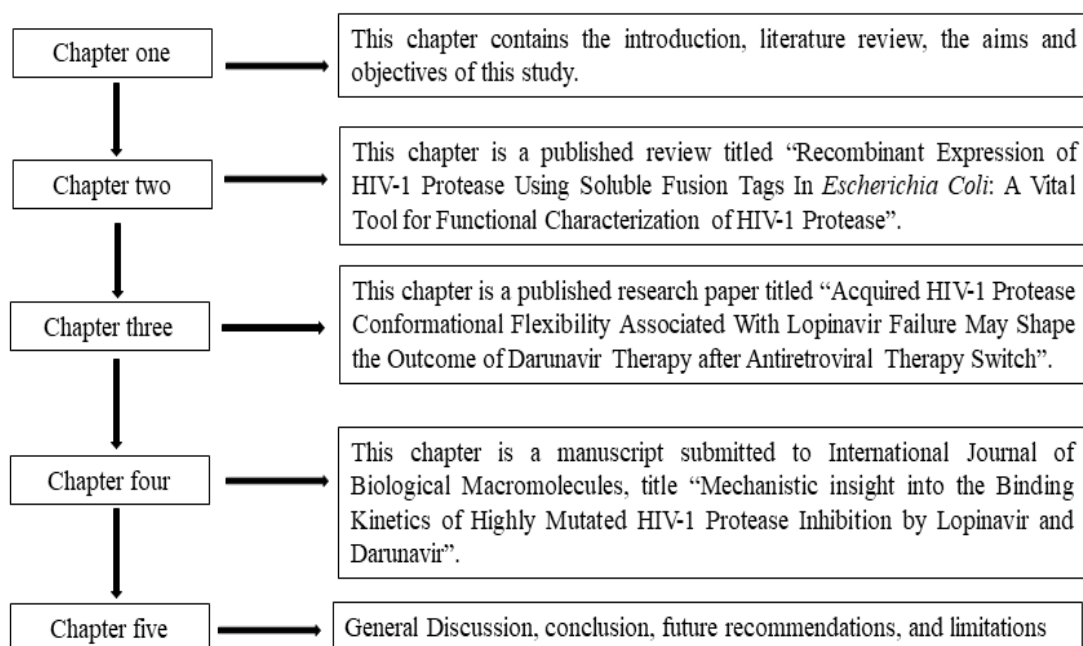
Figure 1.13: Thesis outline

Bridging between chapter one and chapter two

Chapter one gives a general overview of the HIV-1 PR structure, life cycle, and clinical management with the aid of antiretroviral agents with different targets such as the HIV-1 PR gene. HIV-1 PR is an important therapeutic target. Hence the functional characterization of HIV-1 PR will provide crucial knowledge for the development of novel PIs. Chapter two reviewed and summarized the characteristic properties of different fusion tags for easy expression of soluble and pure HIV-1 PR to assist researchers with their choice of protein fusion tag. This chapter is published in the journal *Virus Research* under closed access (the cover page is attached in the appendix).

CHAPTER TWO

Recombinant expression of HIV-1 protease using Soluble fusion tags in *Escherichia Coli*: A vital tool for functional characterization of HIV-1 protease

Abstract

HIV-1 protease expression in the laboratory is demanding because of its high cytotoxicity, making it difficult to express in bacterial expression systems such as *Escherichia coli*. To overcome these challenges, HIV-1 protease fusion with solubility enhancing tags helps to mitigate its cytotoxic effect and drive its expression as a soluble protein. Therefore, this review focuses on the expression of bioactive HIV-1 protease using solubility-enhancing fusion tags in *Escherichia coli* and summarises the characteristic features of the different common fusion tags that have been used in the expression of HIV-1 protease. This review will assist researchers with their choice of protein fusion tag for HIV-1 protease expression.

Keywords: HIV-1 Protease, *Escherichia coli*, soluble Fusion tags

1. Introduction

HIV-1 protease (PR) plays a crucial role in viral maturation. It is a retroviral aspartic PR and cleaves Gag, Gag-Pol, and Nef polyprotein precursors during the virion assembly phase (Perez *et al*., 2010). These cleavages result in the production of active viral enzymes and structural proteins needed for viral replication (Perez *et al*., 2010). The Gag-Pol polyprotein processing is a critical step in the viral lifecycle and therefore makes it an ideal drug target as inhibition of this step will prevent viral maturation (Coman *et al*., 2007). The introduction of protease inhibitors (PIs) as part of highly active antiretroviral therapy (HAART) has helped to reduce the morbidity and mortality associated with HIV-1 infection (Thompson *et al*., 2010). Nevertheless, the development of drug resistance mutations to HIV-1 PIs has jeopardized its effectiveness in HIV/AIDS management (Weber and Agniswamy, 2009). Information obtained from the structural analysis of HIV-1 proteases with mutations in and outside the active site can assist with the characterization of HIV-1 PR and the design of novel PIs (Ghosh *et al*., 2018).

A prerequisite to obtaining protein structures of high quality in the laboratory is producing a functional target protein at the right concentration and purity (Hewitt *et al*., 2011). In the past, HIV-1 PR was synthesized chemically (Nutt *et al*., 1988). However, with the advancement in technology, there has been a switch to the use of recombinant technology to express HIV-1 PR (Hui *et al*., 1993). Different means of expression have been used for the recombinant production of HIV-1 PR, including expression from large precursors (Karacostas *et al*., 1993), expression in different fusion forms (Wan *et al*., 1995), and recovery by refolding of HIV-1 PR from inclusion bodies accumulating in host expression cells (Cheng *et al*., 1990).

A major drawback of most of these methods is the accumulation of HIV-1 PR in insoluble aggregates since the chemicals used in the recovery of HIV-1 PR from the insoluble protein aggregates may compromise its structure and enzymatic function. Furthermore, the methods used to purify the recovered protein from inclusion bodies may be tedious (Hwang *et al*., 2014, Karacostas *et al*., 1993). In addition, the protein refolding process from inclusion bodies is not standard, often needing to be optimized in a manner that is mainly trial-and-error (Idicula-Thomas and Balaji, 2007). The renaturation of aggregated proteins to their natural conformation involves several steps that require large quantities of reagents; this may result in dilution of the protein, affecting the yield (Strandberg *et al*., 1991). Another critical point to note with the recovery of proteins from inclusion bodies is that several proteins do not readily assume their native conformation after chemical denaturation (Thomas and Baneyx, 1996). Therefore, it is more suitable to adopt an *in vivo* technique that will produce the protein of choice in a soluble, pure and active conformation to overcome the challenges of *in vitro* systems (Costa *et al*., 2014).

The cost-effective production of recombinant proteins in *E. coli* provides a means of high-level expression and scalable production of recombinant proteins for academic and pharmaceutical uses (Gupta and Shukla, 2016). Knowledge interchange between academia and industry has led to the development of improved techniques for the expression of HIV-1 PR, thus enabling the characterization of the HIV-1 PR and the development of new PIs (Huang and Chen, 2013). This review thus focuses on the use of fusion tags for the expression and production of bioactive recombinant HIV-1 PR. Studies have shown that when solubility-enhancing fusion tags are introduced in the expression process, HIV-1 PR can be expressed in a highly soluble and easily purified form suitable for biochemical and structural characterization (Azarnezhad *et al*., 2016, Volontè *et al*., 2011). This review also summarizes the characteristic properties of the different common fusion tags that have been employed in the expression of HIV-1 PR to assist researchers with their choice of protein fusion tag.

2. Laboratory production of HIV-1 Protease proteins

Laboratory production of proteins and peptides can be achieved through one of the following means: chemical synthesis; isolation of the target protein from its natural source; the use of recombinant DNA technology; or a combination of all these techniques (Jaradat, 2018). Chemical synthesis of proteins is a technique used to produce proteins *in vitro* from a known amino acid sequence without an *in vivo* system. This technique takes less time to produce functional proteins. Proteins with multiple post-translational modifications that may be difficult to produce using recombinant technology can be made using this method (Hou *et al*., 2017). However, there are some drawbacks to chemical protein synthesis; it can only be used for the synthesis of small to medium peptides (Fanny *et al*., 2007). In addition, proteins synthesized chemically may contain non-native components, such as thioesters that are used to replace the native peptide bond at ligation sites. These thioesters prevent the proteins from taking their native tertiary conformation (Kent, 2009). On the other hand, numerous studies have used recombinant DNA technology for the production of HIV-1 PR (Azarnezhad *et al*., 2016, Do Hyung Kim, 1994, Louis *et al*., 1991, Maseko *et al*., 2016, Volontè *et al*., 2011, Zondagh *et al*., 2018). This method requires the initial amplification of the HIV-1 PR gene, either from patient samples or viral clones, using the polymerase chain reaction (PCR) (Andersson *et al*., 2003, Cao *et al*., 2017).

Several molecular cloning techniques have been used in the generation and assemblage of recombinant DNA molecules *in vitro*, such as PCR cloning, ligation-dependent cloning, seamless cloning, and recombinational cloning (Bertero *et al*., 2017). Irrespective of the method, the amplified gene is cloned into a cloning vector that provides a backbone for stabilizing, purifying, and propagating the target gene in bacterial host cells. However, simple cloning vectors cannot allow gene transcription and translation into functional proteins (Carter and Shieh, 2015), and the HIV-1 PR gene has to be sub-cloned into an expression vector (Clark and Pazdernik, 2015). Expression vectors contain strong regulatory promoters located upstream of the cloned gene that increase the rate of transcription (Clark and Pazdernik, 2015),

as well as a ribosome binding site that facilitates the correct positioning of the mRNA for the initiation of translation (Clark and Pazdernik, 2015). Alternatively, HIV-1 PR PCR products can be cloned directly into expression vectors (McKinstry *et al*., 2014), which saves time. Often, the PCR product is digested with restriction enzymes to facilitate ligation into the expression vector (Dubey *et al*., 2016). Once cloned into an expression vector, the plasmid is then transformed into the host cells for recombinant expression of HIV-1 PR. Commonly used protein expression host systems include yeast cells, mammalian cells, and bacterial cells. These different host systems have their pros and cons (Rosano and Ceccarelli, 2014).

2.1 Expression of HIV-1 PR in yeast and mammalian cells

The expression of HIV-1 PR in yeast and mammalian cells has been optimized for the study of HIV-1 PI drug susceptibility. Both yeast and mammalian cells behave in a similar manner when used as expression host cells for HIV-1 PR in that the cytotoxic nature of HIV-1 PR affects yeast and mammalian host cell integrity during expression, causing cell death. However, HIV-1 PIs block the cytotoxic effect of HIV-1 PR, making these cells a viable platform for the study and determination of HIV-1 PI susceptibility and propagation of HIV-1 infectious clones (full-length genome) (Benko *et al*., 2016, Blanco *et al*., 2003, Ravaux *et al*., 2014).

There are numerous yeast host expression systems available, including *Saccharomyces cerevisiae, Hansenula polymorpha, Pichia pastoris, Yarrowia lipolytica, and Schizosaccharomyces pombe* (Kim *et al*., 2015); however *S. cerevisiae* and *P. pastoris* are most commonly used. Protocols involving the use of *S. cerevisiae* for the production of human therapeutics is well established. This expression host has both the characteristic unicellular and eukaryotic systems, making it easily adaptable for research and industrial protein expression (Huang *et al*., 2014a). One problem associated with the use of *S. cerevisiae* is the low secretion of target proteins into the growth medium. Hence other yeast species, such as *P. pastoris* known for its increased ability to secrete properly folded and functional proteins into the growth medium, are good alternatives to *S. cerevisiae* (Kim *et al*., 2015, Tripathi *et al*., 2019). In addition to poor secretion of the expressed protein into the growth medium thus resulting in intracellular retention of the target protein, yeast host systems are also susceptible to hyperglycosylation of the target proteins, resulting in protein misfolding and lower yields (Gomes *et al*., 2016).

The mammalian protein expression systems, on the other hand, address most of the drawbacks associated with the yeast expression systems. The expression of recombinant proteins using mammalian host cells makes it possible to produce target proteins in a milieu close to that found in nature (Hunter *et al*., 2019). There are several mammalian cells for recombinant protein expression, including Chinese hamster ovary cells (CHO), NS0, and Sp2/0, HEK293, and CAP cells (Tripathi *et al*., 2019), with the CHO and HEK 293 cell lines most commonly used for research and industrial purposes. These two cell

lines have been genetically improved to adapt to several protein expression vectors. They can also produce functional post-translationally modified proteins in high concentrations (Hunter *et al.*, 2019). Mammalian cells have been widely used for studying HIV-1 PR inhibitor susceptibility and screening new inhibitors *in vivo* (Hilton and Wolkowicz, 2010, Rajakuberan *et al.*, 2012, Buzon *et al.*, 2011). Still, there is a paucity of information on recombinant protein concentration and purity *in vitro* when mammalian host expression cells are used for recombinant HIV-1 PR expression. Notwithstanding the sophistication of this method and its merits, the major drawback is the high cost and complexity (Khan, 2013). This review focuses on the expression of HIV-1 PR in E. coli, providing the possibility of accommodating its cytotoxic effect as a heterologous protein (Volontè *et al.*, 2011).

2.2 Expression of HIV-1 PR in *E. coli*

Regardless of the availability of other protein expression systems, *E.coli* still ranks as one of the most used microbial host systems in the academic and industrial sectors to produce recombinant proteins (Selas Castiñeiras *et al.*, 2018). The *E. coli* host is commonly used to propagate cloned genes because of its unrivaled rapid growth and ease of transformation using exogenous DNA (Pope and Kent, 1996, Sezonov *et al.*, 2007). Recombinant protein expression using *E. coli* as a host cell is quite simple. It allows for easy gene manipulation and is a good host for protein expression on a large scale for structural characterization (Yadav *et al.*, 2016). These benefits of the *E. coli* expression system have led to the availability of many molecular biology tools and protocols, such as the massive collection of expression plasmids, numerous genetically engineered *E. coli* strains, and diverse cultivation strategies (Chaudhary and Lee, 2015).

A commonly used *E. coli* strain for HIV-1 PR expression is BL21(DE3) (Azarnezhad *et al.*, 2016, Nguyen *et al.*, 2015, Volontè *et al.*, 2011, Zondagh *et al.*, 2018). It has been shown that the yield of HIV-1 PR in the BL21 (DE3) strain was higher than in the other strains (C41(DE3), C41(DE3)pLysS, C43(DE3), C43(DE3)pLysS, and KRX)) (Volontè *et al.*, 2011). Protein expression using BL21(DE3) is under the control of a T7 promoter, which increases the efficiency of expression and, at the same time, reduces the chance of expressing the target protein in an insoluble form (Joseph *et al.*, 2015).

Though there exist several protocols and modifications to the method of expressing HIV-1 PR in *E. coli*, the common workflow involves the incubation of a single *E. coli* colony containing recombinant plasmids in rich Luria-Bertani (LB) broth overnight at 37°C. The overnight culture containing the plasmids is used to inoculate the protein expression liquid medium in a ratio of approximately 1:100 and incubated at 37°C for an average of 3 to 5 hours (protocol-dependent). The culture is then supplemented with Isopropyl β- d-1-thiogalactopyranoside (IPTG) and incubated at 37°C for an additional time to induce the expression of HIV-PR (Nguyen *et al.*, 2015). Following IPTG induction, the cells are harvested by centrifugation and sonicated.

In some instances, HIV-1 PR may not be in the soluble fraction but as insoluble inclusion bodies and needs to be recovered from the latter (Nguyen *et al*., 2015). Recombinantly expressed proteins can be recovered from insoluble inclusion bodies by the solubilization of the isolated inclusion body. This is followed by refolding the protein obtained from inclusion bodies and purification via several chromatographic methods (Nguyen *et al*., 2015, Volontè *et al*., 2011). Though active HIV-1 PR could be recovered from insoluble inclusion bodies after the expression process (Maseko *et al*., 2016, Nguyen *et al*., 2015), recovery of recombinant proteins from inclusion bodies is extensive and laborious and often results in the loss of protein function (Singh *et al*., 2015).

Several modifications have been made to the bacterial host expression system to produce properly folded, active and soluble proteins. One of the improvements to enhance the recombinant expression of HIV-1 PR in bacteria cells is the use of promoters such as the T7-phage and the araBAD promoter systems in expression vectors (Dergousova *et al*., 1996, Taylor *et al*., 1992). Protein expression vectors that contain a highly regulated promoter will enhance the protein expression process and increase the yield and solubility of the expressed protein (Lebendiker and Danieli, 2014). The control of promoter activity ensures that regular host cellular functions are not interrupted by the expression of the target gene (Carrier *et al*., 1983, Fakruddin *et al*., 2012). Some *E. coli* cells, for example, *E. coli* BL21(DE3)pLysS, have been genetically engineered to control protein expression by inhibiting basal expression of T7 RNA polymerase that is under control of the T7 promoter. BL21(DE3)pLysS harbors a prophage that carries the T7 RNA polymerase gene under the control of the lacUV5 promoter (Rosano and Ceccarelli, 2014, Studier and Moffatt, 1986). When this *E. coli* host cell is used, pLysS produces T7 lysozyme, which inhibits the T7 RNA polymerase until the induction of expression by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG). These features have made it the preferred bacterial host for protein expression (Rosano and Ceccarelli, 2014).

As part of the efforts to enhance protein expression, protein fusion tags have been engineered and have found significant application in recombinant protein expression. The incorporation of protein fusion tags at the N- or C-terminal ends of target proteins during cloning enhances target protein expression, preventing its accumulation as insoluble aggregates in the *E. coli* cytoplasm (Singh *et al*., 2015, Gupta and Shukla, 2016). Solubility enhancing fusion tags have been used to overcome HIV-1 PR building up as inclusion bodies in the bacterial expression system (Volontè *et al*., 2011) and is thus the focus of this review.

3. Employing Fusion Tags to Overcome Challenges Associated with HIV-1 PR Expression

Genetic heterogeneity between the target protein and the host cell and cytotoxicity of the target protein may be a burden to the expression host. This can affect the growth and viability of the expression cells, thus negatively impacting protein expression (Labhsetwar *et al*., 2013, Tripathi *et al*., 2019). The

variability in recombinant protein expression associated with the host cell machinery can be minimized using protein fusion tags (Bhattacharya *et al*., 2016).

The use of fusion tags has helped overcome the cytotoxic effect of HIV-1 PR on bacterial expression host cell integrity during heterologous expression of HIV-1 PR, allowing the expression of the latter in relatively high concentration (Volontè *et al*., 2011). Fusion tags improve the solubility and folding of the expressed protein, in addition to increasing the efficiency of purification (Costa *et al*., 2014). Fusion tags can serve a dual purpose as both solubility and affinity tag simultaneously, where they enhance the soluble expression of the target protein and increase the efficiency of protein purification (Esposito and Chatterjee, 2006). Alternatively, a separate solubility enhancing tag and an affinity tag can be used together to exploit their different properties to increase solubility and improve purification of the expressed protein (Esposito and Chatterjee, 2006). Though there are several solubility enhancing fusion tags used in protein expression, this study has summarised the fusion tags (Table1) optimised for expressing HIV-1 PR recombinantly.

3.1 Maltose binding-protein tag

The maltose-binding protein (MBP) tag is derived from bacterial "ATP-binding cassette maltose/maltodextrin transporter" (Jin *et al*., 2017). It has a molecular weight of 43kDa, and its ability to improve protein solubility is linked to the incorporation of chaperones during the folding process of the expressed protein. (Raran-Kurussi and Waugh, 2012). The capacity of Maltose binding-protein (MBP) to support the expression of HIV-1 PR and ease its purification has been demonstrated. A study has reported the recovery of approximately 1 mg/L of pure HIV-1 PR after cleavage of the MBP tag with an activity of ~8.50 μmol/min/mg of protein (Louis *et al*., 1991). MBP helps to stabilize and increase the solubility of HIV-1 PR and facilitate its purification to homogeneity through the use of a series of chromatography steps, with the recovered HIV-1 PR having a purity > 95% (Do Hyung Kim, 1994, Louis *et al*., 1991).

Studies have also shown that HIV-1 PR expressed in *E. coli* as fusion proteins tagged to MBP is autoprocessing deficient. This permits rapid and efficient purification of the HIV-1 PR-MBP precursor by affinity chromatography using amylose columns. Also, this makes available a precursor form of HIV-1 PR that can be studied to screen inhibitors that effectively act on HIV-1 PR precursors (Louis *et al*., 1991, Louis *et al*., 1994). The inability of the MBP fusion tag to induce HIV-1 PR precursor autoprocessing has been suggested to be due to its monomeric nature, thus suppressing the dimerization of HIV-1 PR needed for normal enzymatic function and activity (Huang and Chen, 2013). Though MBP-tagged HIV-1 PR is auto processing deficient when expressed in *E. coli*, it has been reported that this fusion tag can promote HIV-1 PR autoprocessing efficiently when expressed in mammalian cells (Huang and Chen, 2013, Tien *et al*., 2018). These authors showed autoprocessing of MBP-tagged HIV-

1 PR might be host cell or tag sequence-dependent (Huang and Chen, 2013, Tien *et al*., 2018). Biologically active HIV-1 PR can be released from the MBP-fusion precursor by a cycle of *in vitro* denaturing and refolding (Huang and Chen, 2013). Another way that the fusion tag can be removed is by using cleavage proteases. Expression vectors used to express proteins fused to MBP have the enterokinase and Factor Xa protease cleavage site located between the fusion tag and the multiple cloning site to facilitate the tag's removal.

There exists a debate about the best position for the MBP fusion tag in the expression construct. Research favors the positioning of the MBP tag at the N-terminus to better enhance the solubility of the expressed protein, rather than positioning the fusion tag at the C-terminus (Raran-Kurussi *et al*., 2015). This may be due to the N-terminal MBP tag undergoing folding before the newly synthesized protein emerges from the host cell ribosome. As a result, the folded MBP tag acts as a *cis* chaperone for the newly formed protein, shielding it from factors that promote inclusion body formation (Raran-Kurussi *et al*., 2015, Sachdev *et al*., 1998). Nevertheless, it has been reported in another study that the position of the MBP tag has no impact on the solubility of the expressed protein, as the MBP tag can induce protein solubility whether positioned at the N or C-terminal (Dyson *et al*., 2004).

Expression of proteins in *E. coli* tagged to MBP can either be in the cytoplasm or periplasm. The choice of expressing proteins in the periplasm or cytoplasm depends on the structural characteristics of the protein to be expressed (Kosobokova *et al*., 2016). The periplasmic expression method is preferable for proteins rich in disulfide bonds (Kosobokova *et al*., 2016) since the high oxidizing environment in the periplasm is essential for disulfide bond formation (Volontè *et al*., 2011). Examples of these vectors are the pMAL series vectors from New England Biolabs (e.g., pMAL c2X, pMAL p2X, pMAL c5X, pMAL p5X) and pIVEX vectors from Roche (RTS pIVEX MBP Fusion vector) (Young *et al*., 2012).

The MBP also possesses the characteristics of affinity tags, making it a suitable fusion partner for resolving a wide range of challenges associated with protein expression and purification, such as loss of proteins during the purification process (Lebendiker and Danieli, 2011). The MBP tag can enhance the solubility and expression of several difficult to express proteins because of its large hydrophobic cleft that is able to alter its shape to accommodate different target proteins, promoting the latter's proper folding (Costa *et al*., 2014). Despite these merits of the MBP tag, its large size is a limitation. It places a heavy metabolic burden on the expression host cell, increasing metabolic energy utilization compared to smaller tags. This can make its cleavage result in the target protein becoming susceptible to aggregation and negatively impacting the protein yield (Zhao *et al*., 2013b).

3.2 Glutathione-S-transferase tag

Glutathione S-transferase (GST) is derived from *Schistosoma japonicum,* and is a fusion tag widely used for recombinant expression of proteins in the *E. coli* cytoplasm. This fusion tag can be used as an efficient fusion precursor component to initiate the dimerization of the expressed protein (Tudyka and Skerra, 1997). Glutathione-S-transferase possesses the characteristics of a solubility enhancing tag as well as that of an affinity tag. GST is a dimer with a molecular weight of 26 kDa and prevents proteolytic degradation of the target protein. This makes the GST fusion tag an ideal comparatively sized, secretion-competent fusion tag (Yadav *et al*., 2016, Tudyka and Skerra, 1997). The GST tag is a well-characterized non-viral component used in expression constructs containing HIV-1 PR precursor for the study of HIV-1 PR precursor autoprocessing and also the matured form of HIV-1 PR (Huang and Chen, 2013).

The GST fusion tag's tight dimer stability, and slow rate of unfolding and refolding enhances the dimerization of the GST fusion precursor and its autoprocessing (Kaplan *et al*., 1997). When HIV-1 PR is expressed fused to GST, the HIV-1 PR precursor can undergo autoprocessing in vivo to produce the matured form of HIV-1 PR. This provides a system for the study of HIV-1 PR in its natural environment (Huang and Chen, 2013). Regardless of the positioning of the GST tag (N- or C-terminal), it can still enhance the production of a soluble and active protein with increased yield in bacterial, mammalian, and yeast cells (Tudyka and Skerra, 1997, Young *et al*., 2012). All articles reviewed in this study expressing HIV-1 PR fused to the GST fusion tag had the tag positioned at the N-terminus (Volontè *et al*., 2011, WAN *et al*., 1996, Huang *et al*., 2011, Huang *et al*., 2009). The N-terminus GST fusion tag trans frame region, when HIV-1 PR is expressed in *E. coli*, has also been shown to undergo "N-terminal autocatalytic maturation" (Huang *et al*., 2009). This modulates autoprocessing and enhances the release of mature HIV-1 PR (Huang *et al*., 2009). The incorporation of small peptide epitopes (like HA, Myc, and Flag) into the GST fusion precursor at various positions allows for the detection of the active HIV-1 PR (Huang and Chen, 2013).

This tag makes identification of the expressed protein relatively easy through the use of immunoassay or chromogenic enzymatic assay (Yadav *et al*., 2016). The expression of HIV-1 PR fused to the GST tag can be accomplished using the pGEX vector series (pGEX-2T, pGEX-4T, pGEX-3X, pGEX-5X) manufactured by GE Healthcare. In pGEX vectors, between the GST fusion tag and the multiple cloning site are protease cleavage sites, either the thrombin (pGEX-2T and pGEX-4T vectors) or factor Xa cleavage sites (pGEX-3X and pGEX-5X). This makes the removal of the GST tag from the expressed protein relatively simple after expression (Young *et al*., 2012). In addition, the pET-41a-c(+) (with thrombin and Enterokinase cleavage sites) and pET-42a-c(+) (with thrombin and factor Xa cleavage sites) protein expression vectors from Novagen can be used for the expression of the GST-tagged proteins.

An advantage of the GST tag over other tags is that it results in the production of active HIV-1 PR upon expression in relatively high concentrations (Volontè *et al*., 2011). A study by Volonte *et al*., 2011 developed an optimized protocol for the expression of HIV-1 PR in *E. coli* fused with GST fusion proteins (Volontè *et al*., 2011). This study demonstrated that using this optimized expression and purification protocol leads to the production of a fairly high concentration of HIV-1 PR of good quality and making this protein available on a laboratory-scale for downstream biochemical studies (Volontè *et al*., 2011). The major limitation of this method is the low yield associated with leakage of the expressed protein from the column matrix during the purification step (Yadav *et al*., 2016). A relatively low yield of pure HIV-1 PR has been reported using the GST tag by Volontè *et al*. (2011) (0.15mg/L of culture) and Maseko *et al*. (2016) (0.25 mg/L of culture). These studies found the specific activities of the expressed HIV-1 PR to be 1.22 μmol/min /mg of protein (Volontè *et al*., 2011) and 2.20 μmol/min /mg protein of respectively (Maseko *et al*., 2016).

3.3 Thioredoxin fusion tag

Thioredoxin (Trx) fusion tag has been shown to mitigate the cytotoxic effect of HIV-1 PR to the expression host cell, enhancing expression, solubility as well as folding of recombinant HIV-1 PR (Azarnezhad *et al*., 2016, Zondagh *et al*., 2018). The Trx fusion tag is derived from *E. coli* and has a molecular weight of 12 kDa, making it extremely soluble and effectively inhibits the aggregation of insoluble protein masses through its intrinsic oxidoreductase action. This property of Trx decreases disulfide bond formation, thus promoting protein folding (LaVallie *et al*., 2000), which is essential for the normal functioning of HIV-1 PR (Zondagh *et al*., 2018). Besides proper protein folding, thioredoxin confers high thermal stability to the expressed protein. This leads to the expression of proteins with increased conformational stability (LaVallie *et al*., 2003).

In addition to the robust folding property of Trx, which contributes to its high efficiency as a solubility enhancing fusion tag, Trx is highly translated by the bacterial host cell. This property is also conferred on the fused protein to be expressed (LaVallie *et al*., 2003). Research has shown that the Trx fusion tag also aids in the crystallization of the protein been expressed. It has been reported that the ligand-binding domain of human estrogen receptor (Cura *et al*., 2008) and U2AF homology motif domain of splicing factor Puf60 fused to Trx enhanced their crystallization (Corsini *et al*., 2008). Due to its inability to bind to sorbents, thioredoxin is used in combination with affinity tags such as the His-tag for efficient purification after protein expression (Kosobokova *et al*., 2016). The thioredoxin fusion tag can be fused at the N or C-terminus of the target protein to be expressed, as both termini of this fusion tag are surface accessible, thus making them viable positions for linkage to the target protein (LaVallie *et al*., 2000, Katti *et al*., 1990). This fusion tag effectively initiates the translation of the target proteins, particularly when fused at the N-terminus, where protein translation is initiated (LaVallie *et al*., 2000). To the best of our knowledge, no study has been done with Trx fused at the C-terminal, while a few studies have

expressed HIV-1 PR using the Trx fusion tag positioned at the N-terminal (Azarnezhad *et al*., 2016, Zondagh *et al*., 2018).

The combination of the Trx and His tags has been shown to act synergistically to facilitate the expression and purification of catalytically active HIV-1 PR without affecting its dimerization (Azarnezhad *et al*., 2016, Zondagh *et al*., 2018). The concentration of HIV-1 PR recovered from this combination has been reported to be about 2mg/L; the authors also stated that the protein was active but did not report the activity level (Zondagh *et al*., 2018). The thioredoxin moiety in the expressed HIV-1 PR-Trx fusion precursor has been shown to mimic the structure of gag, but it does not facilitate autocatalysis of the expressed HIV-1 PR (Zondagh *et al*., 2018). The Trx fusion tag does not initiate the dimerization of HIV-1 PR, which is essential for HIV-1 PR precursor autoprocessing. This may also be due to the formation of a high order oligomeric state by HIV-1 PR with the Trx-His-tag combination, which results in stearic hindrance effects, inhibiting the autocatalytic nature of HIV-1 PR (Zondagh *et al*., 2018). The pET Trx Fusion System 32 expression vector from Novagen and pTrxFus vector from Invitrogen are commercially available expression vectors that can be used to express proteins fused to the Trx tag.

3.4 Poly histidine-tag

The poly-histidine tags usually contain six or more histidine residues. The His-tag is relatively small, with a molecular weight of approximately 2.5kDa. Its small size is an advantage as it does not interfere with the function and structure of the target protein with which it is fused for expression (Booth *et al*., 2018). The poly-histidine tag (His-tag) though an affinity tag used to increase purification efficiency after protein expression (Booth *et al*., 2018), can also increase the solubility of the expressed protein (Gräslund *et al*., 2008). The His-tag has been widely used in the expression and purification of recombinant HIV-1 PR alone or in combination with solubility enhancing fusion tags (Azarnezhad *et al*., 2016, Maseko *et al*., 2019, Nguyen *et al*., 2015, Volontè *et al*., 2011, Zondagh *et al*., 2018).

The expression of HIV-1 PR fused to the His-tag has been optimized for efficient recovery and purification of catalytically active HIV-1 PR (Maseko *et al*., 2019, Nguyen *et al*., 2015). However, this has limitations as the His-tag mainly enhances the purification of HIV-1 PR with little capacity to increase its solubility, thus resulting in the accumulation of HIV-1 PR in inclusion bodies. This results in a reduced yield of the expressed protein (Maseko *et al*., 2019, Nguyen *et al*., 2015). To overcome this challenge, the His-tag can be combined with solubility enhancing tags (like MBP, GST, and Trx) to increase the solubility of HIV-1 PR and at the same increase the purification efficiency (Azarnezhad *et al*., 2016, Maseko *et al*., 2019, Volontè *et al*., 2011, Zondagh *et al*., 2018). The poly-histidine tag can be placed at either the N- or the C-terminus of the expression construct for protein expression (Terpe,

2006). However, research has shown that placing the poly-histidine tag at the N-terminus results in a greater yield than when placed at the C-terminus (Aslantas *et al*., 2019, Park *et al*., 2015).

When the His-tag is used, a very high concentration of pure protein can be obtained in just one chromatographic step from *E. coli* (Jia and Jeon, 2016). To express HIV-1 PR fused to the His-tag, the commercially available pET expression vector series (e.g., pET-14b, pET-15b, pET-14b) from Novagen and pTrcHis expression vector series (pTrcHis A, B, & C) from Invitrogen can be used. Purification of histidine-tagged protein is quite efficient due to the capacity of Histidine to readily form bonds with immobilized transition metal ions such as Co^{2+}, Cu^{2+}, Ni^{2+}, Zn^{2+}, Ca^{2+}, and Fe^{3+} used in Immobilized metal ion affinity chromatography (IMAC) matrices for the purification of His-tagged proteins (Kimple *et al*., 2013). The wide application of the His-tag in protein expression and purification is also a result of the ease of in vitro histidine-tagged protein detection using anti-His antibodies (Kimple *et al*., 2013). A limitation of the His-tag is that native protein contaminants from the expression host cell could be co-purified with the target proteins (Yadav *et al*., 2016).

Table 1. A summary of studies on the expression of HIV-1 PR using solubility enhancing fusion tag, their advantages and limitations.

* HIV-1 PR yield when expressed using a combination of Trx tag and His tag

Fusion Tag	Tag Type	Molecular Weight (kDa)	Pure HIV-1 PR yield /Litre of culture after tag removal	Specific activity expressed HIV-1 PR after tag removal	Protein expression vectors with reviewed solubility enhancing tags	Advantages	Limitations	Conclusion of the reviewed Studies	References
Maltose-binding protein (MBP)	Solubility and affinity	43	~ 1 mg/L	~8.50 µmol/min /mg protein.	pMAL series pIVEX series	Improved solubility and purification of the expressed proteins. promotes proper protein folding	The large molecular weight. may interfere with protein yield	• HIV-1 PR fused with MBP stabilises its expression in *E. coli* and facilitates its purification to homogeneity • HIV-1 PR expression using MBP results in high yield and easy purification	(Do Hyung Kim, 1994) (Louis *et al.*, 1991)
Thioredoxin (Trx)	Solubility	12	~ 2 mg/L	-	pET Trx pTrxFus	Improved solubility and aids refolding of the target protein as well as protein crystallization	Unable to bind sorbents. Must be used in with affinity tags to improve protein purification	• Increased rate of expression of catalytically active HIV-1 PR • Increased specificity and sensitivity of HIV-1 PR expressed as thioredoxin-hexahistidine fusion when subjected to immunoassay tests	(Zondagh *et al.*, 2018) (Azarnezhad *et al.*, 2016)
Glutathione-S transferase (GST)	Solubility and affinity	26	~ 0.25 mg/L	~ 2.20 µmol/min /mg protein.	pGEX series pET-41a-c(+) pET-42a-c(+)	Improved solubility and purification of the target protein. It enhances protein dimerization	It can leak out from the column	• Expression of high concentration of HIV-1 PR as GST-fusion proteins • Production pure and high concentration of HIV-1 PR in a lab-scale as GST-fusion proteins for further biochemical studies	(Maseko *et al.*, 2016) (Volontè *et al.*, 2011)
Poly- histidine (His-tag)	Affinity	2.5	~ 2 mg/L [#]	~ 1.22 µmol/min /mg protein *	pET series pTrcHis2 series	It interferes minimally with the expressed protein, aids purification	It may cause the co-purification of protein contaminants	• HIV-1 PR efficiently expressed and purified when His-tag is used in combination with GST and Trx fusion tag	(Volontè *et al.*, 2011, Zondagh *et al.*, 2018)

[#] Specific activity of HIV-1 PR when expressed using a combination of GST tag and His tag

3.5 Removal of fusion tags after protein expression

Fusion tags may affect the structural and functional properties of the expressed protein if not removed (Waugh, 2011). However, it is important to consider which further experiments will be done downstream after protein expression before removing the fusion tag (Malik, 2016). This may be biochemical or structural characterization such as X-ray crystallography and nuclear magnetic resonance (NMR) of the matured HIV-1 PR. (Waugh, 2011). When expressing HIV-1 PR for either biochemical or structural characterization, removing the fusion tag is essential as most of the discussed fusion tags are larger than HIV-1 PR and may interfere with its structure and function. HIV-1 PR is only functional when in the right conformation; the presence of additional amino acid residues may prevent its dimerization (Zondagh *et al*., 2018). Fusion tags, if not completely removed or extra amino acid residues left behind as a result of fusion protein cleavage, will impact its dimerization, conformation, and normal function (Zondagh *et al*., 2018).

Fusion tags can be removed either through enzymatic or chemical means. There are enzyme recognition sites incorporated in most expression vectors, found between the target protein and fusion tag for cleavage of fusion tags from the protein of interest (Zhu *et al*., 2017). The choice of the cleavage protease should be guided by its compatibility with the target protein and fusion tag. In addition, the buffer condition and temperature at which the cleavage protease functions optimally must be compatible with the expressed protein. The chosen cleavage protease should also have a specific cleavage site between the fusion tag and the target protein, so it does not cut at unspecific sites (Waugh, 2011, Pina *et al*., 2014). Non-specificity of cleavage enzyme may result in inaccurate tag removal, affecting the protein sequence, structure, and function (Mahmoodi *et al*., 2019).

In the past, traditional enzymatic tag removal included the use of the proteases: thrombin and factor Xa. Advances in protein expression in recent times have led to the use of "Tobacco Etch Virus (TEV), Rhinovirus 3C Protease, SUMO protease, and enterokinase," which have been shown to have high specificity and high stability (Lebendiker and Danieli, 2014). Specifically, TEV protease is resistant to reducing agents, has high specificity, is inexpensive, and is easy to prepare in large amounts. Usually, TEV protease cleaves target proteins in a way that the native protein remains intact (Esposito and Chatterjee, 2006, Jia and Jeon, 2016).

Tag removal by chemical cleavage involves chemical treatment of the expressed protein to remove the tag. The harsh reaction conditions of chemical cleavage limit its use mainly to recombinant proteins obtained from inclusion bodies as they are likely to cause undesired modifications to the protein. On the other, the

ease of elimination of the chemicals used in this process is an advantage and is quite cheap (Rais-Beghdadi *et al*., 1998).

4. Impact of HIV-1 PR expression using fusion tags on its functional characterization

This review has discussed several studies utilizing fusion tags to overcome challenges associated with recombinant expression of HIV-1 PR in bacterial hosts. Fusion tags help in the development of laboratory techniques that are highly reproducible and scalable. When employed, recombinant HIV-1 PR can be produced at a scale suitable for academic and research purposes (Volontè *et al*., 2011). This also makes HIV-1 PR available for the development of new techniques that can be used to combat AIDS in the form of novel HIV-1 PIs (Broglia *et al*., 2006). For example, recombinantly expressed HIV-1 PR and peptide-specific fluorescent substrates have been used to develop a novel fluorometric assay to phenotypically differentiate mutant forms of HIV-1 PR (Zhu *et al*., 2015).

In addition, the use of fusion tags will help to study the HIV-1 PR autoprocessing mechanism leading to the production of active HIV-1 PR inside the cell in an atmosphere almost the same as the exact biological milieu (Huang and Chen, 2013). This will provide detailed information about drug interactions with the active HIV-1 PR and its precursor form at the same time, thus aiding the development of novel inhibitors that selectively target HIV-1 precursor autoprocessing. (Huang and Chen, 2013). As mentioned earlier, the position of the fusion tag may influence HIV-1 PR autoprocessing (Huang *et al*., 2009), as well as the amino acid sequences that are upstream of the gag derived protein p6* (trans-frame region (TFR)) between the fusion tag and the expressed HIV-1 PR enzyme (Tien *et al*., 2018).

Undertaking a comprehensive biochemical analysis of HIV-1 PR has been made possible by employing fusion tags. Solubility enhancing fusion tags in expressing recombinant HIV-1 PR allows for *in vitro* study of the catalytic mechanism of HIV-1 PR and determine the influence of drug resistance mutations on the enzyme kinetics and inhibitor binding capacity (Maseko *et al*., 2017, Louis *et al*., 1991). Data from the kinetic properties of drug-resistant HIV-1 PR mutants provide information on assessing and predicting drug-resistance patterns to PIs. This may also help in the choice of HIV-1 PI combinations to tackle drug resistance (Gulnik *et al*., 1995).

The discovery of novel potent, broad-spectrum HIV-1 PIs with ideal thermodynamic properties can be achieved through the combination of enzymologic, thermodynamic, and structural evaluation of mutant drug-resistant HIV-1 PR (Kožíšek *et al*., 2014). Enzyme Interaction kinetics and thermodynamic studies will make available useful information in the design of better inhibitors and a better understanding of interactions at the molecular level. In addition, this may be used in combination with structural data for the

characterization of HIV-1 PIs for drug discovery, and determining the mechanisms of resistance to HIV-1 PIs (Shuman *et al*., 2004). All these are made possible by the availability of active HIV-1 PR for analysis, through which most HIV-1 PR structures have been resolved and characterized from inclusion bodies.

Studies have utilized structural analysis of HIV-1 PR crystal structure in the design and development of novel HIV-1 PIs active against multidrug-resistant HIV-1 variants (Ghosh and Chapsal, 2013, Ghosh *et al*., 2018). Structure-based drug design involves several steps, after which an optimized lead will go into the first phase of clinical trials (Anderson, 2003). This could be done using either homology modeling, X-ray crystallography, or Nuclear Magnetic Resonance Spectroscopy (Anderson, 2003). The commonly used method in structure-based drug design is X-ray crystallography. This may probably be due to its superiority in producing high atomic resolution structures and accommodating proteins of various sizes from very small to large ones (Acharya and Lloyd, 2005).

In modern times, novel drug design strategies based on protein X-ray structure has become a reliable and commonly deployed means (Ghosh *et al*., 2018). X-ray crystallography of HIV-1 PR will furnish researchers with details of existing mutations in the expressed viral enzyme. In addition, the mutations will be taken into consideration to design appropriate inhibitors that can bind to and inhibit the target protein (Ghosh *et al*., 2018). Also, molecular insights can be obtained from X-ray structural analysis of HIV-1 PI bound form (Ghosh *et al*., 2017a). This will provide detailed information about the interaction of PIs with the enzyme. These interactions may reflect the potency of the inhibitors (Ghosh *et al*., 2017b, Ghosh *et al*., 2018). Drug design strategies utilizing protein-X-ray structures have been exploited to promote extensive interaction of ligands with the backbone atoms of the HIV-1 PR active site. These studies have shown that enhancing the ligand backbone interaction results in the development of HIV-1 PIs with potent antiviral activity (Ghosh *et al*., 2012, Ghosh *et al*., 2010, Ghosh *et al*., 2007).

5. Conclusion

This review provides researchers with a quick start in expressing HIV-1 PR recombinantly for the first time, as it highlights the current knowledge of recombinant HIV-1 PR expression using fusion tags. This review has also shown that the different fusion tags used for HIV-1 PR expression have their peculiarities with regard to their advantages and limitations. The GST fusion tag has been widely used for the expression of mature and precursor forms of HIV-1 PR. Its ability to facilitate the dimerization and autoprocessing of the expressed HIV-1 PR fusion precursor to release the mature active HIV-1 PR gives it an edge over the MBP and Trx fusion tags. However, the reported low yield compared to the other fusion tags may be a limitation. The MBP tag is also a reliable fusion tag for expressing the mature form of HIV-1 PR and its precursor

forms with relatively high concentrations and specific activity of HIV-1 PR recovered compared to the GST tag. The need for optimized protocols to produce active HIV-1 PR in high purity and concentration cannot be overemphasized.

Bridging between chapter two and chapter three

The availability of pure and active HIV-1 PR makes its functional characterization less cumbersome and straightforward. Chapter two reviewed the various fusion tags used in HIV-1 PR expression. In chapter three, HIV-1 PR was expressed as a soluble protein using the maltose-binding protein (MBP) fusion tag. The expressed protein was characterized using enzyme kinetic analysis, fluorescence spectroscopy, and molecular dynamics simulation to determine the impact of drug resistance mutations on the interaction of LPV with HIV-1 PR and how well DRV would fare in the case of ART switch to DRV. This chapter is published in the journal *Biomolecules* (the cover page is attached in the appendix).

CHAPTER THREE

Acquired HIV-1 protease conformational flexibility associated with lopinavir failure may shape the outcome of darunavir therapy after antiretroviral therapy switch

Abstract: Understanding the underlying molecular interaction during a therapy switch from lopinavir (LPV) to darunavir (DRV) is essential to achieve long-term virological suppression. We investigated the kinetic and structural characteristics of multidrug-resistant South African HIV-1 subtype C protease (HIV-1 PR) during therapy switch from LPV to DRV using enzyme activity and inhibition assay, fluorescence spectroscopy, and molecular dynamic simulation. The HIV-1 protease variants were from clinical isolates with a combination of drug resistance mutations; MUT-1 (M46I, I54V, V82A, L10F), MUT-2 (M46I, I54V, L76V, V82A, L10F, L33F), and MUT-3 (M46I, I54V, L76V, V82A, L90M, F53L). Enzyme kinetics analysis shows an association between increased relative resistance to LPV and DRV with the progressive decrease in the mutant HIV-1 PR variants' catalytic efficiency. A direct relationship between high-level resistance to LPV and intermediate resistance to DRV with intrinsic changes in the three-dimensional structure of the mutant HIV-1 PR as a function of the multidrug-resistance mutation was observed. *In silico* analysis attributed these structural adjustments to the multidrug-resistance mutations affecting the LPV and DRV binding landscape. Though DRV showed superiority to LPV, as a lower concentration was needed to inhibit the HIV-1 PR variants, the inherent structural changes resulting from mutations selected during LPV therapy may dynamically shape the DRV treatment outcome after the therapy switch.

Key words: HIV-1 protease, HIV-1 protease inhibitor, Lopinavir, Darunavir, conformational flexibility

1. Introduction

Globally, HIV-1 infection remains a serious public health problem, with about 38 million infected people at the end of 2019 (UNAIDS, 2020). The global HIV epidemic burden rests heavily on countries in sub-Saharan Africa (Dwyer-Lindgren *et al*., 2019). South Africa remains the global epicenter of the HIV-1 epidemic, with the pandemic dominated by HIV-1 subtype C (Hodes and Morrell, 2018, Mosebi *et al*., 2008). The standard treatment of HIV infection is highly active antiretroviral therapy (HAART) (Henes *et al*., 2019a); HAART has greatly improved the clinical outcome of HIV-infected persons since its introduction (Aoki *et al*., 2018). However, the emergence of drug-resistant HIV-1 variants has significantly contributed to the failure to control HIV-1 replication in some HIV-1 infected individuals (Aoki *et al*., 2018). The increasing cases of virological failure associated with first and second-line antiretroviral treatment (ART) present a significant clinical challenge for patient management in resource-constrained settings. (Grinsztejn *et al*., 2019). The ability to provide effective and sustained virological suppression using ART is crucial; thus, HIV-1 protease inhibitors (PIs) with a high genetic barrier to the evolution of drug resistance forms second and last-line ART in many settings globally to achieve virological suppression (Aoki *et al*., 2018, Grinsztejn *et al*., 2019).

HIV-1 protease (PR) is a key drug target that plays a crucial role in cleaving newly synthesized viral polyprotein into functional proteins needed for the maturation of nascent viral particles (Weikl *et al*., 2019). The HIV-1 PR is a 99 amino acid homodimer, organized into six structural segments (Figure 1) namely: the flap region, made up of residues 43–58/43'–58'; the flap elbow, consisting of residues 35-42/35'-42'; the fulcrum, which comprises of residues 11–22/11'–22'; the cantilever, which is made up of residues 59–75/59'–75'; the dimer interface, made up of residues 1-5/1'-5', 95-99/95'-99'; and the catalytic site which comprises residues 23–30/23'–30' (Perryman *et al*., 2004, Harte *et al*., 1990). The flap covers the HIV-1 PR active site, and it regulates the entry of substrates and protease inhibitors (PIs) into the catalytic site (Yu *et al*., 2017). HIV protease inhibitors (PIs) are non-cleavable substrate analogs designed to bind to the active site of HIV-1 PR. The binding of HIV-1 PIs in the enzyme active site inhibits its normal enzymatic activity by preventing it from cleaving its natural substrate (Huang and Chen, 2013). However, the antiviral capacity of PIs and the affinity of HIV-1 PR for PIs is diminished by the presence of drug resistance mutations in the HIV-1 PR gene (Henes *et al*., 2019b, Wu *et al*., 2003).
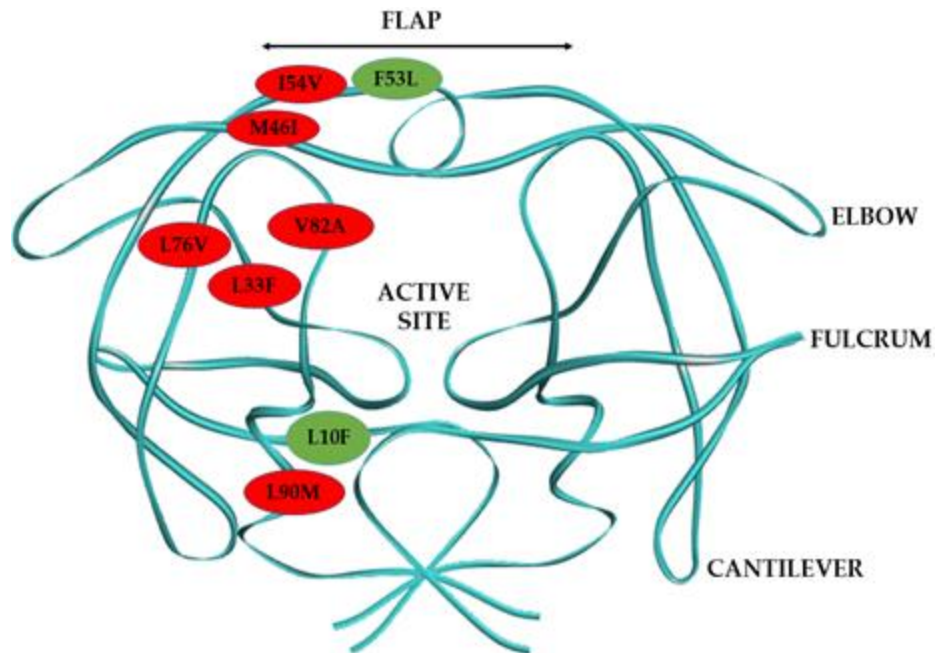
Figure 1. Diagrammatic representation of HIV-1 PR dimer structure and the mutations harbored by the mutant HIV-1 PR variants in this study. The major HIV-1 PR mutations are represented in red balls, and the minor mutations in green balls.

Drug-resistance mutations could be found in the active site of HIV-1 PR and directly impact the binding affinity and interaction of the PIs with HIV-1 PR (Johnson *et al*., 2010). In contrast, non-active site mutations may not directly affect the interaction of HIV-1 PR with inhibitors but may indirectly influence the molecular interaction of inhibitors with the HIV-1 PR through alteration of the protein flexibility and stability (Bastys *et al*., 2020, Piana *et al*., 2002, Chang and Torbett, 2011). The accumulation and interplay between active and non-active site drug-resistance mutations arising from drug pressure may cause structural changes, leading to HIV-1 PR variants with altered protein conformations (de Vera *et al*., 2013, Wang *et al*., 2012). The alteration in the conformations of HIV-1 PR confers on it adaptability and flexibility, thus affecting its interaction with different PIs (Oehme *et al*., 2011, Kumar and Hosur, 2003). The gain in HIV-1 PR flexibility due to changes in inter-residue connections alter the HIV-1 PI binding landscape, resulting in the inability of HIV-1 PIs to bind firmly in the active site (Chetty *et al*., 2016).

In low and middle-income countries, boosted LPV has been the common backbone of the second-line ART regimen (WHO, 2014). The combination of ritonavir-boosted lopinavir (LPV/r) with nucleoside reverse transcriptase inhibitors (NRTIs) has been a well-utilized, cost-effective treatment regimen for the management of HIV infection (Wangpatharawanit and Sungkanuparph, 2016). However, the development of resistance to an LPV-based regimen affects its use in patient management. Its continuous use during virological failure may result in significant cross-resistance to other PIs (WHO, 2014). The major HIV-1

PI resistance mutations that affect the efficacy of boosted LPV regimen are V32I, L33F, M46I/L, I47V/A, I50V, I54V/T/A/L/M, L76V, V82A/F/T/S, I84V and L90M (Wensing *et al*., 2019). Research shows that the emergence of V32I, L33F, I47A, I50V, L76V, 184V under drug pressure during LPV therapy may confers cross-resistance to DRV (Tang and Shafer, 2012, Johnson *et al*., 2010). Where LPV treatment fails, DRV is an effective salvage remedy (Santos *et al*., 2012) because of its high genetic barrier to drug resistance, making it the preferred antiviral agent used in numerous HIV treatment plans for therapy-naive and experienced HIV-1 infected individuals (Dierynck *et al*., 2007). In addition, DRV's dual mechanism of action by directly inhibiting HIV-1 PR, as well as inhibiting HIV-1 PR dimerization, puts it in a better position as the PI of choice for salvage therapy (Aoki *et al*., 2018).

The South African HIV-1 subtype C PR differs from HIV-1 subtype B PR at 8 different amino acid positions; T12S, I15V, L19I, M36I, R41K, H69K, L89M and I93L (Velázquez-Campoy et al., 2003), and its difference with other non-B HIV-1 PR is presented in chapter one (table 1). These polymorphisms in the South African HIV-1 subtype C PR such as the I15V, M36I, and L89M I93L located in its hydrophobic core have been shown to confer increased structural flexibility compared to those in its hydrophobic core to the subtype B HIV-1 PR (Naicker et al., 2014). Currently, there is no comparative analysis of the biochemical and structural implications of this regimen switch in HIV-1 infected individuals harboring multidrug-resistant South African HIV-1 subtype C PR. This study aims to gain mechanistic insight into the development of resistance to LPV by drug-resistant subtype C clinical isolates at the point of regimen switch to DRV, using biochemical and in-silico analysis.

This study also seeks to determine how well DRV would fare in managing HIV-1 infected individuals failing LPV-based therapy when their treatment regimen is changed. The South African HIV-1 subtype C PR variants studied harbored major and minor PI drug-resistant mutations (Figure 1). The major PI mutations are M46I and I54V found in the flap region, L76V located in the central β-sheet that forms one of the active site's borders, the V82A mutation, which is found in the active site, and the L90M and L33F mutations in the HIV-1 PR hydrophobic core. The minor drug resistance mutations are L10F located around the fulcrum, and the F53L – a flap mutation. Research has shown that the most reported major HIV-1 PI resistance mutations in HIV-1 infected individuals failing second-line therapy in South Africa are I54V, V82A, and M46I (Obasa *et al*., 2020, Steegen *et al*., 2016). Therefore, in this study, we have chosen mutant HIV-1 PR variants harboring different combinations of the common HIV-1 PI resistance mutations reported in the south African HIV-1 subtype C. This study will improve the available knowledge to develop more efficient PIs considering the unique characteristics of multidrug-resistant South African HIV-1 subtype C PR.

2. Materials and Methods

2.1 Ethical approval

This study was approved by the University of KwaZulu-Natal Biomedical Research Ethics Committee (BREC NO. 413/17).

2.2 Amplification of HIV-1 Protease gene sequence

Blood samples were obtained from consenting HIV-1 infected individuals failing second-line lopinavir (LPV) inclusive regimen and switching to darunavir (DRV) treatment from an observational cohort study in Durban, South Africa. Reverse Transcription Polymerase Chain Reaction (RT - PCR) and nested PCR were used to amplify HIV-1 gag-protease region from plasma RNA extracted using QiAmp Viral RNA Kit (Qiagen, CA, USA). RT-PCR and nested PCR conducted as described previously (Wright *et al*., 2010).

Briefly, Superscript III One-Step RT-PCR System with Platinum Taq DNA Polymerase (Invitrogen, Carlsbad, CA, USA) was used to perform the RT-PCR using HIV-1 gag – protease specific primers; **Gag+1**: 5'-GAGGAGATCTCTCGACGCAGGAC-3' as forward primer and **3'rvp**: 5'_GGAGTGTTATATGGATTTTCAGGCCCAATT_3' as reverse primer. The PCR conditions were 1 cycle of 55°C for 30 min to generate cDNA, followed by a denaturation cycle at 94°C for 2 min, 35 cycles at 94°C for 15 sec, 55°C for 30 sec and 68°C for 2 min a final extension at 68°C for 5 min. Using the first round PCR products as template, nested PCR was carried out using TaKaRa Ex Taq HS enzyme kit (Takara, Shiga, Japan) with the following primers **Long_fwd**: 5'_GACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTGGTGAGTACGC CAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGG_3' and **Long_rev:** 5'_GGCCCAATTTTTGAAATTTTTCCTTCCTTTTCCATTTCTGTACAAATTTCTACTAATGCTTT TATTTTTTCTTCTGTCAATGGCCATTGTTTAACTTTTG_3' under the following PCR conditions: an initial denaturation at 94°C for 2 min then 40 cycles at 94°C for 30 sec, 60°C for 30 sec and 72°C for 2 min), with a final extension at 72°C for 7 min.

HIV-1 PR gene was then amplified from the HIV-1 gag-protease PCR product using the primers **pMal_Fwd:** 5'_CAGCGGCCGCGGAGAAGAAAGACAGGGAACC_3' with a *Not*1 restriction site attached to the 5' and **MdfINPR_Rev:** 5'_TACGAATTCCCTGGCTTTAATTTTACTGGTACAG_3' with an *Eco*R1 restriction site at the 5' end using the following PCR conditions: an initial denaturation cycle at 94°C for 2 min, followed by 30 cycles at 94°C for 15 sec, 58°C for 30 sec and 72°C for 45 sec, and a

final cycle of extension (72°C for 7 min). The amplified products were bulked sequenced using the ABI Prism Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, CA, USA).

2.3 Cloning, expression, and purification of HIV-1 Protease

HIV-1 PR PCR product and pMAL-c5X expression vector (New England BioLabs, MA, USA) were both digested using FastDigest restriction enzymes *Not*1 and *Eco*R1 (Thermo Fischer Scientific, MA, USA) and ligated into the pMAL-c5X expression vector according to the manufacturer's instructions. The ligation mixture was transformed to One-shot Top10 competent cells (Invitrogen, CA, USA), plated, and incubated in Luria Bertani (LB) ampicillin (100 μg/ml) agar plate overnight at 37°C. The presence of HIV-1 PR in clones was confirmed using colony PCR, and the positive clones were bulked sequenced using the ABI Prism Big Dye Terminator v3.1 Cycle Sequencing Kit. Plasmid DNA isolated from a single positive clone was then used to transform NEBExpress E. coli cells (New England BioLabs, MA, USA), cultured on ampicillin agar plates, and incubated overnight 37°C. The colonies were screened for positive clones using PCR and used for HIV-1 PR expression.

To express and purify HIV-1 PR, a single positive clone was inoculated overnight in LB ampicillin (100 μg/ml) media at 37°C shaking at 230 rpm. 10ml of the overnight culture was inoculated to l liter of LB ampicillin media containing glucose (0.2%) and then induced with 0.3 mM IPTG after 4 hours (OD_{600}=0.5). The cells were harvested after 3 h by centrifuging at 4000 x g for 20 minutes and resuspended in 25 ml of buffer A (20mM Tris-HCl, pH 7.4, 200 mM NaCl, 1 mM EDTA, 1mM DTT, 1 mM azide). The cells were sonicated (CML-4, Thermo Fisher, CA, USA) in short pulses of 15 seconds in an ice-water bath and the supernatant containing the HIV-1 PR fusion protein was collected by centrifugation at 20,000 x g for 20 minutes. The MBP tagged HIV-1 PR was then purified using the 5 ml MBP Trap HP column (GE Healthcare Life Sciences, NJ, USA) according to the manufacturer's protocol. The MBP tag was cleaved from the fusion by treating it with factor Xa (New England BioLabs, USA) followed by dialysis with buffer B (20 mM Tris-HCl, 25 mM NaCl, pH 8.0) and using the HiTrap Q FF column (GE Healthcare Life Sciences, NJ, USA). Factor Xa cleavage protease was removed using the HiTrap Benzamidine column (GE Healthcare Life Sciences, NJ, USA). The expressed HIV-1 PR samples were folded by diluting 10-fold with buffer C (0.05 M Na-acetate, 5% ethylene glycol, 10% glycerol, and 5 mM DTT, pH 5.5) (Lockbaum *et al*., 2019). For every experiment, fresh protein samples were refolded. The protein expression and purity were checked by SDS-PAGE (Laemmli, 1970) at every step, and the concentration of the protein obtained using absorbance spectroscopy at 280 nm. The free HIV-1 PR protein was further confirmed using western blot.

2.4 sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) for confirmation of HIV-1 PR expression

The purity and presence of expressed HIV-1 PR in the protein solution sample were confirmed using SDS-PAGE. Briefly, 200ng of total protein was loaded onto 12% Mini Protean® TGXTM pre-cast SDS polyacrylamide gel (BioRad Laboratories, Inc. USA) and allowed to run at 110V for 45 minutes. After electrophoresis, the gel was washed with deionized water on a shaker for 5 minutes three times. The gels were stained Biosafe Coomassie stain (BioRad Laboratories, Inc. USA), on a shaker 1 hour. The gel was destained with deionized water for 30 minutes and visualized.

2.5 Western blotting to detect the presence of HIV-1 PR

The purified HIV-1 samples (200ng) were loaded into the wells of the SDS-PAGE gel and allowed to run at 110V for 45 minutes. After SDS electrophoresis, the protein bands after SDS electrophoresis were transferred to the nitrocellulose membrane using the Trans-Blot Turbo Transfer System (Bio-Rad Laboratories, Inc. USA). The nitrocellulose membrane was blocked in 5% BSA (containing 0.1% Tween 20) for 2 hours on a shaker. The 5% BSA was discarded, and a 1:1000 dilution of the primary antibody (Anti-HIV protease, Exbio, Czech Republic) was added to the nitrocellulose membrane, left on the shaker for 1 hour, and then stored overnight at $4°C$. The nitrocellulose membrane was then subsequently washed 5 times in wash buffer (10X Tris-Buffered Saline (TBS), and a 1:1000 dilution of the secondary antibody (Anti-Human IgG H&L, HRP, Abcam, United Kingdom) was added and placed allowed on the shaker for 2 hours. The nitrocellulose membrane was then washed a second time. The Pierce ECL Western Blotting Substrate (Thermo Fisher Scientific, MA, USA) was added to the membrane and visualized in a light-sensitive film.

2.6 Enzyme activity assay and inhibition studies

HIV-1 PR activity was measured by adding purified enzyme (100 to 300nM) to 300μM of chromogenic substrate Lys-Ala-Arg-Val-Nle-p-nitro-Phe-Glu-Ala-Nle amide (Sigma-Aldrich, MI, USA) dissolved in buffer D (50 mM sodium acetate buffer containing 100 mM NaCl (pH 5.0)) at 37°C. The change in absorbance upon hydrolysis of the substrate by HIV-1 PR was monitored and an extinction coefficient of 1,800 M^{-1} cm^{-1} at wavelength 300 nm was used to convert the absorbance change to reaction rates (Velazquez-Campoy *et al*., 2001, Velazquez-Campoy *et al*., 2002). The inhibition of HIV-1 PR activity by the inhibitors was studied by measuring the enzyme activity in the presence of 0 - 500nM of LPV and DRV, using three substrate concentration: 100, 200, and 300 μM respectively, in buffer D at 37°C. The data obtained were analyzed by plotting the reciprocal of substrate hydrolysis against substrate concentration

(Lineweaver-Burk method (Lineweaver and Burk, 1934) to determine the K_m and K_{cat}) and inhibitor concentration (Dixon method (Dixon, 1953) to determine the K_i). respectively.

2.7 Fluorescent spectroscopy

Fluorescent spectra were recorded as described previously (Dash and Rao, 2001) using a PerkinElmer LS 55 spectrometer with a 1.0 cm quartz cell (Waltham, MA, USA) and connected to a thermostat-controlled bath. This assay utilized the intrinsic tryptophan (Trp) fluorescence by causing the excitation of the π–π* transition in the tryptophan residues. Two Trp residues are found at positions 6 and 42 in each monomer of HIV-1 PR. Trp 6 is located close to the active site, and Trp 42 is found close to the flap of the HIV-1 PR. The position of the Trp residues on the enzyme surface makes them good probes in monitoring changes in the HIV-1 PR tertiary structure (Dash and Rao, 2001). The protein was excited at 295 nm, and fluorescence measurements were recorded from 300 to 420 nm at room temperature (25°C). The excitation and emission slit widths were set at 5 nm, and the fluorescence spectra were acquired at 500 nm/min. The obtained fluorescence data was corrected and smoothened by running control samples of buffer. For the Fluorescence quenching study, 300nm of HIV-1 PR and varying inhibitor concentration from 10 – 500nM were used. For every reaction, a new enzyme solution was used. The decrease in intrinsic Trp fluorescence (F_0–F) at each concentration of inhibitor was fitted to the equation (F_0–F) $= \Delta Fmax/(1+(K_i/[I])$ to determine K_i and $\Delta Fmax$ values using the Origin(Pro), 2019.(OriginLab Corporation, Northampton, MA, USA). The Stern-Volmer quenching constants (K_{sv}) were also calculated by fitting the data into the equation $F_0/F = 1 + K_{sv}$ [Q]. The Stern-Volmer constant reports the accessibility of fluorophores to a quencher and the solvent accessibility of the fluorophore. Thus, it is an essential tool that can be used to probe the conformational changes around a fluorophore in proteins (Ronda *et al*., 2018). It is also an indication of the inhibitors' quenching capacity, the higher the K_{sv} value, the greater the quenching. The inner filter effect was corrected by using the formula F_c=Fantilog[(A_{ex}+A_{em})/2], "where F_c is the corrected measurement and F is the measured fluorescence intensities, respectively, A_{ex} is solution absorbance at the excitation and A_{em} emission wavelengths (Lakowicz, 2013).

2.8 Computational Methods

2.8.1 HIV-1 PR enzyme and PIs system Preparation and molecular docking

The monomeric form of wild type (WT) South African HIV-1 subtype C PR x-ray crystal structure (3U71) was retrieved from the RSCB Protein Data Bank (Burley *et al*., 2018) and converted to a dimeric structure using the UCSF Chimera software (Yang *et al*., 2012b). The mutant South African HIV-1 PR (MUT-1, MUT-2, and MUT-3) structures were obtained through homology modeling performed on the SWISS-

MODEL web server. The wild type South African HIV-1 subtype C PR x-ray crystal structure (3U71) as a template. The structure of FDA-approved protease inhibitors (PIs) DRV and LPV were obtained from PubChem (Kim *et al*., 2016), and the Avogadro software package was used to prepare the 3-D structures of the PIs (Hanwell *et al*., 2012). Molecular docking was utilized to predict the ligands' best geometric conformation within the HIV-1 PR active site. The Autodock Vina Plugin available on Chimera software was used for molecular docking (Yang *et al*., 2012b), with default docking parameters. Prior to molecular docking, Gasteiger charges were added to the HIV-1 PIs; DRV and LPV, also the non-polar hydrogen atoms, were merged to carbon atoms. The HIV-1 PIs were then docked into the HIV-1 PR binding pocket and subsequently subjected to molecular dynamic (MD) simulations.

2.8.2 MD Simulations

MD simulations were performed using the GPU version provided with AMBER 18 package. The AMBER 18 package Leap module was used for the addition of Na^+ and Cl^- ions to neutralize the system. Atomic partial charges for the ligand was generated using ANTECHAMBER, by utilizing the Restrained Electrostatic Potential (RESP) and the General Amber Force Field (GAFF) procedures. The systems were described using the AMBER ff18SB force field parameters (Nair and Miners, 2014). Amino acid residues of the proteins were renumbered based on the dimeric form of the enzyme from 1 to 198. All the systems were suspended implicitly within a orthorhombic box of TIP3P water molecules in such a manner that all atoms were within 8 Å of any box edge. Initial minimization of 2000 steps with an applied restraint potential of 500 kcal/mol for both solutes were carried out. This was performed for 1000 steps using the steepest descent method and then followed by 1000 steps of conjugate gradients. In addition, full minimization of 1000 steps were further performed by conjugate gradient algorithm without restraint. MD simulation was performed with gradual heating from 0K to 300K, executed for 50ps, such that the systems maintained a fixed number of atoms and fixed volume as previously described (Kehinde *et al*., 2019). MD simulations were performed for 700ns. Post dynamic analysis was done using CPPTRAJ modules implemented in Amber18 for analysis of the root mean square fluctuation (RMSF), root mean square deviation (RMSD), solvent accessible surface area (SASA), and radius of gyration (ROG) as described by (Shunmugam and Soliman, 2018). The active site to flap tip distances (Cα D25 – I50 (chain A), Cα D25' – I50' (chain B)) and the chain A flap tip to chain B flap tip (Cα I50 – I50') distances for the wild type (WT) and mutant PRs bound to DRV and LPV was explored. These distances are often used to determine the vertical and horizontal movement of the HIV-1 PR flap (Karnati *et al*., 2019, Chen *et al*., 2014). The Origin Pro, 6.0 (OriginLab Corp, Northampton, MA, USA) data analysis software was used to plot all the graphs (Seifert, 2014).

2.8.3 Binding Free Energy Calculations

The binding free energies of the systems were determined using the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) method (Ylilauri *et al.*, 2013). Binding free energy was averaged over 100000 snapshots extracted from the last 100ns trajectory. The free binding energy (ΔG) was computed for different molecular species (complex, ligand, and receptor) as described by the equations below (Hayes, 2011):

$$\Delta G_{bind} = G_{complex} - G_{receptor} - G_{ligand} \quad (1)$$
$$\Delta G_{bind} = E_{gas} + G_{sol} - TS \quad (2)$$
$$E_{gas} = E_{int} + E_{vdw} + E_{ele} \quad (3)$$
$$G_{sol} = G_{GB} + G_{SA} \quad (4)$$
$$G_{SA} = \gamma SASA \quad (5)$$

Where E_{gas} is the gas-phase energy and encompasses the internal energy E_{int}; Coulomb energy E_{ele} as well as the van der Waals energies E_{vdw}. The E_{gas} was determined directly from the FF14SB force field terms. The solvation free energy, G_{sol}, is a combination of the energy contribution from the polar states, G_{GB}, and the non-polar states, G. The non-polar solvation energy G_{SA}, was determined from the solvent-accessible surface area (SASA), using a water probe radius of 1.4 Å, while the polar solvation, G_{GB}, was calculated using the GB equation. S and T respectively indicate the total entropy of the solute and temperature.

2.9 Statistics analysis

Enzyme kinetic parameters were evaluated by computer fitting the obtained data using the Origin (Pro), 2019 (OriginLab Corporation, Northampton, MA, USA) and SigmaPlot version 14.5 (Systat Software, Inc., San Jose California USA). The enzyme kinetics experiments were carried out in triplicate, and the average values were used in this study

3. Results

3.1 Cloning and expression of HIV-1 PR

The amplified HIV-1 PR PCR product was visualized using agarose (1%) gel electrophoresis (Figure 2A and Figure S1) and confirmed to be approximately 297 base pairs (bp). The wild type and mutant HIV-1 PR genes were successfully cloned into the pMAL expression vector fused to the MBP tag. Colony PCR

and Sanger sequencing results confirmed the amplified HIV-1 PR gene and recombinant colonies containing the variants of HIV-1 PR. Furthermore, the HIV-1 PR sequences were uploaded onto the HIV drug resistance database (Shafer *et al.*, 2000) for confirmation of the mutations harbored. The expressed and purified MBP tagged HIV-1 PR had a molecular weight of approximately 55kDa (Figure 2B and Figure S2). After the fusion tag's cleavage (Figure 2C and Figure S3), the free HIV-1 PR had an approximate molecular weight of 11kDa (Figure 2D). The western blot confirmation of the expressed HIV-1 PR is shown in Figure 2E (full blot picture in Figure S4).
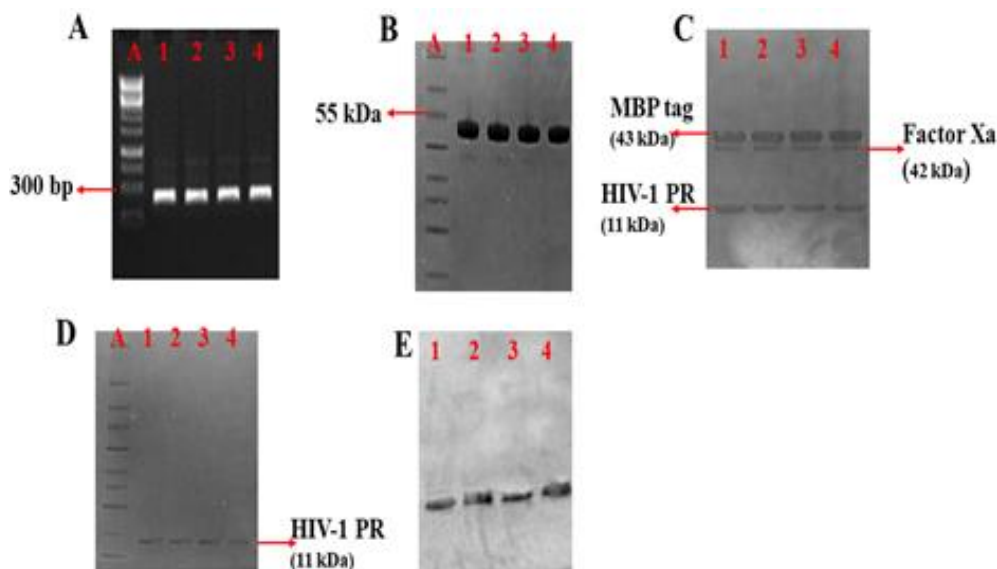
Figure 2. PCR amplification, expression, and purification of HIV-1 PR. (A) Amplified HIV-1 PR gene; DNA marker (Lane A), amplified HIV-1 PR wild type (WT) (Lane 1), amplified mutant HIV-1 PR variants (Lane 2 = MUT-1, Lane 3 = MUT-2, and Lane 4 = MUT-3). (B) SDS-PAGE showing purified MBP tagged HIV-1 PR variants: Lane A (protein marker), Lane 1 (WT), Lane 2 (MUT-1), Lane 3 (MUT-2), and Lane 4 (MUT-3). (C) SDS-PAGE showing cleavage products after Factor Xa cleavage of HIV-1 PR from the MBP tag: Lane 1 (WT), Lane 2 (MUT-1), Lane 3 (MUT-2), and Lane 4 (MUT-3). (D) SDS-PAGE showing purified HIV-1 PR variants: Lane A (protein marker), Lane 1 (WT), Lane 2 (MUT-1), Lane 3 (MUT-2), and Lane 4 (MUT-3). (E) Western blot picture for the HIV-1 PR variants, Lane 1 (WT), Lane 2 (MUT-1), Lane 3 (MUT-2), and Lane 4 (MUT-3).

3.2 Enzyme activity of wild type HIV-1 PR and mutants

The enzyme activity assay results obtained from the hydrolysis of the chromogenic substrate showed that the wild type (WT) HIV-1 PR had about a 2-fold higher ($K_m = 37.49 \pm 0.63$ µM) affinity than the mutant PRs (K_m for MUT-1 = 67.78 ± 1.22 µM, K_m for MUT-2 = 67.46 ± 1.48 µM, and K_m for MUT-3 = 70.59 ± 1.01 µM) for the chromogenic substrate (Table 1, Figure 3A – D). In addition, the catalytic constant (K_{cat}) of the wild type HIV-1 PR ($K_{cat} = 0.79 \pm 0.11$ s^{-1}) was almost 2-fold higher than the mutants (K_{cat} for MUT-1 = 0.48 ± 0.10 s^{-1}, K_{cat} for MUT-2 = 0.44 ± 0.01 s^{-1}, K_{cat} for MUT-3 = 0.39 ± 0.01 s^{-1}). The higher K_m and lower K_{cat} values for the mutant HIV-1 PR resulted in a lower catalytic efficiency (K_{cat}/K_m) than the WT

HIV-1 PR. The catalytic efficiency (K_{cat}/K_m) of the WT HIV-1 PR (0.021 ± 0.003 $_s{}^{-1}\mu M^{-1}$) was markedly higher than in the mutant variants (K_m/K_{cat} for MUT-1 = 0.0071 ± 0.001 $_s{}^{-1}\mu M^{-1}$, K_m/K_{cat} for MUT-2 = 0.0065 ± 0.001 $_s{}^{-1}\mu M^{-1}$, and K_m/K_{cat} for MUT-3 = 0.0055 ± 0.001 $_s{}^{-1}\mu M^{-1}$).

3.3 Inhibition of wild type HIV-1 PR and mutants by LPV and DRV

The inhibition data for the WT and mutant HIV-1 PR by LPV and DRV are shown in Table 1 and Figures 3E – H. Both LPV and DRV effectively inhibited the wild type and mutant proteases. However, it was observed that DRV (K_i =1.58 ± 0.11 nM) was more potent than LPV (K_i =2.13 ± 0.23 nM) in inhibiting the activity of wild type protease. Both the drugs showed higher K_i values for mutants as compared to the wild type. The K_i values for the inhibition of the mutants by LPV (MUT-1 = 46.50 ± 0.14 nM, MUT-2 = 52.63 ± 0.65 nM, and MUT-3 = 76.26 ± 0.09 nM) were about 6 – 8-fold higher than the K_i value for DRV (MUT-1 = 5.53 ± 0.09 nM, MUT-2 = 7.80 ± 0.71 nM MUT-3 = 11.53 ± 1.09 nM).

Table 1. Wild type and mutant HIV-1 PR variants enzyme kinetic parameters (K_m, K_{cat}) and inhibition constant (K_i) calculated using the chromogenic substrate.

HIV-1 PR Variants	K_m (µM)	K_{cat} ($_s{}^{-1}$)	K_{cat}/K_m ($_s{}^{-1}\mu M^{-1}$)	LPV		DRV	
				K_i (nM)	Relative resistance to LPV	K_i	Relative resistance to DRV
WT	37.49 ± 0.63	0.79 ± 0.11	0.021 ± 0.003	2.13 ± 0.23	1.00	1.58 ± 0.11	1.00
MUT-1	67.78 ± 1.22	0.48 ± 0.10	0.0071 ± 0.001	46.50 ± 0.14	21.83	5.53 ± 0.09	3.50
M46I, I54V, V82A, L10F							
MUT-2	67.46 ± 1.48	0.44 ± 0.01	0.0065 ± 0.001	52.63 ± 0.65	24.71	7.80 ± 0.71	4.94
M46I, I54V, L76V, V82A, L33F, L10F							
MUT-3	70.59 ± 1.01	0.39 ± 0.01	0.0055 ± 0.001	76.26 ± 0.09	35.80	11.53 ± 1.09	7.30
M46I, I54V, L76V, V82A, L90M, F53L							

Relative resistance = Ki of mutant / Ki of WT
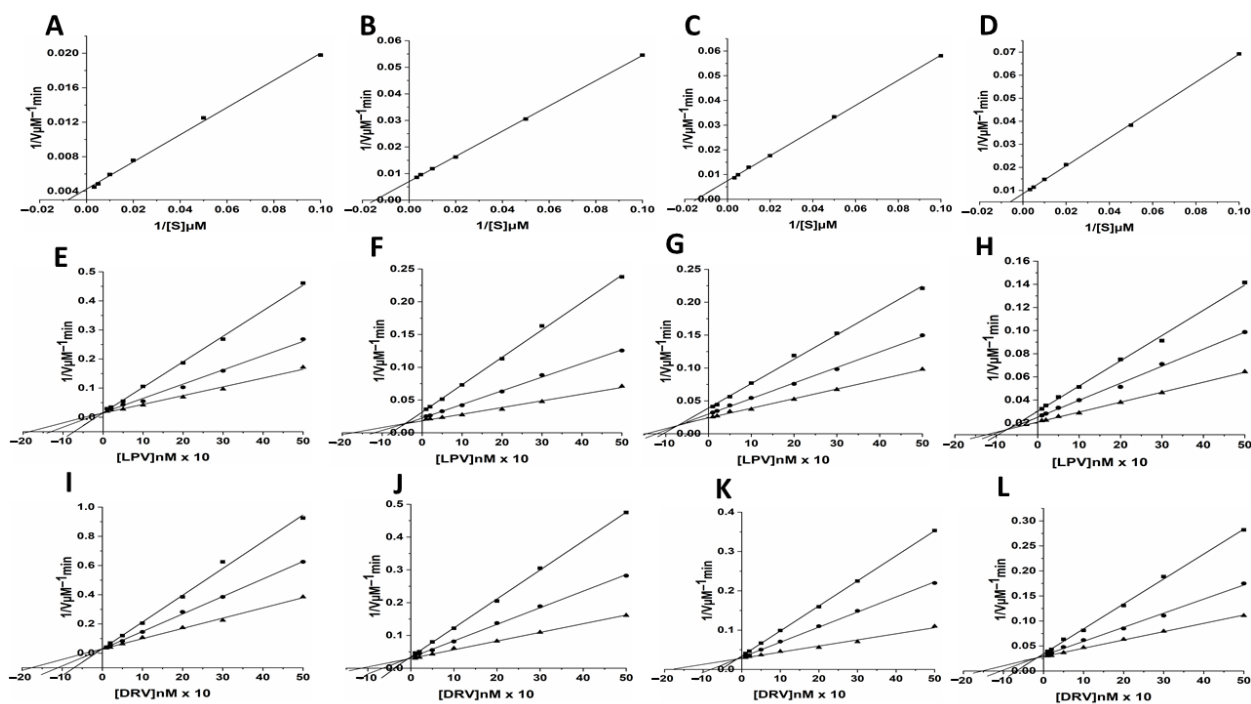
Figure 3. (A–D) Enzyme kinetic activity of WT (A) and mutant HIV-1 PR variants (B = MUT-1, C = MUT-2, and D = MUT-3) was determined following the hydrolysis of the chromogenic synthetic substrate (Lys-Ala-Arg-Val-Nle-p-nitro-Phe-Glu-Ala-Nle amide). (E–H) The activity of the HIV-1 PR variants (E = WT, F = MUT-1, G = MUT-2, and H = MUT-3) measured in the presence of 10–500 nM lopinavir (LPV) using three substrate concentration: 100 (■), 200 (●), and 300 μM (▲) respectively. (I–L) The activity of the HIV-1 PR variants (I = WT, J = MUT-1, K = MUT-2, and L = MUT-3) measured in the presence of 10–500 nM darunavir (DRV) using three substrate concentration: 100 (■), 200 (●), and 300 μM (▲) respectively.

3.4 Fluorescence spectroscopy

Fluorescence spectroscopy was used to investigate secondary and tertiary conformational changes induced in HIV-1 PR due to the binding of the inhibitors LPV and DRV. The WT and mutant HIV-1 PR variants exhibited maximal fluorescence emission spectra (λ_{max}) at 351 nm due to the radiative decay associated with the π–π* transition state of HIV-1 PR Trp residues indicating that the tryptophan residue environment is hydrophilic in nature (Figure 4). The intrinsic Trp fluorescence intensity of the mutant HIV-1 PR variants was increased (14% to 26%) relative to that of the wild type. A concentration-dependent tryptophanyl fluorescence quenching was observed upon titration of HIV-1 PR with the inhibitors LPV and DRV (figures 5 A – D and 6 A – D). There was no red or blue shift in the λ_{max} observed upon increase in the concentration of either of the inhibitors, indicating that the enzymes' secondary structures remain intact (Dash and Rao, 2001). In addition, the gradual decrease in fluorescence intensity of the HIV-1 PR variants observed as a result of the increase in the concentration of LPV and DRV (Figure 4 and 5 A – D) is due to the formation of the enzyme-inhibitor complex (Bekale *et al*., 2014).
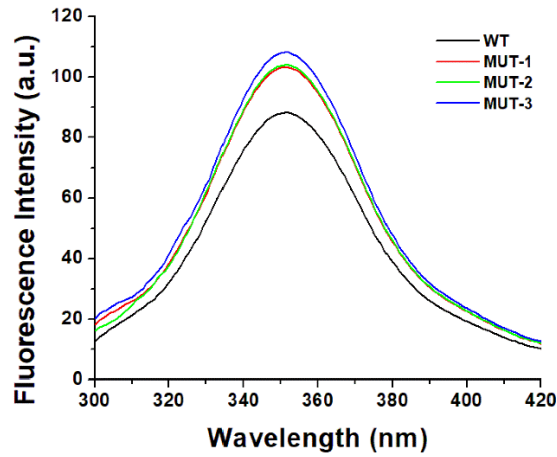
Figure 4. Intrinsic tryptophan fluorescent graph for wild type and mutant HIV-1 PR variants.
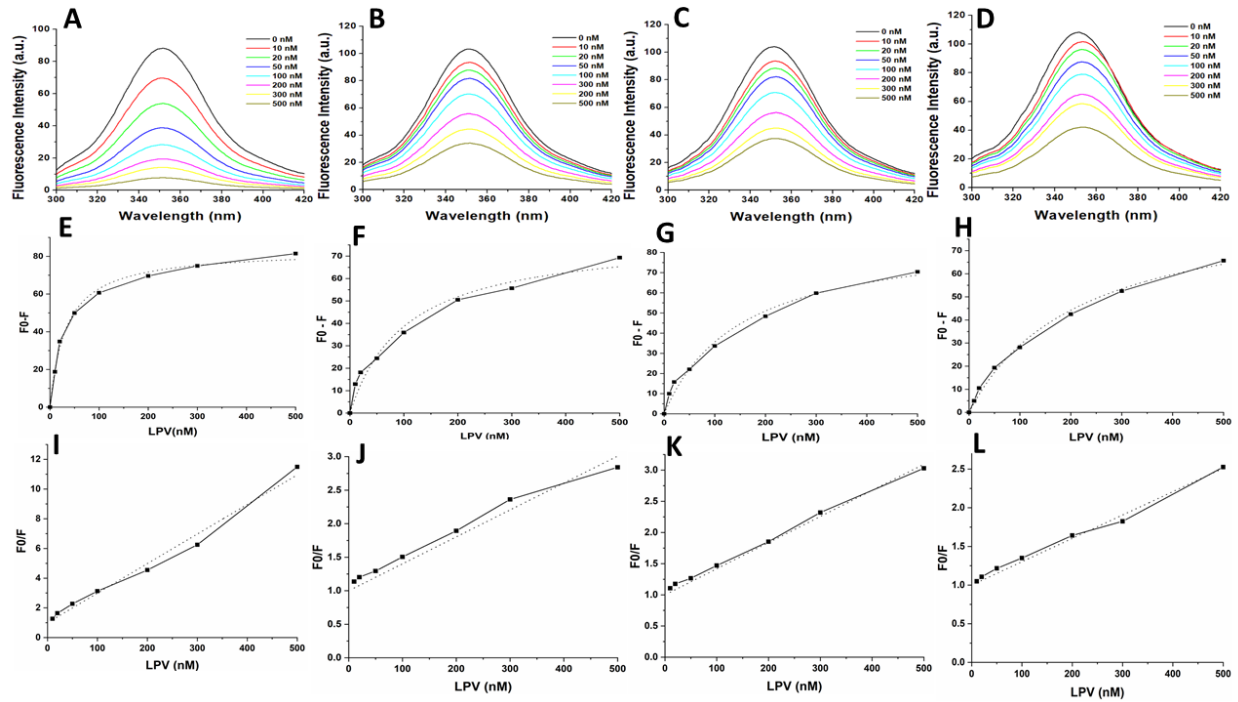
Figure 5. (A–D) Intrinsic tryptophan fluorescence quenching using LPV for the WT, MUT-1, MUT-2, and MUT-3 re-spectively. (E–H) Change in intrinsic tryptophan fluorescence for determination of Ki using LPV for the WT, MUT-1, MUT-2, and MUT-3 respectively. (I–L) Stern–Volmer plot to determine the quenching constants (Ksv) using LPV for the WT, MUT-1, MUT-2, and MUT-3 respectively.
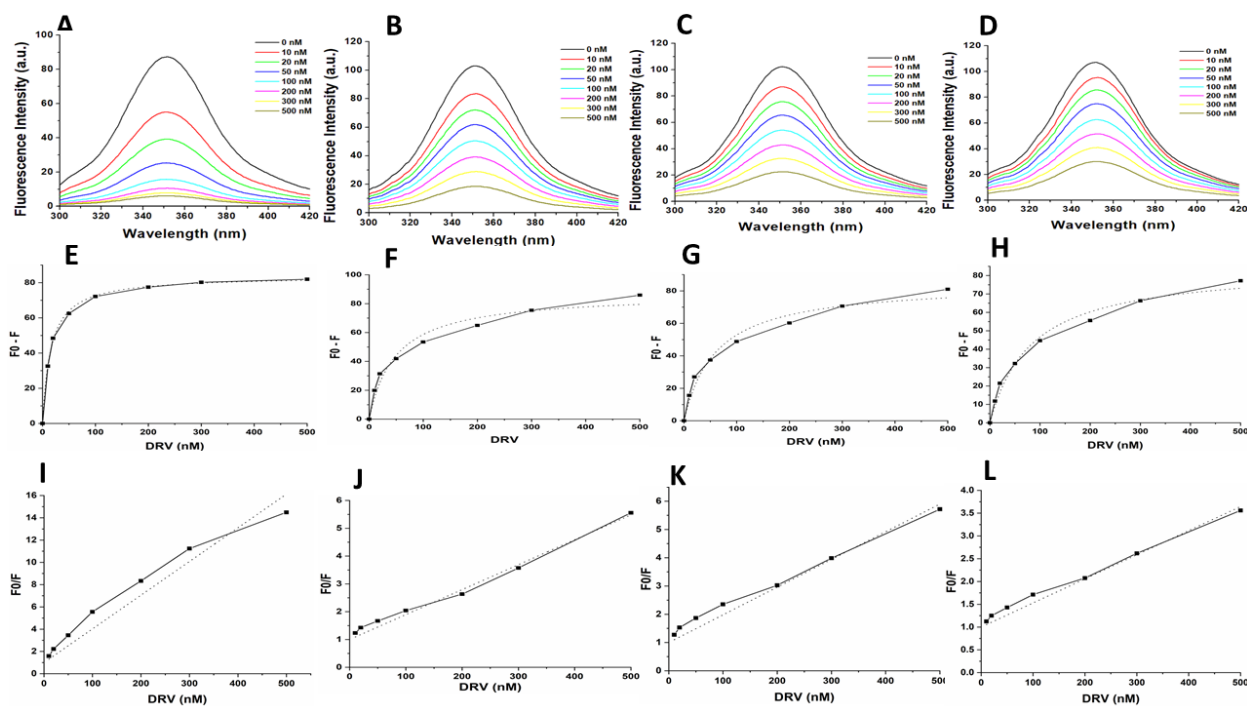
Figure 6. (A–D) Intrinsic tryptophan fluorescence quenching using DRV for the WT, MUT-1, MUT-2, and MUT-3 re-spectively. (E–H) Change in intrinsic tryptophan fluorescence for determination of Ki using DRV for the WT, MUT-1, MUT-2, and MUT-3 respectively. (I–L) Stern–Volmer plot to determine the quenching constants (Ksv) using DRV for the WT, MUT-1, MUT-2, and MUT-3 respectively.

The calculated K_i values for the WT and mutant PRs from the obtained florescence data are shown in Table 2 and figures 5 and 6 E – H. The WT HIV-1 PR K_i value (17.25 nM) for LPV was found to be about 3 – 7-fold lower than in the mutant HIV-1 PR variants (MUT-1 = 54.74 nM, MUT-2 = 79.47 nM, and MUT-3 = 113.16 nM). The WT HIV-1 PR K_i value (8.12 nM) for DRV was approximately 3 – 6-fold lower than the mutants (MUT-1 = 26.34 nM, MUT-2 = 32.85 nM, and MUT-3 = 44.70 nM). The fluorescence data obtained was also analyzed to calculate the Stern-Volmer constant (K_{sv}) by plotting linear Stern-Volmer plots (F_0/F vs inhibitor concentrations), (Figure 5 and 6 I – L) (Maseko *et al.*, 2017). It was observed that the K_{sv} value (Table 2) for the interaction of LPV with the WT (0.02 nM^{-1}) was higher than the mutant HIV-1 PRs (MUT-1 = 0.004 nM^{-1}, MUT-2 = 0.004nM^{-1}, and MUT-3 = 0.003 nM^{-1}) and similar results were observed for DRV interaction with the WT (0.03nM) and mutants (0.01nM^{-1}, 0.009nM^{-1}, and 0.005 nM^{-1} for MUT-1, MUT-2, and MUT-3, respectively (Table 2).

Table 2. Inhibition constant (K_i) and Stern-Volmer quenching constants (K_{sv}) calculated from fluorescence quenching assay.

HIV-1 PR Variants	LPV		DRV	
	K_i (nM)	K_{sv} (nM^{-1})	K_i (nM)	K_{sv} (nM^{-1})
WT	17.25	0.02	8.12	0.030
MUT-1	54.74	0.004	26.34	0.01
M46I, I54V, V82A, L10F				
MUT-2	79.47	0.004	32.85	0.009
M46I, I54V, L76V, V82A, L10F, L33F				
MUT-3	113.16	0.003	44.70	0.005
M46I, I54V, L76V, V82A, L90M, F53L				

3.5 Molecular dynamic simulation

3.5.1 Stability of WT, MUT-1, MUT-2 and MUT-3-inhibitor complex

The dynamic stability of the MD simulation was evaluated using the root mean square deviation (RMSD) of backbone carbon atoms of the different HIV-1 PR variants in complex with LPV and DRV (Figure 7A and B). The lower the RMSD, the more stable the protein complex. The RMSD for the HIV-1 PR-DRV complexes was relatively stable compared to the HIV-1 PR-LPV complexes. The fluctuation in HIV-1 PR amino acid residues as they interact with LPV and DRV throughout the trajectory was monitored using root mean square fluctuation (RMSF) of Cα atoms (figures 7C and D). This gives an insight into the structural flexibility of the different regions of the HIV-1 PR variants. Interestingly, marked fluctuation in the flap residues (residues 45 – 55/ 45' – 55') of the mutant HIV-1 PR variants in complex with LPV and DRV was observed compared to the wild type. A similar fluctuation was seen around the 80's loop in the mutant HIV-1 PR-LPV complexes.

3.5.2 Solvent exposure and radius of gyration of WT and mutant HIV-1 PRs

This study determined the solvent-accessible surface area (SASA) of WT and different mutant HIV-1 PRs to LPV and DRV to gain clarity about the hydrophobic core compactness of HIV-1 PR-LPV and DRV complexes. In the HIV-1 PR-LPV complexes (Figure 7E), the SASA for the WT PR (8620.53A^2) was lower compared to the SASA of the different mutants (9562.60, 9796.39, and 10090.94 A^2 for MUT-1, MUT-2, and MUT-3 respectively). Similarly, the mean SASA for the WT-DRV complex (Figure 7F) (8899.49 A^2) was lower than the different mutants (9961.57, 9605.59, and 9786.64 A^2 for MUT-1, MUT-2, and MUT-3, respectively). This high SASA in the mutant LPV and DRV HIV-1 PR complexes may be due to a

destabilization of the hydrophobic core. Further confirmation of the instability and gain in flexibility of the mutant HIV-1 PR variants was obtained from the ROG. There was an increase in the ROG for the mutants compared to the WT-LPV and DRV complex (Figure 7G and H). The mean ROG value of the WT HIV-1 PR complexed to LPV (Figure 7G) was found to be 17.28 ± 0.14 Å, and 18.39 ± 0.38, 18.73 ± 0.45, and 18.64 ± 0.34 Å, respectively, for MUT-1, MUT-2, and MUT-3. In the HIV-1 PR-DRV complexes (Figure 7H), the mean ROG value for the WT was found to be 16.82 ± 0.09Å, and 17.94 ± 0.1, 18.16 ± 0.15, and 18.25 ± 0.29 Å, respectively, for MUT-1, MUT-2, and MUT-3.
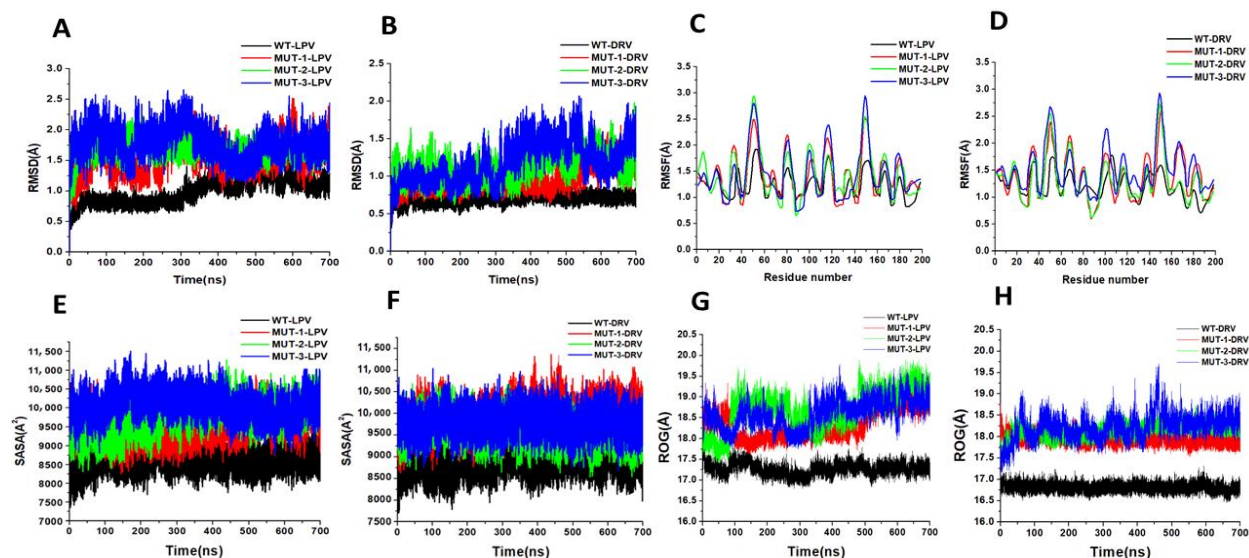
Figure 7. (A) Root mean square deviation (RMSD) for WT and mutant HIV-1 PR variants bound to LPV. (B) RMSD for WT and mutant HIV-1 PR variants bound to DRV. (C) RMSF Figure 1. PR variants bound to LPV. (D) RMSF for WT and mutant HIV-1 PR variants bound to DRV. (E) Solvent-accessible surface area (SASA) for WT and mutant HIV-1 PR variants bound to LPV. (F) SASA for WT and mutant HIV-1 PR variants bound to DRV. (G) Radius of gyration (ROG) for WT and mutant HIV-1 PR variants bound to LPV. (H) ROG for WT and mutant HIV-1 PR variants bound to DRV.

3.5.3 HIV-1 PR Flap dynamics

3.5.3.1 Distance between active site residue to flap tip residue

The distance frequency distribution between the Cα D25 – I50 (chain A) in the WT and mutant HIV-1 PR-LPV complexes is plotted in Figure 8A, and the highest peak values are 12.01, 7.82, 7.73, and 8.40 Å for the WT, MUT-1, MUT-2 and MUT-3, respectively. The distance frequency distribution between the Cα D25 – I50 (chain A) in the HIV-1 PR-DRV complexes is plotted in Figure 8B, and the highest peak values are 11.33, 7.86, 8.0, and 8.8 Å for the WT, MUT-1, MUT-2, and MUT-3, respectively. The distance between active site to flap tip distance for chain B (D25' – I50') for the HIV-1 PR-LPV complexes (Figure 8C) is 11.60, 8.67, 7.87, and 8.53 Å for WT, MUT-1, MUT-2 and MUT-3, respectively. The frequency

distribution of the distance between Cα D25' – I50' (chain B) for the HIV-1 PR-DRV complexes are shown in Figure 8D. These values are 12.13, 9.73, 9.20, and 8.13 Å for WT, MUT-1, MUT-2, and MUT-3. The result obtained showed the Cα D25 – I50 and Cα D25' – I50' distance distribution for the mutant HIV-1 PRs are significantly narrower than in the WT for both the DRV and LPV WT complexes, which is an indication that the presence of these mutations in HIV-1 PR impact the binding of LPV and DRV by causing a compression of the hydrophobic cavity, thus reducing the active site volume of the mutant HIV-1 PRs.
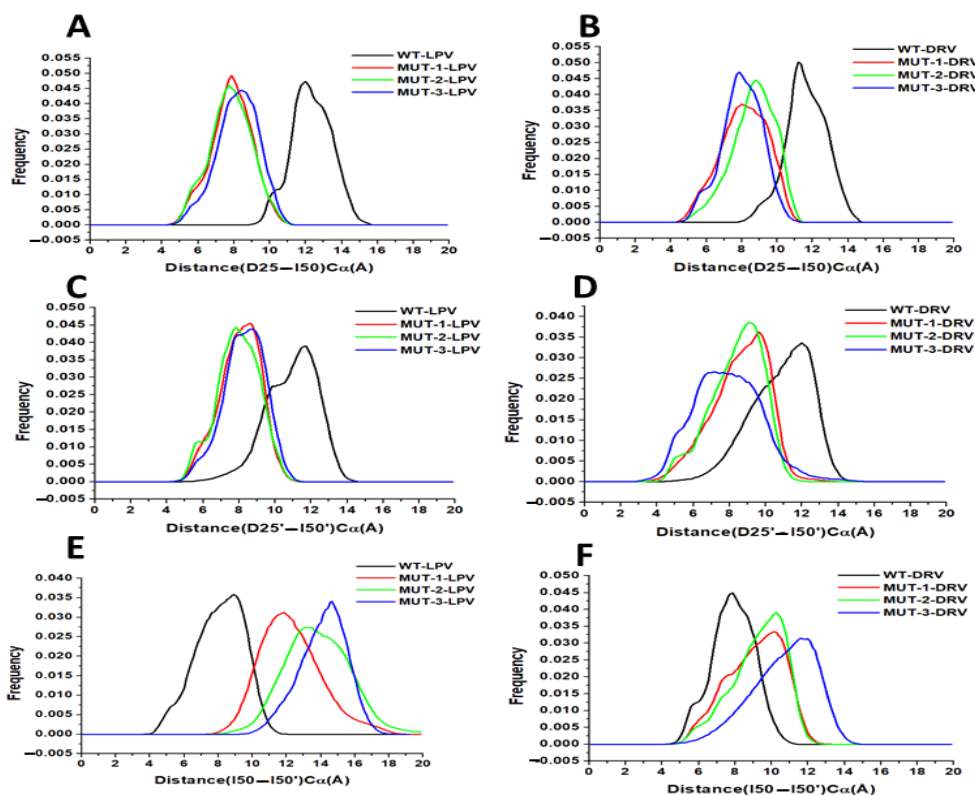
Figure 8. (A) Histogram distribution of D25-I50 distance for WT and mutant HIV-1 PR variants interaction with LPV (B) Histogram distribution of D25-I50 distance for WT and mutant HIV-1 PR variants interaction with DRV (C) Histogram distribution of D25'-I50' distance for WT and mutant HIV-1 PR variants interaction with LPV (D) Histogram distribution of D25'-I50' distance for WT and mutant HIV-1 PR variants interaction with DRV (E) Histogram distribution of I50-I50' distance for WT and mutant HIV-1 PR variants interaction with LPV (F) Histogram distribution of I50-I50' distance for WT and mutant HIV-1 PR variants interaction with DRV.

3.5.3.2 Flap tip to flap tip distance

This study explored the relative motion of the flap tips; this is the distance between Cα I50 – I50'. The distance frequency distribution plot between flap tips in the HIV-1 PR LPV complexes (Figure 8E) is 8.93, 11.73, 13.33, and 14.53 Å for WT, MUT-1, MUT-2, and MUT-3, respectively, and 7.87, 10.13, 10.40, and 11.73 Å for WT, MUT-1, MUT-2 and MUT-3 respectively for the HIV-1 PR DRV complexes (Figure 8F). The narrower flap tip to flap tip distance seen between the WT HIV-1 PR-LPV and DRV complexes suggests that these inhibitors bind tightly. The large distances between flap tips seen in the mutant PRs

indicate open movements in the flap tips and loose binding of these PIs to the mutant HIV-1 PRs. The observed decrease in the Cα D25 – I50 and Cα D25' – I50' distance and increase in Cα I50 – I50' in MUT-1, MUT-2 and MUT-3 DRV/LPV complexes is in agreement with the earlier observation showing high RMSF around the flap residues (residues 45 – 55/ 45' – 55') in the mutant HIV-1 PR-LPV and DRV complexes.

3.5.4 Structural comparison of WT and mutant HIV-1 PR

Structural comparison of the structures of HIV-1 PR WT and mutants, when bound to LPV and DRV, is shown in Figure 9 (generated during the last 20ns of the simulation). It can be observed that the flaps of the mutant variants (MUT-1, MUT-2, MUT-3) when bound to LPV and DRV were in an open conformation compared to the closed conformation in the wild type. In addition, the overall structures of the mutant versus the wild seem altered, which may be due to the impact of the mutations, causing a reorganization of the HIV-1 PR structure. This finding may be associated with the fluctuation in the flap region seen in the RMSF and the high RMSD observed in the mutant HIV-1 PR variants compared to the wild type.
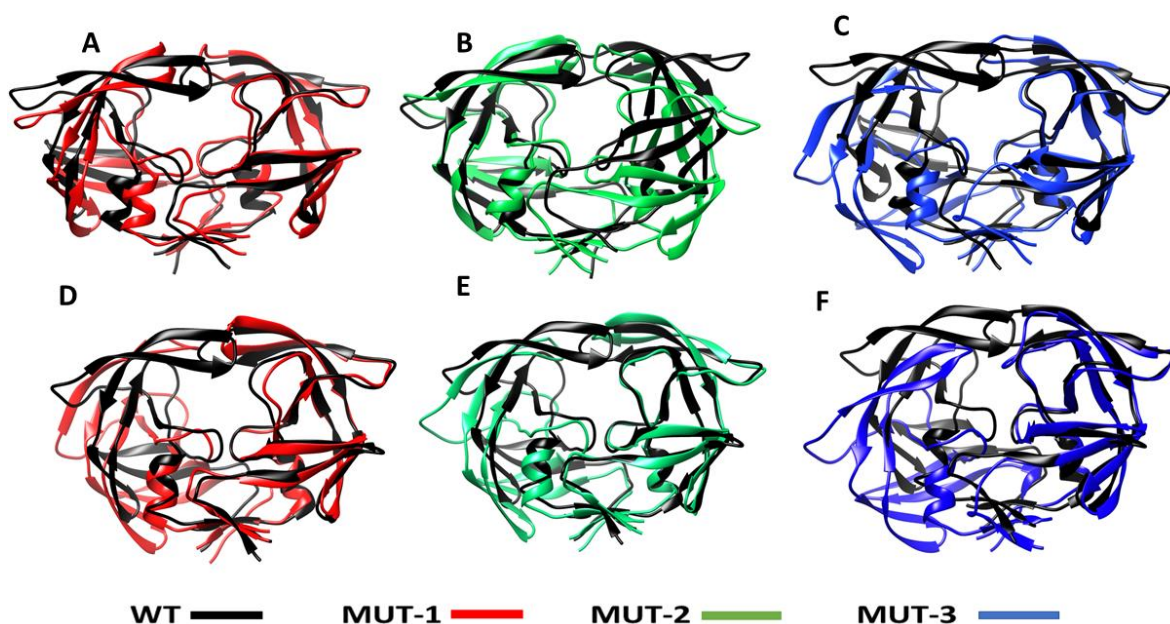
Figure 9. Superimposed structure of WT and the different mutants LPV and DRV complexes. (A) Superimposed structure of WT-LPV and MUT-1-LPV complex (B) Superimposed structure of WT-LPV and MUT-2-LPV complex (C) Superimposed structure of WT-LPV and MUT-3-LPV complex (D) Superimposed structure of WT-DRV and MUT-1-DRV complex (E) Superimposed structure of WT-DRV and MUT-2-DRV complex (F) Superimposed structure of WT-DRV and MUT-3-DRV complex.

3.5.5 HIV-1 PR binding profile calculated from MMGBSA

In order to determine the impact of the mutations on the HIV-1 PR binding landscape, binding free energy of LPV and DRV to the WT and mutant HIV-1 PRs was calculated using the MM-GBSA method (Table 3). The binding energies of LPV (-43.25 kcal/mol) and DRV (-48.19 kcal/mol) to WT were high compared to the binding energies of these inhibitors to the different mutant PRs. The binding energies of DRV to the different mutant PRs (-31.51, -24.43, and -21.58 kcal/mol for MUT-1, MUT-2, and MUT-3, respectively) were found to be higher than the binding energies of LPV to mutants (-25.39, -26.67 and -20.28 kcal/mol for MUT-1, MUT-2 and MUT-3, respectively). The increased binding energy in the HIV-1 PR-DRV complexes may be attributed to the relatively low solvation energy and increased electrostatic interaction compared to the HIV-1 PR-LPV complexes.

Table 3. Binding free energies of DRV and LPV to HIV-1 PR variants

Energy components	PI	WT-DRV	MUT-1-DRV	MUT-2-DRV	MUT-3-DRV
ΔE_{vdw}	LPV	-53.35 ± 7.97	-34.67 ± 5.23	-34.46 ± 4.39	-32.96 ± 8.03
	DRV	-51.27 ± 5.51	-36.10 ± 5.66	-32.46 ± 5.23	-33.13 ± 8.31
ΔE_{elec}	LPV	-26.98 ± 4.65	-25.89 ± 8.30	-24.48 ± 5.06	- 25.12 ± 6.21
	DRV	-29.75 ± 8.55	-27.32 ±7.74	-28.91 ± 6.68	-24.38 ± 8.30
ΔG_{gas}	LPV	-80.33 ± 13.33	-60.56 ± 9.64	-58.94 ± 8.28	-59.08 ± 5.29
	DRV	-81.02 ± 11.63	-65.04 ± 10.43	-61.37 ± 9.80	-57.53 ± 14.07
ΔG_{solv}	LPV	37.08 ± 8.12	35.17 ± 5.76	35.27 ± 7.93	38.80 ± 6.11
	DRV	31.83 ± 5.68	34.03 ± 7.17	35.92 ± 5.21	35.93 ± 7.96
ΔG_{bind}	LPV	-43.25 ± 12.30	-25.39 ± 7.76	-23.67 ± 4.49	-20.28 ± 5.53
	DRV	-48.19 ± 9.28	-31.51 ± 6.81	-24.43 ± 6.05	-21.58 ± 7.58

ΔE_{vdw} = van der Waals free energy; ΔE_{elec} = electrostatic free energy; ΔG_{gas} = gas phase Gibbs free energy; ΔG_{solv} = solvation energy

3.5.5 Hydrogen bond interaction analysis

To further determine the level of interaction and stability between HIV-1 PR-LPV and DRV complexes, hydrogen bond analysis of the snapshots from the last 20ns was analyzed using discovery studio. Table 4 and Figure 10 shows the key hydrogen bond interactions that were found between LPV and DRV complexes of the WT, MUT-1, MUT-2, and MUT-3 PRs, respectively. The hydrogen bond distances are also presented in Table 4. In the HIV-1 PR-LPV complexes, it was observed that the LPV formed hydrogen bonds with residues ARG 8, ARG 107, ASP 25, and ASP 29 in the wild type. Most of these bonds were lost in the mutant HIV-1 PR variants. LPV formed hydrogen bonds with ARG 8 and ILE 50 in the MUT-1-LPV complex, GLY 48 in the MUT-2-LPV complex, and ASP 25 and GLY 150 in the MUT-3-LPV complex.

In the WT-DRV complexes, DRV formed hydrogen bonds with residues ASP 124, VAL 32, VAL 82, and ILE 149. In the mutant HIV-1 PR-DRV complexes, fewer hydrogen bonds were formed between the drug and the protein. DRV formed hydrogen bonds with GLY 48, ILE 50, and PRO 79 in the MUT-1-DRV complex, ARG 8 and ILE 50 in the MUT-2-DRV complex, and ASP 124, GLY 48, ILE 50 in the MUT-3-DRV complex.

Table 4. Key hydrogen bond interactions between active site residues of WT, MUT-1, MUT-2 and MUT-3 with LPV and DRV.

Hydrogen bond interaction	Distance (Å)			
	WT-LPV	MUT-1-LPV	MUT-2-LPV	MUT-3-LPV
ARG8:HH21 – LPV:O5	2.59	1.93	-	-
ARG107:HH12 – LPV:O3	1.77	-	-	-
ARG107:HH22 – LPV:O3	2.57	-	-	-
ASP25:OD2 – LPV:H27	1.95	-	-	-
ASP25:OD1 – LPV:H26	-	-	-	2.08
ASP29:OD1 – LPV:H26	2.02	-	-	-
ASP29:H – LPV:OD3	2.04	-	-	-
GLY 48:H – LPV:O2	-	-	2.30	-
ILE50:H – LPV:O1	-	2.20	-	-
GLY 150:H – LPV:O4	-	-	-	1.93
Hydrogen bond interaction	**Distance (Å)**			
	WT-DRV	MUT-1-DRV	MUT-2-DRV	MUT-3-DRV
ARG8:NH1 – DRV:O2	-	-	2.12	-
ASP124:OD2 – DRV:H14	2.17	-	-	-
ASP124:OD2 – DRV:H20	1.69	-	-	-
ASP124:OD2 – DRV:H36	-	-	-	1.82
VAL32:O – DRV:H36	2.65	-	-	-
GLY 48:H – DRV:O6	-	2.73	-	-
GLY 48:O – DRV:H14	-	-	-	1.91
ILE50:H – DRV: O2	-	2.06	2.04	2.31
PRO79:O – DRV:H36	-	2.03	-	-
VAL82:O – DRV:H36	1.78	-	-	-
ILE 149:H – DRV:O7	2.95	-	-	-
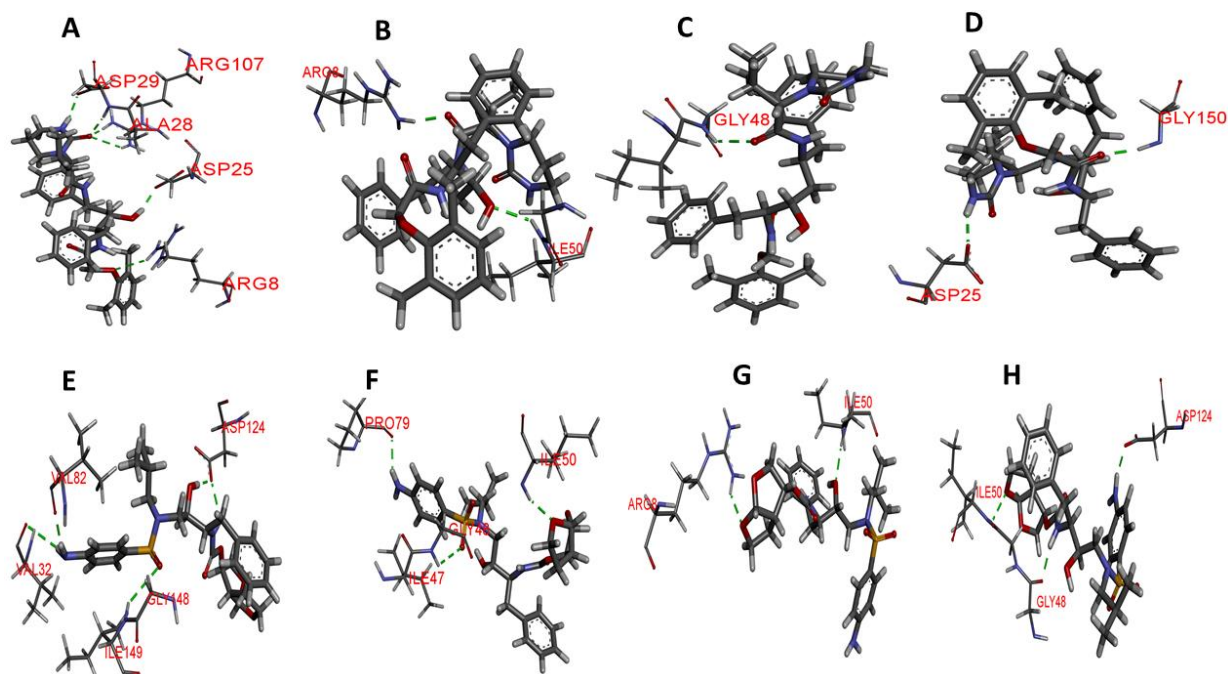
Figure 10. Hydrogen bond interaction of wild type and mutant HIV-1 PR variants with LPV and DRV. The green broken lines represent the hydrogen bond between amino acid residues of HIV-1 PR and the inhibitors and (A – D) Hydrogen bond interaction of WT, MUT-1, MUT-2, and MUT-3 with LPV, respectively. (E – H) Hydrogen bond interaction of WT, MUT-1, MUT-2, and MUT-3 with DRV, respectively.

4. Discussion

This study describes the biochemical and structural characteristics of multidrug-resistant HIV-1 PR, cloned from clinical isolates obtained from HIV-1 infected individuals at the point of switching from LPV to DRV-based regimen. This study used a prokaryotic host (Escherichia Coli) expression system to express and characterize the HIV-1 PR variants. The eukaryotic host expression systems (yeast and mammalian cells) have been used basically to study HIV-1 PR-PI drug susceptibility. However, the Escherichia Coli expression system has been extensively used for functional and structural characterization of HIV-1 PR (see review (Eche and Gordon, 2021)). HIV-1 PR was expressed fused to the MBP tag in this study. The choice of the MBP tag for the expression of the HIV-1 PR used in this study was due to the reported high yield and increased specific activity of HIV-1 PR expressed fused to MBP tag compared to other tags as reported in previous studies (see table 1 in chapter two). We found that the combination of the mutations harbored by the mutant HIV-1 PR variants significantly impacted enzyme catalytic activity. The Kcat/Km progressively decreased with an increase in the number of drug-resistant mutations. Despite the reduced affinity and catalytic efficiency, the mutants could still cleave the substrate. This can be explained by an earlier study by Prabu-Jeyabalan *et al.*; these authors showed that the interaction between the active site

residues and the substrate is relatively conserved even in the presence of mutations, thus favoring substrate cleavage, but the binding landscape of HIV-1 PIs is altered (Prabu-Jeyabalan *et al.*, 2006).

The inhibitory constants (K_i) show that the mutations severely affected the K_i for both LPV and DRV (Table1, Figure 3 E – L). In contrast to the progressive decrease observed in the K_{cat}/K_m, there was a marked increase in the K_i associated with an increase in the number of mutations for LPV and DRV. This interdependence of the K_i value on the K_{cat}/K_m may play a role in the altered recognition of PIs and their binding to the active site. The results of similar studies also show a marked decrease in the catalytic efficiency of the multidrug-resistant HIV-1 PR variants relative to an increase in the K_i of PIs (Šašková *et al.*, 2008, Ohtaka *et al.*, 2003, Gulnik *et al.*, 1995). This indicates the trade-off of catalytic efficiency for decreased inhibitory activity of PIs may be a means of responding to drug selection pressure, which may confer an evolutionary advantage on the mutated HIV-1 PR (Fernàndez *et al.*, 2007). The error-prone nature of HIV-1 replication places HIV-1 under strong selective pressure resulting in the rapid accumulation of drug resistance mutations (Parera *et al.*, 2007). Under selective pressures, proteins evolve to become tolerant to mutations and remain catalytically viable despite the impact of the mutations on their catalytic efficiency (Guo and Bi, 2002, Parera *et al.*, 2007). The catalytic efficiency of the mutant HIV-1 PR variants in this study was between 26 – 34% of the WT HIV-1 PR. This result agrees with the findings of similar studies (Gulnik *et al.*, 1995, Lin *et al.*, 1995). A study has shown that mutant HIV-1 PR variants with catalytic efficiency as low as 7 % of the WT HIV-1 PR are catalytically viable (Gulnik *et al.*, 1995).

The results of this study indicate that both LPV and DRV successfully inhibited the mutant HIV-1 PR variants. However, a lower dose of DRV was needed to inhibit the HIV-1 PR variants showing the superiority of DRV over LPV. In addition, this finding supports the switch to a DRV-based regimen after LPV failure. The high resistance to LPV and intermediate resistance to DRV showed an association with the decreased K_{cat}/K_m and increased K_i in the mutant HIV-1 PR variants. As has been described in a previous study, the lower resistance to DRV compared to LPV may be due to the improved design of DRV, which allows it to fit tightly in the active site, coupled with the increased hydrogen bond interaction with the HIV-1 PR backbone, conferring on it a high binding affinity for the active site (Lefebvre and Schiffer, 2008). In a clinical trial that evaluated the long-term effectiveness of ART, the results showed that a DRV-based regimen was superior to an LPV-based regimen in both ART-naive and treatment-experienced HIV-1 infected individuals. The chances of virological failure were lower for DRV compared to LPV when PI therapy was initiated either as a salvage regimen or a switching strategy in treatment-experienced HIV-1 infected individuals (Santos *et al.*, 2018). Another study also showed that the use of low dose DRV boosted with ritonavir (RTV) is an efficient switch option to suppress virological failure in HIV-1 infected individuals failing LPV based treatment (Venter *et al.*, 2019). In the current study, though DRV better

inhibited the HIV-1 PR variants than LPV, the relative resistance (Table 1) to LPV and DRV followed a similar trend. This may be associated with the similarities in the chemical and structural signatures of the inhibitors, which are design to mimic the HIV-1 PR substrate transition state; thus, the evolution of resistance to HIV-1 PI usually may follow a similar pattern (Lv *et al*., 2015). In addition, most of the mutations harbored by the mutant HIV-1 PR variants in this study are LPV resistance specific mutations except L33F (in MUT-2) and L76V (in MUT-2 and 3); this may have caused the pattern of resistance to LPV and DRV to be similar as the presence of more DRV specific mutations may have changed the dynamics or pattern of resistance to DRV seen in this study.

The K_i values obtained in this study followed a similar pattern to those obtained from other studies. The K_i values in this study for the wild type South African HIV-1 subtype C PR ($K_i = 2.13 \pm 0.23$ nM for LPV and $K_i = 1.58 \pm 0.11$ nM for DRV) is close to that reported in a previous study ($K_i = 2.1 \pm 0.2$ nM for LPV and $K_i = 1.4 \pm 0.2$ nM for DRV) (Williams *et al*., 2019). However, there is no other study reporting the kinetic properties of South African HIV-1 subtype C PR harboring multidrug-resistant mutations and their interaction with LPV and DRV. Extensive research has been done on the kinetic characterization of HIV-1 subtype B PR, and the K_i values of the wild type and different multidrug-resistant forms are known. The LPV K_i value of the wild-type subtype B HIV-1 PR is approximately 0.02 nM. The multidrug-resistant forms harboring numerous major and minor mutations showed K_i values of 0.44nM – 260nM (Šašková *et al*., 2008, Park *et al*., 2016, Kneller *et al*., 2020). The DRV K_i value for the wild type subtype B is around 0.005 nM and the multidrug-resistant forms harboring numerous major and minor mutations with K_i values ranging from 1.8 – 41 nM (Weber *et al*., 2015). Variation in the affinity of subtypes B and C HIV-1 PR for PIs has been attributed to the signature polymorphisms in HIV-1 subtype C. The *in vitro* characterization of HIV-1 PR has shown that PIs better inhibit the wild-type and mutant forms of HIV-1 subtype B PR compared to HIV-1 PR variants from the subtype C (Velazquez-Campoy *et al*., 2001, Velazquez-Campoy *et al*., 2002). In another study, HIV-1 infected individuals infected with HIV-1 subtype C receiving ritonavir-boosted regimens developed secondary virological failure within a shorter time compared to HIV-1 infected individuals infected with HIV-1subtype B (Häggblom *et al*., 2016). This current study is essential as it addresses the paucity of data regarding the biochemical interaction of multidrug-resistant South African HIV-1 subtype C PR variants with LPV and DRV, commonly used in the formulation of salvage regimens.

The findings discussed thus far highlights the broad impact of multidrug-resistant mutations on the molecular interaction and binding landscape of HIV-1 PR variants with the substrate, LPV, and DRV. Though this presents a critical picture of the evolution of resistance to LPV and DRV, it does not address other key factors at the molecular level that may shape the interaction and binding landscape of LPV and

DRV with the HIV-1 PR variants. These factors include the changes in the HIV-1 PR variants three-dimensional structure due to mutation harbored, the influence of these conformational changes on the micro kinetic constants and the extent to which the conformational flexibility impact the enzyme catalytic efficiency (Dash and Rao, 2001, de Vera et al., 2013). This is essential to understand the link between an enzyme structure and function and any diversion from normal enzyme characteristics that may influence ligand binding (Pabis et al., 2018). As a result, we explored the use of intrinsic Trp fluorescence to assess conformational changes in the three-dimensional structure of HIV-1 PR variants associated with these mutations and the binding of LPV and DRV (Dash and Rao, 2001, Fidy *et al*., 2001). The increased intrinsic fluorescence emission spectra in the mutant HIV-1 PR variants compared to the wild type in the absence of an inhibitor (Figure 4) indicates that the Trp residues in the mutant PR variants are more solvent-exposed. This finding highlights the impact of the drug-resistant mutations on the tertiary structural changes in the mutant HIV-1 PR variants (Mosebi *et al*., 2008).

The more solvent-exposed the HIV-1 PR dimer is, the more unstable the enzyme-inhibitor complex becomes (Chetty *et al*., 2016). These changes in the intrinsic HIV-1 PR conformational flexibility associated with resistance mutations from LPV-induced drug pressure may cause permanent changes in the HIV-1 PR structure leading to an unfavorable binding landscape to the DRV-based regimen after therapy switch. Research has shown that the resistance mutations selected during ART introduce conformational changes to the HIV-1 PR structure, which may lead to the evolution of HIV-1 PR variants resistant to several PIs since most of the PIs are similar in their chemical nature (Chen *et al*., 1995, Schock *et al*., 1996, van Maarseveen and Boucher, 2006). While the changes in HIV-1 PR structural conformation selected during a particular PI therapy has been shown to cause cross-resistance to several PIs, it has not been demonstrated for the switch from LPV to DRV therapy using HIV-1 PR from clinical isolates.

Further supporting the suggestion that the tertiary structural changes observed in the mutant HIV-1 PR variants are a product of the drug-resistant mutations harbored are the results obtained from the calculation of the K_{sv} value. The low K_{sv} values from the interaction of the mutant HIV-1 PR variants with LPV and DRV compared to the WT (Table 2) indicate that the structures of the mutant HIV-1 PR variants are in a more open state rather than in a closed conformational state. This may have emerged from structural reorganization caused by the mutations harbored, resulting in a less tight binding and accessibility of Trp residues, thus the resultant low quenching capacity of the inhibitors observed with the mutant HIV-1 PR variants. This finding is corroborated by the highly open flap and open protein conformation observed in the molecular dynamic simulation (Figure 8 and 9). The K_{sv} value also throws more light on the efficiency of LPV and DRV to inhibit the HIV-1 PR variants. The overall observation from the K_{sv} value shows that DRV had a relatively higher quenching capacity than the LPV. This may be due to the tight-fitting and high

binding affinity of DRV to the active site compared to LPV, as mentioned earlier (Lefebvre and Schiffer, 2008). This further confirms the superiority of DRV to LPV

From the preceding discussion, it can be inferred that increased resistance to LPV, intermediate resistance to DRV, decreased catalytic efficiency, and altered tertiary structure of HIV-1 PR observed in this study evolved due to the mutations selected from LPV-induced drug pressure. Earlier studies have shown that intrinsic changes in the HIV-1 PR conformation associated with the accumulation of drug-resistant mutations may cause geometric alteration in the HIV-1 PR structure affecting the active site and other domains (Ohtaka *et al*., 2003, Chen *et al*., 1995). These intrinsic changes may cause sensitization and cross-resistance to HIV-1 PIs by altering the molecular interaction of the former with the protein during binding (Chen *et al*., 1995, Bastys *et al*., 2020). The evolution of these characteristics in proteins is often associated with the loss of protein stability (Matange *et al*., 2018). The alteration in the HIV-1 PR tertiary structure, observed from fluorescent spectroscopy results, agrees with the root mean square deviation (RMSD), root mean square fluctuation (RMSF), the radius of gyration (ROG), and solvent accessibility area (SASA) results (Figure 7A – H) obtained from the molecular dynamic simulation.

The increased fluctuation in the RMSF (Figure 7C and D) values, the unstable RMSD (Figure 7A and B) of the mutant HIV-1 PR LPV, and the DRV complexes confirm the intrinsic conformational instability induced by these drug-resistant mutations. This signifies a disruption in protein structure, leading to a loss of compactness affecting LPV and DRV binding. This finding can also be associated with the high SASA (Figure 7E and F) and ROG (Figure 7G and H) in the mutant HIV-1 PR variants when bound to LPV and DRV. These results agree with the findings of a similar study by Chetty *et al.,* (2016) that investigated the impact of multidrug-resistant mutations on the HIV-1 subtype B PR resistance profile and molecular dynamic characteristics. The authors of this study showed that the accumulation of multidrug-resistant mutations causes inherent changes in the structures of HIV-1 PR, resulting in HIV-1 PR variants with increased conformational flexibility and an open conformation. Thus, causing an increase in the rate of dissociation of PIs bound to the active site (Chetty *et al*., 2016). The increased SASA and ROG of the mutant HIV-1 PR, when bound to LPV and DRV in this study, indicate a loss of hydrophobic core compactness, and this has been shown to modulate the activity of HIV-1 PR as well the drug binding landscape (Mittal *et al*., 2012).

Alteration in the HIV-1 PR hydrophobic core has been proposed as a mechanism by which mutations distant from the active site cause drug resistance (Mittal *et al*., 2012). These mutations do not interact directly with the inhibitor but cause an alteration in the balance between substrate recognition and inhibitor binding, in a way that favors interaction with the natural substrate and alters the drug binding landscape (Mittal *et al*.,

2012). The L76V mutation found in MUT-2 and MUT-3 in this study has been shown previously to confer resistance to several PIs through the local rearrangement of the hydrophobic core (Wong-Sam *et al.*, 2018). It is also associated with decreased stability of the HIV-1 PR dimer (Wong-Sam *et al.*, 2018) and emerges during treatment with an LPV-based regimen (Tang and Shafer, 2012). The L76V is also commonly seen in highly mutated HIV-1 PR variants resistant to DRV (Ragland *et al.*, 2014). The L90M observed in MUT-3 is another mutation associated with the alteration of the HIV-1 PR hydrophobic core flexibility and decreased dimer stability. Though this mutation does not make contact with PIs in the active site, it can cause cross-resistance to most PIs except DRV and tipranavir (TPV) (Mahalingam *et al.*, 2004, Ragland *et al.*, 2014, Mittal *et al.*, 2012). The L90M mutation is not commonly seen in HIV-1 subtype C PR; this mutation may have evolved from previous exposure of the HIV-1 infected individuals to PIs like nelfinavir (NFV) and saquinavir (SQV) before treatment with LPV containing regimen (Sugiura *et al.*, 2002). A comparative analysis of the genotypic variation in the HIV-1 PR from subtype B and C from laboratory-generated sequences and publicly available database showed the L90M occur more in HIV-1 subtype B PR than HIV-1 PR variants from subtype C (Grossman *et al.*, 2001).

The hydrophobic interactions between the hydrophobic amino acid residues in the core are vital in maintaining the conformational flexibility of all HIV-1 PR regions, including the flap dynamics. Thus, the hydrophobic core state also modulates HIV-1 PR flap behavior when binding to PIs (Maphumulo *et al.*, 2018, Mittal *et al.*, 2012). This, therefore, shows that the open and increased flap flexibility (Figure 8A – F) in the mutant HIV-1 PR variant may be a product of the distorted hydrophobic core, as shown by the increased SASA in addition to the presence of the flap mutations M46I and I54V present in all the mutants studied. Flap flexibility contributes to ligand stability in the HIV-1 PR active site and influences the binding affinity of HIV-1 PR to PIs (Halder and Honarparvar, 2019, Huang *et al.*, 2014b, Mahanti *et al.*, 2016). Therefore, the increased flap flexibility of the mutant HIV-1 PR variants in this study contributed to the high-level resistance to LPV and intermediate resistance DRV. The M46I and I54V flap mutations found in the mutants in this study have been shown in a previous study to be associated with a decreased binding affinity of HIV-1 PR to PIs, even though they are not active site mutations (Clemente *et al.*, 2004). These mutations, in the presence of active site mutations, like V82A and I84V, cause high-level cross-resistance to several HIV-1 PIs (Ohtaka *et al.*, 2003). In addition, the F53L mutation in MUT-3, though a minor mutation in the flap, is associated with the loss of the hydrophobic bond interaction between phenylalanine's side chain at position 53 of chain A, with Isoleucine at position 50 on the second PR subunit (chain B). Thus, resulting in a wider gap between the two flaps, keeping the HIV-1 PR variants in an open conformation (Liu *et al.*, 2006). Enzyme kinetics models have been used to show that HIV-1 PR mutations that distort the balance between a closed and open conformation to favor an open conformation keep the

enzyme catalytically active regardless of a PI in the active site. This confers high-level resistance on HIV-1 PR to PIs (Weikl *et al*., 2013).

Analysis of the molecular interaction of the HIV-1 PR variants showed that the reduced capacity of the LPV and DRV (compared to the wild type) emanates from the loss of hydrogen bond contacts (Table 4, Figure 10) between the inhibitors and active site residues. The hydrogen bonds between residues in the HIV-1 PR active site and PIs contribute to the latter's tight binding. The loss of these interactions promotes the dissociation of inhibitors from the enzyme-inhibitor complex and is related to the level of resistance to HIV-1 PIs (Chen *et al*., 2014, Yu *et al*., 2015). The loss in hydrogen bonds and changes induced in HIV-1 PR conformations observed in this study together translate into the low binding energy between the mutant HIV-1 PR variants with LPV and DRV observed (Table 3). The high binding energy in the HIV-1 PR-DRV complexes shows that DRV interacts better with the HIV-1 PR variants, and this may be the reason for the better stability of the HIV-1 PR-DRV complexes. The presence of the V82A mutation in all the HIV-1 PR mutant variants may have contributed significantly to the reduced binding affinity of the mutant HIV-1 PR variants to LPV and DRV. The V82 amino acid residue is located in a region critical to drug and substrate binding as it makes direct contact with both PIs and substrate. The V82A mutation is common in HIV-1 infected individuals failing LPV therapy and emerges from the sustained use of LPV during virological failure (Tang and Shafer, 2012). This confers resistance on HIV-1 PR through a structural shift of residues in the 80's loop (Agniswamy *et al*., 2019). Taken together these findings, from applying a multidimensional approach in studying the evolution of drug resistance to LPV and DRV during ART switch, is critical as it gives an insight into the efficiency of the switch from LPV and DRV, which is a common practice.

5. Conclusion

This study's findings provide mechanistic insight into the link between acquired conformational flexibility, associated with resistance mutations selected during an earlier treatment, and its impact on the outcome of a therapy switch. The HIV-1 PR structural changes associated with mutations that emerge due to drug pressure from LPV treatment during virological failure may shape the DRV binding landscape, affecting the switch to a DRV-based regimen. Though the drug-resistant HIV-1 PR variants showed intermediate resistance to DRV, the latter has proven to be more effective than LPV in inhibiting the wild type and mutant HIV-1 PR variants. The capacity of DRV to be effective against HIV-1 PR with already altered conformations from earlier LPV treatment makes it the drug of choice over LPV. This study provides essential information for the development of future inhibitors to address the impact of the altered binding landscape that evolves from initial treatments to help achieve long-term virological suppression.

Bridging between chapter three and chapter four

In the chapter three, the mutations that evolve due to drug pressure associated with LPV treatment caused changes in the HIV-1 PR structure that may affect subsequent ART switch. These changes resulted in the distortion of the HIV-1 PR structure favoring a more open conformation. Thus, making the HIV-1 PR dimer more solvent-exposed and affecting the stability of the enzyme-inhibitor complex. Chapter four evaluated the binding kinetics of HIV-1 PR interaction with LPV and DRV to gain insight into the enzyme kinetic mechanism and the lifetime of the enzyme-inhibitor complex. This chapter is under review in *International Journal of Biological Macromolecules* (manuscript number: IJBIOMAC-D-21-03271).

CHAPTER FOUR: Manuscript one

Mechanistic insight into the binding kinetics of highly mutated HIV-1 protease inhibition by lopinavir and darunavir

Abstract

Highly mutated HIV-1 protease compromises the efficacy of lopinavir and darunavir. Here we report the kinetics of lopinavir and darunavir inhibition of highly mutated South African HIV-1 subtype C protease. Enzyme inhibition assays and fluorescence spectroscopy were used to determine the binding kinetics of lopinavir and darunavir with the wild-type and mutant HIV-1 protease variants. This study shows that just like darunavir, lopinavir has a mixed-type inhibition mechanism, indicating the possibility of a second binding site on HIV-1 protease. Both inhibitors poorly inhibited the highly mutated HIV-1 protease variants with a markedly increased dissociation rate constant compared to the wild-type. The fast dissociation of these inhibitors translated into a short residence time of the inhibitor bound to the mutant HIV-1 protease variants. Fluorescent spectroscopy showed that the changes in the tertiary structure of the mutant HIV-1 protease variants were associated with a more open conformation and an expanded active site. This resulted in the loss of tight binding and rapid dissociation of the inhibitors. This study's findings provide insight into the mechanism of resistance to lopinavir and darunavir by highly mutated HIV-1 protease. It supports the use of binding kinetics measurement in understanding HIV-1 protease inhibitor drug resistance evolution.

Keywords: HIV-1 protease; HIV-1 protease inhibitor; binding kinetics

1. Introduction

The key role played by HIV-1 protease (PR) in viral maturation makes it an important therapeutic target. HIV-1 PR cleaves gag and gag-Pol polyprotein precursors to produce viral enzymes and structural proteins needed for the development of nascent viral particles to full infectious viruses (Weikl *et al*., 2019). Thus, the effective inhibition of HIV-1 PR by protease inhibitor (PIs) suppresses viral replication (Huang and Chen, 2013). Inhibitors of HIV-1 PR are competitive inhibitors designed to bind in the active site to inhibit the normal enzymatic activity of HIV-1 PR by mimicking the HIV-1 PR substrate (Weber *et al*., 2015). However, the lack of proofreading function of the HIV reverse transcriptase and the persistence of some residual viral activity during active antiretroviral therapy due to non-adherence or poor drug absorption may lead to the emergence of drug-resistant strains that may harbor mutations in the HIV-1 PR gene (Kožíšek *et al*., 2014, Goldfarb *et al*., 2015). These PI mutations can be classified as major (cause resistance on their own) or minor (cause resistance in combination with other PI mutations) drug resistance mutations (Kneller *et al*., 2020).

HIV-1 PI-based ART in sub-Saharan Africa depends greatly on LPV due to its affordability (Venter *et al*., 2019). However, there is also an increase in the number of HIV-1 infected individuals failing LPV-based treatment regimens in sub-Saharan Africa after 12–18-month of treatment initiation (Edessa *et al*., 2019). The efficacy of ritonavir-boosted LPV is affected by the appearance of three or more of the following mutations: "L10F/I/R/V, K20M/N/R, L24I, L33F, M36I, I47V, G48V, I54L/T/V, V82A/C/F/S/T, and I84V" as discussed in chapter three (Wensing *et al*., 2019). HIV-1 infected individuals are switched from LPV inclusive therapy to DRV-based regimens where LPV fails. This is due to the high genetic barrier to the evolution of drug resistance to DRV, making it the HIV-1 PI of choice to treat experienced HIV-1 infected individuals and therapy naïve in first-world countries (Venter *et al*., 2019). However, research has shown that despite the superiority of DRV to LPV in combatting highly mutated HIV-1 PR, the structural changes arising from the impact of mutations selected during LPV therapy may impact the outcome of DRV after an antiretroviral therapy switch (Eche *et al*., 2021). The poor inhibitory effect of LPV and DRV on highly mutated HIV-1 PR stems from the loss of molecular interaction between these inhibitors and HIV-1 PR. It has been demonstrated that the loss of interaction is a product of the conformational changes in the HIV-1 PR structure arising from the interplay of the active site and non-active site mutations (Liu *et al*., 2013, Ohtaka *et al*., 2003, Kožíšek *et al*., 2014, Kneller *et al*., 2020).

Different strategies are being used to develop antiviral compounds targeting the highly mutated HIV-1 PR, including developing inhibitors that targ*et al*ternative binding sites on HIV-1 PR asides from the active site. (Weber *et al*., 2015). Though the interaction of highly mutated HIV-1 subtype B PRs with LPV and DRV

has been characterized using enzymology and structural analyses (Kožíšek *et al*., 2014, Liu *et al*., 2013, Kneller *et al*., 2020), the same cannot be said of highly mutated HIV-1 subtype C PR, which makes this current study timely. In addition, knowledge about the binding kinetics and mechanism of drug binding and its association with HIV-1 PR conformational changes is still poorly understood (Miao *et al*., 2018). The measurement of HIV-1 PR binding kinetics with HIV-1 PIs will help determine the rate of association and dissociation of the latter (Dierynck *et al*., 2007) and the lifetime of the enzyme-inhibitor complex, which is influenced by the balance between the association and dissociation rate of inhibitors (Schuetz *et al*., 2018). In addition to determining the rate of association and dissociation, the inhibitor's residence time, which is the time a drug spends bound to its target, can also be determined from the dissociation rate constant (Pan *et al*., 2013). Therefore, the kinetic constants need to be considered when designing new HIV-1 PIs to ensure effective inhibition

There is a paucity of data for most HIV-1 PIs with regards to the steady-state binding kinetics, the rate of association and dissociation, and the lifetime of the HIV-1 PR-inhibitor complexes. This study reports the steady-state kinetic measurements of the inhibition of the wild type and two highly mutated South African HIV-1 subtype C PR variants failing a LPV-based regimen (MUT-1 with the mutations: M46I, I54V, L76V, V82A, I84V, Q58E, and MUT-2 with the mutations: V32I, M46I, I54V, L76V, V82A, L90M, L33F). In a previous study using isolates with similar PI mutations, we used biochemical and *in silco* analysis to show how acquired HIV-1 PR conformational flexibility associated with LPV failure may shape the outcome of DRV therapy after ART Switch (Eche *et al*., 2021). This current study shows how the conformational changes associated with drug-resistance mutations affect the enzyme-inhibitor interaction and the evolution of resistance to LPV and DRV, in addition to the possibility of a second binding site to LPV using enzyme kinetic analysis.

2. Materials and Methods

2.1 Ethical approval

Ethical approval for this study was obtained from the Biomedical Research Ethics Committee (BREC NO. 413/17), University of KwaZulu-Natal.

2.2 HIV-1 PR gene amplification, expression and purification, and enzyme kinetic analysis

The HIV-1 PR gene was expressed as previously described in chapter three (Eche *et al*., 2021) Briefly viral RNA was extracted from plasma samples obtained from HIV-1 infected individuals failing a LPV-inclusive regimen. The amplified HIV-1 PR PCR products were cloned into the pMAL-c5X expression vector (New

England BioLabs, Ipswich, MA, USA), and the recombinant HIV-1 PRs were expressed in NEBExpress Escherichia coli cells (New England BioLabs, Ipswich, MA, USA) as described previously (Eche *et al*., 2021). Enzyme activity assay and inhibition constants for LPV and DRV were determined using the chromogenic substrate Lys-Ala-Arg-Val-Nle-p-nitro-Phe-Glu-Ala-Nle amide (Sigma-Aldrich, St. Louis, MO, USA) as described previously in chapter three (Eche *et al*., 2021).

2.2 Evaluation of the enzyme kinetic parameter

The initial rate analysis was done assuming a reversible inhibition using the following equations for Lineweaver-Burk plot (Eq. 1a) and Dixon plot (Eq.1b).

$$1/v = (\alpha K_m / V_{max}) (1/[S]) + \alpha'/V_{max} \tag{1a}$$

$$1/v = [I] / V_{max} K_i + 1/V_{max} (1 + K_m / [S] \tag{1b}$$

In the equations above, $\alpha = 1 + 1/K_i$ and $\alpha' = 1 + 1/ K_i'$, K_m is the Michaelis constant, the maximal catalytic rate achieved at a saturating concentration of the enzyme is V_{max}, and I is the concentration of the inhibitor. The dissociation constants for the first reversible enzyme–inhibitor or enzyme–inhibitor–substrate complexes are $K_i = (k_4/k_3)$ and $K_i' = (k_9/k_8)$, respectively. If the affinity of the free enzyme (E) and the enzyme-substrate complex (ES) to bind the inhibitor (I) is equal, then $\alpha = \alpha'$, and the value of the apparent K_m will not change from the K_m' of the reaction with an inhibitor added to the reaction. The only parameter affected by the presence of the inhibitor is the V_{max} (Cha, 1975). The values of K_i and K_i' were determined using Eq. (1c) and the previously suggested method (Cha, 1975, Cleland, 1963).

$$slope_0 / (slope_i - slope_o) = K_i (1 + K_i'/ I) / K_i' - K_i) \tag{1c}$$

In Eq. (1c), when [I]=0, the slope of the straight line on the Lineweaver–Burk is $slope_o$, and $slope_i$ is the slope for the straight line on the Lineweaver–Burk plot when [I]≠0 (Cha, 1975). Analysis of the progress curves for LPV and DRV (0, 5, 10, 20, and 40 nM) interaction with HIV-1 PR was done according to Scheme 1B using Eq. (2) (Morrison *et al*., 1985).

$$[P] = v_s t + (v_o - v_s) [1 - \exp(-kt)] / k \tag{2}$$

Where [P] is the product concentration at any time (*t*), the initial and steady-state velocities are v_o and v_s, respectively. The apparent first-order rate constant for the establishment of the final steady-state equilibrium is *k*. A correction was made for the decrease in the inhibitor concentration associated with the formation of the EI complex since its concentration is not negligible compared to the inhibitor

concentration. The free inhibitor concentration is not equal to the added concentration of the inhibitor for tight-binding inhibition. The difference in the steady-state velocity with the inhibitor concentrations was corrected by applying Eqs. (3) and (4) as described previously (Morrison and Walsh, 1988).

$$v_s = (k_7 [S] Q / 2([S] + K_m)) \tag{3}$$

$$Q = [K_i' + I_t - E_t) + 4 K_i'' E_t]^{1/2} - [(K_i'' + I_t - E_t)] \tag{4}$$

Where $K_i'' = K_i * (1 + [S]/K_m)$, the rate constant for the formation of product is k_7, I_t, and E_t represent the total inhibitor and enzyme concentrations, respectively. The association between the rate constant of enzymatic reaction k and the association constant (k_5), as well as the dissociation constant (k_6) of the enzyme and inhibitor, was determined using Eq. (5).

$$k = k_6 + k_5 ([1]_0 / K_i) / (1 + [S]_0 / K_m + [I]_0 / K_i) \tag{5}$$

The Eqs. (2) and (5) were used to analyze the progress curves. Non-linear least-square parameter minimization was used to determine the best-fit values with the necessary corrections for the tight-binding inhibition. The mode of inhibition was evaluated using the plot of 1/(k–k6) as a function of [S] (Shapiro and Riordan, 1984). The results derived from Eq. (2) were fitted into Eq. (6) to get the overall inhibition constant.

$$v_s = v_{max} [S]_0 / (\alpha'[S]_0 + K_m (1 + [I]_0 / K_i*)) \tag{6}$$

The time-dependent enzyme inhibition analysis showed a time range in the enzymatic reaction progress curves where EI* formation was small. It is possible to determine the K_i directly within this time range (the impact of the inhibitor on v_0). Dixon analysis and fitting of the data derived from Eq. (2) at a constant concentration of substrate [S] as described in Eq. (7) were used to calculate the K_i values

$$v_0 = v_{max} [S]_0 / (\alpha'[S]_0 + \alpha K_m) \tag{7}$$

2.2.1 Enzyme kinetic scheme
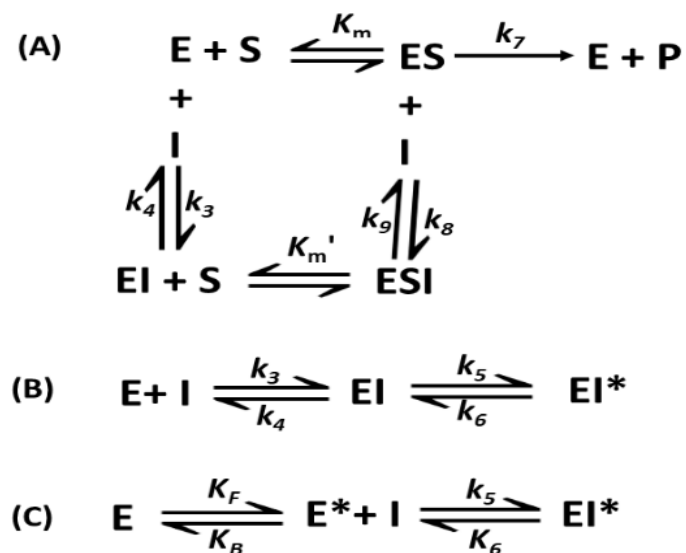
Scheme 1. Enzyme kinetic scheme

In the scheme above, E represents the free enzyme in the enzyme kinetic scheme above, and I is the free inhibitor. The rapidly forming preequilibrium enzyme-inhibitor complex is EI, and EI* represents the final enzyme-inhibitor complex. The free enzyme, E, may undergo inter-conversion into E*. This form binds with the inhibitor at a fast step, and k_F and k_B represent the rate constants for the forward and reverse reactions, respectively, to convert the enzyme. K_m and K_m' are the Michaelis constant for the reaction of the substrate with the enzyme and the reaction of the substrate with enzyme-inhibitor complex, respectively. The rate constants k_3, k_4, k_5, k_6, k_8, and k_9 are described in the text, and k_7 represents the rate constant for product formation from the ES complex.

2.3 Fluorescent spectroscopy

The fluorescent spectra were recorded using the PerkinElmer LS 55 spectrometer with a 1.0 cm quartz cell (Waltham, MA, USA) connected to a thermostat-controlled bath. HIV-1 PR samples were excited at 295 nm, and the fluorescence emitted was recorded from 300 to 420 nm at 25°C as described in chapter three (Eche *et al*., 2021). The data obtained from the decrease in intrinsic Trp fluorescence (F_0–F) at each concentration of LPV and DRV was fitted to the equation $(F_0–F) = \Delta F_{max}/(1+(K_i/[I])$ to determine K_i and ΔF_{max} values. The Stern-Volmer quenching constants (K_{sv}) of LPV and DRV were calculated by fitting the data into the equation $F_0/F = 1 + K_{sv} [I]$. The slow loss of intrinsic Trp fluorescence was used to calculate the first-order rate constants (K_{obs}) at each LPV and DRV concentration. The obtained data after 1 s was

computer fitted to the first-order equation y=a+b*exp($-K_{obs}$ *t) to obtain K_{obs} values which were subsequently analyzed using the equation K_{obs} =k_5[I]/(K_i+[I]) to determine the K_5 based on the assumption that during the onset of the slow loss of intrinsic Trp fluorescence, the value of K_6 can be considered negligible for tight-binding inhibitors (Houtzager *et al*., 1996). The time courses of HIV-1 PR fluorescence upon addition of LPV and DRV were evaluated for 5 min at a data acquisition time of 0.1 s. The excitation wavelength fixed at 295nm and the emission wavelength set at 351nm, respectively. The Origin(Pro), 2019 software (OriginLab Corporation, Northampton, MA, USA) was used to computer fit the data used to obtain the K_{obs}, K_i, K_5, and K_{sv} values. The inner filter effect during fluorescence measurement was corrected as described in the previous chapter (chapter three).

2.4 Statistics

Enzyme kinetic parameters were evaluated by computer fitting the obtained data using the Origin(Pro), 2019 (OriginLab Corporation, Northampton, MA, USA) and SigmaPlot version 14.5 (Systat Software, Inc., San Jose California USA). The enzyme kinetic, as well as the determination of the rate constants, were carried out in triplicate, and the average values were used in this study.

2.5 Modeling and molecular dynamic simulation of the wild type and mutant HIV-1 PR structures

The wild type (WT) South African HIV-1 PR (3U71) structure was retrieved from the RSCB Protein Data Bank (Burley *et al*., 2018). The mutant South African HIV-1 PR (MUT-1 and MUT-2) structures in this study were obtained through homology modeling performed on the SWISS-MODEL web server with the WT South African HIV-1 subtype C PR x-ray crystal structure (3U71) as a template. The HIV-1 PR structural dynamics associated with the impact of the mutations were probed using Molecular (MD) dynamic simulation performed using the GPU version provided with the AMBER 18 package. The amino acid residues were renumbered based on the dimeric form of HIV-1 PR from 1 to 198. The systems were suspended implicitly within an orthorhombic box of TIP3P water molecules. An initial minimization of 2000 steps with an applied restraint potential of 500 kcal/mol was carried out, followed by a full minimization of 1000 steps by a conjugate gradient algorithm without restraint. The MD was performed for a total time of 10ns. This was carried out with gradual heating from 0K to 300K, executed for 50ps, such that the systems maintained a fixed number of atoms and fixed volume as previously described (Kehinde *et al*., 2019). The post-dynamic analysis was done using CPPTRAJ modules implemented in Amber18 for analysis of the root mean square fluctuation (RMSF), root mean square deviation (RMSD) as previously described (Shunmugam and Soliman, 2018). The HIV-1 PR structures obtained after MD simulation were also analyzed for differences in the protein conformation.

2.5.1 Measurement of the HIV-1 PR active site surface area and volume

The WT and mutant HIV-1 PR variants active site surface area and volume were determined using protein structures generated at the last 2ns of MD simulation. The HIV-1 PR structures were uploaded to the Computed Atlas of Surface Topography of Proteins (CASTp 3.0) online server (Tian *et al*., 2018). A radius probe of 1.4 Å was used for the active site cavity surface area and volume measurement. The CASTp 3.0 server calculates the surface area and volume of binding site cavities in proteins based on the principles of Alpha Shape Theory (Liang *et al*., 1998).

3. Results

3.1 Analysis of Kinetic parameters using Lineweaver-Burk and Dixon method

3.1.1 Enzyme kinetics and Inhibition of wild type and mutant HIV-1 PR by LPV and DRV

The wild type (WT) HIV-1 PR had almost a 2-fold lower K_m (36.95 ± 0.93 µM) for the chromogenic substrate (Table S1, Fig. 1A – C) compared to the mutant HIV-1 PR variants (K_m for MUT-1 = 61.29 ± 0.77 µM and K_m for MUT-2 = 64.44 ± 0.21 µM). The catalytic constant (K_{cat}) of the WT HIV-1 PR (K_{cat} = 0.81 ± 0.08 s^{-1}) was higher compared to the K_{cat} of the mutant HIV-1 PRs (K_{cat} for MUT-1 = 0.54 ± 0.13 s^{-1} and K_{cat} for MUT-2 = 0.44 ± 0.15 s^{-1}). The WT had a catalytic efficiency (K_{cat}/K_m) of 0.021 $s^{-1}µM^{-1}$ which was over 2-fold higher than in the mutant HIV-1 PR variants (K_m/K_{cat} for MUT-1 = 0.0088 $s^{-1}µM^{-1}$ and K_m/K_{cat} for MUT-2 = 0.0068 $s^{-1}µM^{-1}$).

Analysis of the enzyme inhibition by LPV and DRV is presented in Table S1 and Figures 1D – I. The results showed DRV was more effective in inhibiting the WT and mutant HIV-1 PR variants. DRV showed a K_i value of 1.74 ± 0.27 nM for the WT, which is lower than the K_i of LPV for the WT (2.09 ± 0.18 nM). The LPV K_i values for the inhibition of the mutant HIV-1 PR variants (MUT-1 = 67.72 ± 0.33 nM and MUT-2 = 83.45 ± 0.89 nM) were about 6 – 7-fold higher than the K_i value for DRV (MUT-1 = 8.93 ± 0.10 nM and MUT-2 = 13.27 ± 0.05 nM).
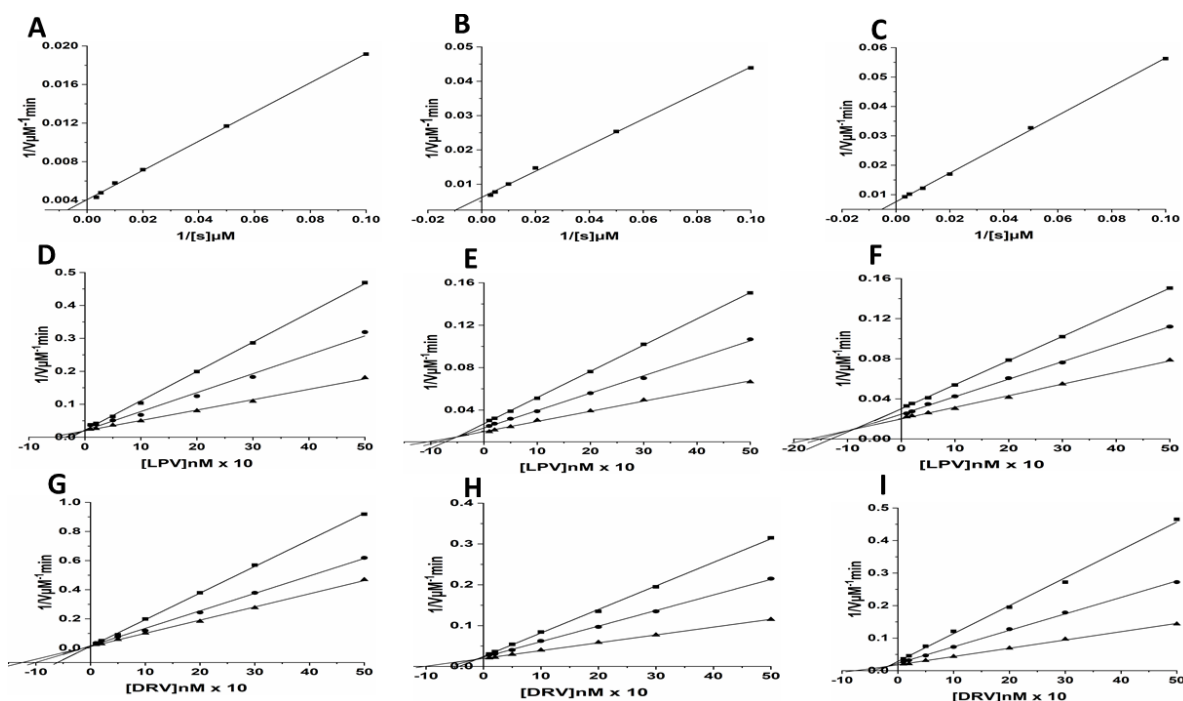
Figure 1. (A – C) Enzyme kinetic activity of the WT (A) and mutant HIV-1 PR variants (B=MUT-1, C=MUT-2) determined following the hydrolysis of the chromogenic synthetic substrate (Lys-Ala-Arg-Val-Nle-p-nitro-Phe-Glu-Ala-Nle amide). The activity of the HIV-1 PR variants (D=WT, E=MUT-1, F=MUT-2) was measured in the presence of 10-500nM LPV (D – F) and DRV (G – I) using three substrate concentration: 100 (■), 200 (●), and 300 μM (▲) respectively.

3.1.2 Analysis of enzyme kinetics and inhibition progress curve

The initial inhibition kinetic evaluation showed that both LPV and DRV inhibited the wild type and mutant HIV-1 PR variants in a mixed type inhibition manner (Figure 1D – I). Three mechanisms in scheme 1 can be used to illustrate the enzyme inhibition mechanism. The results further confirmed the inhibition mechanism was not a slow-tight binding one as it did not vary in the region of K_i. Thus, both k_3I and k_4 values were not low. As a result, the simple second-order interaction between E and I and high rates of association and dissociation would result in a quick binding of the inhibitors (Scheme 1a). Scheme 1a shows that the I bind to ES complex to form ESI complex, this then dissociates into EI and S. The dissociation constant K_i' for ESI to EI and S is greater than the K_i when E and ES complex does not have an equal binding affinity with I. Alternatively, the binding may also involve a two-step model which portrays the rapid formation of an initial collisional EI complex, that isomerizes to form a tightly bound and fast dissociating EI* complex (Scheme 1b). This kind of inhibition mechanism can also arise due to the initial fast inter-conversion of the free enzyme E into another form E*, which subsequently binds to I at a fast step (Scheme 1c). Understanding the basis of the EI complex's isomerization to EI* complex from the application of enzyme kinetics will provide quantitative information needed to design structures that permit

the titration of the lifetime of an EI* complex and the development of future novel tight-binding inhibitors (Kumar and Rao, 2010). The quantitative data obtained from the isomerization of EI and EI* includes the level of EI* formation from the EI complex and the relative rates of formation of EI* and its subsequent relaxation to EI (Kumar and Rao, 2010).

The steady-state rate of catalytic activity of HIV-1 PR was rapidly reached in the absence of LPV and DRV (Figure 2 A – F). In the presence of the inhibitors, the steady-state rate decreased in a time-dependent manner as a function of the concentration of the inhibitor. In order to determine the enzyme kinetic mechanism and the kinetic constants, the initial velocity v_0, steady-state velocity (v_s), and the rate constant (k) data were obtained using varying concentrations of LPV and DRV. Mixed inhibition experiments were used to determine the K_i values (Eq. (5) for the HIV-1 PR variants using calculations done with data obtained from 0 to 150s where EI conversion to EI* was minimal (Figure 2). The mode of HIV-1 PR inhibition was also determined by plotting $1/(k–k_6)$, by fitting the data in Eq. (5) as a function of [S], and this showed a hyperbolic relation to the substrate concentration.
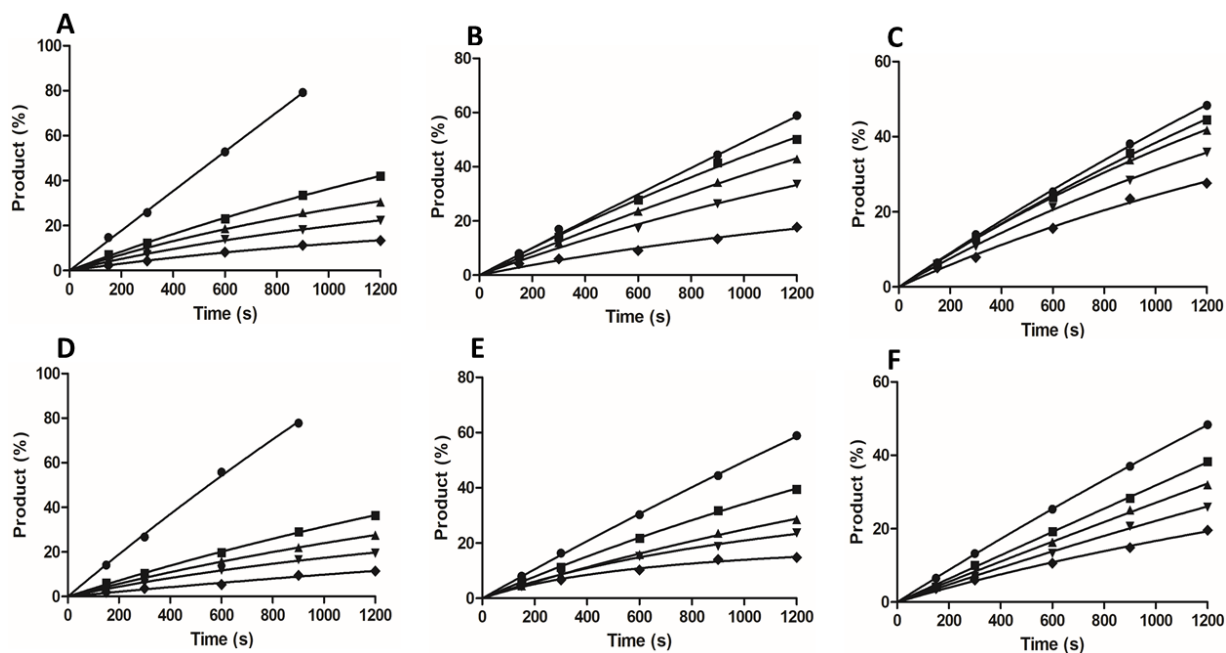
Figure 2. Least-squares fit of progress curves from the inhibition of HIV-1 PR (200nM) by LPV (WT=A, B=MUT-1, C=MUT-2) and DRV (WT=D, E=MUT-1, F=MUT-2) to Eq. (3) using an initial substrate concentration of 200μM. The products released as a function time is represent by the different points. The concentration of LPV and DRV used were 0 nM (●), 5 (■), 10 (▲), 20 (▼), and 40nM (♦) respectively. The lines are the lines of best fits of the data obtained from Eqs. (2) and (5), and corrections were made according to Eqs. (3) and (4).

The progress curves obtained at 0, 5, 10, 20, and 40 nM LPV and DRV were analyzed and individually fitted to Eq. (2). Tables 1 and 2 present the best-fit values of adjustable parameters for LPV and DRV

concentrations. As shown in scheme 1(A), the initial velocity v_0 must be constant for a one-step enzyme inhibition mechanism (Kumar and Rao, 2010). In contrast, for a two-step enzyme inhibition mechanism (Scheme 1(B)), v_0 decreases with an increase in the inhibitor concentration which follows a typical binding curve (Kumar and Rao, 2010). In addition, the apparent rate constant depends on $[I]_0$ as a biphasic hyperbola. The data presented in Tables 2 and 3 favors a two-step enzyme inhibition mechanism since the initial velocity decreases with an increase in the concentration of the inhibitor, as represented by Eq. (7). Likewise, the apparent rate constant increase with an increase in $[I]_0$ is hyperbolic and not linear.

Table 1. The best fit values of adjustable parameters, gotten from fitting the progress curves of HIV-1 PR inhibition by LPV

	WT			MUT-1			MUT-2		
$[LPV]_0$ nM	v_s (µmol/min)	v_0 (µmol/min)	k (s⁻¹)	v_s (µmol/min)	v_0 (µmol/min)	k (s⁻¹)	v_s (µmol/min)	v_0 (µmol/min)	k (s⁻¹)
0	0.083	0.45	0.16	0.048	0.25	0.54	0.039	0.47	0.33
5	0.034	0.44	0.27	0.042	0.20	0.60	0.037	0.45	0.34
10	0.025	0.40	0.36	0.035	0.14	0.64	0.034	0.41	0.36
20	0.018	0.14	0.54	0.026	0.10	0.69	0.031	0.38	0.38
40	0.011	0.13	0.64	0.013	0.07	0.72	0.022	0.34	0.40

Table 2. The best fit values of adjustable parameters, gotten from fitting the progress curves of HIV-1 PR inhibition by DRV

	WT			MUT-1			MUT-2		
$[LPV]_0$ nM	v_s (µmol/min)	v_0 (µmol/min)	k (s⁻¹)	v_s (µmol/min)	v_0 (µmol/min)	k (s⁻¹)	v_s (µmol/min)	v_0 (µmol/min)	k (s⁻¹)
0	0.082	1.63	0.15	0.048	0.70	0.47	0.040	0.42	0.57
5	0.029	0.99	0.56	0.032	0.46	0.58	0.032	0.33	0.78
10	0.022	0.94	0.63	0.023	0.36	0.66	0.027	0.22	0.90
20	0.016	0.85	0.68	0.018	0.25	0.76	0.021	0.18	0.94
40	0.010	0.30	0.74	0.011	0.16	0.79	0.015	0.16	1.00

The results obtained by fitting of the rate constant (k) to Eq. (5) (Table 3 and Table S2, Fig. 3) shows the isomerization rate constants k_5 and k_6 values for the inhibition of HIV-1 PR by LPV for the WT was $k_5 = 0.74$ s⁻¹ and $k_6 = 0.15$ s⁻¹, for MUT-1 the k_5 and k_6 values are 0.19 and 0.42 s⁻¹ respectively, while for MUT-2 the k_5 and k_6 values are 0.13 and 0.33 respectively. The k_5 and k_6 values (Table S2) for DRV obtained from Eq. (5) for the WT was $k_5 = 0.43$ s⁻¹ and $k_6 = 0.15$ s⁻¹, for MUT-1, the k_5 and k_6 values are 0.16 and 0.46 s⁻¹ respectively, while for MUT-2 the k_5 and k_6 values are 0.14 and 0.59, respectively. The association and dissociation constants of EI* for the HIV-1 PR variants indicates a tight-binding inhibition mechanism.

The fitting of v_s for the inhibition of HIV-1 PR by LPV to Eq. (6) resulted in an overall dissociation constant (K_i*) of 0.75nM for the WT, 1.94nM, and 3.16 nM for MUT-1 and MUT-2 respectively (Table 3). The overall HIV-1 PR enzyme inhibition constant K_i* is a function of $k_6/(k_5+k_6)$; this is equal to the product of K_i as well as this function. The K_i* for DRV obtained from Eq. (6) was 0.66, 1.29, and 2.0 for the WT, MUT-1, and MUT-2, respectively (Table 3). The dissociation constant K_i for HIV-1 PR inhibition by LPV obtained from fitting v_0 to Eq. (7) was 2.07 nM, 1.79 nM, and 4.60 nM for the WT, MUT-1, and MUT-2 respectively. For the inhibition of HIV-1 PR by DRV, the K_i from fitting v_0 to Eq. (7) was 2.39 nM, 1.55 nM, and 2.58 nM for the WT, MUT-1, and MUT-2, respectively (Table 3).

Table 3. Best-fit values obtained by fitting k, v_s, and v_0 to Eq. (5) – (7).

Parameters	WT-LPV	WT-DRV	MUT-1 LPV	MUT-1 DRV	MUT-2 LPV	MUT-2 DRV
			Equation 5			
k_6	0.15	0.15	0.42	0.46	0.33	0.59
k_5	0.74	0.42	0.19	0.16	0.13	0.14
K_i	~ 2.11	~ 2.47	~ 2.7	~ 1.22	~ 2.56	~ 2.31
S	= 200.0	= 200.0	= 200.0	= 200.0	= 200.0	= 200.0
K_m	~ 20.26	~ 16.63	~ 15.87	~ 32.11	~ 16.33	~ 17.87
			Equation 6			
Vmax	~ 0.34	~ 0.34	~ 0.27	~ 0.26	~ 0.24	~ 0.23
S	= 200.0	= 200.0	= 200.0	= 200.0	= 200.0	= 200.0
α'	~ 3.34	~ 3.29	~ 4.873	~ 4.75	~ 5.70	~ 5.53
K_m (µM)	~ 149.4	~ 158.1	~ 100.0	~ 132.8	~ 68.71	~ 102.5
K_i* (nM)	~ 0.75	~ 0.66	~ 1.94	~ 1.29	~ 3.16	~ 1.99
			Equation 7			
Vmax	~ 0.82	~ 1.27	~ 0.60	~ 0.95	~ 0.81	~ 1.69
S	= 200.0	= 200.0	= 200.0	= 200.0	= 200.0	= 200.0
α'	~ 1.48	~ 0.67	~ 2.095	~ 1.17	~ 1.67	~ 0.35
K_m (µM)	~ 37.84	~ 27.58	~ 58.18	~ 39.73	~ 15.06	~ 12.41
K_i (nM)	~ 2.07	~ 2.39	~ 1.79	~ 1.55	~ 4.60	~ 2.58

The calculation of the inhibitor residence time which is obtained using the formula $t_R = 1/k_6$ (Pan et al., 2013), was done for LPV and DRV when bound to HIV-1 PR (Table S2). The residence time for LPV (6.66 s) and DRV (6.66 s) when bound to the WT was similar. However, the residence time of LPV for mutant HIV-1 PR variants (2.38 s and 3.03 s for MUT-1 and MUT-2, respectively) was significantly lower compared to the WT. The residence time of DRV for the mutant HIV-1 PR variants was also significantly lower (2.17 s and 1.70 s for MUT-1 and MUT-2 respectively) was also significantly lower compared to the WT.
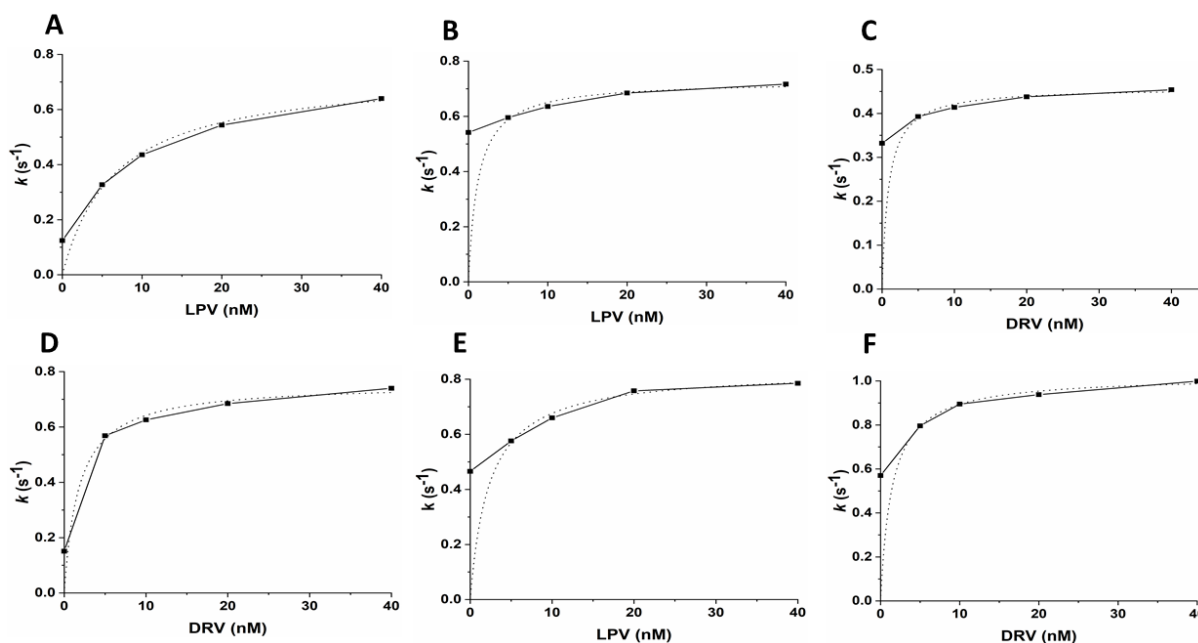
Figure 3. The non-linear least-squares fit of the rate constant (k) to Eq. (5) to calculate the association (k5) and dissociation (k6) constant of LPV (WT=A, B=MUT-1, C=MUT-2) and DRV (WT=D, E=MUT-1, F=MUT-2) from the enzyme kinetic progress curve (Dotted lines represent the best line of fit).

3.2 Fluorescence changes of HIV-1 PR due to binding of LPV and DRV and the dependence of fluorescence on LPV and DRV concentration and time

Analysis of the HIV-1 PR inhibition by LPV and DRV revealed a two-step inhibition mechanism, characterized by the isomerization of the EI complex to a tightly bound and slow dissociating EI* complex. The intrinsic Trp emission spectrum in the presence and absence of LPV and DRV was analyzed to evaluate and correlate the isomerization to the conformational changes in HIV-1 PR associated with the binding of LPV and DRV. The HIV-1 PR variants displayed emission maxima (λmax) at 351 nm due to the radiative decay of the π–π* transition from the Trp residues (Fig. S1). A concentration-dependent fluorescence quenching was observed upon the titration of HIV-1 PR with LPV and DRV. There was no blue or redshift in the λ_{max} (figures 4 A – C and 5 A – C). There was no red or blue shift in the λ_{max} upon an increase in the concentration of either inhibitor. This indicates that the changes in the three-dimensional structure of the enzyme are not due to the binding of the inhibitors but are associated with localized conformational changes in the enzyme. The changes in the HIV-1 PR conformation in the course of EI isomerization to EI* were assessed by evaluating the tryptophanyl fluorescence of the enzyme-inhibitor complexes with time. The binding of LPV and DRV caused an exponential decrease in fluorescence intensity; this is shown by the marked reduction in the fluorescence quantum yield accompanied by a slow decrease to a stable value.

87

Additionally, the titration of HIV-1 PR with LPV and DRV showed the extent of the initial rapid loss in intrinsic Trp fluorescence (F_0–F) increased in a saturation-type manner (figures 4 D – F and 5 D – F). This supports the two-step tight-binding inhibition of HIV-1 PR by LPV and DRV. The data in Fig. 5 and 6 shows that the degree of the rapid loss of intrinsic Trp fluorescence at a particular concentration of LPV and DRV was close to the total fluorescence quenching; this signifies that the EI and EI* complexes may have similar intrinsic Trp fluorescence. The changes in the environment of the different Trp residues may cause an alternation of fluorescence characteristics like the emission wavelength, the quantum yield, and exposure to quenching (Pawagi and Deber, 1990). Fluorescence quenching may also result from the transfer of energy to an acceptor molecule that has an overlapping absorption spectrum (Cheung and Lakowicz, 1991). To rule out the possibility of the fluorescence quenching being associated with energy transfer between the inhibitors and the tryptophan residues, the inhibitors' fluorescence capacity was evaluated, and they were found not to have absorption in the region of 290–420 nm (Dash *et al.*, 2001).
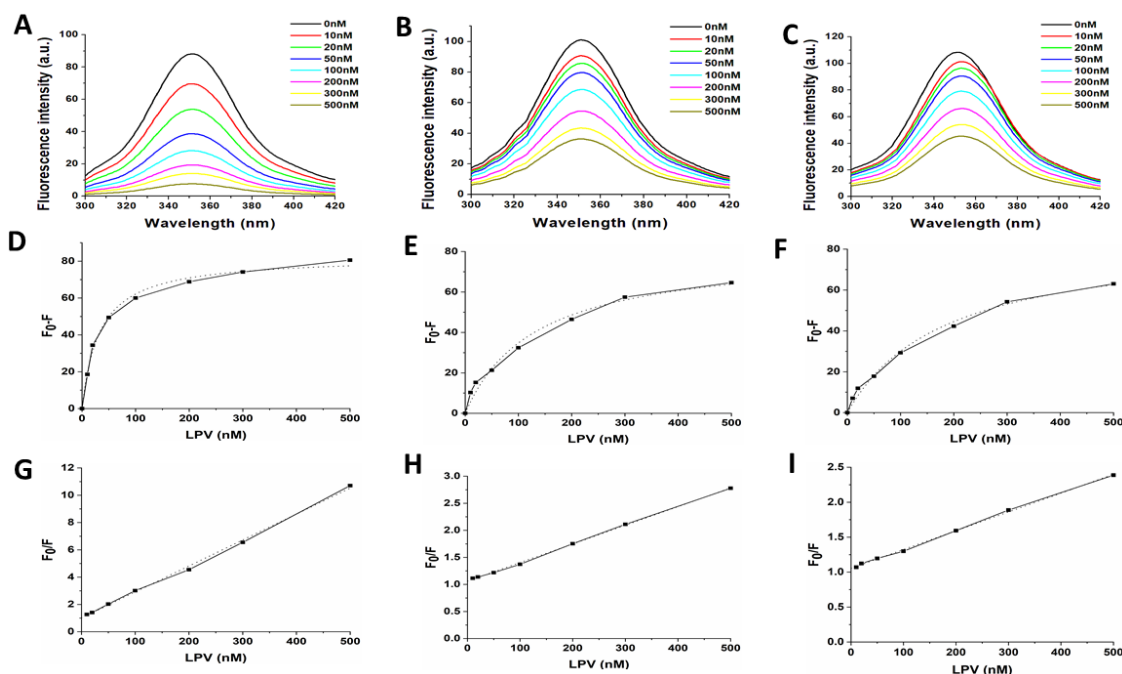
Figure 4. (A – C) Intrinsic tryptophan fluorescence quenching using LPV for the WT, MUT-1, and MUT-2, respectively. (D – F) The effect of LPV on the tryptophanyl fluorescence of HIV-1 PR (WT, MUT-1, and MUT-2, respectively). (G – I) Stern-Volmer plot to determine the quenching constants (Ksv) of LPV for the WT, MUT-1, and MUT-2 respectively.
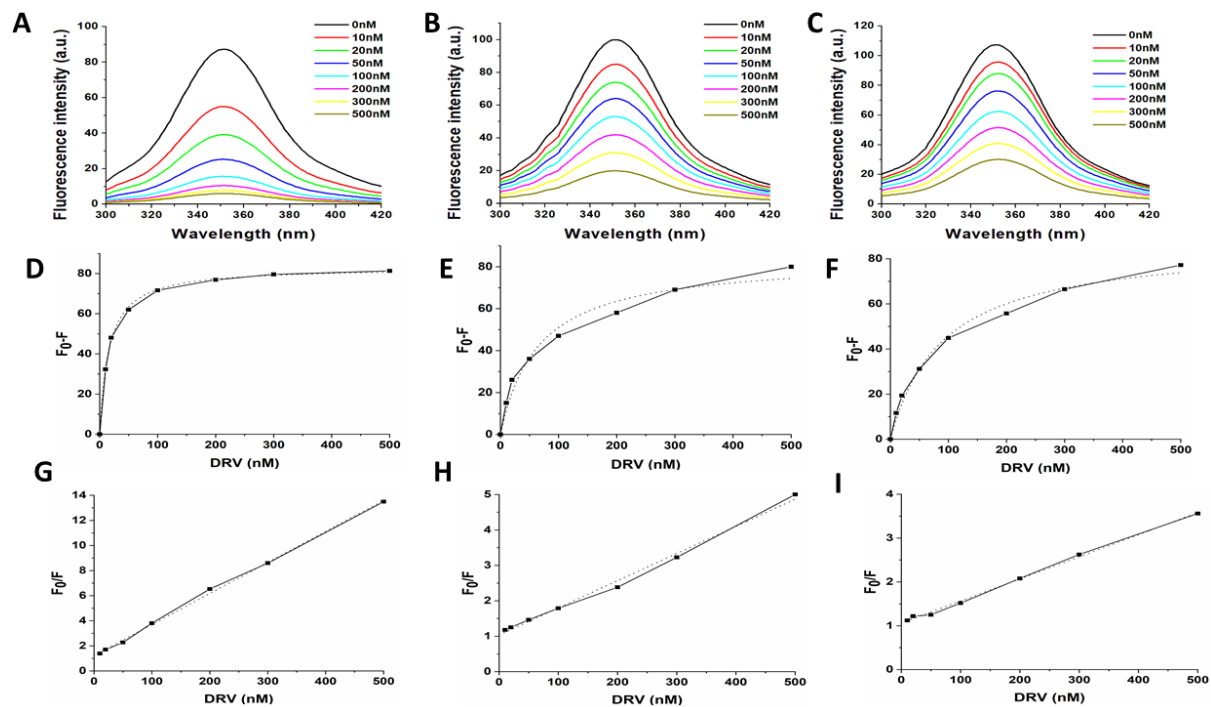
Figure 5. (A – C) Intrinsic tryptophan fluorescence quenching using DRV for the WT, MUT-1, and MUT-2, respectively. (D – F) The effect of DRV on the tryptophanyl fluorescence of HIV-1 PR (WT, MUT-1, and MUT-2, respectively). (G – I) Stern-Volmer plot to determine the quenching constants (Ksv) of DRV for the WT, MUT-1, and MUT-2 respectively.

The value of the K_i of LPV and DRV determined by fitting the data for the magnitude of the rapid fluorescence decrease (F_0–F) for the different HIV-1 PR variants is presented in Table S3 (Figure 4 D – F and 5 D – F). The WT HIV-1 PR K_i value for LPV (19.28 ± 3.02 nM) was found to be about 4 – 6-fold lower than in the mutant HIV-1 PR variants (MUT-1 = 78.05 ± 5.22 nM and MUT-2 = 123.68 ± 7.46 nM). The WT HIV-1 PR K_i value for DRV (8.67 ± 0.54 nM) was approximately 4 – 6-fold lower than in the mutants (MUT-1 = 39.73 ± 4.07 nM and MUT-2 = 50.11 ± 3.19 nM). This followed a similar pattern as the K_i values obtained from the enzyme inhibition analysis.

The K_i and k_5 values determined from the slow reduction of fluorescence are presented in Table S4. The WT HIV-1 PR K_i value for LPV (3.13 ± 0.55 nM) was found to be about 3 – 7-fold lower than in the mutant HIV-1 PR variants (MUT-1 = 49.69 ± 3.27 nM and MUT-2 = 79.24 ± 4.48 nM). The WT HIV-1 PR K_i value for DRV (2.67 ± 0.38 nM) was approximately 3 – 6-fold lower than the mutants (MUT-1 = 13.21 ± 1.23 nM and MUT-2 = 19.89 ± 3.02 nM). This followed a similar pattern as the K_i values obtained from the enzyme inhibition analysis. The k_5 values of the WT for LPV is 0.17 s^{-1}, 0.11 s^{-1} and 0.10 s^{-1} for MUT-1 and MUT-2 respectively. The k_5 value of the WT for DRV is 0.14 s^{-1}, 0.11, and 0.12 s^{-1} for MUT-1 and MUT-2, respectively (Table S4, Figure 6 and 7 A – F). The kinetic constant obtained from the slow loss of intrinsic Trp fluorescence agrees with the ones obtained from the Kinetic analysis. This shows the initial

rapid fluorescence decrease can be associated with the formation of the reversible complex EI. In contrast, the time-dependent decrease indicates the accumulation of the tight bound slow dissociating complex EI*.
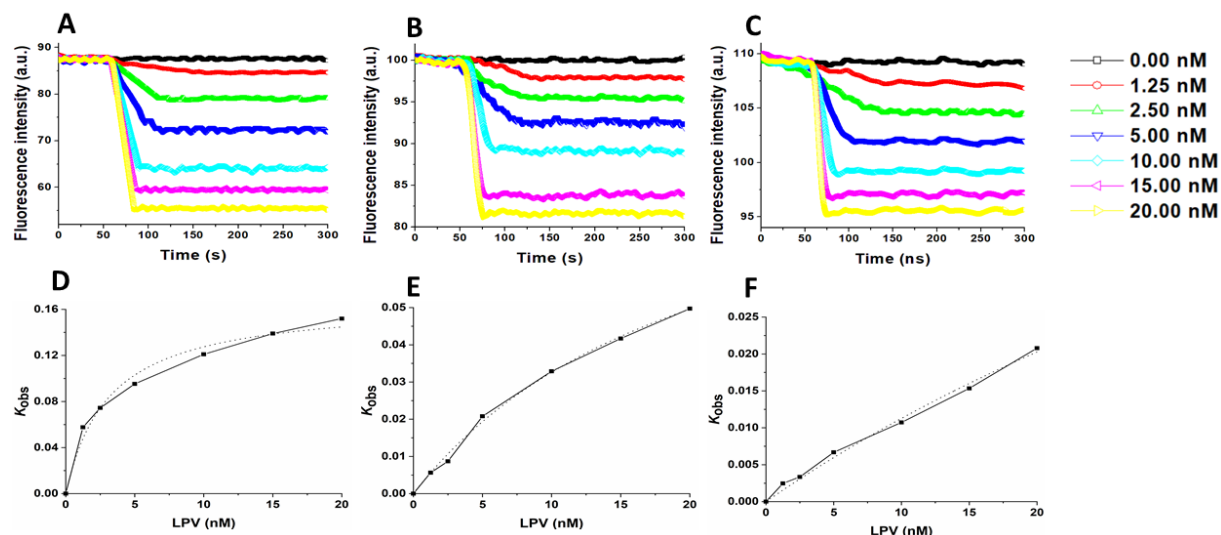
Figure 6. (A – C) Time-dependent effect of LPV on the intrinsic fluorescence of HIV-1 PR (WT=A, B=MUT-1, C=MUT-2). (D – F) The decay of HIV-1 PR fluorescence at different concentrations of LPV (WT=A, B=MUT-1, C=MUT-2) used to calculate the Ki and k5 (Dotted lines represent the best line of fit).
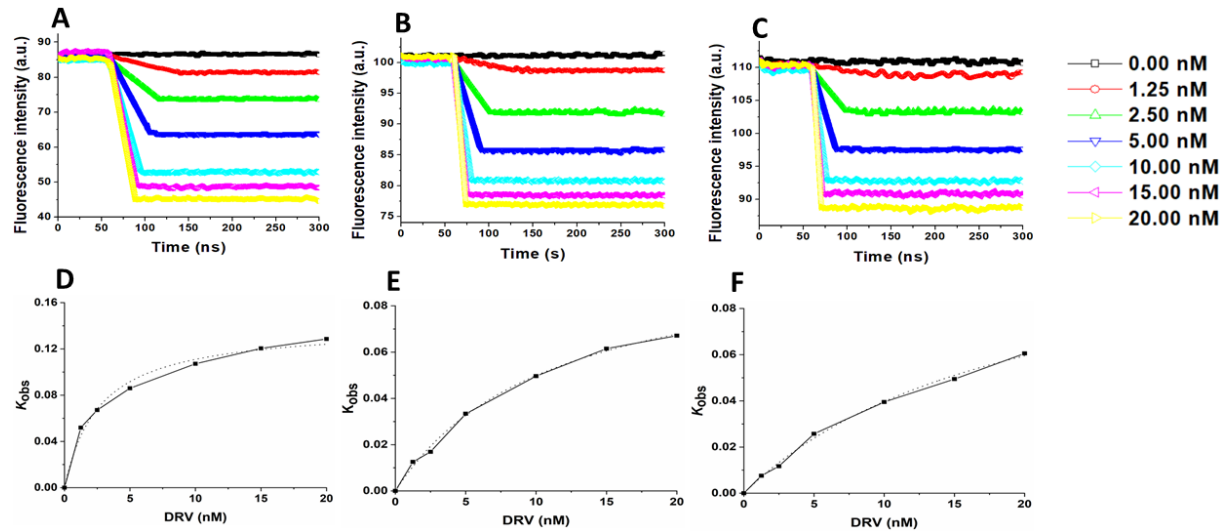
Figure 7. (A – C) Time-dependent effect of DRV on the intrinsic fluorescence of HIV-1 PR (WT=A, B=MUT-1, C=MUT-2). (D – F) The decay of HIV-1 PR fluorescence at different concentrations of DRV (WT=A, B=MUT-1, C=MUT-2) used to calculate the Ki and k5 (Dotted lines represent the best line of fit).

The fluorescence data obtained were further analyzed to determine the Stern-Volmer constant (K_{sv}) by plotting linear Stern-Volmer plots (F_0/F vs inhibitor concentrations) (Figures 4 G – I and 5 G – I). The Stern-Volmer quenching constants (K_{sv}) obtained from analysis of the fluorescence data measures the

accessibility of fluorophores to LPV and DRV and the solvent accessibility of the fluorophores to probe the conformational changes in the HIV-1 PR variants (Eche *et al*., 2021). The K_{sv} values (Table S3) for the interaction of LPV with the WT (0.019 nM^{-1}) was higher compared to the mutant HIV-1 PRs (0.003 nM^{-1} and 0.003 nM^{-1} for MUT-1 and MUT-2 respectively), and similar results were observed for the interaction of DRV with the WT (0.025nM) and mutants (0.008 nM^{-1} and 0.005 nM^{-1} for MUT-1 and MUT-2 respectively). The K_{sv} values of DRV interaction with the wild-type and mutant HIV-1 PR variants were higher compared to the K_{sv} values of LPV.

3.3 HIV-1 PR Structure analysis

The WT and the mutant HIV-1 PR variant structures were stable throughout the simulation trajectory of 10ns. The root mean square deviation (RMSD) of the backbone carbon atoms (Figure S2A) of the WT was 1.33 Å, slightly lower than the average RMSD of MUT-1 (1.56 Å) and MT-2 (1.80 Å). The root mean square fluctuation (RMSF) of Cα atoms (Figure S2B) showed increased fluctuation around the flap residues (residues 45 – 55/ 45' – 55') of MUT-1 and MUT-1 compared to the wild type. The difference in the structure of the WT and mutant HIV-1 PR variants were analyzed to determine the impact of the accumulated drug resistance mutations on the HIV-1 PR structure to gain further insight into its influence on the binding kinetics. The flaps of the mutant HIV-1 PR variants (MUT-1 and MUT-2) were in an open conformation compared to the closed conformation in the WT (Figure 8A).

3.3.1 HIV-1 PR active site cavity surface area and volume measurement

The impact of the open conformation and open flap on the HIV-1 PR active site cavity was shown by the measurements obtained from the CASTp 3.0 server (Figure 8B – C). The active site cavity surface area (392.59 $Å^2$) and volume (295.66 $Å^3$) of the WT were significantly lower than the mutant HIV-1 PR variants. MUT-1 and MUT-2 had active site cavity surface area of 834.39 $Å^2$ and 896.82 $Å^2$, respectively, with active site volume of 1240.40 $Å^3$ and 1259.98 $Å^3$, respectively. This finding suggests a decrease in the contact surface area of LPV and DRV with the active site residues and may be associated with the low association and increased dissociation rate constants and the low residence time of LPV and DRV for the mutant HIV-1 PR variants compared to the wild type observed in this study.
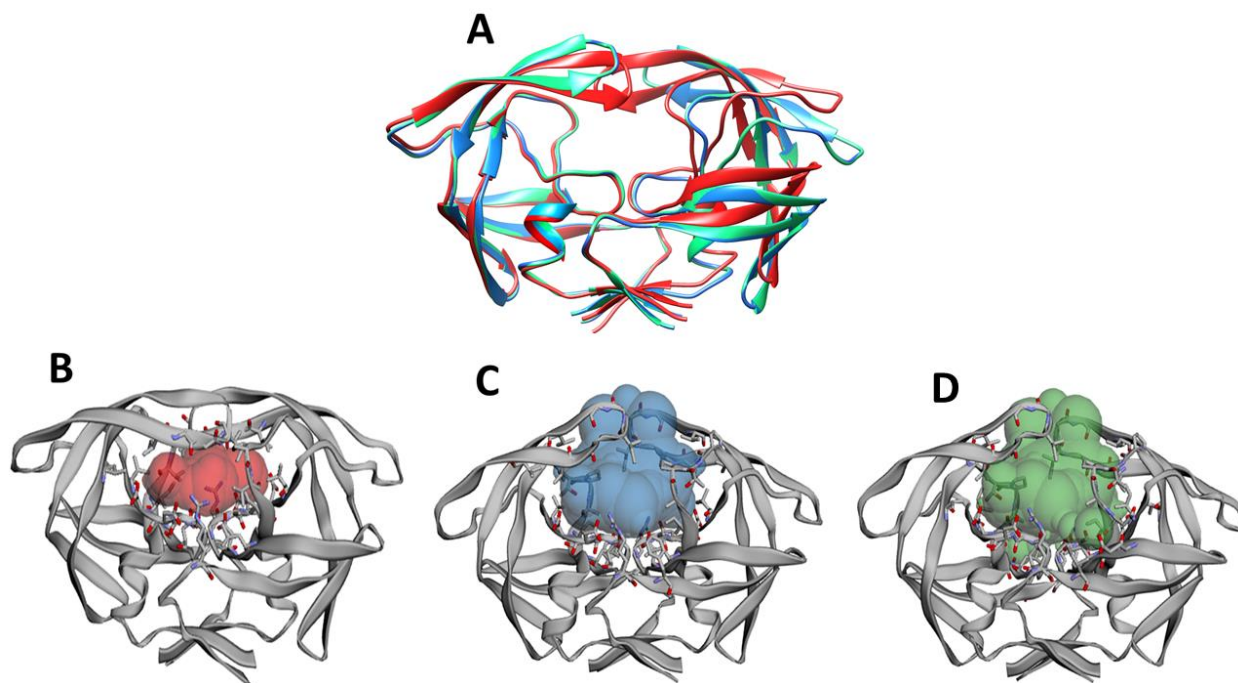
Figure 8. (A) Superimposed structure of the WT and the different mutant HIV-1 PR variants (B – D) Diagrammatic representation of the active site cavity area and volume for the HIV-1 PR WT (B), MUT-1 (C), and MUT-2 (D).

4. Discussion

This study describes the binding kinetics of LPV and DRV with the WT and highly mutated HIV-1 PR variants from clinical isolates obtained from HIV-1 infected individuals at the point of failing a LPV-inclusive regimen and being switched to DRV-based therapy. The enzyme kinetic analysis (Table S1) showed that the mutations harbored by these variants caused a marked decrease in the catalytic activity and catalytic efficiency of the mutant HIV-1 PR variants. Furthermore, the inhibitory constants (K_i) show that the mutations severely affected the K_i for both LPV and DRV. The K_i values obtained in this study followed a similar trend as those we reported in a previous study (chapter three) for the wild type and mutated South African HIV-1 subtype C PR (Eche *et al.*, 2021). The variation in the K_i values of the WT and mutant HIV-1 PR may be associated with the difference in the association and dissociation rate constants and the inhibitor residence time, as will be discussed subsequently.

In order to gain more insight into the binding kinetics of LPV and DRV with highly mutated HIV-1 subtype C PR, asides from the information obtained based on the optimization of the K_i values, the determination of rate constants for HIV-1 PR inhibition was essential and required the use of mathematical models (Pargellis *et al.*, 1994). In addition to determining the association and dissociation rate constants, this current study also explored the mode of inhibition of HIV-1 PR by LPV and DRV. The initial kinetic

analysis showed a tight-binding and mixed inhibition of HIV-1 PR by LPV and DRV. Tight-binding inhibitors are inhibitors that have a binding constant (K_i) at /or below the enzyme concentration used in the reaction ($K_i \leq [E]_0$) (Murphy, 2004). The data showing the lifetimes of enzyme-inhibitor complexes, which did not show a slow dissociation rate constant (k_6), indicating that the enzyme inhibition mechanism was not a slow-tight binding inhibition mechanism, confirmed the tight-binding inhibition. In contrast, a slow-tight binding inhibition mechanism involves enzyme kinetic reactions where the dissociation rate constant (k_6) is relatively slow and long regardless of the association rate constant (k_5) (Kumar and Rao, 2010).

The K_i and K_i' values obtained from the double-reciprocal plots at the different inhibitor concentration [I] by fitting the data in Eq. (1c) shows $K_i > K_i'$ and the double-reciprocal plots intersect above $1/v$ axis, supporting the mixed inhibition mechanism (Ramsay and Tipton, 2017). This inhibition mechanism occurs when the inhibitor binds to the active site of the free enzyme (E) and distant sites outside the active site. An inhibitor's potential to bind to the target outside the active site is the determinant for the formation of the ESI complex, resulting in the observed mixed inhibition. If the inhibitors do not have a second binding site outside the active site, the ESI complex will not form; thus, the inhibition will not be a mixed one (Silverman, 2007). This inhibition mechanism has also been demonstrated in a study to involve two non-mutually exclusive sites for inhibitor binding on the enzyme (Lorey *et al*., 2003).

The mixed-type inhibition of the HIV-1 PR variants by LPV and DRV is described in scheme 1 and may be attributed to the binding of these inhibitors at the active site and at sites distant to the active site. This has been demonstrated in previous studies for DRV and saquinavir (SQV) (Kovalevsky *et al*., 2008, Kovalevsky *et al*., 2006). The authors suggest that this phenomenon, which is rare for competitive inhibitors like DRV and SQV is due to the presence of a second binding site on HIV-1 PR for these inhibitors (Kovalevsky *et al*., 2008, Kovalevsky *et al*., 2006). To the best of our knowledge, we are showing this mixed inhibition mechanism for LPV for the first time and using highly mutated HIV-1 PR from clinical isolates obtained from HIV-1 infected individuals who are failing an all-inclusive LPV regimen. The presence of binding sites on antiviral drug targets outside the active site indicates that the drug target can accommodate inhibitors with different inhibition mechanisms (Hang *et al*., 2009). The binding of PIs outside the active site also interferes with normal enzyme function; this mechanism of inhibition has been proposed to develop novel PIs (Zhang *et al*., 2014, Chen *et al*., 2018). This is because inhibitors that bind at allosteric sites are not affected by the increase in substrate concentration during virion assembly compared to inhibitors that bind to the active site through a competitive binding mechanism (Chang *et al*., 2010).

Aside from the determination of the mode of inhibition, the association (k_5) and dissociation (k_6) rates analysis of HIV-1 PR inhibition provides a unique possibility of determining the affinities of PIs for HIV-1 PR. The results of the kinetic profiling of HIV-1 PR inhibition by LPV and DRV in the current study have shown that both LPV and DRV had a higher k_5 for converting EI to EI* when inhibiting the WT compared to the mutant variants. Similarly, the k_6 value was markedly higher in the dissociation of EI* to EI when LPV and DRV inhibit the mutant HIV-1 PR variants compared to the WT (Table S2). This finding may be due to the impact of the mutations harbored by the mutant HIV-1 PR variants, causing an increase in the dissociation of the inhibitor bound to the enzyme. The decreased association rate and the increased dissociation rate of LPV and DRV with the mutant HIV-1 PR variants may be a product of the distortion of the HIV-1 PR conformation caused by these mutations. These changes may occur at the inhibitor binding site, affecting the inhibitor's recognition and binding kinetics (Huang *et al.*, 2017). The association and dissociation rate constants determine the lifetime of the interaction of an inhibitor with its target. Generally, the enzyme-inhibitor complex's lifetime is extended when the drug binds rapidly and with a slow dissociation rate from the enzyme-inhibitor complex (Copeland, 2016). The current study's finding agrees with a similar study that showed that drug resistance mutations cause decreased association and increased dissociation rates of LPV and DRV when bound to HIV-1 PR. However, a different technique from initial kinetic analysis was used (Dierynck *et al.*, 2007). This shows that considering the rates of association and dissociation in the design of drugs is important. This will help design inhibitors with the desired kinetic characteristics such as the mode of binding, association, and dissociation from the enzyme-inhibitor complex (Holdgate *et al.*, 2018).

The low rate of association and increased dissociation in mutants studied may have been highly influenced by the active site mutations harbored. The V82A mutation is located at a region critical to inhibitor and substrate binding since it makes direct contact with both PIs and substrate. This mutation causes resistance to HIV-1 PIs through a structural shift of residues in the 80's loop, impacting the interaction and tight binding of inhibitors in the active site (Tang and Shafer, 2012, Agniswamy *et al.*, 2019). The I84V mutation also found in the active site seen in MUT-2 in this study causes cross-resistance to several PIs through the loss of van der Waals interaction between active site residues and the HIV-1 PIs promoting the dissociation of PIs when bound to HIV-1 PR (Ragland *et al.*, 2017). Aside from the loss of van der Waals interaction, the I84V mutation also causes an alteration in the hydrophobic core sliding mechanism (26, 34), impacting the binding affinities of HIV-1 PIs (Rhee *et al.*, 2010, Clemente *et al.*, 2004).

The determination of the association rate constant (k_5) of HIV-1 PR may be necessary; however, research has shown that the dissociation rate constant (k_6) is more therapeutically significant (Pargellis *et al.*, 1994). This is because the dissociation rate constant (k_6) is a means for the calculation of the residence time (t_R) of

an inhibitor when bound to a protein (Pan *et al*., 2013). Research has shown that the residence time of an enzyme-inhibitor complex is a better predictor of drug-receptor interaction than the measurement of binding affinity (Copeland *et al*., 2006, Guo *et al*., 2012). In addition, the residence time of an inhibitor may be a reflection of the desired in vivo pharmacological action (Copeland *et al*., 2006, Copeland, 2011). The results of the residence time (Table S2) calculated in this study showed that the residence time for the interaction of LPV and DRV with the wild type was markedly higher compared to the residence time of these inhibitors with the mutant HIV-1 PR variants. This may be associated with the conformational changes caused by these mutations affecting the lifetime of the enzyme-inhibition complex due to the more open conformation of the mutant HIV-1 PR variants, as will be described subsequently. Research favors the measurement of residence time from the rate constant determination over other in-vitro methods because the residence time also accounts for the impact of the conformational dynamics of the drug target on inhibitor binding and dissociation (Copeland, 2016).

Interestingly, we did not find any difference in the residence of LPV and DRV for the WT HIV-1 PR, while the residence time of LPV showed almost a 2-fold increase for MUT-1 compared to DRV and for MUT-2 the residence time of LPV was slightly higher to that of DRV. This contrasts with the knowledge that DRV is more potent than LPV and binds tightly in the active site due to more hydrogen bond interaction with the HIV-1 PR backbone than other PIs (Lefebvre and Schiffer, 2008). Furthermore, it is important to note that though DRV showed a much lower K_i compared to LPV in the inhibition of the WT and mutant HIV-1 PR variants (Table S1) but has a higher dissociation constant compared to LPV. This is because the K_i only gives information about the strength of an inhibitor's interaction with its target, describing the percentage of an enzyme bound to an inhibitor at equilibrium. It does not give information about the rate at which the binding of an inhibitor to its target occurs at equilibrium, which is only obtained from the association and dissociation rate constants (de Witte *et al*., 2018). The higher rate constant for conversion of EI to EI* (k_5) and the lower rate constant for the dissociation of EI* to EI (k_5) together with the higher t_R ($1/k_6$) of the WT HIV-1 PR compared to the mutant HIV-1 PR variants gives a clue about the effective inhibition of former by these inhibitors. In addition, it throws more light on the impact of the mutations harbored by the mutant variants on the evolution of resistance to LPV and DRV.

Further confirming the two-step inhibition and time-dependent HIV-1 PR inhibition by LPV and DRV is the observed intrinsic Trp fluorescence quenching pattern. The obtained fluorescence data show that the loss of fluorescence is attributed to the formation of the reversible HIV-1 PR-LPV or DRV complexes. The slow decrease in fluorescence intensity resulted from the increased formation of the tightly bound EI* complex. The absence of any shift (blue or red) in the intrinsic Trp fluorescence maxima because of the increase in the concentration of LPV and DRV (Fig. 4 and 5 A – C) is an indication that there was no

significant change in the three-dimensional structure of the HIV-1 PR variants associated with the isomerization of the EI complex to EI*. This agrees with the values of the rate constant determined from the enzyme kinetic experiments, indicating the localized changes in conformation in the enzyme-inhibitor complex is associated with the isomerization of the EI to EI*. The gradual decrease in fluorescence intensity of the HIV-1 PR variants related to the increase in the concentration of LPV and DRV is an indication that the quenching of HIV-1 PR variants by these inhibitors is due to the formation of the enzyme-inhibitor complex (Bekale *et al.*, 2014).

The impact of the mutations harbored on the conformation of the mutant HIV-1 PR variants and their effects on binding kinetics is highlighted by the data obtained from the analysis of intrinsic Trp fluorescence. This has helped clarify the impact of the conformational flexibility of the mutant HIV-1 PR variants on the binding kinetics. Drugs and their target proteins are dynamic; as a result, their conformational flexibility may influence the binding kinetics in subtle ways, thus controlling how the inhibitor binds to the active site (Miller *et al.*, 2012). The L76V mutation seen in both mutants in this study is associated with conferring flexibility to HIV-1 PR through the local rearrangement of the HIV-1 PR hydrophobic core and alteration of the dimer stability, casing the inhibitors to be unstable when they bind to HIV-1 PR (Wong-Sam *et al.*, 2018). Furthermore, the conformational changes in the tertiary structure of the mutant HIV-1 PR may be due to the flap mutations harbored. The M46I and I54V mutations confer high flexibility on the HIV-1 PR flap. This affects the stability of the ligand when bound to the HIV-1 PR active site (Halder and Honarparvar, 2019, Huang *et al.*, 2014b, Mahanti *et al.*, 2016). This explains why these mutations contribute to the decreased binding affinity of HIV-1 PR to PIs, even when they are not found in the active site (Clemente *et al.*, 2004).

The increased intrinsic Trp fluorescence seen in the mutant HIV-1 PR variants compared to the WT (Fig S1) and the observed low K_{sv} values from the interaction of the mutants with LPV and DRV is an indication that the mutant HIV-1 PR is in a more open state rather than in a semi-open or closed conformational state (Table S3, Fig. 4 and 5 G – 1). This open conformational state in the mutant variants arising from the structural reorganization caused by the mutations harbored leads to the LPV and DRV not easily accessing and quenching the Trp fluorescence. This finding may be attributed to the low rate of association and high dissociation rates in the interaction of the mutant HIV-1 PR with LPV and DRV observed from the initial kinetic analysis, thus lending support to the increased resistance to LPV and DRV. Enzyme kinetics models have been used to demonstrate how conformational changes in HIV-1 PR affect catalysis (Weikl *et al.*, 2013). In addition, mutations that cause disbalance between a closed and open conformation to favor an open conformation keep the enzyme catalytically active while conferring resistance on HIV-1 PR to PIs (Weikl *et al.*, 2013).

The increased dissociation rate, low residence time, as well as low K_{sv} of LPV and DRV when bound to the mutant HIV-1 PR variants agree with the finding showing an expansion in the active site cavity of the mutant HIV-1 PR variants (Fig. 8C – D). The open conformation of the highly mutated HIV-1 PR variants results in the expansion of the active site, as shown by the marked increase in the active site surface area and volume, leading to reduced contact of LPV and DRV with the active site residues. This agrees with the findings of a similar study that showed that the expansion of the active site is a consequence of mutated HIV-1 PR failing to attain a closed state when bound to an inhibitor, thus leading to a loss of the potency of the inhibitor (Goldfarb et al., 2015). Research has shown that the expansion of the HIV-1 PR active site leads to the loss of tight binding and low binding affinity of HIV-1 PIs due to the loss of molecular interaction like the hydrogen bonds and van der Waals interactions between the inhibitors and the active site residues (Goldfarb et al., 2015, Sk et al., 2021).

The reduced contact surface area and increased volume of the active site of the mutant HIV-1 PR variants may be likely responsible for increased dissociation rate and low residence time of LPV and DRV when bound to the highly mutated HIV-1 PR variants. The V32I mutation harbored by MUT-2 and I84V harbored by MUT-1 in this study may have contributed significantly to the alteration of the active site surface are, and volume observed. Both the V32I and I84V mutation have been implicated in the pathway to the evolution of resistance to HIV-1 PIs by altering the hydrophobic core sliding mechanism (Ragland et al., 2014, Mittal et al., 2012). The hydrophobic core allows the anchoring of the flap and participates in regulating the opening and closing of the flap (Naicker et al., 2014). Therefore, the perturbation of the hydrophobic sliding mechanism leads to the failure of HIV-1 PR to attain a fully closed state. This causes the mutated HIV-1 PR to have an expanded active site leading to the loss of molecular interactions and low binding affinity for HIV-1 PIs (Goldfarb et al., 2015).

5. Conclusion

The results of this study provide insight into the binding kinetics and molecular mechanism of the inhibition of highly mutated HIV-1 PR by LPV and DRV. The results suggest that LPV and DRV follow a two-step tight binding and mixed inhibition mechanism in which LPV and DRV bind to HIV-1 PR at both the active site and at an allosteric site. The possibility of LPV and DRV to bind to an allosteric site signals the existence of a second binding site for these inhibitors, highlighting the potential for antiviral compounds with the capacity to bind to HIV-1 PR at both the active site and sites distant from the active site. This is promising for antiviral drug development as a combination of drugs with different enzyme inhibition mechanisms will raise the barrier to drug resistance evolution. In addition, the results of the binding kinetics experiments show the importance of measuring the rate constants of inhibition and the determination of the

inhibitor residence time as this provides valuable information needed to optimize the lifetime of the enzyme-inhibitor complex leading to a more improved inhibitor residence time and binding affinity. Thus, the information provided in this study will add to existing knowledge for the design of novel HIV-1 PR inhibitors with the potential to evade the impact of drug resistance mutations at the active site that affect the binding affinity of HIV-1 PIs, taking into consideration the mechanism of mixed inhibition.

CHAPTER FIVE

General Discussion, conclusion, future recommendations, and limitations

General discussion

The development of resistance to antiretroviral therapy (ART) is a threat to the treatment of HIV-1 infected individuals (WHO, 2019). The use of HIV-1 PIs in formulating HAART has helped improve the life expectancy of HIV-1 infected individuals (Venter *et al*., 2019, Teply *et al*., 2011). However, the benefits of using PIs can be transitory as HIV-1 infected individuals may experience virological failure due to characteristic PI mutations that emerge in the HIV-1 PR gene (Henderson *et al*., 2012). The impact of drug-resistance mutations on the inhibitory kinetics and structural conformation of HIV-1 PR has been well studied. However, there is a paucity of literature on the binding kinetics of HIV-1 PIs with HIV-1 PR. In the literature, more commonly found is information on the optimized inhibitory constant (K_i) values for HIV-1 PIs. As such, the principles underlying the binding affinity of inhibitors for their target are relatively well known, but the same cannot be said for the binding kinetics (Pan *et al*., 2013).

Understanding the molecular determinants that regulate drug-receptor binding kinetics is essential to determining the mechanism of action of inhibitors and the pathway to the evolution of resistance to a drug, thus providing the knowledge needed to guide the design of novel inhibitors (Lee *et al*., 2019). These molecular determinants that control the binding kinetics include the rate of association and dissociation, as well as the inhibitor residence time (Lee *et al*., 2019). The current study investigated the kinetic and structural properties of highly mutated HIV-1 subtype C PR. In addition, the rate of association and dissociation of the lifetime of the enzyme-inhibitor complex, drug residence time, and mode of inhibition of DRV and LPV was determined using enzyme kinetic modeling analysis (initial kinetic analysis).

Firstly, this current research reviewed relevant literature on the use and impact of different fusion tags on the expression of HIV-1 PR using recombinant DNA technology and its purification (chapter 2). The purpose of this review is to provide a quick start to scientists who have not expressed HIV-1 PR before. The availability of functional proteins in the right concentration and purity will aid the biochemical or structural characterization of HIV-1 PR, providing invaluable information that can assist with the design of novel tight-fitting and binding HIV-1 PIs (Hewitt *et al*., 2011, Ghosh *et al*., 2018). The methods of expressing HIV-1 PR in the past are associated with a common drawback of the expressed protein accumulating as insoluble aggregates in the expression host cell (Karacostas *et al*., 1993, Wan *et al*., 1995, Cheng *et al*., 1990). The non-standardized method of refolding proteins from insoluble protein aggregates is a significant disadvantage of using these methods (Idicula-Thomas and Balaji, 2007). In addition, the use

of large volumes of reagents in the renaturation of proteins from insoluble aggregates to their natural conformation may lead to the dilution of the protein (Strandberg *et al.*, 1991), and the proteins may not assume their natural conformation after chemical treatment (Thomas and Baneyx, 1996). This, therefore, supports the need to employ an in vivo method that will facilitate the production of a protein in a "soluble, pure, and active conformation to overcome the problems associated with *in vitro* systems" (Eche and Gordon, 2014).

Fusion tags have significantly helped to address the problems that accompany the recombinant expression of HIV-1 PR (Volontè *et al.*, 2011). The solubility enhancing fusion tags reviewed all showed great potential for the expression and purification of HIV-1 PR. The fusion tag with the most application in the expression of HIV-1 PR is the GST tag. These GST tags' applications range from the expression of the precursor to matured form and studying of HIV-1 PR auto-processing (Huang and Chen, 2013). The characteristics mentioned above gives the GST fusion tag an advantage over the other tags (MBP and Trx fusion tags) that have also been used in the recombinant expression of HIV-1 PR. However, the decreased yield and activity of HIV-1 PR expressed fused to the GST tag are significant drawbacks when a high protein concentration is needed (Volontè *et al.*, 2011). On the other hand, the MBP tag is also a reliable solubility enhancing tag that can be used to express the precursor and matured HIV-1 PR. When used to express HIV-1 PR, the concentration of the recovered protein and the specific activity is fairly high compared to when the GST tag is used (Eche and Gordon, 2021). This current study utilized the MBP fusion tag for the expression of the HIV-1 PR variants used.

For the first time, we have characterized highly mutated South African HIV-1 subtype C PR from clinical isolates showing the inhibitory constants (K_i) of their interaction with LPV and DRV. Unlike the HIV-1 subtype C PR, extensive studies have been carried out on the kinetic characterization of HIV-1 subtype B PR, and the K_i values of the wild type and different multidrug-resistant forms are known (Kneller *et al.*, 2020, Park *et al.*, 2016, Šašková *et al.*, 2008, Weber *et al.*, 2015). However, there is a paucity of data regarding the biochemical interaction of highly mutated HIV-1 subtype C PR variants with LPV and DRV, which this study has shown.

This study findings, for the first time, showed the comparative analysis of the biochemical and structural implications of the regimen switch from LPV to DRV during virological failure using the former (chapter 3). This is important as it is common for HIV-1 infected individuals failing LPV based regimens to be switched to DRV-based therapy (Santos *et al.*, 2012). The results of the comparative analysis of the biochemical and structural implications of this regimen switch revealed inherent changes in the HIV-1 structure of highly mutated HIV-1 PR associated with drug pressure from earlier LPV therapy which

affected the outcome of the DRV interaction with HIV-1 PR in a pattern similar to what was observed with LPV. Research shows enzymes are not static but dynamic, and changes in their conformational flexibility will affect their binding kinetics with both their substrate and inhibitors alike (Miller *et al*., 2012). This may further explain why the DRV binding landscape may have been affected. In addition, since HIV-1 PIs are similar in their chemical signature, mutations in HIV-1 PR that emerge due to drug pressure from one HIV-1 PI and distorting the HIV-1 PR structure may easily confer cross-resistance to other HIV-1 PIs (Chen *et al*., 1995, Schock *et al*., 1996, van Maarseveen and Boucher, 2006).

Mechanistic insight into the impact of these mutations revealed showed that the inherent distortion of the hydrophobic core and increased solvent exposure of the HIV-1 PR dimer associated with the mutations that emerge due to earlier therapy might affect the ART switch. The distortion of the hydrophobic core highlighted by the increased solvent accessibility area, radius of gyration, and low Stern Volmer constant (K_{sv}) points to these mutations causing the mutant HIV-1 PR dimer to assume an open conformation. This thus lends support to the decreased lifetime of the enzyme-inhibitor complex, promoting the dissociation of LPV and DRV from the binding site as observed in the binding kinetic analysis results of this current study. Distortion of the HIV-1 PR structure favoring an open conformation will result in the loss of tight-binding of inhibitors to HIV-1 PR (Weikl *et al*., 2013). The direct impact of accumulated HIV-1 PR mutations causing the dimer to assume an open conformation has been studied (Chetty *et al*., 2016, Mittal *et al*., 2012). This has been shown to cause inherent changes leading to increased conformational flexibility (Chetty *et al*., 2016) and the loss of protein compactness, thus modulating the HIV-1 PR activity and HIV-1 PI binding landscape (Mittal *et al*., 2012). The L76V mutation found in the mutant HIV-1 PR variants in this study promotes resistance to several HIV-1 PIs through the local rearrangement of the HIV-1 PR hydrophobic core, which results in the loss of protein compactness and high flexibility (Wong-Sam *et al*., 2018). The L76V mutation emerges during treatment with an LPV-based regimen, and is it is a common feature of HIV-1 PR variants highly resistant to DRV (Ragland *et al*., 2014, Tang and Shafer, 2012).

The flap mutations M46I and I54V harbored by the mutant HIV-1 PR variants in this study have been shown to significantly affect the flexibility of the HIV-1 PR flap, influencing the stability of the HIV-1 PR-PI complex (Halder and Honarparvar, 2019, Huang *et al*., 2014b, Mahanti *et al*., 2016). Research has shown that these mutations have a more severe impact on the development of resistance to HIV-1 PIs in the presence of the V82A mutation, causing high-level cross-resistance to several PIs (Clemente *et al*., 2004, Ohtaka *et al*., 2003). We have also shown that the open conformation caused by these mutations may result in the expansion of the active site, as demonstrated by the increased active site cavity surface area and volume. This may promote the dissociation of HIV-1 PIs from the active site through the loss of molecular interaction like the hydrogen bond contacts and van der Waals interactions between the inhibitors

and the active site residues. Research supports the loss of molecular interactions between inhibitors and their target protein due to active site expansion affecting the rate of association and dissociation and the residence time of inhibitors (Pan *et al*., 2013).

The results obtained show the superiority of DRV over LPV. A lower concentration of DRV was needed to inhibit the HIV-1 PR variants, and the better stability of the HIV-1 PR-DRV complexes compared to the HIV-1 PR-LPV complexes further confirms this. The capacity of DRV to be effective in an already altered binding landscape supports the switch from LPV to DRV based regimen. Previous research findings also support the superiority of DRV over LPV. Research attributes the superiority of DRV over LPV to its design, making it fit tightly in the active site with increased molecular interaction with the HIV-1 PR backbone. This confers a high binding affinity on DRV for the HIV-1 PR active site (Lefebvre and Schiffer, 2008). A clinical trial evaluating ART's long-term effectiveness showed that a DRV-based regimen was superior to an LPV-based regimen in both ART-naive and treatment-experienced HIV-1 infected individuals. These authors showed that the chances of virological failure developing with DRV than LPV when initiated either as a salvage regimen or a switching strategy in treatment-experienced HIV-1 infected individuals were lower (Santos *et al*., 2018). Another study also showed that the use of low dose DRV boosted with ritonavir (RTV) is an efficient switch option to suppress virological failure in HIV-1 infected individuals failing LPV based treatment (Venter *et al*., 2019).

Chapter 4 shows the relationship between the rate of association and dissociation and inhibitor residence time with the conformational changes during the isomerization of the EI complex to EI*. The initial kinetic analysis of the interaction of LPV and DRV with the HIV-1 PR variants showed tight-binding and mixed inhibition mechanism. The observed mixed inhibition mechanism indicates two non-mutually exclusive sites of inhibitor binding for LPV and DRV in HIV-1 PR. Though an earlier study has shown this for DRV (Kovalevsky *et al*., 2008, Kovalevsky *et al*., 2006), this is the first time it is reported for LPV. In addition, we have shown that LPV and DRV inhibited the HIV-1 PR variants in a two-step mechanism, as shown by the results obtained from the enzyme kinetics and fluorescent spectroscopy. This two-step binding mechanism is characterized by the isomerization of the EI complex to a tightly bound EI* complex. It has been shown in a previous study that the interaction of drugs designed to have affinity for their targets often follows a multi-step binding process, resulting in higher steric complementarity between the inhibitor and the target enzyme (Copeland, 2016). This may explain the reason for the two-step inhibition mechanism of LPV and DRV. Understanding the basis of the EI complex's isomerization to EI* complex from the application of enzyme kinetics will provide the quantitative information needed to design structures that permit the titration of the lifetime of an EI* complex and the development of future novel tight-binding inhibitors (Kumar and Rao, 2010).

Measuring the association (k_5) and dissociation (k_6) rate constants provided insight into the lifetime of the enzyme-inhibitor complex when LPV and DRV are bound to the HIV-1 PR variants. It also showed how mutations affected the association (k_5) and dissociation (k_6) rate constants and the overall residence time (t_R) of the inhibitors when bound to HIV-1 PR. Measuring the dissociation rate constant has been suggested in previous studies to be the most crucial parameter in determining the lifetime of the enzyme-inhibitor complex. This is because it is a means for estimating the inhibitor residence time (t_R = 1/dissociation constant) when bound to a target protein (Copeland, 2016, Copeland, 2011, Copeland *et al.*, 2006). Research supports the use of the inhibitor residence time as a better predictor of drug-receptor interaction than the measurement of binding affinity (Copeland *et al.*, 2006, Guo *et al.*, 2012, Copeland, 2016). This is because conventional methods that directly measure the binding affinity of drug-target interactions provide results solely in terms of the equilibrium affinity. They do not consider the changes in the target enzyme's conformational dynamics, affecting the rates of association and dissociation, unlike the residence time model (Copeland, 2016).

The current study, has shown using binding kinetics and fluorescence spectroscopy the relationship between the rates of the enzyme reaction and conformational changes in highly mutated HIV-1 PR and how this impacts the binding of LPV and DRV. Inhibitors are designed to bind to their targets at a specific conformation. Changes in the conformations of target enzymes arising from mutations will affect the affinity of inhibitors for their targets (Rusnak *et al.*, 2001, Copeland, 2011). This may explain the impact of the conformational changes in the highly mutated HIV-1 PR variants in this study on the level of resistance seen to LPV and DRV in addition to the fast rate of dissociation when bound to the mutant HIV-1 PR variants.

Interestingly LPV and DRV had almost the same residence time when bound to the WT HIV-1 PR. However, LPV had a longer residence time than DRV when bound to the mutant HIV-1 PR variants. This finding confirms the importance of measuring the association and dissociation rate to get detailed information about the binding kinetics. This is because the inhibitory constant (K_i) will only show how potent an inhibitor is and will not provide information on the rate at which the inhibitor binds to its target or dissociates from the enzyme-inhibitor complex (de Witte *et al.*, 2018). Though DRV is more potent than LPV, the longer residence time exhibited by LPV in this study can be leveraged upon as a platform to develop future HIV-1 PR inhibitors with the capacity to bind for a longer time to their target. Research shows that the current advances being made in the development of future HIV-1 PIs involve incorporating novel ligands into existing HIV-1 PIs for higher binding affinity and increased inhibitory capacity (Ghosh *et al.*, 2016). An example is incorporating "cyclopentyltetrahydrofuran and substituted Bis-tetrahydrofuran as a P2 ligand" (Ghosh *et al.*, 2016) into existing inhibitors. This method has been used to develop a novel

HIV-1 PI with similarities to DRV in chemistry and structure currently undergoing clinical development (Ghosh *et al*., 2016, Ghosh *et al*, 2018).

Conclusions

For the first time, this study provided a link between the conformational changes in highly mutated HIV-1 PR selected during LPV therapy with the outcome of DRV therapy after a switch from LPV to DRV. As shown in this current study, the efficacy of DRV after ART switch is dependent on the changes in HIV-1 PR acquired as a result of selection pressure during LPV therapy. However, this study's results support the switch from LPV to DRV as only a low concentration of DRV was needed to effectively inhibit the highly mutated HIV-1 PR variants with an already altered binding landscape. This makes it the HIV-1 PI of choice to treat HIV-1 PI experienced HIV-1 infected individuals. From the results obtained from enzyme kinetic and structural analysis, this study proposes the expansion of the active site resulting in loss of tight fit of inhibitors, fast dissociation rate, and short residence time may be a mechanism of evolution resistance to LPV and DRV by highly mutated HIV-1 PR. This study also highlights the importance of determining the residence time of HIV-1 PIs. Binding kinetic analysis showed that both LPV and DRV are tight-binding inhibitors, and their inhibition mechanism involves a two-stage mixed inhibition mechanism. The mixed enzyme inhibition mechanism displayed by LPV and DRV suggests a second binding site on the HIV-1 PR for LPV and DRV. This finding supports the strategy for developing allosteric inhibitors to tackle HIV-1 PI resistance.

Future recommendations and limitations

From the results presented in this thesis, further research is needed to understand the development of resistance to LPV and DRV. There is a need for the impact of the individual mutations on the structural and kinetic characteristics of HIV-1 PR to be studied using site-directed mutagenesis. Furthermore, research is needed to gain insight into the structural changes occurring during the lifetime of the enzyme-inhibitor complex. Understanding the transition phase between the rate of association and dissociation using X-ray crystallography, NMR, and computational methods of resolving protein structures will provide more information about the mode of interaction of HIV-1 PIs and the necessary information to develop novel HIV-1 PIs with the required residence time.

From the afore mentioned, this study could be extended by using the method of HIV-1 PR expression described in chapter three to make available soluble pure HIV-1 PR for characterization and probing its interaction with different lead compounds. This approach can be used to elucidate the structure-activity relationships of HIV-1 PR and a different set of drug-like compounds. The data obtained from the

biophysical characterization of the HIV-1 PR with these new compounds can be used to identify novel inhibitors with the desired mechanisms. This will aid the identification of compounds with the possibility of translation through clinical trials for clinical treatment. In addition to fluorescence spectroscopy used in this current study, other biophysical methods such as Isothermal titration calorimetry and Differential scanning calorimetry to determine the thermodynamic parameters of interactions of HIV-1 PR with PIs can be incorporated to get complete details of the molecular interaction of PIs with HIV-1 PR.

References

Acharya, K. R. & Lloyd, M. D., 2005. The advantages and limitations of protein crystal structures. *Trends in Pharmacological Science*, 26, 10-14.

Adamson, C. S., 2012. Protease-mediated maturation of HIV: inhibitors of protease and the maturation process. *Molecular biology international*, 2012.

Agniswamy, J., Kneller, D. W., Brothers, R., Wang, Y.F., Harrison, R. W. & Weber, I. T., 2019. Highly Drug-Resistant HIV-1 Protease Mutant PRS17 Shows Enhanced Binding to Substrate Analogues. *ACS omega*, 4 (5), 8707-8719.

Ali, A., Bandaranayake, R. M., Cai, Y., King, N. M., Kolli, M., Mittal, S., Murzycki, J. F., Nalam, M. N., Nalivaika, E. A. & Özen, A., 2010. Molecular basis for drug resistance in HIV-1 protease. *Viruses*, 2, 2509-2535.

Ambrose, Z. & Aiken, C., 2014. HIV-1 uncoating: connection to nuclear entry and regulation by host proteins. *Virology,* 454, 371-379.

Anderson, A. C., 2003. The process of structure-based drug design. *Chemistry & biology,* 10, 787-797.

Anderson, J., Schiffer, C., Lee, S.K. & Swanstrom, R., 2009. Viral protease inhibitors. *Antiviral strategies.* Springer. 85-110

Andersson, H. O., Fridborg, K., Löwgren, S., Alterman, M., Mühlman, A., Björsne, M., Garg, N., Kvarnström, I., Schaal, W. & Classon, B., 2003. Optimization of P1–P3 groups in symmetric and asymmetric HIV-1 protease inhibitors. *European journal of biochemistry*, 270, 1746-1758.

Aoki, M., Das, D., Hayashi, H., Aoki-Ogata, H., Takamatsu, Y., Ghosh, A. K. & Mitsuya, H., 2018. Mechanism of darunavir (DRV)'s high genetic barrier to HIV-1 resistance: a key V32I substitution in protease rarely occurs, but once it occurs, it predisposes HIV-1 to develop DRV resistance. *MBio*, 9(2).

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

Aslantas, Y., Surmeli, N. B., 2019. Effects of N-Terminal and C-Terminal Polyhistidine Tag on the Stability and Function of the Thermophilic P450 CYP119. 2019. *Bioinorganic chemistry and applications*, 2019.

Azarnezhad, A., Sharifi, Z., Seyedabadi, R., Hosseini, A., Johari, B. & Fard, M. S., 2016. Cloning and expression of soluble recombinant HIV-1 CRF35 protease-HP thioredoxin fusion protein. *Avicenna journal of medical biotechnology,* 8, 175.

Bastys, T., Gapsys, V., Walter, H., Heger, E., Doncheva, N. T., Kaiser, R., De Groot, B. L. & Kalinina, O. V., 2020. Non-active site mutants of HIV-1 protease influence resistance and sensitisation towards protease inhibitors. *Retrovirology,* 17, 1-14.

Bekale, L., Chanphai, P., Sanyakamdhorn, S., Agudelo, D. & Tajmir-Riahi, H., 2014. Microscopic and thermodynamic analysis of PEG–β-lactoglobulin interaction. *RSC advances*, 4(5), 31084-31093.

Benko, Z., Elder, R. T., Li, G., Liang, D. & Zhao, R. Y., 2016. HIV-1 Protease in the Fission Yeast Schizosaccharomyces pombe. *PLoS One,* 11, e0151286.

Bertero, A., Brown, S. & Vallier, L., 2017. Methods of Cloning. *Basic Science Methods for Clinical Researchers.* Elsevier.

Bessong, P. O. & Iweriebor, B., 2016. A putative HIV-1 subtype C/CRF11_cpx unique recombinant from South Africa. *SpringerPlus*, 5 (1), 1-15.

Bhattacharya, S., Reddy, D., Reddy, R., Sharda, A., Bose, K. & Gupta, S., 2016. Incorporation of a tag helps to overcome expression variability in a recombinant host. *Biotechnology Reports*, 11, 62-69.

Blanco, R., Carrasco, L. & VentosO, I. 2003., Cell killing by HIV-1 protease. *Journal of Biological Chemistr*y, 278, 1086-1093.

Blood, G.A.C., 2016. Human immunodeficiency virus (HIV). *Transfusion Medicine and Hemotherapy*, 43, 203.

Booth, W. T., Schlachter, C. R., Pote, S., Ussin, N., Mank, N. J., Klapper, V., Offermann, L. R., Tang, C., Hurlburt, B. K. & Chruszcz, M. 2018. Impact of an N-terminal Polyhistidine Tag on Protein Thermal Stability. *ACS omega,* 3, 760-768.

Briggs, J. A. & Kräusslich, H.G., 2011. The molecular architecture of HIV. *Journal of molecular biology*, 410, 491-500.

Brik, A., Wong, C.H., B., 2003. HIV-1 protease: mechanism and drug discovery. *Organic & biomolecular chemistry*, 1, 5-14.

Broglia, R., Provasi, D., Vasile, F., Ottolina, G., Longhi, R. & Tiana, G., 2006. A folding inhibitor of the HIV-1 protease. *Proteins: Structure, Function, and Bioinformatics,* 62, 928-933.

Burley, S. K., Berman, H. M., Christie, C., Duarte, J. M., Feng, Z., Westbrook, J., Young, J. & Zardecki, C., 2018. RCSB Protein Data Bank: Sustaining a living digital data resource that enables breakthroughs in scientific research and biomedical education. *Protein Science,* 27, 316-330.

Buzon, M. J., Erkizia, I., Pou, C., Minuesa, G., Puertas, M. C., Esteve, A., Castello, A., Santos, J. R., Prado, J. G., Izquierdo-Useros, N., Pattery, T., Van Houtte, M., Carrasco, L., Clotet, B., Ruiz, L. & Martinez-Picado, J., 2011. A non-infectious cell-based phenotypic assay for the assessment of HIV-1 susceptibility to protease inhibitors. *Journal of Antimicrobial Chemotherapy,* 67, 32-38.

Cao, L., Zhou, Y., Huang, L., Dong, S. & Ma, Y., 2017. Development of a dual-expression vector facilitated with selection-free PCR recombination cloning strategy. *AMB Express,* 7, 98.

Carrier, M., Nugent, M., Tacon, W. & Primrose, S., 1983. High expression of cloned genes in E. coli and its consequences. *Trends in Biotechnology,* 1, 109-113.

Carter, M. & Shieh, J. 2015. Chapter 10 - Molecular Cloning And Recombinant Dna Technology. *In:* Carter, M. & Shieh, J. (eds.) *Guide to Research Techniques in Neuroscience (Second Edition).* San Diego: Academic Press.

Cha, S., 1975. Tight-binding inhibitors—I: Kinetic behavior. Biochemical pharmacology, 24 (23), 2177-2185.

Chang, M. W., Giffin, M. J., Muller, R., Savage, J., Lin, Y. C., Hong, S., Jin, W., Whitby, L. R., Elder, J. H. & Boger, D. L., 2010. Identification of broad-based HIV-1 protease inhibitors from combinatorial libraries. *Biochemical Journal*, 429 (3), 527-532.

Chang, M. W. & Torbett, B. E., 2011. Accessory mutations maintain stability in drug-resistant HIV-1 protease. *Journal of molecular biology,* 410(4), 756-760.

Chaudhary, A. K. & Lee, E. Y., 2015. Tightly regulated and high level expression vector construction for Escherichia coli BL21 (DE3). *Journal of Industrial and Engineering Chemistry,* 31, 367-373.

Chen, J., Liang, Z., Wang, W., Yi, C., Zhang, S. & Zhang, Q., 2014. Revealing origin of decrease in potency of darunavir and amprenavir against HIV-2 relative to HIV-1 protease by molecular dynamics simulations. *Scientific reports*, 4(1), 6872.

Chen, J., Peng, C., Wang, J., Zhu, W., 2018. Exploring molecular mechanism of allosteric inhibitor to relieve drug resistance of multiple mutations in HIV-1 protease by enhanced conformational sampling. *Proteins: Structure, Function, and Bioinformatics*, 86 (12), 1294-1305.

Chen, Z., Li, Y., Schock, H. B., Hall, D., Chen, E. & Kuo, L. C., 1995. Three-dimensional structure of a mutant HIV-1 protease displaying cross-resistance to all protease inhibitors in clinical trials. *Journal of Biological Chemistry*, 270(37), 21433-21436.

Cheng, Y.S.E., Mcgowan, M. H., Kettner, C. A., Schloss, J. V., Erickson-Viitanen, S. & Yin, F. H., 1990. High-level synthesis of recombinant HIV-1 protease and the recovery of active enzyme from inclusion bodies. *Gene,* 87, 243-248.

Chetty, S., Bhakat, S., Martin, A. J., Soliman, M., 2016. Multi-drug resistance profile of PR20 HIV-1 protease is attributed to distorted conformational and drug binding landscape: molecular dynamics insights. Journal of Biomolecular Structure and Dynamics, 34 (1), 135-151.

Cheung, H. & Lakowicz, J., 1991. Topics in Fluorescence Spectroscopy. Vol. 2: Principles. Plenum Press, New York.

Clark, D. P. & Pazdernik, N. J., 2015. Biotechnology: Applying the Genetic Revolution. Elsevier Science.

Clavel, F. & Hance, A. J., 2004. HIV drug resistance. *New England Journal of Medicine,* 350, 1023-1035.

Cleland, W.W., 1963. The kinetics of enzyme-catalyzed reactions with two or more substrates or products: II. Inhibition: Nomenclature and theory. *Biochimica Biophysica Acta*, 67, 173-187.

Clemente, J. C., Moose, R. E., Hemrajani, R., Whitford, L. R., Govindasamy, L., Reutzel, R., Mckenna, R., Agbandje-Mckenna, M., Goodenow, M. M. & Dunn, B. M., 2004. Comparing the accumulation of active-and nonactive-site mutations in the HIV-1 protease. *Biochemistry,* 43, 12141-12151.

Coman, R. M., Robbins, A., Goodenow, M. M., Mckenna, R. & Dunn, B. M., 2007. Expression, purification and preliminary X-ray crystallographic studies of the human immunodeficiency virus 1 subtype C protease. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications,* 63, 320-323.

Copeland, R. A., 2011. Conformational adaptation in drug-target interactions and residence time. *Future medicinal chemistry,* 3(12), 1491-501.

Copeland, R. A., Pompliano, D. L. & Meek, T. D., 2006. Drug–target residence time and its implications for lead optimization. *Nature reviews Drug discovery*, 5(9), 730-739.

Copeland, R. A., 2016. The drug–target residence time model: a 10-year retrospective. *Nature Reviews Drug Discovery*, 15(2), 87.

Corsini, L., Hothorn, M., Scheffzek, K., Sattler, M. & Stier, G., 2008. Thioredoxin as a fusion tag for carrier-driven crystallization. Protein Science, 17(2), 2070-2079.

Cosnefroy, O., Murray, P. J. & Bishop, K. N., 2016. HIV-1 capsid uncoating initiates after the first strand transfer of reverse transcription. *Retrovirology*, 13, 1-17.

Costa, S., Almeida, A., Castro, A. & Domingues, L., 2014. Fusion tags for protein solubility, purification and immunogenicity in Escherichia coli: the novel Fh8 system. *Frontiers in microbiology,* 5, 63.

Craigie, R. & Bushman, F. D., 2012. HIV DNA Integration. *Cold Spring Harbor perspectives in medicine,* 2, a006890.

Cura, V., Gangloff, M., Eiler, S., Moras, D., Ruff, M., 2008. Cleaved thioredoxin fusion protein enables the crystallization of poorly soluble ERα in complex with synthetic ligands. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 64(1), 54-57.

Dash, C., Phadtare, S., Deshpande, V. & Rao, M., 2001. Structural and mechanistic insight into the inhibition of aspartic proteases by a slow-tight binding inhibitor from an extremophilic Bacillus sp.: correlation of the kinetic parameters with the inhibitor induced conformational changes. *Biochemistry*, 40(38), 11525-11532.

Davis, D. A., Soule, E. E., Davidoff, K. S., Daniels, S. I., Naiman, N. E. & Yarchoan, R., 2012. Activity of human immunodeficiency virus type 1 protease inhibitors against the initial autocleavage in Gag-Pol polyprotein processing. *Antimicrobial agents and chemotherapy,* 56, 3620-3628.

De Oliveira, T., Engelbrecht, S., Van Rensburg, E. J., Gordon, M., Bishop, K., Zur Megede, J., Barnett, S. W. & Cassol, S. 2003. Variability at human immunodeficiency virus type 1 subtype C protease cleavage sites: an indication of viral fitness? *Journal of virology,* 77, 9422-9430.

De Vera, I. M. S., Smith, A. N., Dancel, M. C. A., Huang, X., Dunn, B. M. & Fanucci, G. E., 2013. Elucidating a relationship between conformational sampling and drug resistance in HIV-1 protease. *Biochemistry,* 52(19), 3278-3288.

De Vivo, M., Masetti, M., Bottegoni, G. & Cavalli, A., 2016. Role of molecular dynamics and related methods in drug discovery. *Journal of medicinal chemistry,* 59, 4035-4061.

De Witte, W. E. A., Danhof, M., Van Der Graaf, P. H. & De Lange, E. C. M., 2018. The implications of target saturation for the use of drug-target residence time. *Nature reviews Drug discovery,* 18, 82-84.

Deeks, S. G., Overbaugh, J., Phillips, A. & Buchbinder, S., 2015. HIV infection. *Nature reviews Disease primers,* 1, 1-22.

Dergousova, N. I., Amerik, A., Volynskaya, A. M. & Rumsh, L. D., 1996. HIV-I protease. Cloning, expression, and purification. *Appl Biochem Biotechnol,* 61, 97-107.

Descamps, D., Apetrei, C., Collin, G., Damond, F., Simon, F. & Brun-Vézinet, F., 1998. Naturally occurring decreased susceptibility of HIV-1 subtype G to protease inhibitors. *Aids,* 12, 1109-11.

Dierynck, I., De Wit, M., Gustin, E., Keuleers, I., Vandersmissen, J., Hallenberger, S. & Hertogs, K., 2007. Binding kinetics of darunavir to human immunodeficiency virus type 1 protease explain the potent antiviral activity and high genetic barrier. *Journal of virology*, 81(24), 13845-13851.

Dixon, M., 1953. The determination of enzyme inhibitor constants. *Biochemical journal*, 55(1), 170-171.

Do Hyung Kim, K. J. L., Sung, Y.C., & Choi, K.Y., 1994. Expression and Purification of HIV -1 Protease Utilizing a Maltose Binding Protein. *Molecules and Cells,* Vol. 4, 79-84.

Dubey, A. A., Singh, M. I. & Jain, V., 2016. Rapid and robust PCR-based all-recombinant cloning methodology. *PLoS One,* 11(3), p.e0152106.

Durrant, J. D. & Mccammon, J. A., 2011. Molecular dynamics simulations and drug discovery. *BMC biology,* 9, 1-9.

Dwyer-Lindgren, L., Cork, M. A., Sligar, A., Steuben, K. M., Wilson, K. F., Provost, N. R., Mayala, B. K., Vanderheide, J. D., Collison, M. L. & Hall, J. B., 2019. Mapping HIV prevalence in sub-Saharan Africa between 2000 and 2017. *Nature*, 570(7760), 189-193.

Dyson, M. R., Shadbolt, S. P., Vincent, K. J., Perera, R. L. & Mccafferty, J., 2004. Production of soluble mammalian proteins in Escherichia coli: identification of protein features that correlate with successful expression. *BMC biotechnology*, 4(1), 1-18.

Eagling, V. A., Back, D. J. & Barry, M. G., 1997. Differential inhibition of cytochrome P450 isoforms by the protease inhibitors, ritonavir, saquinavir and indinavir. *British Journal of Clinical Pharmacology,* 44, 190-4.

Eche, S. & Gordon, M. L., 2021. Recombinant expression of HIV-1 protease using soluble fusion tags in Escherichia coli: A vital tool for functional characterization of HIV-1 protease. *Virus Research,* 295, 198289.

Eche, S., Kumar, A., Sonela, N. & Gordon, M. L., 2021. Acquired HIV-1 Protease Conformational Flexibility Associated with Lopinavir Failure May Shape the Outcome of Darunavir Therapy after Antiretroviral Therapy Switch. *Biomolecules,* 11(4),489

Edessa, D., Sisay, M. & Asefa, F. 2019. Second-line HIV treatment failure in sub-Saharan Africa: A systematic review and meta-analysis. *PloS one,* 14, e0220159.

Eisenthal, R., & Danson, M. J. 2002. Enzyme assays: a practical approach, *Practical Approach* (Paperback).

Esposito, D. & Chatterjee, D. K., 2006. Enhancement of soluble protein expression through the use of fusion tags. *Current opinion in biotechnology,* 17, 353-358.

Fakruddin, M., Mohammad Mazumdar, R., Bin Mannan, K. S., Chowdhury, A. & Hossain, M. N., 2012. Critical factors affecting the success of cloning, expression, and mass production of enzymes by recombinant E. coli. *ISRN biotechnology,* 2013.

Fanny, G., Sonia, B. & Andrés, I., 2007. Peptide synthesis: chemical or enzymatic. *Electronic Journal of Biotechnology; Vol 10, No 2 (2007).*

Fernàndez, G., Clotet, B. & Martínez, M. A., 2007. Fitness landscape of human immunodeficiency virus type 1 protease quasispecies. *Journal of virology,* 81, 2485-2496.

Fidy, J., Laberge, M., Ullrich, B., Polgar, L., Szeltner, Z., Gallay, J., Vincent, M., 2001. Tryptophan rotamers that report the conformational dynamics of proteins. *Pure and Applied Chemistry,* 73, 415-419.

Freed, E. O., 1998. HIV-1 gag proteins: diverse functions in the virus life cycle. *Virology*, 251(1), 1-15.

Fun, A., Wensing, A. M., Verheyen, J. & Nijhuis, M., 2012. Human Immunodeficiency Virus Gag and protease: partners in resistance. *Retrovirology*, 9(1), 1-14.

Gajula, M., Kumar, A. & Ijaq, J. 2016. Protocol for molecular dynamics simulations of proteins. *Bio-protocol,* 6, e2051.

Ganser-Pornillos, B. K., Yeager, M. & Sundquist, W. I., 2008. The structural biology of hiv assembly. *Current opinion in structural biology*, 18, 203-217.

Gelderblom, H. R., 1991. Assembly and morphology of HIV: potential effect of structure on viral function. *Aids*, 5(6), 617-638.

Ghisaidoobe, A. B. & Chung, S. J., 2014. Intrinsic tryptophan fluorescence in the detection and analysis of proteins: a focus on Förster resonance energy transfer techniques. *International Journal of Molecular Science,* 15, 22518-38.

Ghosh, A. K., Anderson, D. D., Weber, I. T. & Mitsuya, H., 2012. Enhancing Protein Backbone Binding—A Fruitful Concept for Combating Drug-Resistant HIV. *Angewandte Chemie International Edition,* 51, 1778-1802.

Ghosh, A. K., Brindisi, M., Nyalapatla, P. R., Takayama, J., Ella-Menye, J.-R., Yashchuk, S., Agniswamy, J., Wang, Y.F., Aoki, M. & Amano, M., 2017a. Design of novel HIV-1 protease inhibitors incorporating isophthalamide-derived P2-P3 ligands: Synthesis, biological evaluation and X-ray structural studies of inhibitor-HIV-1 protease complex. *Bioorganic & medicinal chemistry,* 25, 5114-5127.

Ghosh, A. K. & Chapsal, B. D., 2013. Design of the anti-HIV protease inhibitor darunavir. *Introduction to Biological and Small Molecule Drug Research and Development.* Elsevier.

Ghosh, A. K., Chapsal, B. D. & Mitsuya, H., 2010. Darunavir, a New PI with Dual Mechanism: From a Novel Drug Design Concept to New Hope against Drug-Resistant HIV. *Aspartic acid proteases as therapeutic targets,* 45, 205-243.

Ghosh, A. K., Chapsal, B. D., Weber, I. T. & Mitsuya, H., 2007. Design of HIV protease inhibitors targeting protein backbone: an effective strategy for combating drug resistance. *Accounts of chemical research,* 41, 78-86.

Ghosh, A. K., Fyvie, W. S., Brindisi, M., Steffey, M., Agniswamy, J., Wang, Y.F., Aoki, M., Amano, M., Weber, I.T. & Mitsuya, H., 2017b. Design, synthesis, X-ray studies, and biological evaluation of novel macrocyclic HIV-1 protease inhibitors involving the P1′-P2′ ligands. *Bioorganic & medicinal chemistry letters,* 27, 4925-4931.

Ghosh, A. K., R. Nyalapatla, P., Kovela, S., Rao, K. V., Brindisi, M., Osswald, H. L., Amano, M., Aoki, M., Agniswamy, J. & Wang, Y.F., 2018. Design and Synthesis of Highly Potent HIV-1 Protease Inhibitors Containing Tricyclic Fused Ring Systems as Novel P2 Ligands: Structure–Activity Studies, Biological and X-ray Structural Analysis. *Journal of medicinal chemistry*.

Ghosh, a. k., Osswald, H. L. & Prato, G. 2016. Recent progress in the development of HIV-1 protease inhibitors for the treatment of HIV/AIDS. *Journal of medicinal chemistry,* 59, 5172-5208.

Godfrey, C., Thigpen, M. C., Crawford, K. W., Jean-Phillippe, P., Pillay, D., Persaud, D., Kuritzkes, D. R., Wainberg, M., Raizes, E. & Fitzgibbon, J., 2017. Global HIV antiretroviral drug resistance: a perspective and report of a national institute of allergy and infectious diseases consultation. *The Journal of infectious diseases,* 216, S798-S800.

Goldfarb, N. E., Ohanessian, M., Biswas, S., Mcgee, T. D., Mahon, B. P., Ostrov, D. A., Garcia, J., Tang, Y., Mckenna, R., Roitberg, A. & Dunn, B. M., 2015. Defective Hydrophobic Sliding Mechanism and Active Site Expansion in HIV-1 Protease Drug Resistant Variant Gly48Thr/Leu89Met: Mechanisms for the Loss of Saquinavir Binding Potency. *Biochemistry,* 54, 422-433.

Gomes, A. R., Byregowda, S. M., Veeregowda, B. M. & Balamurugan, V., 2016. An overview of heterologous expression host systems for the production of recombinant proteins. *Advances in Animal and Veterinary Sciences,* 4(7), 346-356.

Gräslund, S., Nordlund, P., Weigelt, J., Hallberg, B. M., Bray, J., Gileadi, O., Knapp, S., Oppermann, U., Arrowsmith, C. & Hui, R., 2008. Protein production and purification. *Nature methods,* 5, 135.

Grinsztejn, B., Hughes, M. D., Ritz, J., Salata, R., Mugyenyi, P., Hogg, E., Wieclaw, L., Gross, R., Godfrey, C. & Cardoso, S. W., 2019. Third-line antiretroviral therapy in low-income and middle-income countries (ACTG A5288): a prospective strategy study. *The Lancet HIV*, 6, e588-e600.

Grossman, Z., Vardinon, N., Chemtob, D., Alkan, M. L., Bentwich, Z., Burke, M., Gottesman, G., Istomin, V., Levi, I., Maayan, S., Shahar, E., & Schapiro, J. M., 2001. Genotypic variation of HIV-1 reverse transcriptase and protease: comparative analysis of clade C and clade B. *AIDS,* 15.

Guarna, A. & Trabocchi, A., 2014. *Peptidomimetics in Organic and Medicinal Chemistry: the art of transforming peptides in drugs*, John Wiley & Sons.

Gulnik, S. V., Suvorov, L. I., Liu, B., Yu, B., Anderson, B., Mitsuya, H. & Erickson, J. W., 1995. Kinetic characterization and cross-resistance patterns of HIV-1 protease mutants selected under drug pressure. *Biochemistry,* 34, 9282-9287.

Guo, B. & Bi, Y., 2002. Cloning PCR products: An overview. *PCR Cloning Protocols*, 111-119.

Guo, D., Mulder-Krieger, T., Ijzerman, A. P. & Heitman, L. H., 2012. Functional efficacy of adenosine A2A receptor agonists is positively correlated to their receptor residence time. *British journal of pharmacology*, 166, 1846-1859.

Gupta, S. K. & Shukla, P., 2016. Advanced technologies for improved expression of recombinant proteins in bacteria: perspectives and applications. *Critical reviews in biotechnology*, 36, 1089-1098.

Häggblom, A., Svedhem, V., Singh, K., Sönnerborg, A. & Neogi, U., 2016. Virological failure in patients with HIV-1 subtype C receiving antiretroviral therapy: an analysis of a prospective national cohort in Sweden. *Lancet HIV,* 3, e166-74.

Halder, A. K. & Honarparvar, B., 2019. Molecular alteration in drug susceptibility against subtype B and C-SA HIV-1 proteases: MD study. *Structural Chemistry*, 30, 1715-1727.

Hang, J. Q., Yang, Y., Harris, S. F., Leveque, V., Whittington, H. J., Rajyaguru, S., Ao-Ieong, G., Mccown, M. F., Wong, A. & Giannetti, A. M., 2009. Slow binding inhibition and mechanism of resistance of non-nucleoside polymerase inhibitors of hepatitis C virus. *Journal of Biological Chemistry*, 284, 15517-15529.

Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E. & Hutchison, G. R., 2012. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of cheminformatics,* 4, 17.

Harte, W., Swaminathan, S., Mansuri, M., Martin, J., Rosenberg, I. & Beveridge, D., 1990. Domain communication in the dynamical structure of human immunodeficiency virus 1 protease. *Proceedings of the National Academy of Sciences,* 87, 8864-8868.

Hayes, J. M., Archontis, G., 2011. Molecular Dynamic-Studies of Synthetic and Biological Molecules. (Ed.), ISBN: 978-953-51-0444-5, InTech.

Hemelaar, J., 2013. Implications of HIV diversity for the HIV-1 pandemic. *Journal of Infection,* 66, 391-400.

Hemelaar, J., Elangovan, R., Yun, J., Dickson-Tetteh, L., Fleminger, I., Kirtley, S., Williams, B., Gouws-Williams, E., Ghys, P. D. & Alash'le G, A., 2019. Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. *The Lancet infectious diseases,* 19, 143-155.

Henderson, G. J., Lee, S.-K., Irlbeck, D. M., Harris, J., Kline, M., Pollom, E., Parkin, N. & Swanstrom, R., 2012. Interplay between single resistance-associated mutations in the HIV-1 protease and viral infectivity, protease activity, and inhibitor sensitivity. *Antimicrobial agents and chemotherapy,* 56, 623-633.

Henes, M., Kosovrasti, K., Lockbaum, G. J., Leidner, F., Nachum, G. S., Nalivaika, E. A., Bolon, D. N., Kurt Yilmaz, N., Schiffer, C. A. & Whitfield, T. W., 2019a. Molecular determinants of epistasis in HIV-1 protease: Elucidating the interdependence of L89V and L90M mutations in resistance. *Biochemistry*, 58, 3711-3726.

Henes, M., Lockbaum, G. J., Kosovrasti, K., Leidner, F., Nachum, G. S., Nalivaika, E. A., Lee, S.K., Spielvogel, E., Zhou, S. & Swanstrom, R., 2019b. Picomolar to Micromolar: Elucidating the Role of Distal Mutations in HIV-1 Protease in Conferring Drug Resistance. *ACS chemical biology*, 14, 2441-2452.

Hewitt, S. N., Choi, R., Kelley, A., Crowther, G. J., Napuli, A. J. & Van Voorhis, W. C., 2011. Expression of proteins in Escherichia coli as fusions with maltose-binding protein to rescue non-expressed targets in a high-throughput protein-expression and purification pipeline. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications,* 67, 1006-1009.

Hilton, B. J. & Wolkowicz, R., 2010. An assay to monitor HIV-1 protease activity for the identification of novel inhibitors in T-cells. *PloS one,* 5, e10940-e10940.

116

Hodes, R. & Morrell, R., 2018. Incursions from the epicentre: Southern theory, social science, and the global HIV research domain. *African Journal of AIDS Research*, 17, 22-31.

Holdgate, G. A., Meek, T. D. & Grimley, R. L., 2018. Mechanistic enzymology in drug discovery: a fresh perspective. *Nature reviews Drug discovery,* 17, 115.

Holec, A. D., Mandal, S., Prathipati, P. K. & Destache, C. J., 2017. Nucleotide reverse transcriptase inhibitors: a thorough review, present status and future perspective as HIV therapeutics. *Current HIV research,* 15, 411-421.

Hollingsworth, S. A. & Dror, R. O., 2018. Molecular dynamics simulation for all. *Neuron,* 99, 1129-1143.

Hou, W., Zhang, X. & Liu, C.F., 2017. Progress in chemical synthesis of peptides and proteins. *Transactions of Tianjin University,* 23, 401-419.

Houtzager, V., Ouellet, M., Falgueyret, J. P., Passmore, L. A., Bayly, C. & Percival, M. D., 1996. Inhibitor-induced changes in the intrinsic fluorescence of human cyclooxygenase-2. *Biochemistry,* 35, 10974-84.

Huang, L. & Chen, C., 2013. Understanding HIV-1 protease autoprocessing for novel therapeutic development. *Future medicinal chemistry*, 5, 1215-1229.

Huang, L., Li, Y. & Chen, C., 2011. Flexible catalytic site conformations implicated in modulation of HIV-1 protease autoprocessing reactions. *Retrovirology*, 8, 79.

Huang, L., Sayer, J. M., Swinford, M., Louis, J. M. & Chen, C., 2009. Modulation of human immunodeficiency virus type 1 protease autoprocessing by charge properties of surface residue 69. *Journal of virology*, 83, 7789-7793.

Huang, M., Bao, J. & Nielsen, J. 2014a. Biopharmaceutical protein production by Saccharomyces cerevisiae: current state and future prospects. *J Pharmaceutical Bioprocessing,* 2, 167-182.

Huang, X., Britto, M. D., Kear-Scott, J. L., Boone, C. D., Rocca, J. R., Simmerling, C., Mckenna, R., Bieri, M., Gooley, P. R. & Dunn, B. M., 2014b. The role of select subtype polymorphisms on HIV-1 protease conformational sampling and dynamics. *Journal of Biological Chemistry*, 289, 17203-17214.

Huang, Y.M., Raymundo, M.A., Chen, W. & Chang, C.E., 2017. Mechanism of the association pathways for a pair of fast and slow binding ligands of HIV-1 protease. *Biochemistry*, 56(9), 1311-1323.

Hui, J. O., Tomasselli, A. G., Reardon, I. M., Lull, J. M., Brunner, D. P., Tomich, C.S., & Heinrikson, R. L., 1993. Large scale purification and refolding of HIV-1 protease fromEscherichia coli inclusion bodies. *Journal of protein chemistry,* 12, 323-327.

Hunter, M., Yuan, P., Vavilala, D. & Fox, M., 2019. Optimization of protein expression in mammalian cells. *J Current Protocols in Protein Science,* 95, e77.

Hwang, P. M., Pan, J. S. & Sykes, B. D., 2014. Targeted expression, purification, and cleavage of fusion proteins from inclusion bodies in Escherichia coli. *FEBS letters,* 588, 247-252.

Idicula-Thomas, S. & Balaji, P. V., 2007. Protein aggregation: A perspective from amyloid and inclusion-body formation. *Current Science*, 758-767.

Ikuta, K., Suzuki, S., Horikoshi, H., Mukai, T., Luftig, R. B., 2000. Positive and negative aspects of the human immunodeficiency virus protease: development of inhibitors versus its role in AIDS pathogenesis. *Microbiology and molecular biology reviews*, 64, 725-745.

Jakubowski, H. 2015. Biochemistry Online: An Approach Based on Chemical Logic. College of St. Benedict, St. John's University.

Jaradat, D. M. M., 2018. Thirteen decades of peptide synthesis: key developments in solid phase peptide synthesis and amide bond formation utilized in peptide ligation. *Amino Acids,* 50, 39-68.

Jia, B. & Jeon, C. O., 2016. High-throughput recombinant protein expression in Escherichia coli: current status and future perspectives. *Open biology,* 6, 160196.

Jin, T., Chuenchor, W., Jiang, J., Cheng, J., Li, Y., Fang, K., Huang, M., Smith, P. & Xiao, T. S., 2017. Design of an expression system to enhance MBP-mediated crystallization. *Scientific reports*, 7, 1-11.

Johnson, V. A., Brun-Vézinet, F., Clotet, B., Günthard, H. F., Kuritzkes, D. R., Pillay, D., Schapiro, J. M. & Richman, D. D., 2010. Update of the drug resistance mutations in HIV-1. *Top HIV Med,* 18, 156-63.

Johnson, V. A., Calvez, V., Günthard, H. F., Paredes, R., Pillay, D., Shafer, R., Wensing, A. M. & Richman, D. D., 2011. 2011 update of the drug resistance mutations in HIV-1. *Topics in antiviral medicine*, 19(4), 156.

Joseph, B. C., Pichaimuthu, S., Srimeenakshi, S., Murthy, M., Selvakumar, K., Ganesan, M. & Manjunath, S. R., 2015. An overview of the parameters for recombinant protein expression in Escherichia coli. *Journal of Cell Science & Therapy,* 6, 1.

Kantor, R. & Katzenstein, D., 2003. Polymorphism in HIV-1 non-subtype B protease and reverse transcriptase and its potential impact on drug susceptibility and drug resistance evolution. *AIDS rev*, 5, 25-35.

Kaplan, W., Erhardt, J., Sluis-Cremer, N., Dirr, H., Hüsler, P. & Klump, H., 1997. Conformational stability of pGEX-expressed Schistosoma japonicum glutathione S-transferase: A detoxification enzyme and fusion-protein affinity tag. *Protein Science*, 6(2), 399-406.

Karacostas, V., Wolffe, E. J., Nagashima, K., Gonda, M. A. & Moss, B., 1993. Overexpression of the HIV-1 gag-pol polyprotein results in intracellular activation of HIV-1 protease and inhibition of assembly and budding of virus-like particles. *Virology,* 193, 661-671.

Karnati, K. R., Wang, Y., 2019. Structural and binding insights into HIV-1 protease and P2-ligand interactions through molecular dynamics simulations, binding free energy and principal component analysis. *Journal of Molecular Graphics and Modelling*, 92, 112-122.

Karplus, M. & Mccammon, J. A., 2002. Molecular dynamics simulations of biomolecules. *Nature structural biology,* 9, 646-652.

Katti, S. K., Lemaster, D. M. & Eklund, H., 1990. Crystal structure of thioredoxin from Escherichia coli at 1.68 Å resolution. *Journal of molecular biology*, 212, 167-184.

Kehinde, I., Ramharack, P., Nlooto, M. & Gordon, M., 2019. The pharmacokinetic properties of HIV-1 protease inhibitors: A computational perspective on herbal phytochemicals. *Heliyon*, 5, e02565.

Kent, S. B., 2009. Total chemical synthesis of proteins. *Chemical Society reviews,* 38, 338-351.

Khan, K. H., 2013. Gene expression in mammalian cells and its applications. *J Advanced pharmaceutical bulletin,* 3, 257.

Kim, H., Yoo, S. J. & Kang, H. A., 2015. Yeast synthetic biology for the production of recombinant therapeutic proteins. *FEMS yeast research*, 15(1), 1-16.

Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S. & Shoemaker, B. A., 2016. PubChem substance and compound databases. *Nucleic acids research,* 44, D1202-D1213.

Kimata, J. T., Rice, A. P. & Wang, J., 2016. Challenges and strategies for the eradication of the HIV reservoir. *Current opinion in immunology,* 42, 65-70.

Kimple, M. E., Brill, A. L. & Pasker, R. L., 2013. Overview of affinity tags for protein purification. *Current protocols in protein science*, 73(1), 1-9.

Kirchhoff, F., 2013. HIV life cycle: overview. *Encyclopedia of AIDS*, 1-9.

Kirchner, J. T., 2019. The Origin, Evolution, and Epidemiology of HIV-1 and HIV-2. *Fundamentals of HIV Medicine*, 14.

Kneller, D. W., Agniswamy, J., Harrison, R. W. & Weber, I. T., 2020. Highly drug-resistant HIV-1 protease reveals decreased intra-subunit interactions due to clusters of mutations. *The FEBS journal*, 287(15), 3235-3254.

Kosobokova, E., Skrypnik, K. & Kosorukov, V., 2016. Overview of fusion tags for recombinant proteins. *Biochemistry (Moscow)*, 81, 187-200.

Kovalevsky, A. Y., Ghosh, A. K. & Weber, I. T., 2008. Solution kinetics measurements suggest HIV-1 protease has two binding sites for darunavir and amprenavir. *Journal of medicinal chemistry*, 51, 6599-6603.

Kovalevsky, A. Y., Liu, F., Leshchenko, S., Ghosh, A. K., Louis, J. M., Harrison, R. W. & Weber, I. T., 2006. Ultra-high Resolution Crystal Structure of HIV-1 Protease Mutant Reveals Two Binding Sites for Clinical Inhibitor TMC114. *Journal of Molecular Biology,* 363, 161-173.

Kožíšek, M., Lepšík, M., Grantz Šašková, K., Brynda, J., Konvalinka, J. & Řezáčová, P., 2014. Thermodynamic and structural analysis of HIV protease resistance to darunavir–analysis of heavily mutated patient-derived HIV-1 proteases. *The FEBS journal,* 281, 1834-1847.

Kumar, A. & Rao, M., 2010. A novel bifunctional peptidic aspartic protease inhibitor inhibits chitinase A from Serratia marcescens: Kinetic analysis of inhibition and binding affinity. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1800, 526-536.

Kumar, M. & Hosur, M. V., 2003. Adaptability and flexibility of HIV-1 protease. *European journal of biochemistry,* 270, 1231-1239.

Labhsetwar, P., Cole, J. A., Roberts, E., Price, N. D. & Luthey-Schulten, Z. A., 2013. Heterogeneity in protein expression induces metabolic variability in a modeled Escherichia coli population. *Proceedings of the National Academy of Sciences*,110, 14006-14011.

Laemmli, U. K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227, 680-685.

Lakowicz, J. R., 2013. *Principles of fluorescence spectroscopy*, Springer science & business media.

Lavallie, E. R., Diblasio-Smith, E. A., Collins-Racie, L. A., Lu, Z. & Mccoy, J. M., 2003. Thioredoxin and related proteins as multifunctional fusion tags for soluble expression in E. coli. *E. coliGene Expression Protocols.* Springer.

Lavallie, E. R., Lu, Z., Diblasio-Smith, E. A., Collins-Racie, L. A. & Mccoy, J. M., 2000. [21] Thioredoxin as a fusion partner for production of soluble recombinant proteins in Escherichia coli. *Methods in enzymology.* Elsevier.

Lebendiker, M. & Danieli, T., 2011. Purification of proteins fused to maltose-binding protein. *Protein Chromatography: Methods and Protocols*, 281-293.

Lebendiker, M. & Danieli, T., 2014. Production of prone-to-aggregate proteins. *FEBS letters,* 588, 236-246.

Lee, K. S. S., Yang, J., Niu, J., Ng, C. J., Wagner, K. M., Dong, H., Kodani, S. D., Wan, D., Morisseau, C. & Hammock, B. D., 2019. Drug-Target Residence Time Affects in Vivo Target Occupancy through Multiple Pathways. *ACS central science,* 5, 1614-1624.

Lefebvre, E. & Schiffer, C. A., 2008. Resilience to resistance of HIV-1 protease inhibitors: profile of darunavir. *AIDS reviews*, 10, 131.

Lemey, P., Pybus, O. G., Wang, B., Saksena, N. K., Salemi, M. & Vandamme, A.M., 2003. Tracing the origin and history of the HIV-2 epidemic. *Proceedings of the National Academy of Sciences*, 100, 6588-6592.

Levy, J. A. 2007. *HIV and the pathogenesis of AIDS*, Washington.

Liang, J., Woodward, C. & Edelsbrunner, H., 1998. Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Protein science,* 7, 1884-1897.

Lin, Y., Lin, X., Hong, L., Foundling, S., Heinrikson, R. L., Thaisrivongs, S., Leelamanit, W., Raterman, D. & Shah, M., 1995. Effect of point mutations on the kinetics and the inhibition of human immunodeficiency virus type 1 protease: relationship to drug resistance. *Biochemistry*, 34, 1143-1152.

Lineweaver, H. & Burk, D., 1934. The determination of enzyme dissociation constants. *Journal of the American chemical society*, 56, 658-666.

Liu, F., Kovalevsky, A. Y., Louis, J. M., Boross, P. I., Wang, Y.-F., Harrison, R. W. & Weber, I. T., 2006. Mechanism of drug resistance revealed by the crystal structure of the unliganded HIV-1 protease with F53L mutation. *Journal of molecular biology*, 358, 1191-1199.

Liu, F., Kovalevsky, A. Y., Tie, Y., Ghosh, A. K., Harrison, R. W. & Weber, I. T., 2008. Effect of flap mutations on structure of HIV-1 protease and inhibition by saquinavir and darunavir. *Journal of molecular biology*, 381(1), 102-115.

Liu, Z., Yedidi, R. S., Wang, Y., Dewdney, T. G., Reiter, S. J., Brunzelle, J. S., Kovari, I. A. & 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.

Locatelli, S. & Peeters, M., 2012. Cross-species transmission of simian retroviruses: how and why they could lead to the emergence of new diseases in the human population. *Aids*, 26(6), 659-673.

Lockbaum, G. J., Leidner, F., Rusere, L. N., Henes, M., Kosovrasti, K., Nachum, G. S., Nalivaika, E. A., Ali, A., Yilmaz, N. K. & Schiffer, C. A., 2019. Structural Adaptation of Darunavir Analogues against Primary Mutations in HIV-1 Protease. *ACS Infect Dis,* 5, 316-325.

Lopes, C. A., Soares, M. A., Falci, D. R. & Sprinz, E. 2015., The Evolving Genotypic Profile of HIV-1 Mutations Related to Antiretroviral Treatment in the North Region of Brazil. *Biomed Res Int,* 2015, 738528.

Lorey, S., Stöckel-Maschek, A., Faust, J., Brandt, W., Stiebitz, B., Gorrell, M. D., Kähne, T., Mrestani-Klaus, C., Wrenger, S. & Reinhold, D., 2003. Different modes of dipeptidyl peptidase IV (CD26) inhibition by oligopeptides derived from the N-terminus of HIV-1 Tat indicate at least two inhibitor binding sites. *European journal of biochemistry*, 270(10), 2147-2156.

Los Alamos HIV database., HIV-1 Gene map: http://www.hiv.lanl.gov. [Accessed 15th May, 2020].

Louis, J. M., Mcdonald, R. A., Nashed, N. T., Wondrak, E. M., Jerina, D. M., Oroszlan, S. & Mora, P. T., 1991. Autoprocessing of the HIV-1 protease using purified wild-type and mutated fusion proteins expressed at high levels in Escherichia coli. *European journal of biochemistry*, 199, 361-369.

Louis, J. M., Nashed, N. T., Parris, K. D., Kimmel, A. R. & Jerina, D. M., 1994. Kinetics and mechanism of autoprocessing of human immunodeficiency virus type 1 protease from an analog of the Gag-Pol polyprotein. *Proceedings of the National Academy of Sciences*, 91(17), 7970-7974.

Lu, D. Y., Wu, H. Y., Yarla, N. S., Xu, B., Ding, J. & Lu, T. R., 2018. HAART in HIV/AIDS treatments: future trends. *Infectious Disorders-Drug Targets (Formerly Current Drug Targets-Infectious Disorders),* 18, 15-22.

Lv, Z., Chu, Y. & Wang, Y. 2015. HIV protease inhibitors: a review of molecular selectivity and toxicity. *Hiv/aids (Auckland, NZ),* 7, 95-104.

Maartens, G., Celum, C. & Lewin, S. R., 2014. HIV infection: epidemiology, pathogenesis, treatment, and prevention. *The Lancet,* 384, 258-271.

Mahalingam, B., Wang, Y. F., Boross, P. I., Tozser, J., Louis, J. M., Harrison, R. W. & Weber, I. T., 2004. Crystal structures of HIV protease V82A and L90M mutants reveal changes in the indinavir-binding site. *European journal of biochemistry*, 271, 1516-1524.

Mahanti, M., Bhakat, S., Nilsson, U. J., Söderhjelm, P., 2016. Flap dynamics in aspartic proteases: a computational perspective. *Chemical Biology & Drug Design*, 88, 159-177.

Mahmoodi, S., Pourhassan-Moghaddam, M., Wood, D. W., Majdi, H. & Zarghami, N., 2019. Current affinity approaches for purification of recombinant proteins. *Cogent Biology*, 5, 1665406.

Malik, A., 2016. Protein fusion tags for efficient expression and purification of recombinant proteins in the periplasmic space of E. coli. *3 Biotech,* 6, 44.

Maphumulo, S. I., Halder, A. K., Govender, T., Maseko, S., Maguire, G. E., Honarparvar, B., Kruger, H. G., 2018. Exploring the flap dynamics of the South African HIV subtype C protease in presence of FDA-approved inhibitors: MD study. *Chemical biology & drug design,* 92, 1899-1913.

Martinez-Cajas, J. L. & Wainberg, M. A., 2007. Protease inhibitor resistance in HIV-infected patients: molecular and clinical perspectives. *Antiviral research*,76, 203-221.

Maseko, S. B., Govender, D., Govender, T., Naicker, T., Lin, J., Maguire, G. E. & Kruger, H. G., 2019. Optimized Procedure for Recovering HIV-1 Protease (C-SA) from Inclusion Bodies. *The protein journal*, 38, 30-36.

Maseko, S. B., Natarajan, S., Sharma, V., Bhattacharyya, N., Govender, T., Sayed, Y., Maguire, G. E., Lin, J. & 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.

Maseko, S. B., Padayachee, E., Govender, T., Sayed, Y., Kruger, G., Maguire, G. E. & Lin, J., 2017. I36T↑ T mutation in South African subtype C (C-SA) HIV-1 protease significantly alters protease-drug interactions. *Biological chemistry,* 398, 1109-1117.

Matange, N., Bodkhe, S., Patel, M. & Shah, P., 2018. Trade-offs with stability modulate innate and mutationally acquired drug resistance in bacterial dihydrofolate reductase enzymes. *Biochemical Journal*, 475, 2107-2125.

Mckinstry, W. J., Hijnen, M., Tanwar, H. S., Sparrow, L. G., Nagarajan, S., Pham, S. T. & Mak, J., 2014. Expression and purification of soluble recombinant full length HIV-1 Pr55 Gag protein in Escherichia coli. *Protein expression and purification,* 100, 10-18.

Menéndez-Arias, L., 2013. Molecular basis of human immunodeficiency virus type 1 drug resistance: overview and recent developments. *Antiviral research*, 98, 93-120.

Miao, Y., Huang, Y. M. M., Walker, R. C., Mccammon, J. A. & Chang, C. E. A., 2018. Ligand binding pathways and conformational transitions of the HIV protease. *Biochemistry*, 57, 1533-1541.

Miller, D. C., Lunn, G., Jones, P., Sabnis, Y., Davies, N. L. & Driscoll, P., 2012. Investigation of the effect of molecular properties on the binding kinetics of a ligand to its biological target. *MedChemComm*, 3, 449-452.

Mittal, S., Cai, Y., Nalam, M. N., Bolon, D. N. & Schiffer, C. A., 2012. Hydrophobic core flexibility modulates enzyme activity in HIV-1 protease. *Journal of the American Chemical Society*, 134, 4163-4168.

Morris, L., Williamson, C., Gray, C. & Tiemessen, C., 2000. HIV-1 subtype C as a major determinant of the global AIDS epidemic. *South African Journal of Science,* 96.

Morrison, J. F., Stone, S. R., 1985. Approaches to the study and analysis of the inhibition of enzymes by slow-and tight-binding inhibitors. *Comments on molecular and cellular Biophysics,* 2(6), 347-368.

Morrison, J. F. & Walsh, C. T., 1988. The behavior and significance of slow-binding enzyme inhibitors. *Adv Enzymol Relat Areas Mol Biol,* 61, 201-301.

Mosebi, S., Morris, L., Dirr, H. W. & 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.

Mourez, T., Simon, F. & Plantier, J., 2013. Non-M variants of human immunodeficiency virus type 1. *Clinical Microbiology Reviews*, 26: 448–461.

Müller, T. G., Zila, V., Peters, K., Schifferdecker, S., Stanic, M., Lucic, B., Laketa, V., Lusic, M., Müller, B. & Kräusslich, H.-G. 2021. HIV-1 uncoating by release of viral cDNA from capsid-like structures in the nucleus of infected cells. *Elife*, 10, e64776.

Murphy, D. J., 2004. Determination of accurate KI values for tight-binding enzyme inhibitors: an *in silico* study of experimental error and assay design. *Analytical biochemistry*, 327, 61-67.

Naicker, P., Stoychev, S., Dirr, H. W. & Sayed, Y., 2014. Amide hydrogen exchange in HIV–1 subtype B and C proteases–insights into reduced drug susceptibility and dimer stability. *The Febs Journal*, 281, 5395-5410.

Nair, P. C. & Miners, J. O., 2014. Molecular dynamics simulations: from structure function relationships to drug discovery. *In silico pharmacology*, 2(1), 1-4.

Nguyen, H. L. T., Nguyen, T. T., Vu, Q. T., Le, H. T., Pham, Y., Le Trinh, P., Bui, T. P. & Phan, T. N., 2015. An efficient procedure for the expression and purification of HIV-1 protease from inclusion bodies. *Protein expression and purification,* 116, 59-65.

Niedrig, M., Gelderblom, H. R., Pauli, G., März, J., Bickhard, H., Wolf, H. & Modrow, S., 1994. Inhibition of infectious human immunodeficiency virus type 1 particle formation by Gag protein-derived peptides. *Journal of general virology*, 75, 1469-1474.

Nutt, R. F., Brady, S. F., Darke, P. L., Ciccarone, T. M., Colton, C. D., Nutt, E. M., Rodkey, J. A., Bennett, C. D., Waxman, L. H. & Sigal, I. S., 1988. Chemical synthesis and enzymatic activity of a 99-residue peptide with a sequence proposed for the human immunodeficiency virus protease. *Proceedings of the National Academy of Sciences,* 85, 7129-7133.

Obasa, A. E., Mikasi, S. G., Brado, D., Cloete, R., Singh, K., Neogi, U. & Jacobs, G. B., 2020. Drug Resistance Mutations Against Protease, Reverse Transcriptase and Integrase Inhibitors in People Living With HIV-1 Receiving Boosted Protease Inhibitors in South Africa. *Frontiers in Microbiology,* 11, 438.

Oehme, D. P., Wilson, D. J., Brownlee, R. T., 2011. Effect of Structural Stress on the Flexibility and Adaptability of HIV-1 Protease. *Journal of chemical information and modeling*, 51, 1064-1073.

Ohtaka, H., Schön, A. & Freire, E. 2003. Multidrug resistance to HIV-1 protease inhibition requires cooperative coupling between distal mutations. *Biochemistry,* 42, 13659-13666.

Pabis, A., Risso, V. A., Sanchez-Ruiz, J. M. & Kamerlin, S. C. L. 2018. Cooperativity and flexibility in enzyme evolution. *Current opinion in structural biology*, 48, 83-92.

Pan, A. C., Borhani, D. W., Dror, R. O. & Shaw, D. E. 2013. Molecular determinants of drug-receptor binding kinetics. *Drug Discov Today,* 18, 667-73.

Parera, M., Fernandez, G., Clotet, B., Martínez, M. A., 2007. HIV-1 protease catalytic efficiency effects caused by random single amino acid substitutions. *Molecular biology and evolution,* 24, 382-387.

Pargellis, C. A., Morelock, M. M., Graham, E. T., Kinkade, P., Pav, S., Lubbe, K., Lamarre, D. & Anderson, P. C., 1994. Determination of kinetic rate constants for the binding of inhibitors to HIV-1 protease and for the association and dissociation of active homodimer. *Biochemistry*, 33, 12527-12534.

Park, J. H., Sayer, J. M., Aniana, A., Yu, X., Weber, I. T., Harrison, R. W. & Louis, J. M., 2016. Binding of clinical inhibitors to a model precursor of a rationally selected multidrug resistant HIV-1 protease is significantly weaker than that to the released mature enzyme. *Biochemistry*, 55, 2390-2400.

Park, W.J., You, S.H., Choi, H.-A., Chu, Y.J. & Kim, G.J., 2015. Over-expression of recombinant proteins with N-terminal His-tag via subcellular uneven distribution in Escherichia coli. *Acta Biochimica Biophysica Sinica*, 47, 488-495.

Pawagi, A. B. & Deber, C. M., 1990. Ligand-dependent quenching of tryptophan fluorescence in human erythrocyte hexose transport protein. *Biochemistry*, 29, 950-955.

Peeters, M., Jung, M. & Ayouba, A. 2013. The origin and molecular epidemiology of HIV. *Expert review of anti-infective therapy,* 11, 885-896.

Perez, M., Fernandes, P. & Ramos, M. 2010. Substrate recognition in HIV-1 protease: a computational study. *The Journal of Physical Chemistry B,* 114, 2525-2532.

Perryman, A. L., Lin, J. H. & 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.

Piana, S., Carloni, P. & Rothlisberger, U., 2002. Drug resistance in HIV-1 protease: flexibility-assisted mechanism of compensatory mutations. *Protein Science*, 11, 2393-2402.

Pina, A. S., Lowe, C. R. & Roque, A. C. A., 2014. Challenges and opportunities in the purification of recombinant tagged proteins. *Biotechnology advances*, 32, 366-381.

Pope, B. & Kent, H. M. 1996. High efficiency 5 min transformation of Escherichia coli. *Nucleic acids research,* 24, 536-537.

Prabu-Jeyabalan, M., Nalivaika, E. A., King, N. M. & Schiffer, C. A., 2003. Viability of a drug-resistant human immunodeficiency virus type 1 protease variant: structural insights for better antiviral therapy. *Journal of virology,* 77, 1306-1315.

Prabu-Jeyabalan, M., Nalivaika, E. A., Romano, K. & Schiffer, C. A., 2006. Mechanism of substrate recognition by drug-resistant human immunodeficiency virus type 1 protease variants revealed by a novel structural intermediate. *Journal of virology*, 80, 3607-3616.

Ragland, D. A., Nalivaika, E. A., Nalam, M. N., Prachanronarong, K. L., Cao, H., Bandaranayake, R. M., Cai, Y., Kurt-Yilmaz, N. & Schiffer, C. A., 2014. Drug resistance conferred by mutations outside the active site through alterations in the dynamic and structural ensemble of HIV-1 protease. *Journal of the American Chemical Society,* 136, 11956-11963.

Ragland, D. A., Whitfield, T. W., Lee, S.K., Swanstrom, R., Zeldovich, K. B., Kurt-Yilmaz, N., Schiffer, C. A., 2017. Elucidating the interdependence of drug resistance from combinations of mutations. *Journal of chemical theory and computation*, 13, 5671-5682.

Rais-Beghdadi, C., Roggero, M. A., Fasel, N. & Reymond, C. D., 1998. Purification of recombinant proteins by chemical removal of the affinity tag. *Applied biochemistry and biotechnology,* 74, 95-103.

Rajakuberan, C., Hilton, B. J. & Wolkowicz, R., 2012. Protocol for a mammalian cell-based assay for monitoring the HIV-1 protease activity. *Methods Mol Biol,* 903, 393-405.

Ramirez, B. C., Simon-Loriere, E., Galetto, R. & Negroni, M., 2008. Implications of recombination for HIV diversity. *Virus research*, 134, 64-73.

Ramsay, R. R. & Tipton, K. F., 2017. Assessment of enzyme inhibition: a review with examples from the development of monoamine oxidase and cholinesterase inhibitory drugs. *Molecules*, 22, 1192.

Raran-Kurussi, S., Keefe, K. & Waugh, D. S., 2015. Positional effects of fusion partners on the yield and solubility of MBP fusion proteins. *Protein Expr Purif,* 110, 159-64.

Raran-Kurussi, S. & Waugh, D. S., 2012. The ability to enhance the solubility of its fusion partners is an intrinsic property of maltose-binding protein but their folding is either spontaneous or chaperone-mediated. *PloS one*, 7, e49589.

Ravaux, I., Perrin-East, C., Attias, C., Cottalorda, J., Durant, J., Dellamonica, P., Gluschankof, P., Stein, A. & Tamalet, C., 2014. Yeast cells as a tool for analysis of HIV-1 protease susceptibility to protease inhibitors, a comparative study. *J Virol Methods,* 195, 180-4.

Rhee, S.-Y., Taylor, J., Fessel, W. J., Kaufman, D., Towner, W., Troia, P., Ruane, P., Hellinger, J., Shirvani, V., Zolopa, A. & Shafer, R. W., 2010. HIV-1 Protease Mutations and Protease Inhibitor Cross-Resistance. *Antimicrobial Agents and Chemotherapy,* 54, 4253.

Roberts, J. D., Bebenek, K. & Kunkel, T. A., 1988. The accuracy of reverse transcriptase from HIV-1. *Science,* 242, 1171-3.

Robertson, D. L., Anderson, J., Bradac, J., Carr, J., Foley, B., Funkhouser, R., Gao, F., Hahn, B., Kalish, M. & Kuiken, C., 2000. HIV-1 nomenclature proposal. *Science*, 288, 55-55.

Rogers, A. & Gibon, Y. 2009. Enzyme kinetics: theory and practice. *Plant metabolic networks.* Springer.

Ronda, L., Pioselli, B., Catinella, S., Salomone, F., Marchetti, M. & Bettati, S., 2018. Quenching of tryptophan fluorescence in a highly scattering solution: Insights on protein localization in a lung surfactant formulation. *Plos one*, 13, e0201926.

Rosano, G. L. & Ceccarelli, E. A., 2014. Recombinant protein expression in Escherichia coli: advances and challenges. *Frontiers in microbiology,* 5.

Roskoski, R. 2007. Michaelis-Menten Kinetics. *In:* Enna, S. J. & Bylund, D. B. (eds.) *xPharm: The Comprehensive Pharmacology Reference.* New York: Elsevier.

Roskoski, R. 2015. Michaelis-Menten Kinetics☆. *Reference Module in Biomedical Sciences.* Elsevier.

Roth, G. A., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J. & Abdelalim, A. J., 2018. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*, 392, 1736-1788.

Rusnak, D. W., Lackey, K., Affleck, K., Wood, E. R., Alligood, K. J., Rhodes, N., Keith, B. R., Murray, D. M., Knight, W. B., Mullin, R. J. & Gilmer, T. M., 2001. The effects of the novel, reversible epidermal growth factor receptor/ErbB-2 tyrosine kinase inhibitor, GW2016, on the growth of human normal and tumor-derived cell lines in vitro and in vivo. *Mol Cancer Ther,* 1, 85-94.

Sachdev, D., Chirgwin, J. M., 1998. Order of Fusions between Bacterial and Mammalian Proteins Can Determine Solubility inEscherichia coli. *Biochemical and biophysical research communications*, 244, 933-937.

Samuele, A., Crespan, E., Garbelli, A., Bavagnoli, L. & Maga, G., 2013. The power of enzyme kinetics in the drug development process. *Current pharmaceutical biotechnology,* 14, 551-560.

Santos, J. R., Cozzi-Lepri, A., Phillips, A., De Wit, S., Pedersen, C., Reiss, P., Blaxhult, A., Lazzarin, A., Sluzhynska, M., Orkin, C., Duvivier, C., Bogner, J., Gargalianos-Kakolyris, P., Schmid, P., Hassoun, G., Khromova, I., Beniowski, M., Hadziosmanovic, V., Sedlacek, D., Paredes, R. & Lundgren, J. D., 2018. Long-term effectiveness of recommended boosted protease inhibitor-based antiretroviral therapy in Europe. *HIV Med,* 19, 324-338.

Santos, J. R., Llibre, J. M., Imaz, A., Domingo, P., Iribarren, J. A., Mariño, A., Miralles, C., Galindo, M. J., Ornelas, A. & Moreno, S., 2012. Mutations in the protease gene associated with virological failure to lopinavir/ritonavir-containing regimens. *Journal of antimicrobial chemotherapy*, 67, 1462-1469.

Šašková, K. G., Kožíšek, M., Lepšík, M., Brynda, J., Řezáčová, P., Václavíková, J., Kagan, R. M., Machala, L. & Konvalinka, J., 2008. Enzymatic and structural analysis of the I47A mutation contributing to the reduced susceptibility to HIV protease inhibitor lopinavir. *Protein Science*, 17, 1555-1564.

Sauter, D., Unterweger, D., Vogl, M., Usmani, S. M., Heigele, A., Kluge, S. F., Hermkes, E., Moll, M., Barker, E. & Peeters, M., 2012. Human tetherin exerts strong selection pressure on the HIV-1 group N Vpu protein. *PLoS Pathog*, 8, e1003093.

Schock, H. B., Garsky, V. M. & Kuo, L. C., 1996. Mutational anatomy of an HIV-1 protease variant conferring cross-resistance to protease inhibitors in clinical trials compensatory modulations of binding and activity. *Journal of Biological Chemistry*, 271, 31957-31963.

Schuetz, D. A., Bernetti, M., Bertazzo, M., Musil, D., Eggenweiler, H.-M., Recanatini, M., Masetti, M., Ecker, G. F., Cavalli, A., 2018. Predicting residence time and drug unbinding pathway through scaled molecular dynamics. *Journal of chemical information and modeling,* 59, 535-549.

Seifert, E., 2014. OriginPro 9.1: Scientific Data Analysis and Graphing Software☐ Software Review. ACS Publications.

Selas Castiñeiras, T., Williams, S. G., Hitchcock, A. G. & Smith, D. C., 2018. E. coli strain engineering for the production of advanced biopharmaceutical products. *FEMS microbiology letters,* 365, fny162.

Sezonov, G., Joseleau-Petit, D. & D'ari, R., 2007. Escherichia coli physiology in Luria-Bertani broth. *Journal of bacteriology,* 189, 8746-8749.

Shafer, R. W., Jung, D. R. & Betts, B. J., 2000. Human immunodeficiency virus type 1 reverse transcriptase and protease mutation search engine for queries. *Nature medicine,* 6, 1290-1292.

Shah, S., Alexaki, A., Pirrone, V., Dahiya, S., Nonnemacher, M. R. & Wigdahl, B., 2014. Functional properties of the HIV-1 long terminal repeat containing single-nucleotide polymorphisms in Sp site III and CCAAT/enhancer binding protein site I. *Virology journal*, 11, 92.

Shapiro, R. & Riordan, J. F., 1984. Inhibition of angiotensin converting enzyme: mechanism and substrate dependence. *Biochemistry*, 23, 5225-5233.

Shaw, G. M. & Hunter, E., 2012. HIV transmission. *Cold Spring Harbor perspectives in medicine,* 2, a006965.

Shukla, E. & Chauhan, R., 2019. Host-HIV-1 Interactome: A Quest for Novel Therapeutic Intervention. *Cells,* 8, 1155.

Shum, K.-T., Zhou, J. & Rossi, J. J., 2013. Aptamer-based therapeutics: new approaches to combat human viral diseases. *Pharmaceuticals,* 6, 1507-1542.

Shuman, C. F., Hämäläinen, M. D. & Danielson, U. H., 2004. Kinetic and thermodynamic characterization of HIV-1 protease inhibitors. *Journal of Molecular Recognition*, 17, 106-119.

Shunmugam, L. & Soliman, M. E., 2018. Targeting HCV polymerase: a structural and dynamic perspective into the mechanism of selective covalent inhibition. *RSC advances,* 8, 42210-42222.

Sigaloff, K. C., Hamers, R. L., Wallis, C. L., Kityo, C., Siwale, M., Ive, P., Botes, M. E., Mandaliya, K., Wellington, M. & Osibogun, A., 2011. Unnecessary antiretroviral treatment switches and accumulation of HIV resistance mutations; two arguments for viral load monitoring in Africa. *Journal of Acquired Immune Deficiency Syndromes*, 58, 23-31.

Silverman, R. B., 2007. Enzyme Inhibition. *Wiley Encyclopedia of Chemical Biology*, 1-19.

Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, S. M. & Panda, A. K., 2015. Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. *Microbial cell factories,* 14, 41.

Sk, M. F., Roy, R. & Kar, P., 2021. Exploring the potency of currently used drugs against HIV-1 protease of subtype D variant by using multiscale simulations. *J Biomol Struct Dyn,* 39, 988-1003.

Steegen, K., Bronze, M., Papathanasopoulos, M.A., Van Zyl, G., Goedhals, D., Van Vuuren, C., Macleod, W., Sanne, I., Stevens, W. S., & Carmona, S. C., 2016. Prevalence of antiretroviral drug resistance in patients who are not responding to protease inhibitor–based treatment: results from the first national survey in South Africa. *Journal of Infectious Diseases*, 214, 1826-1830.

Strandberg, L., Enfors, S.O., 1991. Factors influencing inclusion body formation in the production of a fused protein in Escherichia coli. *Applied and environmental microbiology,* 57(6), 1669-1674.

Studier, F. W. & Moffatt, B. A., 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of molecular biology,* 189, 113-130.

Sugiura, W., Matsuda, Z., Yokomaku, Y., Hertogs, K., Larder, B., Oishi, T., Okano, A., Shiino, T., Tatsumi, M., Matsuda, M., Abumi, H., Takata, N., Shirahata, S., Yamada, K., Yoshikura, H. & Nagai, Y., 2002. Interference between D30N and L90M in selection and development of protease inhibitor-resistant human immunodeficiency virus type 1. *Antimicrobial agents and chemotherapy,* 46, 708-715.

Szeltner, Z. & Polgár, L., 1996. Conformational stability and catalytic activity of HIV-1 protease are both enhanced at high salt concentration. *J Biol Chem,* 271, 5458-63.

Talens-Perales, D., Marín-Navarro, J. & Polaina, J., 2016. Enzymes: Functions And Characteristics. *In:* Caballero, B., Finglas, P. M. & Toldrá, F. (eds.) *Encyclopedia of Food and Health.* Oxford: Academic Press.

Tang, M. W. & Shafer, R. W., 2012. HIV-1 antiretroviral resistance. *Drugs*, 72, e1-e25.

Taylor, A., Brown, D. P., Kadam, S., Maus, M., Kohlbrenner, W. E., Weigl, D., Turon, M. C. & Katz, L., 1992. High-level expression and purification of mature HIV-1 protease in Escherichia coli under control of the araBAD promoter. *Appl Microbiol Biotechnol,* 37, 205-10.

Taylor, B. S., Sobieszczyk, M. E., Mccutchan, F. E. & Hammer, S. M., 2008. The challenge of HIV-1 subtype diversity. New England Journal of Medicine, 358, 1590-1602.

Teply, R., Goodman, M. & Destache, C. J., 2011. Lopinavir/Ritonavir: A Review for 2011. C*linical Medicine Insights: Therapeutics*, 3, CMT. S2018.

Terpe, K. 2006. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Applied microbiology and biotechnology,* 72, 211.

Thomas, J. G. & Baneyx, F., 1996. Protein misfolding and inclusion body formation in recombinant Escherichia coli cells overexpressing heat-shock proteins. *Journal of Biological Chemistry,* 271, 11141-11147.

Thompson, M. A., Aberg, J. A., Cahn, P., Montaner, J. S., Rizzardini, G., Telenti, A., Gatell, J. M., Günthard, H. F., Hammer, S. M. & Hirsch, M. S., 2010. Antiretroviral treatment of adult HIV infection: 2010 recommendations of the International AIDS Society–USA panel. *Jama,* 304, 321-333.

Thompson, M. A., Aberg, J. A., Hoy, J. F., Telenti, A., Benson, C., Cahn, P., Eron, J. J., Günthard, H. F., Hammer, S. M. & Reiss, P., 2012. Antiretroviral treatment of adult HIV infection: 2012 recommendations of the International Antiviral Society–USA panel. *Jama*, 308, 387-402.

Tian, W., Chen, C., Lei, X., Zhao, J. & Liang, J., 2018. CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Research,* 46, W363-W367.

Tien, C., Huang, L., Watanabe, S. M., Speidel, J. T., Carter, C. A. & Chen, C., 2018. Context-dependent autoprocessing of human immunodeficiency virus type 1 protease precursors. 13, *PloS one*, e0191372.

Tripathi, N. K., & Shrivastava, A., 2019. Recent developments in bioprocessing of recombinant proteins: expression hosts and process development. *J Frontiers in Bioengineering,* 7.

Trylska, J., Tozzini, V., Chang, C. E. A. & Mccammon, J. A., 2007. HIV-1 protease substrate binding and product release pathways explored with coarse-grained molecular dynamics. *Biophysical journal,* 92, 4179-4187.

Tseng, A., Seet, J. & Phillips, E. J., 2015. The evolution of three decades of antiretroviral therapy: challenges, triumphs and the promise of the future. *British journal of clinical pharmacology,* 79, 182-194.

Tudyka, T. & Skerra, A., 1997. Glutathione S-transferase can be used as a C-terminal, enzymatically active dimerization module for a recombinant protease inhibitor, and functionally secreted into the periplasm of Escherichia coli. *Protein Science*, 6, 2180-2187.

Ullrich, B., Laberge, M., Tölgyesi, F., Fidy, J., Szeltner, Z. & Polgár, L., 2000. Trp42 rotamers report reduced flexibility when the inhibitor acetyl-pepstatin is bound to HIV-1 protease. *Protein Science,* 9, 2232-2245.

UNAIDS 2018. People living with HIV receiving ART. Joint United Nations Programme on HIV/AIDS.

UNAIDS 2020. Global HIV & AIDS statistics—2020 fact sheet. Joint United Nations Programme on HIV/AIDS.

Van Maarseveen, N. & Boucher, C., 2006. Resistance to protease inhibitors. *Antiretroviral resistance in clinical practice.* Mediscript.

Velazquez-Campoy, A., Todd, M. J., Vega, S. & Freire, E., 2001. Catalytic efficiency and vitality of HIV-1 proteases from African viral subtypes. *Proceedings of the National Academy of Sciences*, 98(11), 6062-6067.

Velázquez-Campoy, A., Vega, S., Fleming, E., Bacha, U., Sayed, Y., Dirr, H. W. & Freire, E., 2003. Protease inhibition in African subtypes of HIV-1. *Aids Review*, 5, 165-171.

Velazquez-Campoy, A., Vega, S. & Freire, E., 2002. Amplification of the effects of drug resistance mutations by background polymorphisms in HIV-1 protease from African subtypes. *Biochemistry,* 41, 8613-9.

Venter, W. D., Moorhouse, M., Sokhela, S., Serenata, C., Akpomiemie, G., Qavi, A., Mashabane, N., Arulappan, N., Sim, J. W. & Sinxadi, P. Z., 2019. Low-dose ritonavir-boosted darunavir once daily versus ritonavir-boosted lopinavir for participants with less than 50 HIV RNA copies per mL (WRHI 052): A randomised, open-label, phase 3, non-inferiority trial. *Lancet HIV* ,6(7), e428-e437.

Volontè, F., Piubelli, L. & Pollegioni, L., 2011. Optimizing HIV-1 protease production in Escherichia coli as fusion protein. *Microbial cell factories,* 10, 53.

Wan, M., Takagi, M., Loh, B.-N. & Imanaka, T., 1995. Comparison of HIV-1 protease expression in different fusion forms. *Biochemistry and molecular biology international,* 36, 411-419.

Wan, M., Takagi, M., Loh, B. N., Xu, X. Z. & Imanaka, T., 1996. Autoprocessing: an essential step for the activation of HIV-1 protease. *Biochemical Journal*, 316, 569-573.

Wang, Y., Dewdney, T. G., Liu, Z., Reiter, S. J., Brunzelle, J. S., Kovari, I. A. & Kovari, L. C., 2012. Higher desolvation energy reduces molecular recognition in multi-drug resistant HIV-1 protease. *Biology*, 1, 81-93.

Wangpatharawanit, P. & Sungkanuparph, S., 2016. Switching lopinavir/ritonavir to atazanavir/ritonavir vs adding atorvastatin in HIV-infected patients receiving second-line antiretroviral therapy with hypercholesterolemia: a randomized controlled trial. *Clinical Infectious Diseases,* 63, 818-820.

Waugh, D. S., 2011. An overview of enzymatic reagents for the removal of affinity tags. *Protein expression and purification,* 80, 283-293.

Weber, I. & Agniswamy, J., 2009. HIV-1 protease: structural perspectives on drug resistance. *Viruses*, 1, 1110-1136.

Weber, I. T., Kneller, D. W. & Wong-Sam, A., 2015. Highly resistant HIV-1 proteases and strategies for their inhibition. *Future medicinal chemistry*, 7, 1023-1038.

Weikl, T. R., 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.

Weikl, T. R., Hemmateenejad, B., 2019. Accessory mutations balance the marginal stability of the HIV-1 protease in drug resistance. *Proteins: Structure, Function, and Bioinformatics*, 88(3), 476-484.

Wensing, A. M., Calvez, V., Ceccherini-Silberstein, F., Charpentier, C., Günthard, H. F., Paredes, R., Shafer, R. W. & Richman, D. D., 2019. 2019 update of the drug resistance mutations in HIV-1. *Topics in antiviral medicine,* 27, 111.

Wensing, A. M., Van Maarseveen, N. M. & Nijhuis, M., 2010. Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance. *Antiviral research*, 85, 59-74.

WHO. 2014. Access to antiretroviral drugs in low-and middle-income countries: technical report July 2014. Geneva, World Health Organization

WHO 2016. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach. World Health Organization.

WHO 2019. HIV drug resistance report 2019. Geneva, World Health Organization.

WHO 2020. HIV/AIDS. Geneva, World Health Organization

Williams, A., Basson, A., Achilonu, I., Dirr, H. W., Morris, L. & Sayed, Y., 2019. Double trouble? Gag in conjunction with double insert in HIV protease contributes to reduced DRV susceptibility. *Biochemical Journal*, 476, 375-384.

Wong-Sam, A., Wang, Y.-F., Zhang, Y., Ghosh, A. K., Harrison, R. W. & Weber, I. T., 2018. Drug Resistance Mutation L76V Alters Nonpolar Interactions at the Flap–Core Interface of HIV-1 Protease. *ACS omega*, 3, 12132-12140.

Wright, J. K., Brumme, Z. L., Carlson, J. M., Heckerman, D., Kadie, C. M., Brumme, C. J., Wang, B., Losina, E., Miura, T. & Chonco, F., 2010. Gag-protease-mediated replication capacity in HIV-1 subtype C chronic infection: associations with HLA type and clinical parameters. *Journal of virology*, 84(20), 10820-10831.

Wu, T. D., Schiffer, C. A., Gonzales, M. J., Taylor, J., Kantor, R., Chou, S., Israelski, D., Zolopa, A. R., Fessel, W. J. & Shafer, R. W., 2003. Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments. *Journal of virology*, 77, 4836-4847.

Yadav, D. K., Yadav, N., Yadav, S., Haque, S. & Tuteja, N., 2016. An insight into fusion technology aiding efficient recombinant protein production for functional proteomics. *Archives of biochemistry and biophysics,* 612, 57-77.

Yamaguchi, J., Mcarthur, C., Vallari, A., Sthreshley, L., Cloherty, G., Berg, M. & Rodgers, M., 2019. Complete genome sequence of CG-0018a-01 establishes HIV-1 subtype LJ Acquir. *Immune Defic. Syndr*.

Yan, L., Yu, F., Zhang, H., Zhao, H., Wang, L., Liang, Z., Zhang, X., Wu, L., Liang, H., Yang, S., Tang, Y. & Zhang, F., 2020. Transmitted and Acquired HIV-1 Drug Resistance from a Family: A Case Study. *Infection and drug resistance,* 13, 3763-3770.

Yang, H., Nkeze, J., Zhao, R. Y., 2012a. Effects of HIV-1 protease on cellular functions and their potential applications in antiretroviral therapy. *Cell & bioscience*, 2, 32.

Yang, Z., Lasker, K., Schneidman-Duhovny, D., Webb, B., Huang, C. C., Pettersen, E. F., Goddard, T. D., Meng, E. C., Sali, A. & Ferrin, T. E., 2012b. UCSF Chimera, MODELLER, and IMP: an integrated modeling system. *Journal of structural biology*, 179, 269-278.

Ylilauri, M., Pentikäinen, O. T., 2013. MMGBSA as a tool to understand the binding affinities of filamin–peptide interactions. J*ournal of chemical information and modeling*, 53, 2626-2633.

Young, C. L., Britton, Z. T. & Robinson, A. S., 2012. Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. *Biotechnology journal,* 7, 620-634.

Yu, Y., Wang, J., Chen, Z., Wang, G., Shao, Q., Shi, J. & Zhu, W., 2017. Structural insights into HIV-1 protease flap opening processes and key intermediates. *RSC advances*, 7, 45121-45128.

Yu, Y., Wang, J., Shao, Q., Shi, J. & Zhu, W., 2015. Effects of drug-resistant mutations on the dynamic properties of HIV-1 protease and inhibition by Amprenavir and Darunavir. *Scientific reports,* 5, 10517-10517.

Zhang, Y., Chang, Y.-C. E., Louis, J. M., Wang, Y.-F., Harrison, R. W. & Weber, I. T., 2014. Structures of darunavir-resistant HIV-1 protease mutant reveal atypical binding of darunavir to wide open flaps. *ACS chemical biology*, 9, 1351-1358.

Zhao, G., Perilla, J. R., Yufenyuy, E. L., Meng, X., Chen, B., Ning, J., Ahn, J., Gronenborn, A. M., Schulten, K. & Aiken, C., 2013a. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature,* 497, 643-646.

Zhao, X., Li, G. & Liang, S., 2013b. Several affinity tags commonly used in chromatographic purification. *Journal of analytical methods in chemistry*, 2013.

Zhu, K., Zhou, X., Yan, Y., Mo, H., Xie, Y., Cheng, B. & Fan, J., 2017. Cleavage of fusion proteins on the affinity resins using the TEV protease variant. *Protein expression and purification,* 131, 27-33.

Zhu, Q., Yu, Z., Kabashima, T., Yin, S., Dragusha, S., El-Mahdy, A. F., Ejupi, V., Shibata, T. & Kai, M., 2015. Fluorometric assay for phenotypic differentiation of drug-resistant HIV mutants. *Scientific reports,* 5, 10323.

Zondagh, J., Williams, A., Achilonu, I., Dirr, H. W. & Sayed, Y., 2018. Overexpression, Purification and Functional Characterisation of Wild-Type HIV-1 Subtype C Protease and Two Variants Using a Thioredoxin and His-Tag Protein Fusion System. *The protein journal*, 1-11.

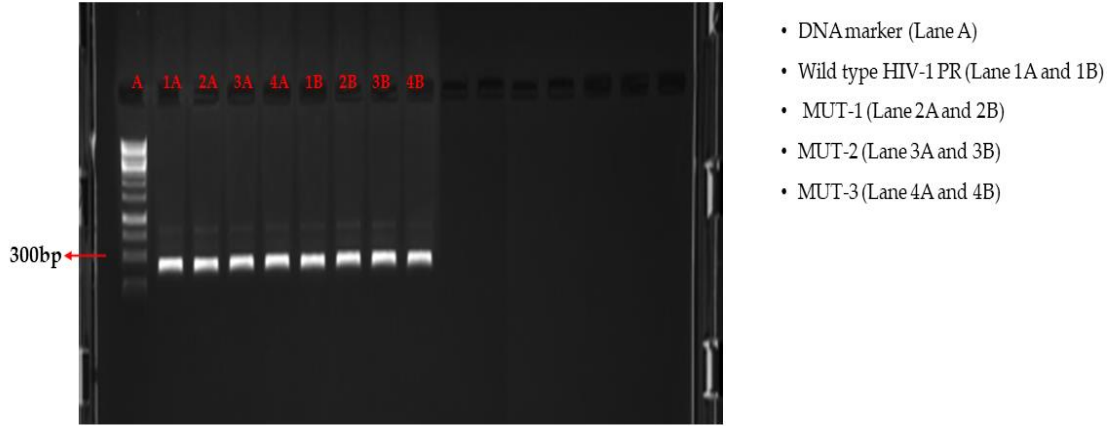# Appendix

Appendix 1: Supplementary data for chapter 3

- DNA marker (Lane A)
- Wild type HIV-1 PR (Lane 1A and 1B)
- MUT-1 (Lane 2A and 2B)
- MUT-2 (Lane 3A and 3B)
- MUT-3 (Lane 4A and 4B)

Figure S1. Unedited gel electrophoresis picture of amplified HIV-1 PR gene

- Purified HIV-1 PR variants : Land A – 4A
  - A: Protein marker
  - 1A: Wild type HIV-1 PR
  - 2A: MUT-1
  - 3A: MUT-2
  - 4A: MUT-3
- Purified MBP-tagged HIV-1 PR: Land B – 4B
  - B: Protein marker
  - 1B: Wild type HIV-1 PR
  - 2B: MUT-1
  - 3B: MUT-2
  - 4B: MUT-3
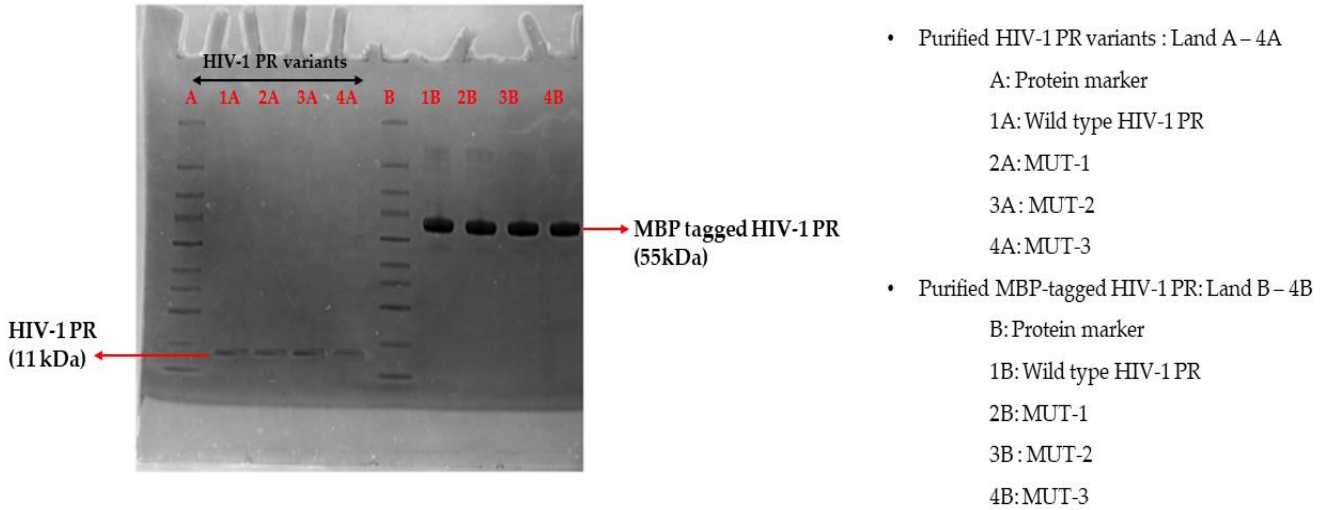
Figure S2. Unedited SDS-PAGE picture of purified HIV-1 PR gene and MBP tagged HIV-1 PR
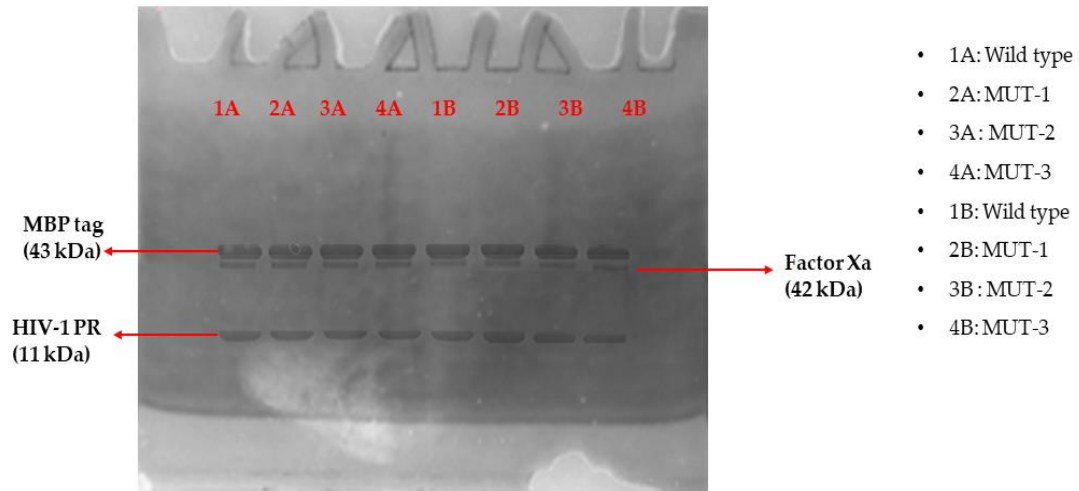
Figure S3. Unedited SDS-PAGE picture showing cleavage products after factor Xa cleavage of HIV-1 PR gene from the MBP tag
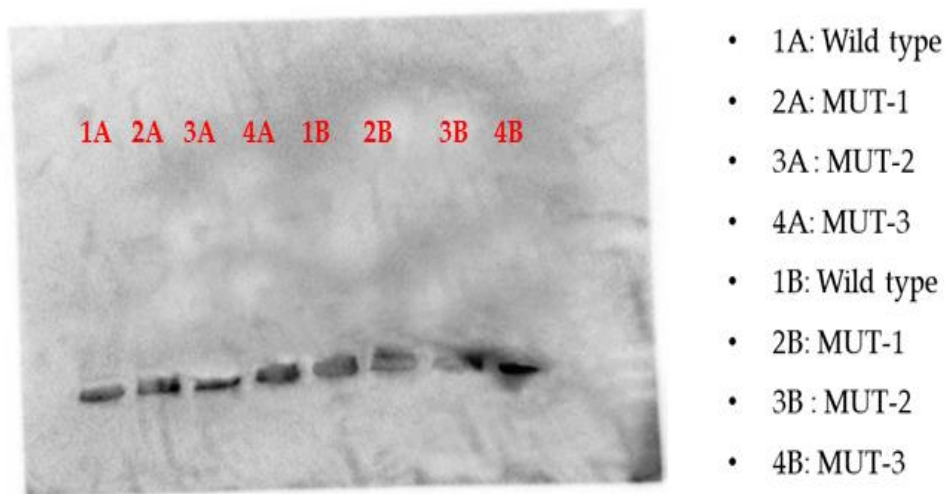

Figure S4. Unedited western blot picture of HIV-1 PR variants

Appendix 2 Supplementary data for chapter 4

Table S1. Wild type and mutant HIV-1 PR variants enzyme kinetic parameters (K_m, K_{cat}) and inhibition constant (Ki) calculated using the chromogenic substrate

HIV-1 PR Variants	K_m (μM)	K_{cat} (s^{-1})	K_{cat}/K_m ($s^{-1}\mu M^{-1}$)	LPV		DRV	
				K_i (nM)	Relative resistance to LPV	K_i	Relative resistance to DRV
WT	36.95 ± 0.95	0.81 ± 0.08	0.022	2.09 ± 0.18	1.00	1.74 ± 0.27	1.00
MUT-1	61.29 ± 0.77	0.54 ± 0.13	0.0088	67.72 ± 0.33	32.40	8.93 ± 0.10	5.13
M46I, I54V, L76V, V82A, I84V, Q58E							
MUT-2	64.44 ± 0.21	0.44 ± 0.15	0.0068	83.45 ± 0.89	39.93	13.27 ± 0.05	7.63
V32I, M46I, I54V, L76V, V82A, L90M, L33F							

Relative resistance = K_i of mutant / K_i of WT

Table S2. The rate constants of association (k_5) and dissociation (k_6) calculated by fitting k to Eq. (5).

HIV-1 PR Variants	LPV			DRV		
	K_5 (s^{-1})	K_6 (s^{-1})	t_R	K_5 (s^{-1})	K_6 (s^{-1})	t_R
WT	0.74	0.15	6.67	0.43	0.15	6.67
MUT-1	0.19	0.42	2.38	0.16	0.46	2.17
M46I, I54V, L76V, V82A, I84V, Q58E						
MUT-2	0.13	0.33	3.03	0.14	0.59	1.70
V32I, M46I, I54V, L76V, V82A, L90M, L33F						

Table S3. The inhibition constant (K_i) calculated by fitting the data for the magnitude of the rapid fluorescence decrease (F_0–F $= \Delta Fmax/(1+(K_i/[I]))$) and the Stern-Volmer quenching constants (K_{sv}) calculated from fluorescence quenching assay ($F_0/F = 1 + K_{sv}$ [I])

HIV-1 PR Variants	LPV		DRV	
	K_i (nM)	K_{sv} (nM^{-1})	K_i (nM)	K_{sv} (nM^{-1})
WT	19.28 ± 3.02	0.02	8.67 ± 0.54	0.030
MUT-1	78.05 ± 5.22	0.004	39.73 ± 4.07	0.01
M46I, I54V, L76V, V82A, I84V, Q58E				
MUT-2	123.68 ± 7.46	0.004	50.11 ± 3.19	0.009
V32I, M46I, I54V, L76V, V82A, L90M, L33F				

Table S4. The inhibition constant (K_i) calculated by fitting the first-order rate constants (K_{obs}) at each LPV and DRV concentration into the first-order equation y=a+b*exp ($-K_{obs}$ *t) using. The K_{obs} values were fitted to the equation $K_{obs} =k_5[I]/(K_i+[I])$.

HIV-1 PR Variants	LPV		DRV	
	K_i (nM)	k_5 (s^{-1})	K_i (nM)	k_5 (s^{-1})
WT	3.13 ± 0.55	0.17	2.67 ± 0.38	0.14
MUT-1 M46I, I54V, L76V, V82A, I84V, Q58E	49.69 ± 3.27	0.11	13.21 ± 1.23	0.11
MUT-2 V32I, M46I, I54V, L76V, V82A, L90M, L33F	79.24 ± 4.48	0.10	19.89 ± 3.02	0.12
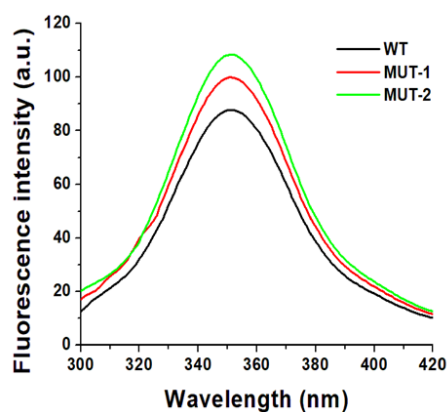
Figure S1. Intrinsic tryptophan fluorescent graph for wild type and mutant HIV-1 PR variants

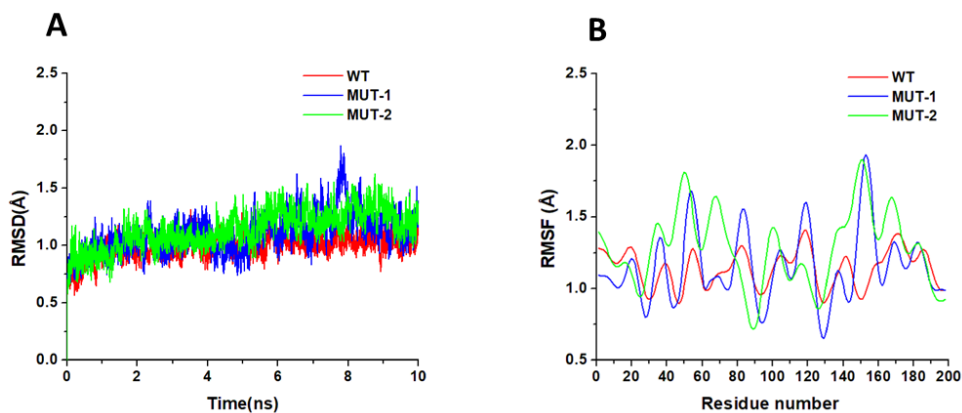

Figure S2. (A) Root mean square deviation (RMSD) for WT and mutant HIV-1 PR variants (B) RMSF for WT and mutant HIV-1 PR variants

Appendix 3: Biomedical Ethics Approval

UNIVERSITY OF ™
KWAZULU-NATAL
INYUVESI
YAKWAZULU-NATALI

24 December 2020

Mr S Eche (215000273)
School of Laboratory Medicine and Medical Sciences
College of Health Sciences
echesimeon@gmail.com

Dear Mr Eche

PROTOCOL: Structural characterization of highly mutated HIV-1subtype C protease
Degree: PhD
BREC Ref No: BE413/17

RECERTIFICATION APPLICATION APPROVAL NOTICE

Approved:	**06 November 2020**
Expiration of Ethical Approval:	**05 November 2021**

I wish to advise you that your application for recertification received on 18 December 2020 has been **noted and approved** by a subcommittee of the Biomedical Research Ethics Committee (BREC). The start and end dates of this period are indicated above.

If any modifications or adverse events occur in the project before your next scheduled review, you must submit them to BREC for review. Except in emergency situations, no change to the protocol may be implemented until you have received written BREC approval for the change.

The committee will be notified of the above approval at its next meeting to be held on 09 February 2021.

Yours sincerely

....

Ms A Marimuthu
(for) Prof D Wassenaar
Chair: Biomedical Research Ethics Committee

Appendix 4: Biomedical Ethics Recertification

UNIVERSITY OF ™
KWAZULU-NATAL
INYUVESI
YAKWAZULU-NATALI

16 July 2020

Mr S Eche (215000273)
School of Laboratory Medicine and Medical Sciences
College of Health Sciences
echesimeon@gmail.com

Dear Mr Eche

PROTOCOL: Structural characterization of highly mutated HIV-1subtype C protease
**Degree: PhD**
BREC Ref No: BE413/17
NEW TITLE: Biochemical and functional characterization of highly mutated South African HIV-1 subtype C protease

We wish to advise you that your response to BREC letter dated 06 July 2020 received on 09 July 2020 has been **noted** by a subcommittee of the Biomedical Research Ethics Committee.
Your correspondence received on 30 June 2020 submitting an application for amendments including changing the title to the new title above for the above study has now been **approved** by a subcommittee of the Biomedical Research Ethics Committee.

The committee will be notified of the above approval at its next meeting to be held on 11 August 2020.

Yours sincerely
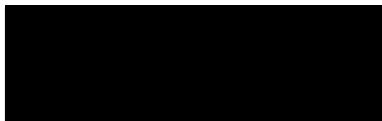
▬▬▬▬▬▬▬▬▬▬
....
Ms A Marimuthu
(for) Prof D Wassenaar
Chair: Biomedical Research Ethics Committee

Appendix 5: Cover page of Published paper (Paper 1)

Review

Recombinant expression of HIV-1 protease using soluble fusion tags in *Escherichia coli*: A vital tool for functional characterization of HIV-1 protease

Simeon Eche *, Michelle L. Gordon *

School of Laboratory Medicine and Medical Sciences, University of KwaZulu-Natal, Durban, 4001, South Africa

ARTICLE INFO

ABSTRACT

HIV-1 protease expression in the laboratory is demanding because of its high cytotoxicity, making it difficult to express in bacterial expression systems such as *Escherichia coli*. To overcome these challenges, HIV-1 protease fusion with solubility enhancing tags helps to mitigate its cytotoxic effect and drive its expression as a soluble protein. Therefore, this review focuses on the expression of bioactive HIV-1 protease using solubility-enhancing fusion tags in *Escherichia coli* and summarises the characteristic features of the different common fusion tags that have been used in the expression of HIV-1 protease. This review will assist researchers with their choice of protein fusion tag for HIV-1 protease expression.

Appendix 6: Cover page of Published paper (Paper 2)

biomolecules

Article

Acquired HIV-1 Protease Conformational Flexibility Associated with Lopinavir Failure May Shape the Outcome of Darunavir Therapy after Antiretroviral Therapy Switch

Simeon Eche [1], Ajit Kumar [2], Nelson Sonela [3,4] and Michelle L. Gordon [1,*]

[1] Discipline of Virology, School of Laboratory Medicine and Medical Sciences, University of KwaZulu-Natal, Durban 4001, South Africa; echesimeon@gmail.com
[2] Discipline of Microbiology, School of Life Sciences, University of KwaZulu-Natal (Westville Campus), Durban 4000, South Africa; ajitkanwal@yahoo.com
[3] School of Medicine, Physical and Natural Sciences, University of Rome Tor Vegata, 1-00133 Rome, Italy; nelsonela@yahoo.fr
[4] Chantal Biya International Reference Center for Research on the Management and Prevention of HIV/AIDS (CIRCB), Yaoundé P.O. Box 3077, Cameroon
[*] Correspondence: Tarinm@ukzn.ac.za; Tel.: +27-031-260-4498

check for updates

Abstract: Understanding the underlying molecular interaction during a therapy switch from lopinavir (LPV) to darunavir (DRV) is essential to achieve long-term virological suppression. We investigated the kinetic and structural characteristics of multidrug-resistant South African HIV-1 subtype C protease (HIV-1 PR) during therapy switch from LPV to DRV using enzyme activity and inhibition assay, fluorescence spectroscopy, and molecular dynamic simulation. The HIV-1 protease variants were from clinical isolates with a combination of drug resistance mutations; MUT-1 (M46I, I54V, V82A, and L10F), MUT-2 (M46I, I54V, L76V, V82A, L10F, and L33F), and MUT-3 (M46I, I54V, L76V, V82A, L90M, and F53L). Enzyme kinetics analysis shows an association between increased relative resistance to LPV and DRV with the progressive decrease in the mutant HIV-1 PR variants' catalytic efficiency. A direct relationship between high-level resistance to LPV and intermediate resistance to DRV with intrinsic changes in the three-dimensional structure of the mutant HIV-1 PR as a function of the multidrug-resistance mutation was observed. In silico analysis attributed these structural adjustments to the multidrug-resistance mutations affecting the LPV and DRV binding landscape. Though DRV showed superiority to LPV, as a lower concentration was needed to inhibit the HIV-1 PR variants, the inherent structural changes resulting from mutations selected during LPV therapy may dynamically shape the DRV treatment outcome after the therapy switch.

Keywords: HIV-1 protease; HIV-1 protease inhibitor; lopinavir; darunavir; conformational flexibility