1-1-2016

Structure Function Studies Of Hiv-1 Protease

Bradley James Keusch
Wayne State University,

STRUCTURE FUNCTION STUDIES OF HIV-1 PROTEASE

By

BRAD KEUSCH

THESIS

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

2016

MAJOR: BIOCHEMISTRY AND
MOLECULAR BIOLOGY

Approved by:

_____

Advisor Date

ACKNOWLEDGEMENTS

I would like to thank my family for their unending support throughout my entire life, much less graduate school (though as a starving student those weekend meals always taste even better). I especially thank my future wife Janelle Stokely for listening to my frustrations and complaints (and there were many) and celebrating my successes along the way.

I thank my mentor Dr. Ladislau Kovari for the hard work he does day in and day out, and for acting as a great guide in both science and life. The support he gives to all the students in the lab is greatly appreciated, and we thank you for all you do! I would also like to thank Iulia Kovari for her management of the lab, and for the thoughtful gifts she has given me over the years (those chocolate oranges during Christmas time always tasted incredible, and those Costco chips were life-saving on more than one occasion!).

I thank Dr. Yang and Dr. Edwards for serving on my committee and for offering their advice and insight to my project, and others in the lab. I especially thank Dr. Yang for the use of his lab to help further our work.

I thank my former lab mate and science sister Dr. Tamaria Dewdney for her amazing guidance, training, and spirit. She always made the lab a great place to be and could always be counted on to brighten up an otherwise frustrating or dull day.

I thank my lab mates (Ben Kuiper, Kyla Ross, Cathy McLeod, Dr. Poorvi Chordia, Kristin Slater, and Prasanna Medapureddy) for their companionship, hard work, and insight into all of our projects. I wish you all the best in the future, and cheer those of you who have already begun to make great strides after leaving the lab!

I thank Josh Holcolmb, and Nick Spellmon for their assistance with protein crystallography and for getting us up and running with our use of the Yang lab equipment, you guys were invaluable!

I would like to thank the BMB chair Dr. Bharati Mitra, and everyone in the BMB department, as well as the BMB office staff (April Wolak, Yanna Marsh, Roselle Cooper, and Joseph Fiore) for their support and for all the times they have helped me out.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

1.1 HIV-1 Epidemiology and life cycle

Human immunodeficiency virus (HIV) was first characterized as the direct cause of Acquired Immune Deficiency Syndrome (AIDS) in 1983[1]. Initially a devastating diagnosis with few options for treatment, major research efforts led to the development of multiple small molecule inhibitors through the mid-1990s and into the mid-2000s [2]. These efforts continue to contribute to a reduction in AIDS related deaths; there was a reduction of ~35% in AIDS death worldwide from 2011 to 2015 (World Health Organization, 2015). However, the number of people living with HIV worldwide has increased by ~10% over this same span, and effective treatment remains an important global concern.

Table 1.1 World Health Organization (WHO) data illustrating worldwide HIV/AIDS prevalence for the year 2015. This report is compiled biannually by a consortium of the WHO, UNAIDS, and Unicef.

Global summary of the AIDS epidemic | 2015

Number of people living with HIV in 2015	**Total**	36.7 million [34.0 million – 39.8 million]
	Adults	34.9 million [32.4 million – 37.9 million]
	Women (15+)	17.8 million [16.4 million – 19.4 million]
	Children (<15 years)	1.8 million [1.5 million – 2.0 million]
People newly infected with HIV in 2015	**Total**	2.1 million [1.8 million – 2.4 million]
	Adults	1.9 million [1.7 million – 2.2 million]
	Children (<15 years)	150 000 [110 000 – 190 000]
AIDS deaths in 2015	**Total**	1.1 million [940 000 – 1.3 million]
	Adults	1.0 million [840 000 – 1.2 million]
	Children (<15 years)	110 000 [84 000 – 130 000]

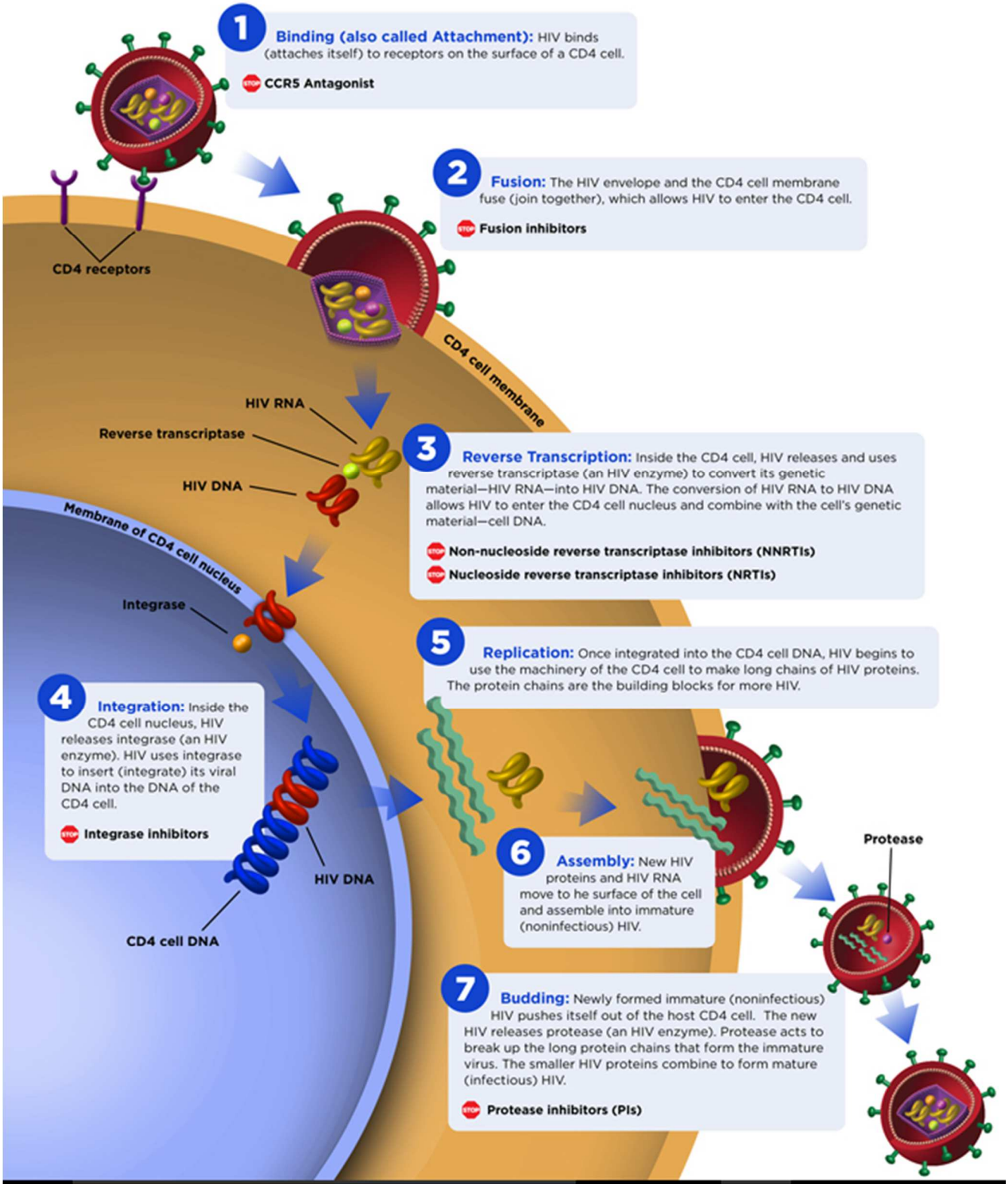
WHO – HIV department | June 15, 2016 World Health Organization UNAIDS unicef

HIV is a Lentivirus (a genus of the Retroviridae family) and human infection is caused by two species: HIV-1 and HIV-2. HIV-2 is mostly local to western Africa and has lower virulence and

infectivity than HIV-1, the latter of which causes the majority of HIV infections throughout the world [3, 4].

The HIV life cycle begins with the attachment of virus to host immune cells, predominantly T4-lymphocytes (**Fig 1.1)**. These cells express the CD4 receptor on their surface, and interactions between this receptor (along with the co-receptors CCR5 and CXCR4) and HIV glycoproteins gp41 and gp120 allow the virus to fuse with and release its contents into the cell. The virus contains two copies of the single-stranded, positive sense RNA genome, as well as the viral enzymes reverse transcriptase and integrase. The former transcribes the RNA genome into DNA, creating a second strand of DNA in the process to create complete double-stranded DNA. Integrase binds and cleaves a dinucleotide group from the 3' ends of the nascent DNA, priming it to react with host DNA. This complex of newly synthesized DNA and bound viral enzymes is called the pre-integration complex, and along with some host proteins it translocates into the nucleus. Once inside, integrase catalyzes the strand-transfer reaction that allows the exposed hydroxyl groups from the newly nicked 3' DNA ends to hydrolyze phosphodiester bonds in actively transcribed host DNA regions. Host cell DNA repair proteins repair the single stranded breaks, integrating the viral DNA into the host genome. The host DNA is then transcribed, and the viral protein containing mRNA is processed by the HIV protease, releasing structural and non-structral viral proteins. Mature virions are then assembled and bud off from the host cell membrane, allowing the cycle to repeat. The time to complete this replication cycle is approximately 36 hours.

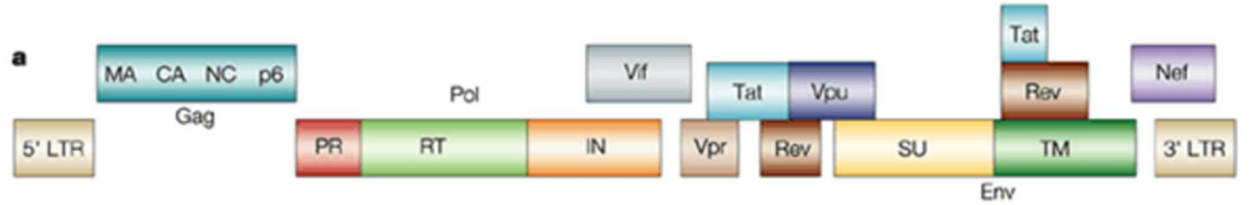
Figure 1.1 The HIV-1 life cycle. The figure covers virus attachment through budding and maturation (aidsinfo.nih.gov)



1.2 *The HIV genome*

HIV possesses two copies of its 9,719 nucleotide containing positive-sense RNA genome. Each RNA component contains a 3' poly(A) tail and a 5'Gppp cap [5]. The viral RNA is tightly associated with viral nucleocapsid proteins to prevent degradation via host cell nucleases. The HIV genome contains 3 major open reading frames, and utilizes differential RNA splicing to produce 9 gene products (**Fig 1.2**). The 3 major genes are *gag*, *pol* and *env*. *Gag* encodes viral structural proteins including MA (p17, the matrix protein), CA (the capsid protein p24) and NC (nucleocapsid protein p7). *Pol* encodes viral enzymes including reverse transcriptase, integrase, and protease. *Env* encodes envelope glycoproteins including gp120 and gp41, which are essential for viral entry[6].

Fig 1.2 The HIV-1 RNA genome. Overlap occurs between some gene products due to differential RNA splicing [7].



The essential and accessory regulatory proteins (many of which are shared by most retroviruses, some of which are unique to HIV) are produced in smaller, spliced mRNA transcripts. While *gag* and *pol* are produced as an unspliced 9.2kb transcript (in addition to a smaller *gag*-only transcript), the essential regulatory proteins Tat and Rev are found on a multiply spliced 2 kb transcript[8]. These proteins function mainly in the nucleus where they serve to stimulate transcription and regulate viral RNA production. Vif, Vpu, and Vpr are found on a singly spliced 4.5 kb transcript. Vif (viral infectivity factor) improves the efficiency of mature virion assembly by degrading a host cell antiretroviral cytidine aminase; Vpu (viral protein u) aids in virion release from the host cell membrane; Vpr (viral protein r) arrests the host cell in G2 and allows for translocation of the pre-integration complex (containing the reverse transcribed viral DNA) into the host cell nucleus [9]. The long terminal repeats found on both the 3' and 5' ends of the genome are essential and serve as recognition sites for integrase.

1.3 *HIV-1 protease*

HIV-1 protease (protease) is a 99-amino acid containing homodimer (**Fig 1.3**). It is an aspartyl protease and is essential to the HIV life cycle, rendering it an attractive and very successful drug target. The enzyme contains the classic aspartyl protease motif (Asp/Thr/Gly) at residues 25-27, with residue 25 acting as the catalytic residue (and both monomers contributing one Asp residue that extends into the active site cavity). Each monomer contains an anti-parallel β-sheet with a turn (referred to as the "flap") that in the homodimer forms the top of the active site cavity (residues 46-56). Crystal structures of both apo and ligand bound protease show that these flaps can move as much as 7 Å upon substrate or inhibitor binding, and their motion is affected by drug-resistance mutations [10]. protease performs 12 proteolytic reactions, 5 on the Gag polyprotein, 6 on the Gag-Pol polyprotein, and one on Nef. The initial auto-cleavage of protease from the Gag-Pol polyprotein occurs via an intramolecular reaction [11]. The MA/CA cleavage site on the Gag polyprotein is the site with the highest affinity for protease, as it strongly prefers binding to the N-terminal side of Pro residues (with either Tyr or Phe residues adjacent). Notably, there are no known human enzymes that cleave at a similar site (an aromatic residue directly adjacent to a Pro residue) [12]. These reactions occur late in the viral life cycle during virion assembly and maturation near the host cell membrane.

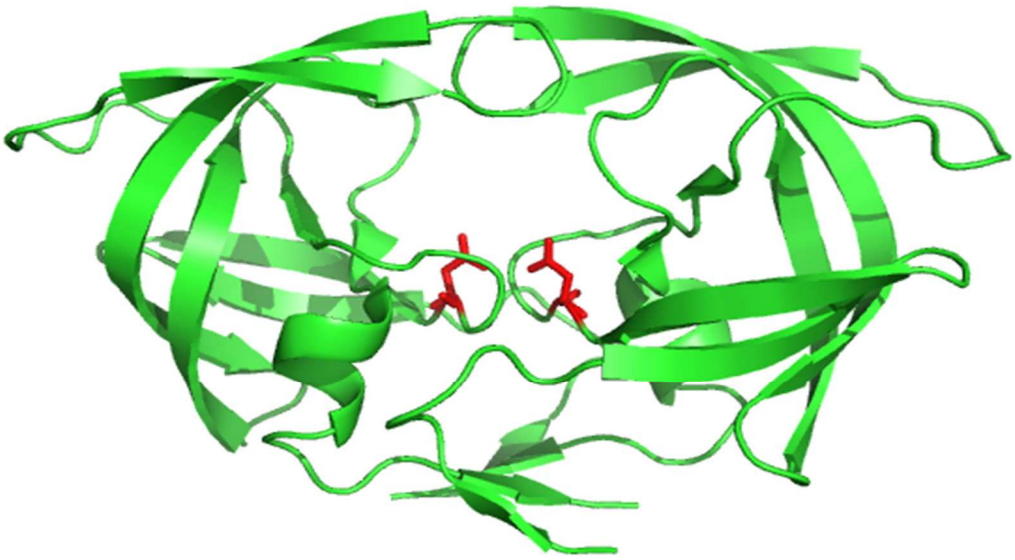


Fig 1.3 Wild type HIV-1 Protease. The catalytic Asp residues are highlighted in red. Note that the flaps are closed in this conformation (PDB 3PHV)

The mechanism of action of the HIV protease has been extensively studied, and is thought to proceed in a similar manner to other aspartyl proteases [13]. Although two Asp residues exist in the active site, only one is protonated at any given time, allowing for the activation of a water molecule held between the two catalytic residues. This activated water molecule is then able to perform nucleophilic attack on the carbonyl carbon of the substrate's scissile bond (**Fig 1.4**).

1.4 *HIV-1 Protease inhibitors and mechanism of action*

There are 9 FDA approved protease inhibitors (PIs), most of which are peptidomimetic and designed to mimic the tetrahedral intermediate cleavage substrate product (**Fig 1.5**). PIs quickly became a mainstay of HIV treatment. Highly Active Anti-Retroviral Therapy (HAART) became the standard of

care throughout the late 90s and into the 2000s, and this treatment involves a combination of drug

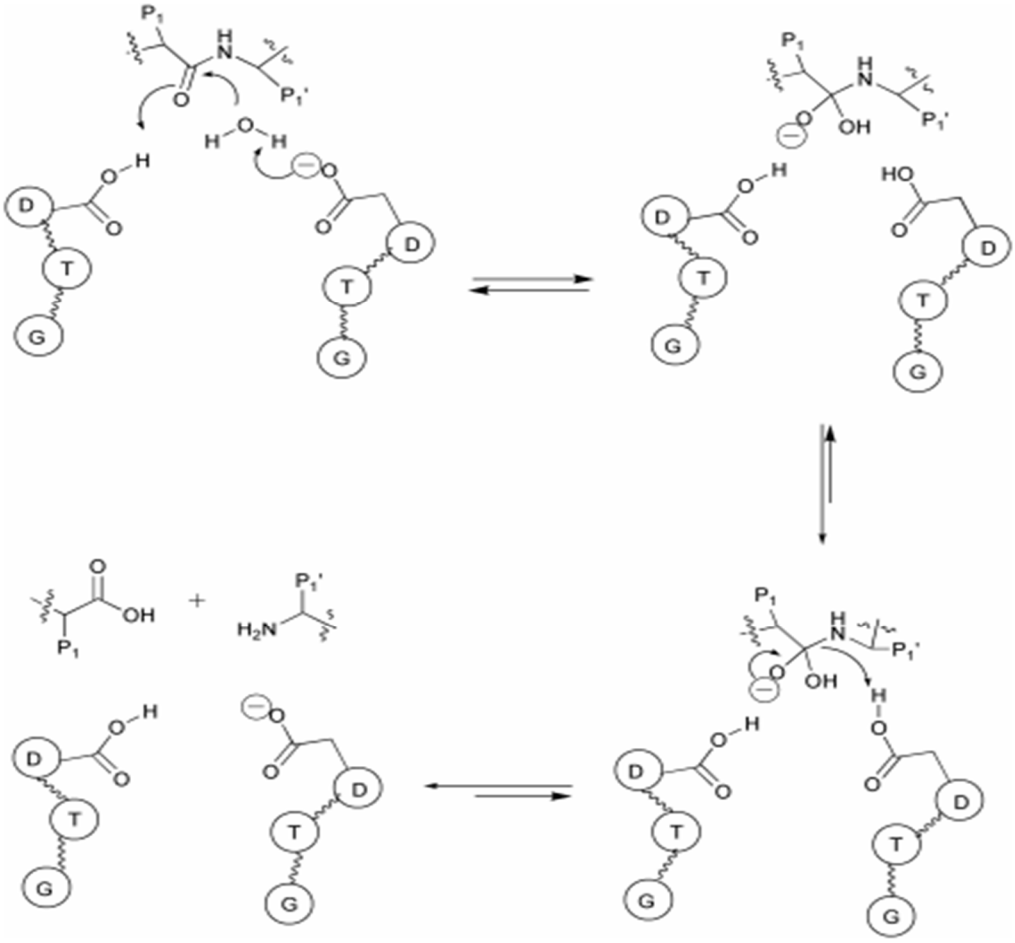


Fig 1.4 General mechanism of action for aspartyl proteases (including HIV-1 protease). D/T/G represent the conserved Aspartic acid, threonine, glycine catalytic triad [13].

therapies, generally including 2 nucleoside reverse transcriptase inhibitors, and one ritonavir-boosted protease inhibitor. Depending on patient need, occasionally the PI is substituted for a non-nucleoside reverse transcriptase inhibitor or integrase strand transfer inhibitor [14]. Saquinavir (SQV) was the first PI to be approved, entering clinical trials in 1989 and reaching the market in 1995. It was initially designed to mimic a proline residue at P1' and a phenylalanine residue at the P1site [15]. As noted above, this is a unique recognition motif for HIV protease. These groups were eventually changed slightly to increase contacts with the protease and improve binding affinity. SQV is generally well tolerated but is unable to achieve high blood serum levels.

Following the release of SQV, three other PIs attained FDA approval within one year: Ritonavir, Indinavir, and Nelfinavir. Ritonavir in particular was an important discovery; although it has significant side effects when administered at a dose effective for protease inhibition, it was discovered that it acts as a potent cytochrome P450-3A4 inhibitor, which is the cytochrome P450 form responsible for metabolism of the other PIs [16]. Thus, a low dose of ritonavir co-administered with another PI allows for effective protease inhibition at a lower dosing concentration, further reducing the chance of severe side effects. Nearly all PIs are now co-administered with ritonavir. Lopinavir (LPV), first approved by the FDA in 2000, was the first PI to be approved exclusively for co-administration with ritonavir [16]. Amprenavir was approved in late 1999, and was supplanted by fosamprenavir a few years later. The latter is a prodrug formulation of amprenavir, allowing for the drug to last longer with each dosing and to require fewer pills overall [15]. Atazanavir was approved in 2003 and was popular not only because of its effectiveness as a PI, but because it is less likely than the other PIs to cause lipodystrophy (abnormal fat deposits, especially in the neck and back) and produce elevated cholesterol levels, significant side effects that often accompany PI treatment. It also has a more favorable resistance profile than earlier PIs [17]. Tipranvir was approved in 2005, and is generally reserved for use as a second-line or "rescue" PI, because it retains effectiveness in the face of resistance to other PIs (it generally requires multiple mutations before its effectiveness wanes as severely) [18]. Additionally, the side effect profile for Tipranavir is worse than most other PIs making it less attractive as an initial therapy. Finally, Darunavir was the most recent HIV PI to be FDA approved, attaining this status in 2006. It was specifically designed to utilize a hydrogen bonding strategy with the backbone structure of both wild type and PI-resistant HIV protease, and is capable of rescuing patients who have developed drug resistance to other PIs [19]. It too is administered with ritonavir as a booster, in a formulation marketed as prezista.

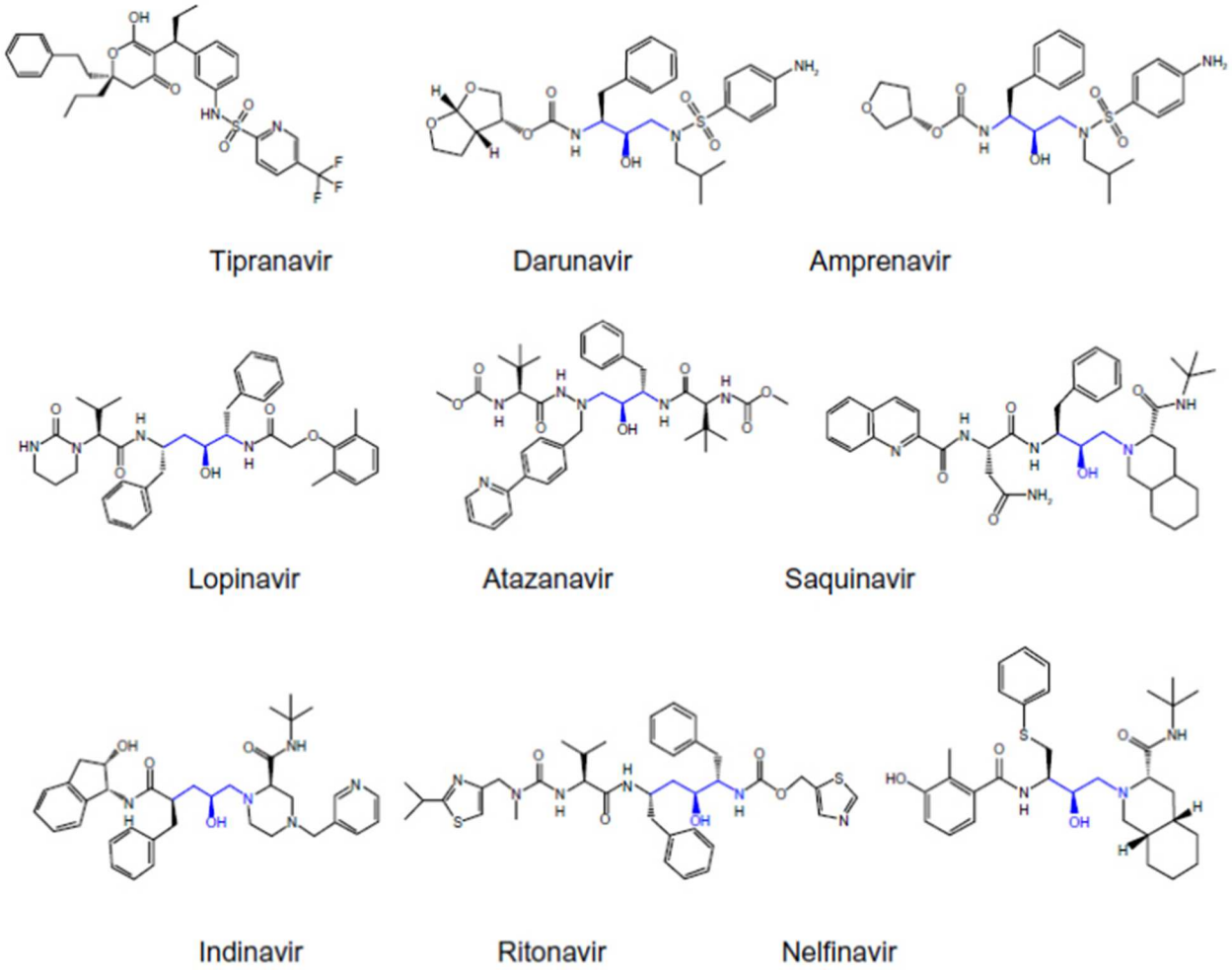


Fig 1.5 FDA approved HIV-1 protease inhibitors. The hydroxyl highlighted in blue interacts with the protease catalytic Asp residues and is conserved across all but one of the HIV PIs [20].

All but one PI contains a hydroxyl attached to the core of the molecule that upon binding to the protease is oriented toward the catalytic Asp residues, forming a hydrogen bond with the carboxylic acid on these residues. Additionally, most PIs form hydrogen bonds via a bridging water molecule with Ile 50 and 50', which helps to tether the flaps in their closed position, increasing contacts with the PI [16]. Crystal structures of protease in complex with all 9 of the FDA approved inhibitors exist. **Fig 1.6**

illustrates the HIV protease – DRV complex at 1.7 Å resolution [21].

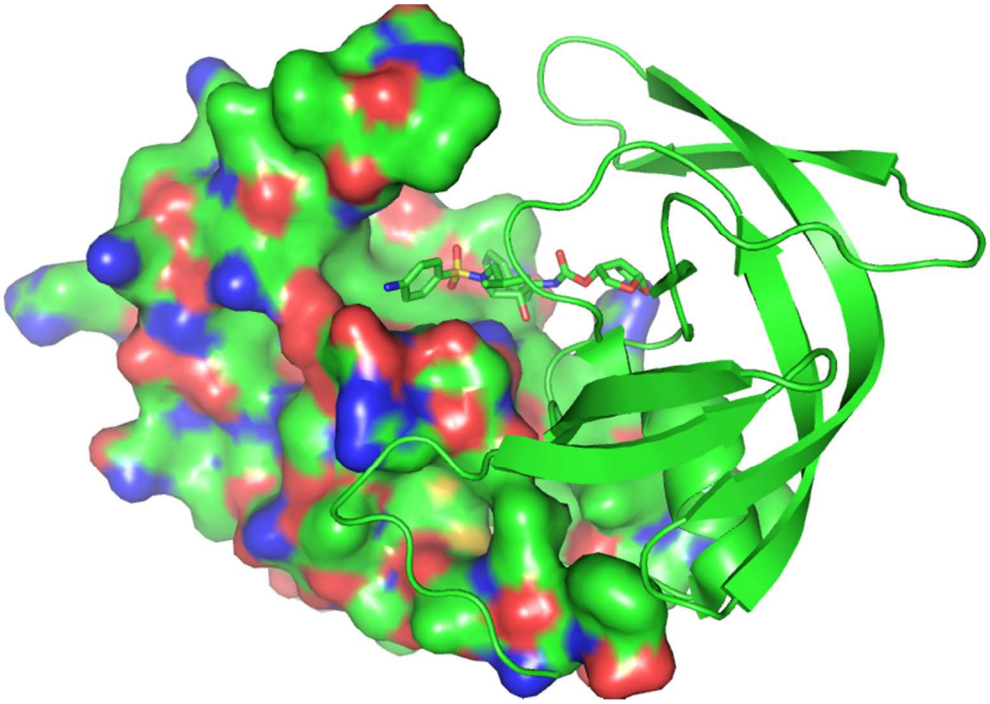


Figure 1.6 HIV-1 protease in complex with Darunavir (pdb: 4YOA). One monomer is shown as a surface filling model to show the fit of DRV in the active site more clearly.

1.5 *HIV Protease inhibitor resistance*

As mentioned above, HIV replication occurs rapidly, and the reverse transcriptase has no proofreading capability and as such is highly prone to error (incorporating an incorrect nucleotide on average once every 1700 nucleotides). This allows various mutations to accumulate in the host HIV population [22]. With HAART/PI treatment, patients can (and often do) develop drug resistance mutations in HIV protease (and other viral enzymes depending on treatment, including Integrase) that reduce the efficacy of treatment over time [23]. This is especially pronounced in patients that are poorly adherent to their treatment regimen. The Stanford HIV Drug Resistance Database reports major mutations appear in treatment experienced HIV protease populations at 13 different residues (30D, 32V, 33L, 46M, 47I, 48G, 50I, 54I, 76L, 82V, 84I, 88N, and 90L) [24]. **Table 1.2** illustrates the wide variety of mutations that can affect treatment, and which PIs they affect. The mutations producing the most marked effect on phenotype are displayed bold in red. All of the noted mutations ultimately reduce the number and strength of contacts between the protease and PIs. As noted earlier, darunavir and tipranavir have a higher barrier

to resistance, and this is evident in the lower number of major mutations reported that interfere with their treatment efficacy.

Major Protease Inhibitor (PI) Resistance Mutations

	30	32	33	46	47	48	50	54	76	82	84	88	90
Cons	D	V	L	M	I	G	I	I	L	V	I	N	L
ATV/r		I	F	IL	V	**VM**	**L**	VTALM		ATFS	**V**	**S**	**M**
DRV/r		I	F		**VA**		**V**	LM	**V**	F	V		
FPV/r		**I**	F	IL	**VA**		**V**	VTA**LM**	**V**	ATS**F**	**V**		**M**
IDV/r		I		IL	V			**VTA**LM	**V**	**AFTS**	**V**	S	**M**
LPV/r		I	F	IL	**VA**	VM	**V**	**VTALM**	**V**	**AFTS**	**V**		M
NFV	**N**		F	IL	V	**VM**		**VTALM**		**AFTS**	**V**	**DS**	**M**
SQV/r						**VM**		VTALM		AT	**V**	S	**M**
TPV/r		I	F	IL	**VA**			**VAM**		**TL**	**V**		

Table 1.2 Stanford HIV Drug Resistance Database reported major resistance mutations. The mutations displayed in bold and red have a more severe effect on treatment efficacy for their respective PI. The "/r" nomenclature indicates concurrent administration with ritonavir as a booster.

Understanding the effects these mutations have on the HIV protease, and thus how to potentially design new and more effective PIs to overcome this drug resistance, is a major focus of ongoing HIV research including the work the forms the bulk of this thesis.

Chapter 2: Materials and Methods

2.1 *HIV-1 protease expression and purification*

MDR769 L33F is based on the previously studied multi-drug resistant variant 769 (MDR769) which contains the mutations Q7K, L10I, M36V, M46L, I54V, I62V, L63P, A71V, V82T, I84V, L90M [25]. This isolate was originally obtained from a patient who was poorly adherent to antiretroviral therapy at the Wayne State University infectious disease clinic in Detroit, MI. MDR769 L33F contains all mutations present in MDR769 as well as the additional mutation L33F. This mutation was initially identified as an accessory (minor) mutation affecting darunavir treatment. Its status was upgraded to a major drug resistance mutation in 2013, concurrent with the release of the work described in chapter 3. The Q7K mutation is not found in patient HIV populations but rather is introduced in order to prevent protease autoproteolysis.

The recombinant MDR769 L33F HIV-1 protease was expressed using a pET-21b T7 promoter expression vector with Escherichia coli BL21 (DE3) as the host (and ampicillin as the selectable marker). A 20 mL liquid culture was inoculated with a single colony of transformed cells, and incubated at 37 °C at 220 rpm for ~2 hrs. These bacteria were then used to inoculate a large 1L culture, which was induced at an absorbance of 0.5 with 1 mM IPTG and allowed to grow for 3 hours after induction. Cells were harvested via centrifugation at 10,000*g* for 5 minutes at 4 °C. The pellet was washed once (50 mM Tris, pH 7.5, 200 mM NaCl, 2 mM β-mercaptoethanol, 2 mM EDTA) and suspended in lysis buffer (50 mM Tris, pH 7.5, 200 mM sodium chloride, and 10 mM β-mercaptoethanol). The bacterial cells were lysed using a one inch diameter French Press operating at 20,000 psi. The HIV protease is fairly insoluble and is deposited into inclusion bodies. These inclusion bodies containing the protease were separated from the soluble fraction by centrifugation at 38,000*g* for 30 minutes (at 4 °C). Inclusion bodies were washed three times with lysis buffer with sequential additives (first, 0.1% Triton X-100; second, 1M NaCl; third, 1M Urea) with a 5 minute centrifugation at 38,000*g* following each wash. Following the third wash, the inclusion bodies were solubilized using lysis buffer with 6M Urea addition. This sample was then allowed to flow over a Q-Sepharose ion exchange column which was equilibrated with the same 6M urea

containing lysis buffer. The protease was contained in the flow-through from the column (**Fig 2.1)**.

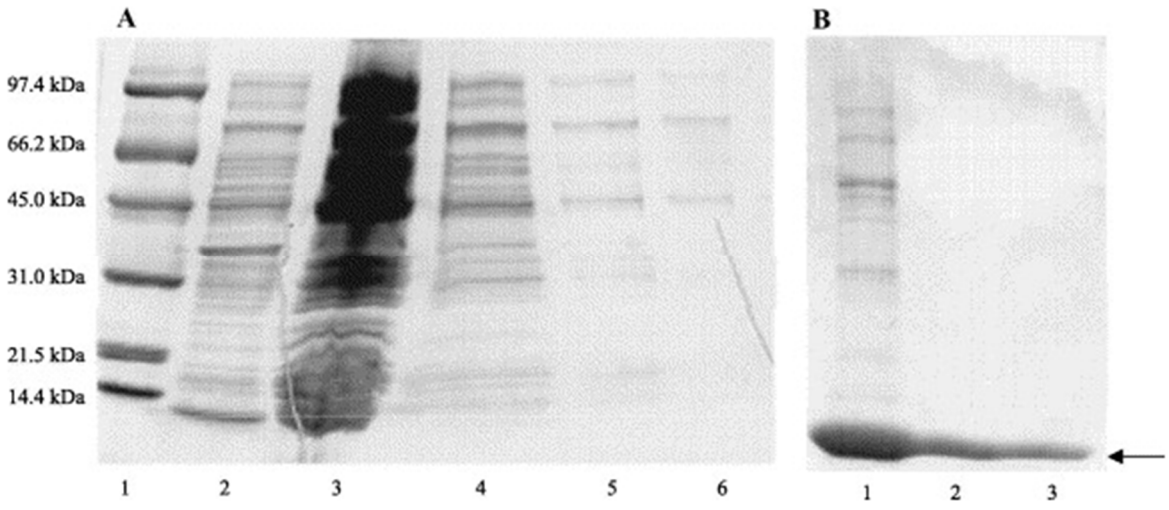


Figure 2.1 Purification of MDR 769 L33F HIV-1 protease. A) Lane 1, marker; lane 2, wash with lysis buffer; lane 3, supernatant; lane 4: wash with lysis buffer containing 0.2% Triton X-100; lane 5, wash with lysis buffer containing 1 M sodium chloride; lane 6, wash with lysis buffer containing 1 M urea. (B) Lane 1, sample of solubilized inclusion bodies applied to Q-Sepharose ion exchange column; lane 2, flowthrough from ion exchange column; and lane 3, wash with buffer A containing 6 M urea. The arrow indicates purified MDR HIV-1 protease (10.8 kDa monomer).

To refold the pure protease, the fractions containing protease were combined and diluted to ~0.2 mg/ml in the 6M urea containing lysis buffer. The dilute protease was dialyzed against 4L of dialysis buffer (pH 7.0, 25 mM NaCl, 20 mM sodium phosphate, 0.2% β-mercaptoethanol, 10% glycerol). The buffer was changed twice using a secondary 4L dialysis buffer omitting the glycerol (pH 7.0, 25 mM NaCl, 20 mM sodium phosphate, 0.2% β-mercaptoethanol), twice using a 4L tertiary dialysis buffer (pH 5.6, 10 mM sodium acetate, 0.2% β-mercaptoethanol), and once using a final 4L dialysis buffer (pH 5.6, 10 mM sodium acetate, 10 mM DTT). protease was concentrated to 5 mg/ml using a Centriprep YM-3 filter concentrator at 3000*g*.

2.2 *HIV-1 Protease crystallization and diffraction data collection*

Apo MDR769 L33F was crystallized using the hanging-drop vapor diffusion method. Two precipitant conditions produced crystals: (2.4 M ammonium sulfate, 0.1 M MES, pH 6.2) and (2.4 M ammonium sulfate, 0.1 M HEPES, pH 6.8). Co-crystallization methods were unable to produce high-

quality crystals. As a result, apo crystals were soaked for 19 h in conditions matching the mother liquor in which they were formed, with the addition of DRV in molar excess (5 mM DRV, 5% DMSO). The crystals were cryoprotected with 30% glucose and were flash frozen in liquid nitrogen. Data were collected at the LS-CAT facility, located within the Advanced Photon Source at Argonne National Laboratory.

2.3 *HIV-1 protease structure refinement and analysis*

The structure of the apo L33F model was determined at a resolution of 1.50 Å. It was phased by molecular replacement using PHASER [26] with PDB entry 1TW7 as the initial search model. Refinement was performed using Phenix [27]. Subsequent structures containing a PI were phased using the apo L33F structure as a search model. The models were built in COOT [28]. After molecular replacement, ligands were added manually into the model after the protein was refined. A round of refinement was performed with PDB-REDO [29] before deposition to the protein data bank (www.wwpdb.org). The final models were analyzed and validated with MolProbity [30]. All images were created using PyMoL [31]. Noncovalent interactions were identified using LigPlot+ [32]. Hydrogen bonds were defined utilizing a donor–acceptor cutoff distance of 3.2 Å; all distances were measured in PyMoL. Crystallographic data are presented in the appendix **Table A1**.

2.4 *Molecular Dynamics*

Molecular Dynamics (MD) is a computational method that allows the evolution of a prepared molecular system to be calculated (approximately, using one of the handful of available molecular force fields written for this purpose) over a given length of time, generally in a nanosecond to microsecond time range. These calculations can often reveal the likely internal motion of proteins, including changes in conformation that may play a critical role in protein function. Simulations in the nanosecond range may reveal atomic fluctuations, side chain motion, and loop movement, whereas those in the microsecond range can allow for identification of helix and whole subunit/domain motions [33]. MD simulations lasting milliseconds (or longer!) currently require very specialized equipment and enormous processing

power, but may be capable of revealing the dynamics behind protein folding and unfolding, as well as the binding and dissociation of protein substrate and ligands [34]. New molecular force fields are continually published and updated, and continue to increase the accuracy of MD output as well. Thus, molecular dynamics will continue to be an important experimental methodology well into the future.

2.4.1 The CHARMM force field

All MD simulations were performed using the CHARMM 36 (Chemistry at HARvard Macromolecular Mechanics) forcefield which was developed in the lab of Dr. Martin Karplus, initially released for public use in the 1980s, and is regularly updated today by the lead developer Dr. Alex MacKerell and his team [35]. All experiments were performed using version c39b1; the latest version is c40b2 and can be downloaded at http://www.charmm.org. CHARMM was one of the first force fields written for MD experiments, and the first to account for atomic partial charges derived from quantum mechanical calculations of the interactions between model atoms and water. CHARMM accounts for torsion (the dihedral angle), bond angle bending and bond stretching, as well as non-bonded interactions. The latter include van der Waals and Coulomb energies [36]. The potential energy and related terms are calculated from the coordinates based on fixed point charges as shown in **Figure 2.2**.

$$U(\vec{R}) = \sum_{bonds} K_b(b-b_0)^2 + \sum_{angles} K_\theta(\theta-\theta_0)^2 + \sum_{Urey-Bradley} K_{UB}(S-S_0)^2 + \sum_{dihedrals} K_\phi(1+\cos(n\phi-\delta)) + \sum_{impropers} K_\omega(\omega-\omega_0)^2 + \sum_{\substack{non-bonded \\ pairs}} \left\{ \varepsilon_{ij}^{\min} \left[\left(\frac{R_{ij}^{\min}}{r_{ij}} \right)^{12} - 2\left(\frac{R_{ij}^{\min}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon r_{ij}} \right\} + \sum_{residues} U_{CMAP}(\phi,\psi)$$

Figure 2.2 Calculation of potential energy in the CHARMM force field [36]

The potential energy, ($U(\vec{R})$), is a summation of terms that represent the non-bonded and internal contributions as a function of the atomic coordinates. Internal terms include valence angle (θ), improper angle (ω), dihedral angle (φ), bond (b), Urey-Bradley (UB, S), and backbone torsional correction (CMAP,

φ,ψ) contributions. The parameters Kb, K φ, KUB, Kθ and Kω are the force constants and all variables with the subscript 0 are their respective equilibrium values. Additionally, n is the multiplicity of the dihedral angle and δ is the phase shift [36].

2.4.2 Molecular Dynamics using NAMD

All MD simulations were performed using Scaling NAno Molecular Dynamics (NAMD) V. 2.9 [37]. The protease models were solvated in a water box utilizing the TIP3P model for all water atoms. The cutoff for non-bonded interactions was 10 Å. The Particle Mesh Ewald (PME) method was used for the calculation of long-range electrostatic interactions [38]. Our system does not possess large fluctuations in density making the PME method more efficient than the fast multipole method used in systems that do. All model systems were initially energy minimized using the conjugate gradient method and heated gradually over 200 picoseconds from a liquid nitrogen temperature of 70K to the physiological temperature of 310K (~98 °C). Simulations were conducted in the NPT ensemble in order to maintain a constant number of atoms, pressure, and temperature. Langevin dynamics were used to keep the system's temperature constant at 310K, and the Nose-Hoover Langevin piston method was utilized to maintain a constant pressure of 1.0 atm. Simulations were performed for 40 ns using the CHARMM force field 36 and a timestep of 2 fs. This simulation length ensures all the conformers of likely biological relevance are sampled during the simulation. A trajectory frame was recorded every 1000 timesteps, producing an output trajectory of 20,000 frames.

MD simulations were performed in parallel on multiple processors using the Wayne State University high performance scientific computing Grid (www.grid.wayne.edu). The WSU Grid currently has the combined processing power of 7,376 cores: 2,480 Intel cores, 4,896 AMD cores, with over 22TB of RAM and 1.2PB of disk space.

2.4.3 Model preparation for MD

The X-ray structures of the WT, MDR, and MDR L33F HIV-1 proteases are solved and were used to determine coordinates for these structures. Homology models for DetMDR1, DetMDR2, and DetMDR3 were created using SWISS-MODEL [39] and utilizing wild type HIV-1 protease as a template.

Based on the known catalytic mechanism of HIV-1 protease, residues Asp 25 and Asp 25' were assigned alternative protonation states (Asp 25 was protonated and Asp 25' was deprotonated). All histidine residues were assigned a neutral charge. Protonation states of other amino acid residues were determined based on their likely state at physiological pH. Crystallographic waters were added to the non-crystallographic MD set ups by merging the water coordinates from the template PDB into the new model. All apo/complexed protease prior to MD were placed into a TIP3P water box and magnesium chloride atoms were added by VMD in order to neutralize any charge on the system. The system was prepared in Visual Molecular Dynamics v1.91 (VMD) [40].

2.4.4 Analysis of molecular dynamics trajectories

All output trajectories of MD experiments were analyzed using VMD. Pymol. Root-mean-square deviation (RMSD) values were calculated using in VMD using the RMSD trajectory wizard. The first frame following energy minimization (frame 21) was used as the top reference frame. The RMSD wizard generated RMSD values for Cα backbone atoms of the protease for each complex. The Timeline wizard in VMD was used to generate RMSD values for each residue side chain utilizing either the first frame after energy minimization or the first frame of the final 10 ns of the simulation as the reference frame. Structure figures were prepared in PyMol.

2.5 Ligand docking and scoring with Swiss Dock

Protein-ligand docking is a computational method that utilizes a molecular mechanics force field (often a CHARMM force field) and a scoring algorithm to predict the binding location and conformation of a ligand to its protein target. This method can additionally provide an estimate of the affinity of binding. Docking is somewhat limited in that the protein target must generally be held rigid (and so large changes in conformation that may occur during ligand binding cannot be predicted or observed), but it is extremely useful as a precursor to more rigorous methods (like molecular dynamics). Experiments can be run with differing degrees of rigor, allowing for high-throughput methods to screen millions of compounds quickly looking for possible molecules to study in more detail, or a more detailed analysis allowing for highly accurate active side binding predictions. HIV-1 PI drug design studies were

performed here using the web based docking program SwissDock to predict the binding mode and affinity of the protease inhibitors.

SwissDock is a protein-ligand docking UI based on EADock DSS that provides rapid docking results and has proven to be highly accurate in its prediction of binding modes [41]. In a test docking of 251 protein-ligand complexes, 57% of the top-scoring ligand binding modes predicted by the software were within at least 2Å of the known crystal structure. This number rose to 70% when considering the top five scoring predictions, and the scoring algorithm is continually be improved upon [42]. Multiple HIV-1 protease structures were used as input for docking experiments, including crystal structures of both wild type (PDB: 2S09) and MDR 769 (PDB: 3SPK). SwissDock prepares input PDB files by converting them to CHARMM format, adding the hydrogen atoms that are likely present at physiological pH (all PDB files are lacking hydrogen atoms), and performing energy minimization on the protein using the steepest decent method. All lopinavir analogs were drawn in ChemDraw, saved as .smi files, and then converted to a three dimensional .mol2 file using ChemBioDraw3D, and uploaded to SwissDock. SwissDock performs energy minimization of all ligands via the Merck Molecular Force Field prior to all docking experiments. Ligands are allowed to remain flexible during docking. Docking performed by SwissDock proceeds over multiple steps. Initially, multiple binding modes for the ligand are generated, and the CHARMM energies associated with these modes are estimated on a grid covering the protein target. Subsequently, the binding modes with the most favorable energies are evaluated with a scoring algorithm and grouped into clusters of similar poses. The predicted binding energies of modes within the most favorable clusters are attached to a pose viewer file, and can be visualized in UCSF Chimera [43]. Rescoring of all docking output was performed manually based on a combination of their predicted ΔG of binding values as well as a visual inspection of the ligand to ensure that it was properly placed in the active site (and for the lopinavir analogs, ensuring that the central hydroxyl is oriented toward the catalytic Asp residues).

Chapter 3: The L33F darunavir resistance mutation acts as a molecular anchor reducing the flexibility of the HIV-1 protease 30s and 80s loops

3.1 *Introduction*

This work was published in the journal Biochemistry and Biophysics Reports in July 2015. The current standard of care for HIV, HAART, often employs a protease inhibitor (PI) containing regimen [44] but mutations in HIV-1 protease (protease) that develop in treatment-experienced patients decrease the efficacy of all current PIs including DRV [44, 45]. Thus, further analysis of key HIV-1 protease resistance mutations is needed to develop more potent antivirals to combat drug resistance.

Clinical isolates previously obtained from the Wayne State University Infectious Disease Clinic in Detroit, MI contain major drug resistance mutations L33F, I47V, I50V, I54M, L76V, V82I/F, and I84F as well as non-polymorphic accessory mutations L10V/G, V11I, I13V, K20T/R, L33I/M, K43T, F53L, A71L, T74P, and L89V. These mutations confer resistance to all FDA approved PIs (http://hivdb.stanford.edu/) [24]. Molecular dynamics simulations with these isolates showed altered protease flap dynamics [46].

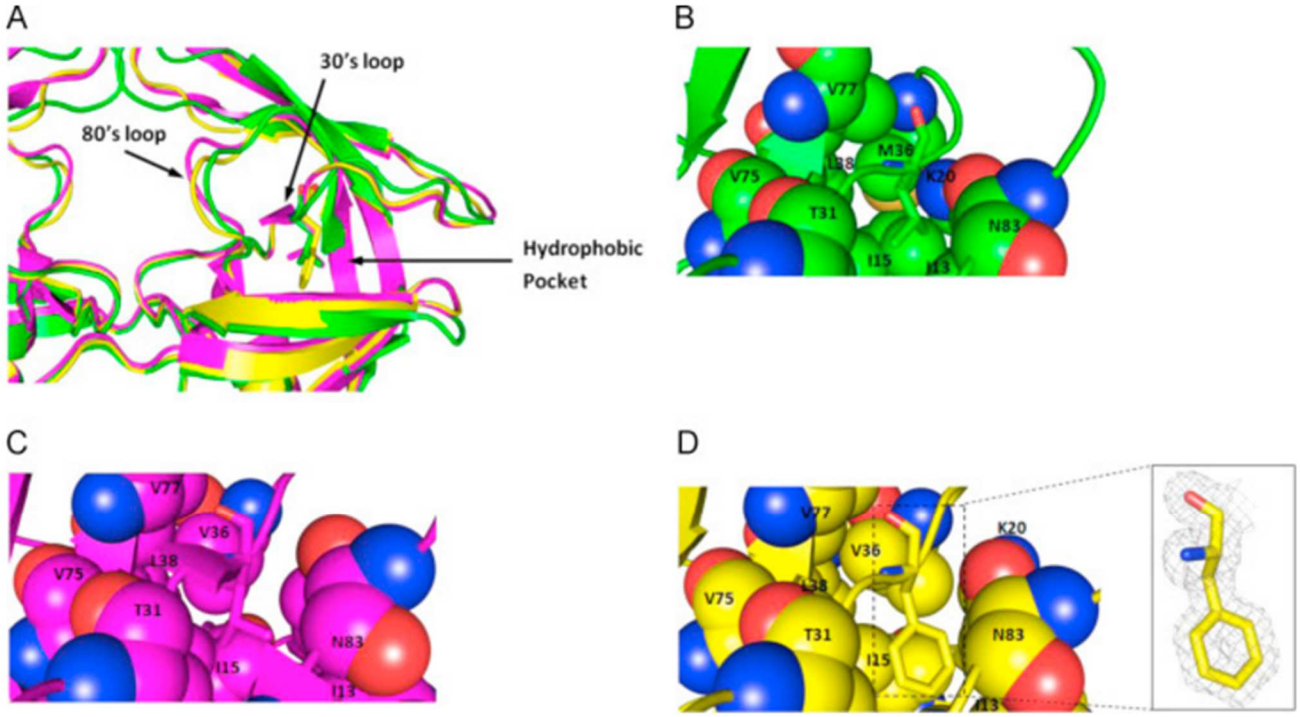
To further investigate the role of the L33F mutation, we created a recombinant MDR769 L33F protease and performed X-ray crystallographic studies. L33F was initially identified as an accessory mutation to I54L/M, V32I+I47V, and I84V/I but is now recognized as a nonpolymorphic major drug resistance mutation [24, 47]. L33F is selected for in patients on a ritonavir pharmacokinetic boosted darunavir (DRV/r) regimen [45], is associated with DRV/r resistance [48], has greatly increased in prevalence since the year 2000, and has direct influence on inhibitor-interacting residues [49]. We hypothesize that reduced flexibility of the 30s and 80s loops due to molecular anchoring properties of L33F may contribute to drug resistance.

3.2 *Results*

3.2.1 Structural features of the residue 33 environment

The side chain of L33F extends 2.2 Å deeper into the hydrophobic pocket compared to L33 in WT protease (**Fig. 1**)

Figure 3.1 Structural features of the HIV-1 protease residue 33 molecular anchor. (A) Superposition of the WT protease (green), MDR769 protease (magenta), and MDR769 L33F protease (yellow) apo structures. The 30s loop, which contains residue 33, is positioned between the 80s loop and the hydrophobic pocket. In (B), (C) and (D) WT, MDR769, and MDR769 L33F are shown, respectively. L33F fills the hydrophobic pocket more completely than L33. The inset in panel (D) shows the 2Fo– Fc map for MDR769 L33F contoured at 1σ



leading to increased hydrophobic interactions between L33F and the hydrophobic pocket. The hydrophobic pocket is defined by residues I13, I15, K20, A22, T31, M/V36, L38, I64, I66, V75, V77, N83, and I85 (**Fig. 3.1B–D**). To visually identify changes in these residues, we aligned and superimposed the WT, MDR769, and MDR769 L33F structures. Although conformational and positional changes in these residues are seen between the WT and MDR769 structures (**Fig. 3.1B and C**), the L33F mutation produces further alterations in many of these residues (**Fig. 3.1D**). The most notable change is in residue I13, which rotates to avoid steric clashes with L33F. Other significant changes due to the L33F mutation

are noted in residues I15, K20, A22, V36, L38, I66, and N83. These changes lead to increased hydrophobic interactions in L33F compared to the WT and MDR769 structures (**Table 3.1**).

The MDR769 L33F structures show increased rigidity compared to the WT and MDR769 structures. Upon drug binding, L33 in the WT–DRV complex shifts 1.1 Å towards the active site and a hydrogen bond is formed between the backbone amide nitrogen of L33 and the backbone carbonyl of

Structure name	WT apo	WT:DRV complex	MDR769 apo	MDR769:DRV	MDR769 33F apo	MDR769 33F:DRV complex
PDB code	3PHV	4LL3	1TW7	3SO9	4YOB	4YOA
Residue 33L/F non-covalent interactions	30	27	24	23	32	31
DRV:non-covalent interactions	N/A	63	N/A	53	N/A	34

Table 3.1 The impact of L33F on interactions with the hydrophobic pocket and darunavir. Top row: non-covalent interactions between residue 33 (L/F) and hydrophobic pocket residues. Bottom row: non-covalent interactions between DRV and active site residues. Interactions involving residue 33L/F were identified using the DimPlot script in LigPlot+; interactions between protease and DRV in the complexed structures were determined using LigPlot+. Mutations present in MDR769 reduce the number of non-covalent interactions with the hydrophobic pocket. Substitution of L33F restores the interactions between residue 33 and the hydrophobic pocket and extends them beyond what is observed in the WT structures. Interactions between DRV and the active site are reduced by mutations present in MDR769 but are further reduced due to the molecular anchoring properties of the non-polymorphic L33F mutation.

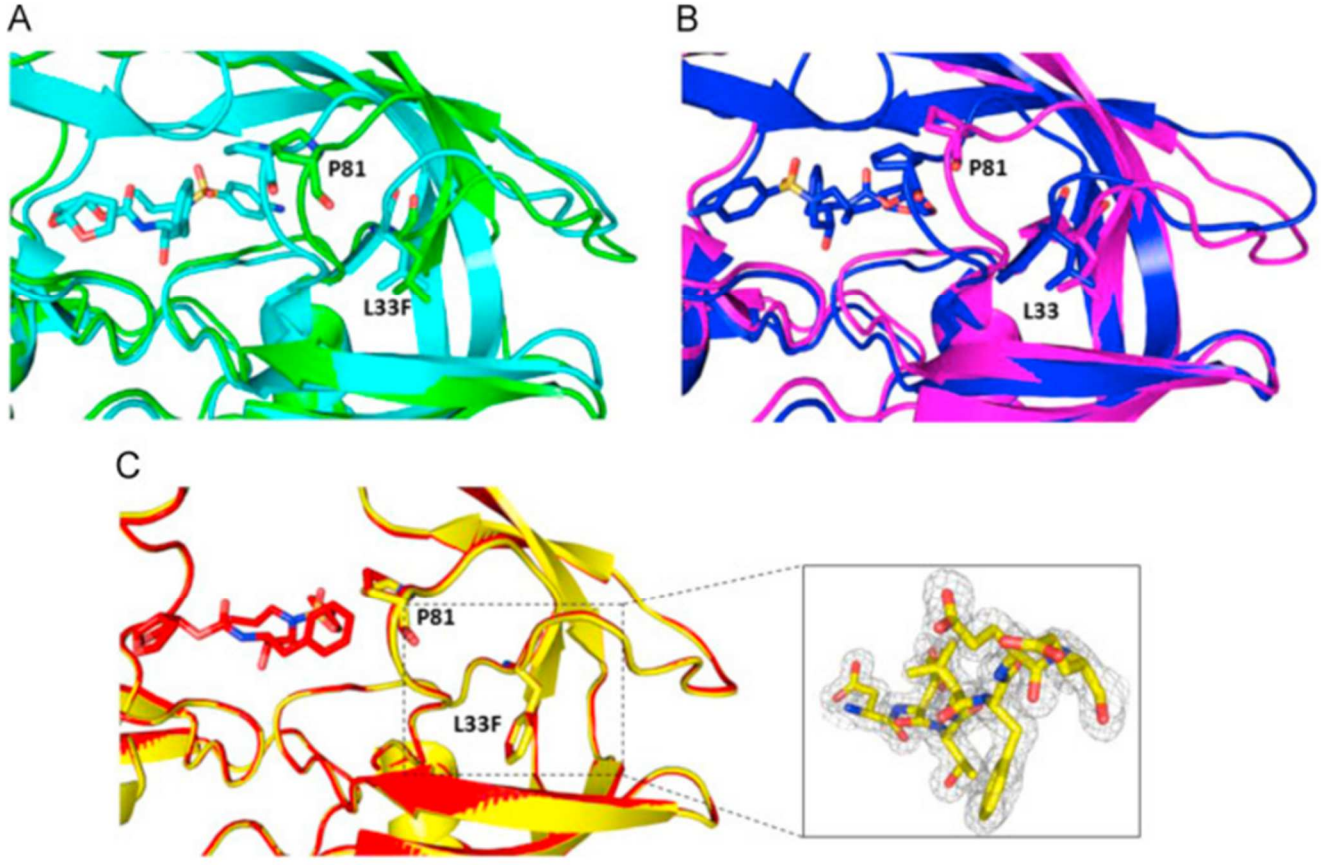
G78. Additionally, the side chains of residues I13, I15, M36, I66, and V75 are rotated, residues I15, K20, M36, L38, and V77 shift 0.4 Å, 1.2 Å, 2.3 Å, and 2.1 Å, respectively, into the hydrophobic pocket, and residues T31, V75, N83, and I85 shift 0.8 Å, 0.7 Å, 1.0 Å, and 0.5 Å away from the pocket. Similar changes occur in the MDR769 structures upon drug binding. However, with the L33F mutation, minimal changes in conformation or position occur in either L33F or residues of the hydrophobic pocket upon drug binding in the MDR769 L33F structure.

3.2.2 L33F as a Molecular Anchor

When the 30s loop (residues 29–35) of protease bears the L33F mutation, flexibility of both the 30s and 80s loops (residues 79–84) is decreased, likely through increased hydrophobic interactions. The 80s loop influences the S1/S1′ binding site [50], and the 30s loop lies between the 80s loop and the

hydrophobic pocket (**Fig. 3.1A**). In the WT–DRV complex, drug binding causes the 30s loop residues to shift, on average, 1.5 Å towards the active site compared to the WT apo structure (**Fig. 3.2A**). In a similar fashion, the adjacent 80s loop residues also shift, on average, 2.0 Å towards the active site (**Fig. 3.2A**). Furthermore, residue 33L in the WT–DRV complex displays fewer noncovalent interactions compared to the WT apo structure (**Table 3.1**).

Figure 3.2 L33F acts as a molecular anchor that restricts movement of the 30s and 80s loops. (A) Superposition of WT apo protease (green) and WT–DRV complex (cyan). (B) Superposition of MDR769 apo (magenta) and MDR769–DRV complex (blue). The 30s loop and 80s loop in (A) and (B) both shift with DRV bound. (C) Superposition of MDR769 33F apo (yellow) and MDR769 L33F–DRV complex (red). L33F prevents movement of the 30s and 80s loops towards the active site as in the WT and MDR769 structures. The inset in (C) shows the 2Fo– Fc map for the 30s loop residues (D29–E35) contoured at 1σ.



Similar to the WT–DRV complex, the 30s and 80s loops of the MDR769–DRV complex shift 1.6 Å and 2.4 Å, respectively, towards the active site compared to the apo structure (**Fig. 3.2B**). Also, the number of noncovalent interactions is severely reduced in the MDR769 structures compared to the WT structures (**Table 3.1**). However, the 30s and 80s loops of the MDR769 L33F–DRV complex show minimal shifting

upon drug binding; the only significant change is in residue P81, which shifts 0.4 Å into to the active site and puckers up when DRV is bound in the active site (**Fig. 3.2C**). In addition to the decreased flexibility of the 30s and 80s loops in the MDR769 L33F structures, the L33F mutation restores the noncovalent interactions with the hydrophobic pocket that were originally lost in the MDR769 complex. These restored interactions are also more extensively maintained than in the WT–DRV complex (**Table 3.1**). The decreased flexibility of both the 30s and 80s loops is likely due to enhanced anchoring by L33F via increased hydrophobic interactions within the hydrophobic pocket.

To further assess the hypothesis of L33F acting as a molecular anchor, 40 ns MD simulations were performed. Differences in protein flexibility due to molecular anchoring of L33F should be more pronounced in apo protease compared to PI-complexed forms. Therefore, the RMSD of the 30s and 80s loops for both chains of apo WT, MDR769, and MDR769 L33F were analyzed over the last 10 ns of the MD trajectory. For WT protease, the average RMSD values of the 30s and 80s loops were 1.53 Å and 1.65 Å, respectively. The 30s loop of the MDR769 structure showed reduced flexibility compared to the WT structure with an RMSD of 1.34 Å whereas the 80s loop of the MDR769 showed similar flexibility compared to the WT (1.66 Å). Flexibility of the 30s and 80s loops in MDR769 L33F were further reduced compared to both WT and MDR769 structures with the 30s loop displaying an RMSD of 1.27 Å and the 80s loop displaying an RMSD of 1.55 Å.

3.2.3 Reduced flap interactions and altered drug conformation

Previous work documented interactions between the 80s loop and the flap tips [51], their importance in substrate recognition and binding [52], and their influence on forming the S1/S1′ subsite [50]. Thus, effects of the L33F mutation on the 80s loop and flap tips are possibly implicated in resistance development.

In the WT protease, the flexibility of the 80s loop and flaps produces favorable interactions allowing for proper formation of the S1/S1′ subsite. In the WT apo structure, a large 7.5 Å gap between the P81 of the 80s loop and I50 of the flap tips exists, preventing any interactions between the two (**Fig. 3.3A**). However, with DRV bound in the active site, the flaps close and the 80s loop shifts 2.0 Å towards

the active site. This brings residue P81 within 3.7 Å of G49′ and T80 within 3.9 Å of I50′ (**Fig. 3.3A**). These distances allow for favorable interactions between the flap and the 80s loop resulting in a closed active

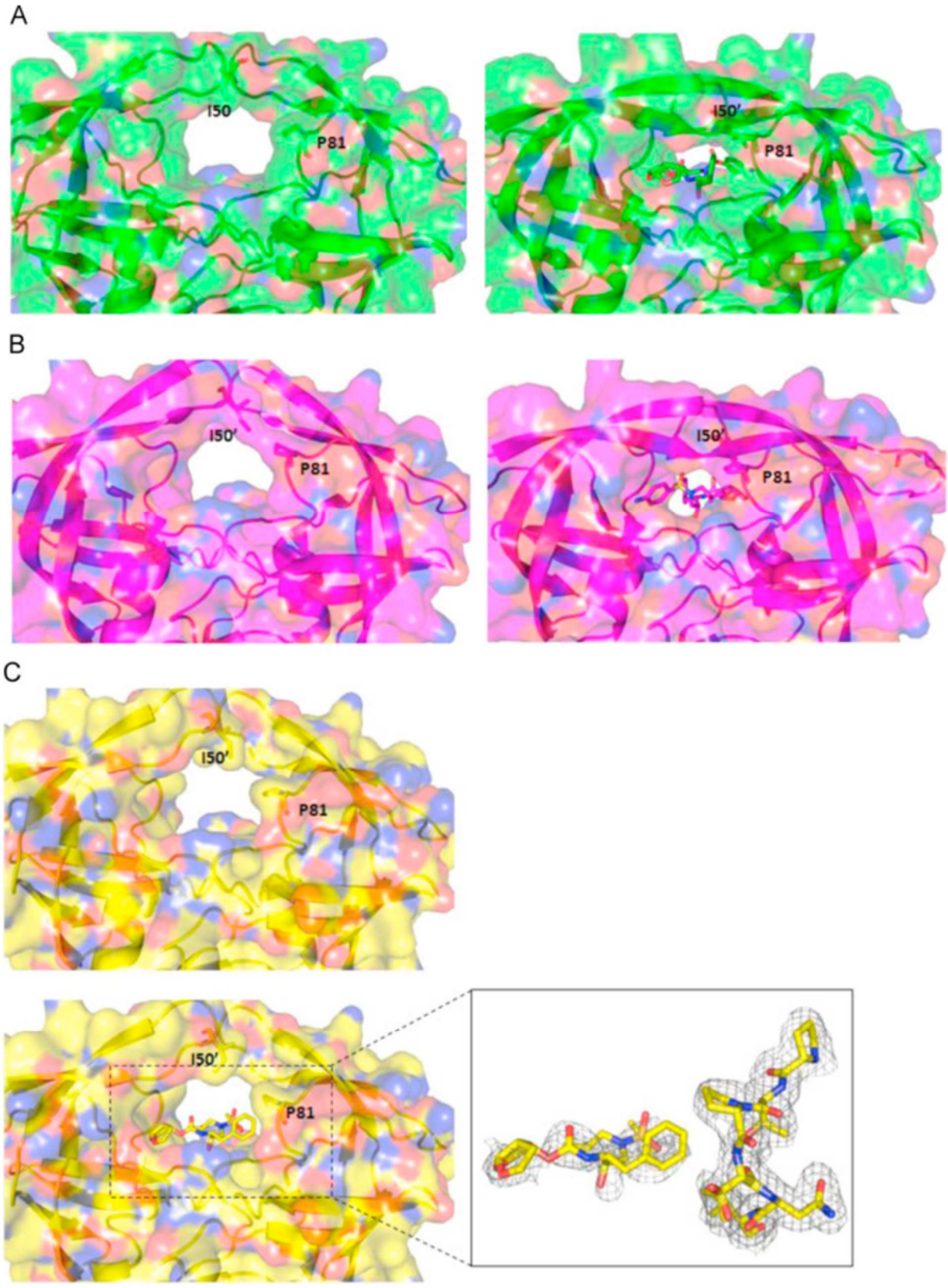

A
I50
P81
I50'
P81
B
I50'
P81
I50'
P81
C
I50'
P81
I50'
P81

Figure 3.3 Reduced flap interactions due to L33F anchoring. (A) WT apo protease and WT protease complexed with DRV. The apo WT shows a 7.5 Å gap between P81 of the 80s loop and I50 of the flap producing an active site in the open conformation. When complexed with DRV, the flaps and active site close. (B) MDR769 apo protease and MDR769–DRV complex. Despite interactions between the 80s loop and flaps in the MDR769 protease, the flaps and active site display an open conformation. With DRV in the active site, the flexibility of MDR769 protease produces a closed active site. (C) MDR769 L33F apo protease and MDR769 L33F complexed with DRV bound resembles the WT apo structure in (A) with a large gap between residues I50′ and P81. The gap between I50′ and P81 is maintained between the apo and complexed structures (5.8 Å and 5.9 Å, respectively) and is too large for interactions to occur. The active site and S1/S1′ subsite remain open. The inset in (C) shows the 2Fo– Fc map for DRV and the 80s loop in the L33F structure contoured at 0.5σ in order that the density around DRV is shown.

site, and more specifically, a properly formed S1/S1′ subsite (**Fig. 3.3A**). As a result, DRV makes extensive noncovalent interactions with the residues lining the active site as indicated in **Table 3.1**.

Similar to the WT protease, the MDR769 protease displays considerable flexibility in the 30s and 80s loops (**Fig. 3.3B**). The gap between residues P81 and I50′ in the MDR apo structure is not large compared to the WT (3.6 Å compared to 7.5 Å). The flaps are in a "wide-open" conformation as reported previously [25], which leaves the active site open. In the MDR769–DRV complex, the flexibility of the 30s and 80s loops (**Fig. 3.2B**) allows the active site to close (**Fig. 3.3B**). Even though the active site is closed, the number of noncovalent interactions between DRV and the active site is decreased, suggesting a slightly distorted active site (**Table 3.1**).

In the MDR769 L33F apo structure, the gap between the 80s loop and the flaps resembles the WT apo structure, producing an open active site and an open S1/S1′ subsite. The apo structure contains a 5.8 Å gap between P81 and I50′; this gap is smaller than the WT but still too large for any significant interactions to occur. In the L33F–DRV complex, P81 shifts 0.4 Å into the active site, but I50′ also shifts and rotates leaving a 5.9 Å gap which resembles the WT apo structure (**Fig. 3.3C**). The 5.9 Å gap between P81 and I50′ leaves an open S1/S1′ subsite and also results in an open active site. The result is a severely reduced number of interactions between DRV and the active site compared to both the WT and MDR769 structures (**Table 3.1**).

Additionally, in the L33F–DRV complex, the side chains of P2, P1, and P1′ of DRV are rotated to compensate for the open S1/S1′ subsite (**Fig. 3.3C**) which alters the hydrogen-bonding network compared

to the WT–DRV complex. Previous reports have indicated a conserved hydrogen bonding network between DRV and backbone and side-chain atoms of residues D25, G27, D29, D30, D25′, and D30′ [50]. In the WT–DRV complex, the P2 bis-THF moiety, hydroxyl, and P2′ amine of DRV are responsible for formation of five hydrogen bonds with residues D25, D29, D30, D25′, and D30′. However, the L33F–DRV complex contains an expanded active site and S1/S1′ subsite which alters the conformation of P2, P1, P1′, and P2′ of DRV. As a result, the P2 bis-THF moiety, hydroxyl, and P2′ amine of DRV form only three hydrogen bonds with residues D25N, D30, and D29′. Additionally, the number of contacts decreased between the P1 group of DRV and the residues of the S1/S1′ subsite (**Table 3.1**).

3.3 *Discussion*

The L33F mutation is selected in patients receiving a DRV/r regimen [45], and is associated with reduced response to DRV/r treatment [24, 48] as it has direct influence on the inhibitor-interacting residues [49]. This work describes the effects of L33F on the structure of HIV protease as well as the effect it has on inhibitor recognition.

Superposition of MDR769 L33F with WT and MDR769 HIV-1 protease with and without DRV in the active site reveals altered conformation of the 30s and 80s loops. The larger side chain of L33F embeds further into the hydrophobic pocket than L33, the latter of which is present in both WT and MDR769 structures (**Fig. 3.1**). As a result, noncovalent interactions are increased in this region compared to the WT, causing the L33F to act as a molecular anchor. MD simulations showed a clear reduction in flexibility of both the 30s and 80s loops for MDR769 L33F compared to both WT and MDR769. The results of these simulations support the hypothesis that L33F may play a role as a molecular anchor within HIV-1 protease.

Enhanced molecular anchoring by the L33F mutation reduces the flexibility of the 30s and 80s loops, thereby inhibiting proper formation of the S1/S1′ subsite and keeping the active site in an open conformation in the MDR769 L33F–DRV complex (**Fig. 3.2** & **3.3**). Conversely, in the WT and MDR769 complex structures, DRV binding causes a shift in the 30s and 80s loops leading to formation of the S1/S1′ subsite and closing of the active site (**Fig. 3.2** & **3.3**).

The structure of the L33F DRV complex reported here shows the protease flaps in an open conformation. To date, all deposited structures of HIV protease L33F complexes have been solved in three different space groups: P212121, P61, and P41212. The majority of these HIV protease L33F DRV complexes have been solved with the flaps in a closed conformation. Other research groups such as the Schiffer group [53] and the Konvalinka group [54] have previously reported L33F DRV complexes with the flaps in a closed conformation by solving the structures in P212121 (PDB: 4QY1) and P61 (PDB: 3GGU), respectively. In contrast, the Weber group [55] and this report describe the L33F DRV complex as a structure with open protease flaps when the structures were solved in P41212. The Weber group [55] (PDB: 4NPT) also utilized the inactivating mutation D25N to facilitate HIV protease expression, purification, and crystallization.

With specific regard to the L33F mutation, the Schiffer group (closed protease flaps) reports L33F may play a role in active site expansion [53]. The Konvalinka group (closed protease flaps) suggests that L33F is possibly implicated in structural changes in the flap and flap hinge regions of protease. Despite structural and symmetry differences between the aforementioned and this report, our results regarding L33F as a molecular anchor are consistent with the previous reports by the the Schiffer and Konvalinka groups. A key finding of our report is the influence of the L33F mutation on the open active site and on the S1/S1′ subsites through anchoring of the 30s and 80s loops independent of space group.

In conclusion, for the first time we here report the molecular mechanisms by which the non-polymorphic protease mutation L33F contributes to DRV resistance. The L33F mutation may contribute to resistance via two mechanisms: one, by restoring noncovalent interactions lost due to other primary mutations, and two, by further reducing interactions between DRV and active site residues. These findings may contribute to our overall understanding of drug resistance as well as future drug design strategies. We propose that modifications to the P1/P1′ groups of existing PIs to fill the open S1/S1′ subsite might result in a greater response by patients who harbor the L33F mutation in HIV-protease.

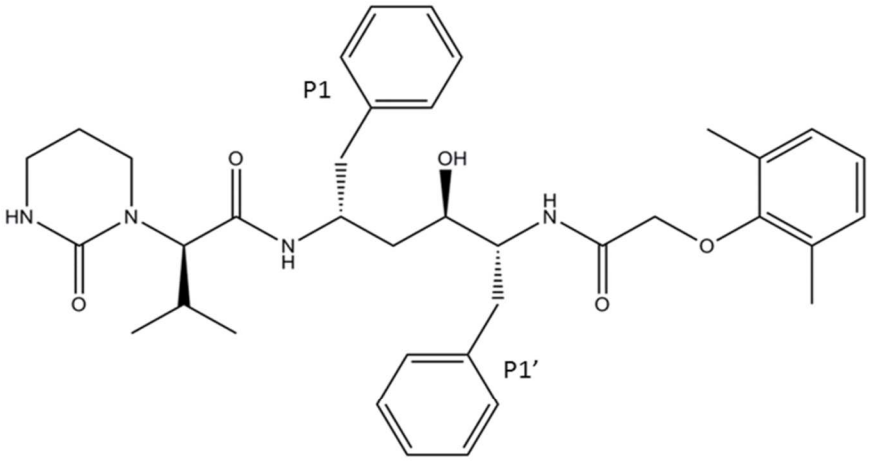
3.4 *Author's contribution*

The author shared first author credits with his lab mate Ben Kuiper on this paper, and they both performed work in equal measure to produce this paper. The author performed protein expression and purification aided by former lab member Dr. Tamaria Dewdney, and the author trained Mr. Kuiper in the use of protein crystallography software for the phasing and refinement of reflection data. The author also trained Mr. Kuiper on the use of VMD/NAMD and both performed and analyzed the molecular dynamics experiments.

Chapter 4: Structure based design of modified lopinavir analogs targeting multi-drug resistant HIV-1 Protease

4.1 *Introduction*

HIV protease is a 99 amino acid aspartyl protease responsible for processing the viral polyprotein, resulting in mature enzymes and structural proteins. Proteolysis of the Gag and Gag/Pol polyproteins by HIV protease is an essential step in the replication and assembly of HIV virions thus making protease an important target for anti-retroviral therapy [13]. Peptidomimetic design based on HIV-1 protease substrates is a popular strategy for the design of HIV-1 protease inhibitors, and has been successful in the past [56]. However, the emergence of drug-resistant HIV variants with cross-resistance to protease inhibitors remains a major drug design challenge. To examine the effect of P1/P1' size diversity and composition on binding affinity to multi-drug resistant (MDR) HIV-1 protease, we produced and examined a series of 36 lopinavir analogs with modifications at the P1/P1' sites designed to increase contacts with the protease. An MDR HIV-1 protease isolate (MDR 769) was obtained from a patient at the Wayne State University infectious disease clinic who was poorly adherent to antiretroviral HAART therapy, and had developed multi-PI drug resistance. This isolate was previously crystallized, and with the set of MDR mutations present the protease adopts a wide-open flap conformation in the absence of ligand [25]. The asymmetric expansion of the MDR protease active site cavity is likely to play a role in the reduced efficacy of PIs against MDR variants [49]. MDR protease also has an expanded active site cavity with an enlarged substrate envelope that reduces the efficacy of FDA approved protease inhibitors including lopinavir, a common first line treatment option [25, 52]. Docking and molecular dynamics experiments were performed utilizing our 36 lopinavir analogs (**Table 4.1**) in order to predict their binding affinity against our MDR protease isolate.



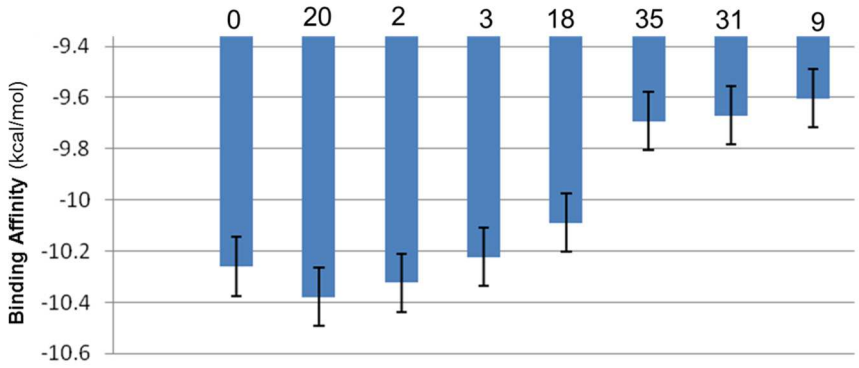
Position		Modification
P1	P1/P1'	
1	19	p-methyl benzyl
2	20	p-flourobenzyl
3	21	isobutyl
4	22	m-dimethyl benzyl
5	23	m-diethyl benzyl
6	24	o-methyl benzyl
7	25	m-isopropyl benzyl
8	26	o-isopropyl benzyl
9	27	m-ethyl benzyl
10	28	o-dimethyl benzyl
11	29	m-di-isopropyl benzyl
12	30	p-ethyl benzyl
13	31	m-methyl benzyl
14	32	o-ethyl benzyl
15	33	napthalene
16	34	tryptophan
17	35	cyclohexyl (chair)
18	36	p-methylflourine benzyl

Table 4.1 Modifications made to LPV to increase binding affinity to HIV-1 protease. The top panel displays LPV, and identifies the P1 and P1' moieties. These were modified via addition to the aromatic ring(s), or replacement of the aromatic ring(s) as indicated in the table.

4.2 *Results*

4.2.1 Fluorinated lopinavir has the highest predicted binding affinity against drug resistant HIV-1 protease

Crystal structures of both wild type (PDB: 2O4S) and multi-drug resistant HIV-1 protease (PDB: 3D20 and 3SPK) were obtained from the Protein Data Bank. Modifications were made to the structure of Lopinavir (LPV) at the P1 and P1' positions using ChemDraw Bio Ultra software [57]. The ligands and target protein were then submitted to the web based docking software SwissDock [41]. The modified LPV analogs were ranked based on their predicted binding affinities and active site conformations, and those with the highest binding affinity were submitted to all atom molecular dynamics simulations using NAMD v2.9. **Fig 4.1** displays the compounds with the highest predicted binding affinity, those selected for MD experiments were (2) P-fluorobenzyl at P1, (3) isobutyl at P1, and (20) P-fluorobenzyl at P1/P1'.



LPV Variant	204S (WT)	3D20 (I54V)	3SPK (MDR 769)	Binding Affinity (kcal/mol)
(0) Control	-10.45	-10.69	-9.64	-10.3 ± 0.6
(20) P-flourobenzyl (P1/P1')	-10.3	-10.75	-10.09	-10.4 ± 0.3
(2) P-flourobenzyl (P1)	-10.52	-10.45	-10	-10.3 ± 0.3
(3) isobutyl (P1)	-10.3	-10.62	-9.75	-10.2 ± 0.4
(18) P-methylflourine benzyl (P1)	-10.19	-9.7	-10.38	-10.1 ±0.4
(35) cyclohexyl (chair) (P1/P1')	-9.19	-10.51	-9.38	-9.7 ± 0.7
(31) M-methyl benzyl (P1/P1')	-9.46	-10.08	-9.47	-9.7 ± 0.4
(9) M-ethyl benzyl (P1)	-8.92	-10.18	-9.71	-9.6 ± 0.6

Figure 4.1 Predicted binding affinity of LPV analogs with an LPV control. The top 3 compounds (2, 3, 20) along with the control were utilized in MD experiments.

4.2.2 P1/P1' fluorinated lopinavir stabilizes the HIV-1 protease

Using 10 ns molecular dynamics simulations, we monitored the RMSD of the protease backbone. After energy minimization using the conjugate gradient method, MD was performed for 10 nanoseconds using the CHARMM36 force field in the NPT ensemble accounting for a constant number of atoms, temperature (300K), and pressure (1.0 atm) with a 2 fs timestep. RMSD analysis of the protease Cα backbone atoms in the resultant MD trajectories showed that while the average RMSD increased for the MDR variants (as expected), the top performing lopinavir analogs reduced the average RMSD in the MDR protease variants. This reduction in RMSD was the strongest in compound (20), P1/P1' p-fluorobenzyl lopinavir (**Fig 4.2**).

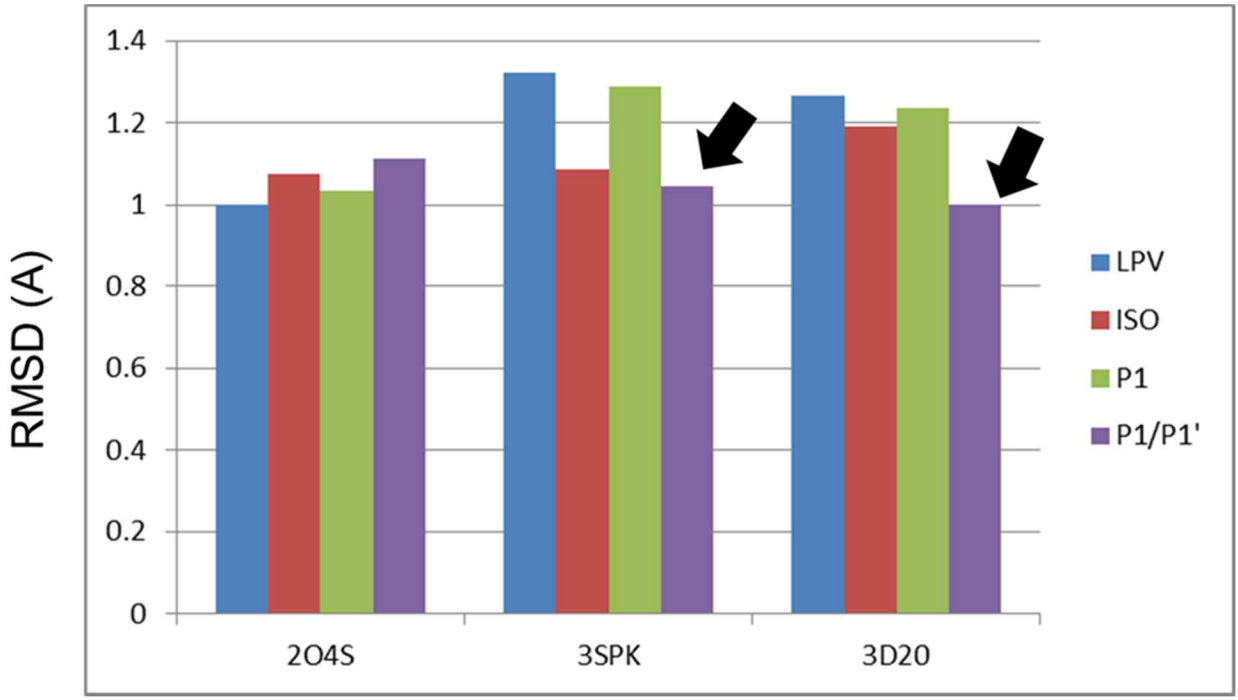


Figure 4.2 Average RMSD of wild type (2O4S) and DR/MDR protease when bound to lopinavir and lopinavir analogs. LPV: control lopinavir. ISO: compound (3), P1 isobutyl. P1: compound (2) P1 p-fluorobenzyl. P1/P1': compound (20), P1/P1' p-fluorobenzyl.

4.2.3 P1/P1' fluorinated lopinavir increases non bonded interactions with HIV-1 protease

The final frame from each MD trajectory was analyzed in LigPlot+ [32]. Non-bonded interactions formed and lost with the protease by each of the top performing lopinavir analogs were compared to those interactions present in each of the protease-LPV complexes. **Table 4.2** displays the list of these contacts

for the best performing lopinavir analog, compound (20): P1/P1' p-fluorobenzyl lopinavir. In both the wild type, and I54V drug resistant protease complexes, this compound gained only slightly more contacts than it lost (in comparison to the contacts lopinavir formed with these protease). But in the MDR 769 protease complex, compound (20) gained many more contacts than those formed with the lopinavir control.

	P1/P1' Flouro			P1/P1' Flouro			P1/P1' Flouro	
	Contacts Gained	Contacts Lost		Contacts Gained	Contacts Lost		Contacts Gained	Contacts Lost
2O4S	D29 D30 D30' I84' L23 R8 V82' *D29	*D29 G48' I50' P29 P30 P81	3SPK	A28' D25 D30' G48' G49' I47 T82 *D29	*G49	3D20	D30' G49' I47 I47' P81 V82 *D30 *I50 *R8 *R8'	D29' D30 G48 G48' R8 R8' L23'

Table 4.2 Non-bonded contacts gained and lost by compound (20) compared to the lopinavir control. 2O4S: wild type protease. 3D20: I54V drug resistant protease. 3SPK: MDR 769 protease.

4.3 *Discussion*

These results suggest that the presence of fluorinated P1 or P1' groups, particularly di-fluorobenzyl at P1/P1', produces greater binding affinity by facilitating an increase in the number of contacts between the ligand and the protease target. Analysis of MD simulation data further suggests that these groups play a role in stabilizing the protein backbone and therefore may increase overall efficacy against MDR protease variants. Interestingly, the fluorinated lopinavir analogs actually caused a slight increase in RMSD when bound to the wild type protease. This suggests that lopinavir is highly optimized for binding to wild type protease, and the extra atoms may cause binding to worsen in the analogs. However, it has been shown that the drug resistance mutations commonly acquired and that are present in the MDR form of protease, cause an expansion of the active site cavity [25]. The additional non-bonded interactions identified in the lopinavir analogs likely result from the ability for these analogs to better fill this increased space. Thus, these analogs may represent a class of PI that could be used as rescue treatment for

patients who have accumulated multiple drug resistance mutations. More research into these compounds is needed.

4.4 *Author's Contribution*

The author performed all work described in this section himself. The project was a continuation of work begun by Dr. Tamaria Dewdney, and was overseen by her.

Chapter 5: The role of mutations at codons 32, 47, 54, and 90 in HIV-1 protease flap dynamics

5.1 *Introduction*

This work was published in the journal Discoveries in December 2014. The high viral replication rate and error-prone reverse transcriptase leads to the emergence of drug resistant human immunodeficiency virus variants under the selective pressure of anti-retroviral therapy [58, 59]. The current standard treatment of HIV utilizes HAART where the PIs represent a key class of drugs [44]. PIs have a higher barrier of resistance relative to other classes of inhibitors, as multiple protease mutations are needed for a patient to develop resistance [60]. Although there are nine FDA approved PIs, some first generation inhibitors (e.g. indinavir, saquinavir, and nelfinavir) are considered obsolete by clinicians due to emergence of drug resistance or from their side effects [60]. Second generation PIs, especially darunavir and tipranavir, remain potent against mutant viral strains that are resistant to first generation PIs [61]. With the availability of second generation PIs, treatment of patients with MDR HIV is possible and can result in full viral suppression [61]. Many factors, including mutations in the viral polyprotein (particularly in Gag and Env) [62], the development of cross-resistance [63], and the high mutability of the HIV-protease [64] continue to produce viral strains that are resistant to second generation PIs [45]. Therefore, novel and potent new drugs are required for the treatment of multidrug resistant HIV infection.

This study utilized molecular dynamics simulation techniques in order to better understand the molecular mechanisms by which MDR-PRs are able to evade inhibition by potent PIs. Clinical MDR-protease isolates were obtained from three patients at the Wayne State University Infectious Disease Clinic in Detroit, MI. On a boosted PI regimen, these patients had a low CD4 cell count and high HIV viral load, indicating treatment failure.

Genotypic and virtual phenotypic analysis of the protease genes showed multiple major and minor drug resistance mutations which imparted resistance to the nine FDA approved PIs. The DetMDRs differ from isolates previously studied by our group in that these contain the major drug resistance mutations L33F, I47V, I50V, I54M, L76V, V82I/F, and I84F not present in the previous isolate (**Table**

5.1).

Table 1. Sequence alignment of WT, MDR-769, DetMDR1-3

NL4-3	PQITLWKRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
MDR-769	PQITLW**K**RP**I** VTIKIGGQLK EALLDTGADD TVLEE**V**NLPG RWKPK**L**IGGI GGF**V**KVRQYD Q**VP**IEICGHK **V**IGTVLVGPT P**T**N**V**IGRNL**M** TQIGCTLNF
DetMDR1	PQITLWKRP**V** **I**T**V**KI**A**GQLK EALLDTGADD T**IF**EEMNLPG RW**T**PK**IV**GGI GGF**M**KVRQYD QI**P**IEICGHK **LV**G**P**VLVGPT P**T**N**V**IGRN**MM** TQ**L**GCTLNF
DetMDR2	PQITLWKRP**V** VT**V**K**V**GGQL**F** EALLDTGADD TV**F**EE**I**NLPG RWKPK**IV**GGI GGF**V**KVRQYD QILIEICG**K**K **I**I**S**TVLVGPT PVN**V**IGRN**T**L TQ**M**GCTLNF
DetMDR3	PQITLWKRP**F** VT**V**KIGGQL**I** EALLDTGADD T**IF**EEMNLPG RWKPK**I**IGGI GGF**L**KVRQYD QILIEICGHK AIGTV**V**VGPT PVN**V**IGRN**M**L TQIGCTLNF

Table 5.1 Sequence alignment of WT, MDR769, and the DetMDR isolates. Drug resistance mutations are highlighted in bold.

The DetMDRs also contain previously identified non-polymorphic accessory mutations L10V/G, V11I, I13V, K20T/R, L33F/I/M, K43T, F53L, A71L, T74P, and L89V. To understand the molecular mechanisms of resistance of the MDR HIV protease, 40 ns molecular dynamics simulations of apo and complexed DetMDRs were performed. These studies provide insight into the structural changes that lead to multidrug resistance; in particular, we identify previously unreported roles for the I47V, V32I, I54M, L90M mutations in protease flap dynamics (**Figure 5.1**).

5.2 *Results*

5.2.1 Mutations in the Detroit isolates reveal alternative protein dynamics

RMSD analysis of the apo structure of the DetMDR isolates identified increased fluctuation of all three isolates compared to the WT protease (**Figure 5.1**). DetMDR1 shows only a brief increase in flexibility while DetMDR2 shows a sharp increase followed by sustained elevation in RMSD for the rest of the trajectory. DetMDR3 displays short bursts of increased flexibility but no sustained changes in

RMSD. These data suggest a large conformational change in DetMDR1 and DetMDR2 but not in DetMDR3 or the WT.

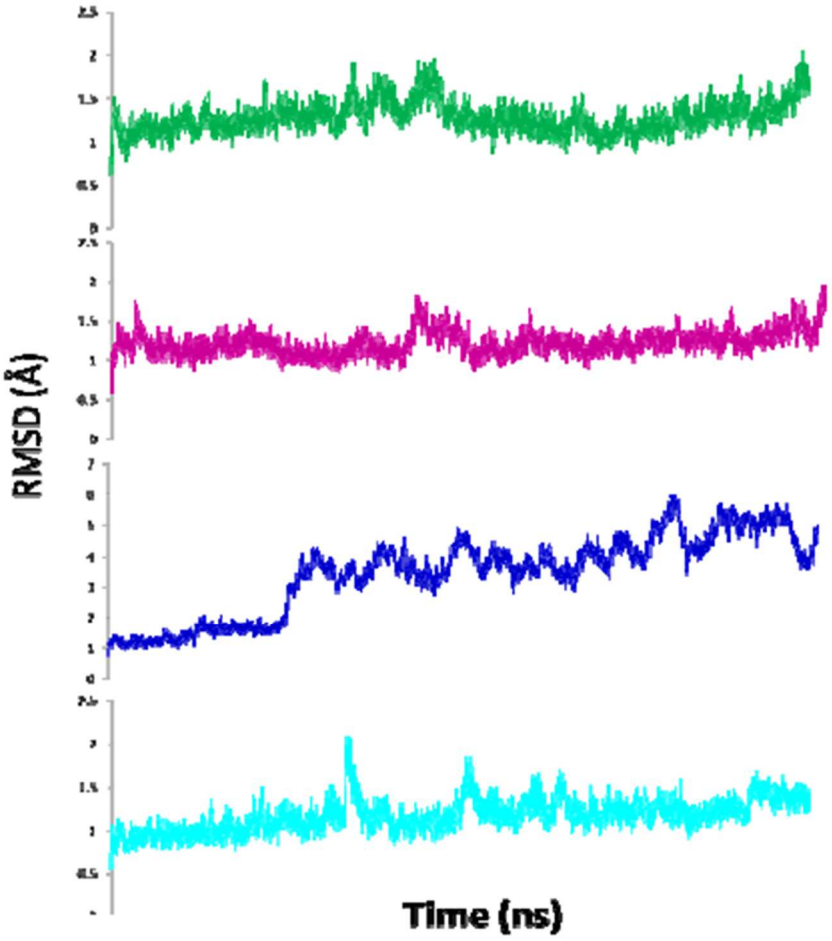


Figure 5.1 Root Mean Square Deviation (RMSD) of Cα backbone atoms of apo HIV-1 protease isolates. WT protease is displayed in green, DetMDR1 in pink, DetMDR2 in blue, and DetMDR3 in cyan. Note that DetMDR2 has an altered y-axis as it displayed a significantly larger increase in RMSD than the other protease.

5.2.2 V32I interacts with I47V to tether the protease flaps in a closed conformation

Visual analysis of the trajectory shows that these increases in RMSD are due to opening of the flaps in DetMDR1 and DetMDR2. In DetMDR1 the flaps open and then close. In contrast, the flaps of DetMDR2 open and stay open for the remainder of the trajectory. The WT and DetMDR3 flaps do not open during these simulations. These differences in flap movement correspond to the presence or absence of particular mutations. I47V likely contributes to flap opening as it is present in both isolates 1 and 2

where the flaps open but not in DetMDR3 where the flaps do not open. However, vdW interactions between the V32I and I47V mutations cause flap closure in HIV-1 protease (**Figure 5.2**). These vdW interactions could explain the flap opening followed by quick flap closure observed in DetMDR1, while the absence of V32I in DetMDR2 removes this interaction and cause the flaps to remain open. Similarly, the WT and DetMDR3 have V32 and I47 at these loci and therefore the vdW interactions in the WT and DetMDR3 are maintained casing the flaps to remain closed.

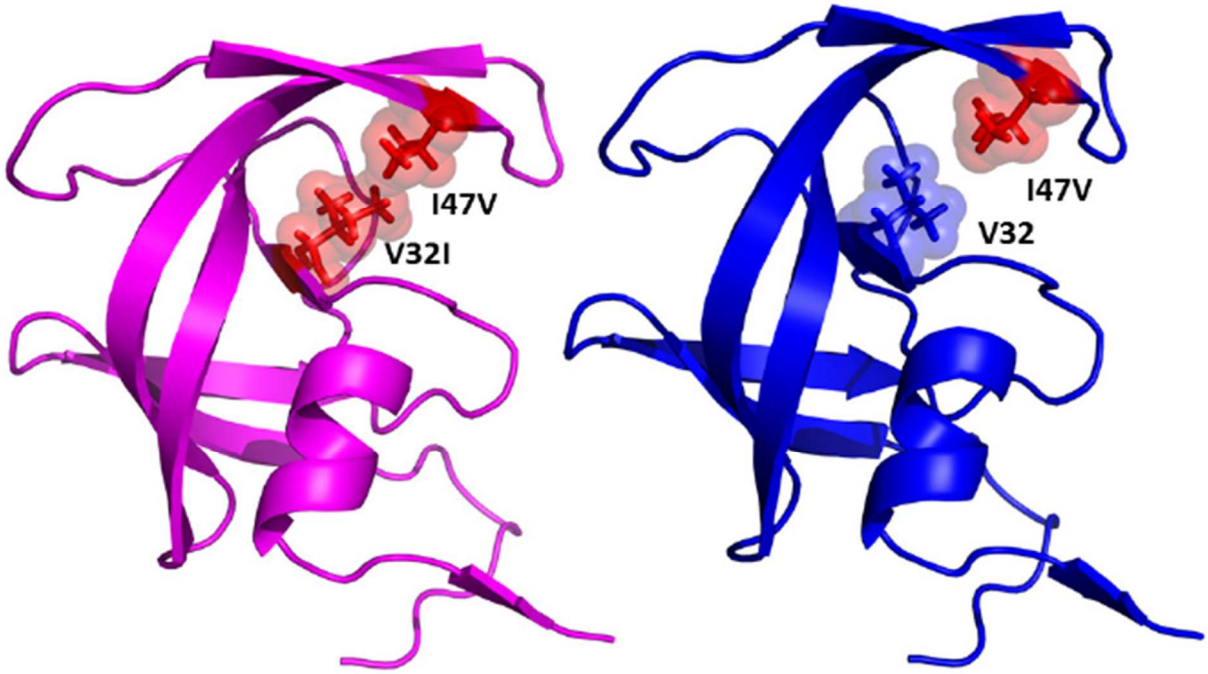


Figure 5.2 Change in the van der Waals volume induced by the drug resistance mutations I47V and V32L. The left panel shows DetMDR1 in pink. The right panel shows DetMDR2 in blue.

5.2.3 I54M in combination with L90M plays a role in asymmetric flap movement

I54M and L90M are associated with asymmetric movement in DetMDR1 corresponding to opening of the flaps. The RMSD of L90M on chain B is different than that of chain A. In DetMDR1 the asymmetric movement of the flaps as indicated by the differences in RMSD at each residue show a difference in residues 47-54 on chain B compared to chain A (**Figure 5.3**). Movement of the residues in

this region was symmetric in DetMDR2 where the residue fluctuation per residue is conserved between both chains of the protease.

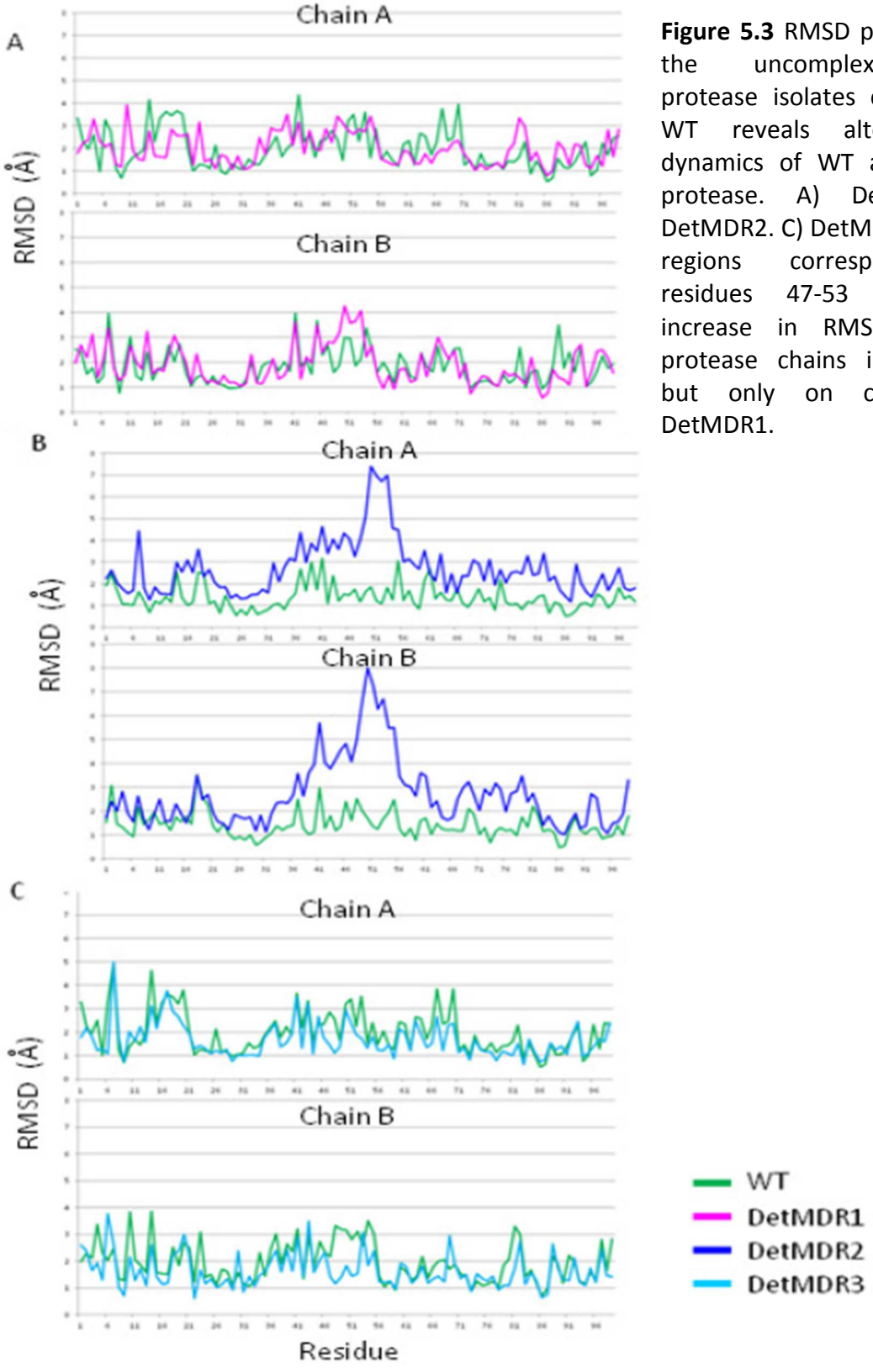


Figure 5.3 RMSD per residue of the uncomplexed HIV-1 protease isolates compared to WT reveals alternate flap dynamics of WT and DetMDR protease. A) DetMDR1. B) DetMDR2. C) DetMDR3. The flap regions corresponding to residues 47-53 display an increase in RMSD on both protease chains in DetMDR2, but only on chain B of DetMDR1.

5.2.4 Darunavir, atazanavir and lopinavir binding stabilize the HIV-1 protease flaps

The opening of the flaps observed in the apo simulations of DetMDR1 and DetMDR2 did not occur in the drug complex simulations. Drug binding maintains a closed flap conformation in DetMDR2. This suggests that the structural mechanism of drug resistance in these isolates is not due to changes in protein flexibility (**Figure 5.4**).

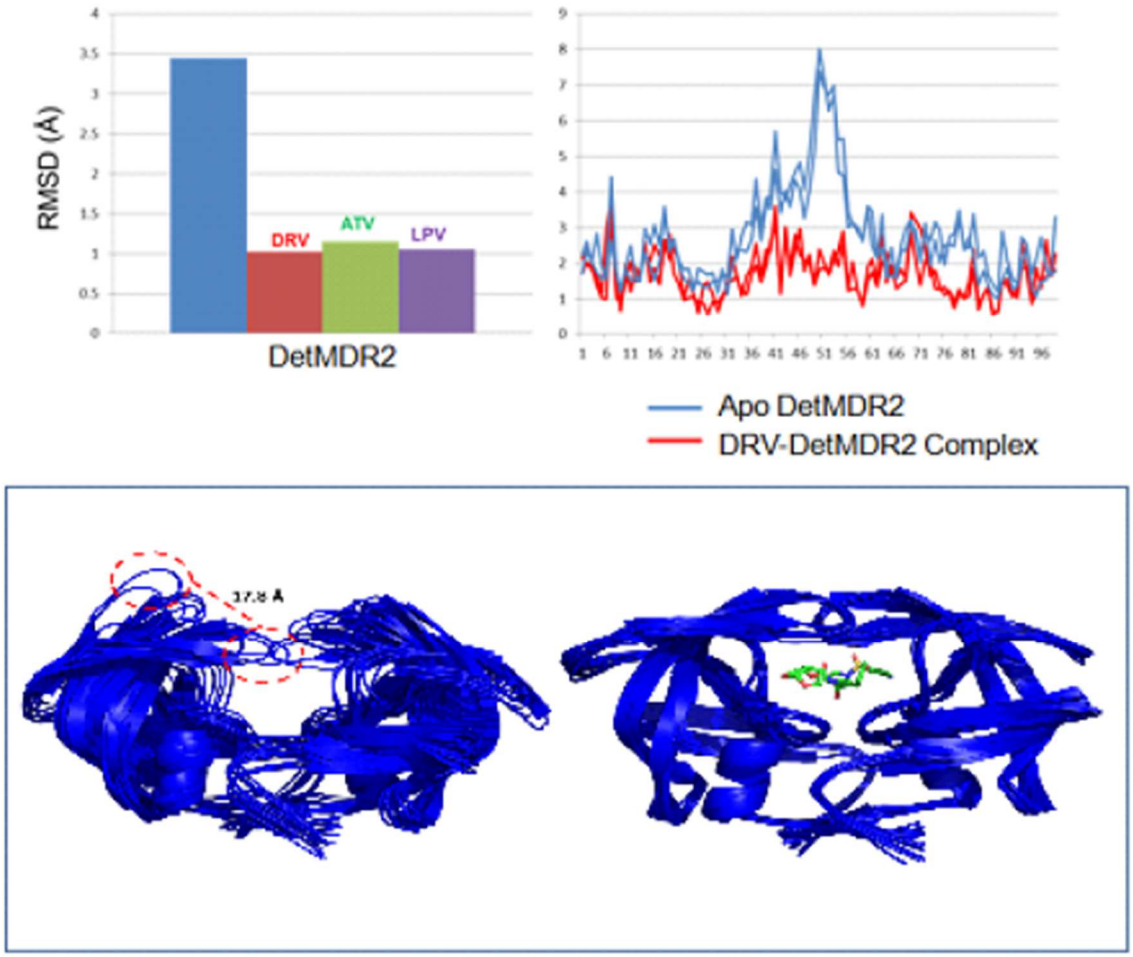


Figure 5.4 Darunavir, atazanavir, and lopinavir stabilize the HIV-1 protease flaps. The top left panel shows the average RMSD of DetMDR2 either alone (blue) or in complex with DRV (red), ATV (green), and LPV (purple). The RMSD per residue is shown in the top right panel. The bottom left panel is representative of the 40ns MD simulation using frames post energy minimization and at 4ns intervals of uncomplexed and DRV complexed DetMDR2, respectively aligned by the Cα positions. The red dashed lines indicate a 17.8Å movement of the Cα of Ile50 of the protease flap. This figure was made in PyMol

5.2.5 Mutation induced changes to the hydrogen bonding network alters inhibitor conformation

In addition to the altered protein dynamics, there is a change in the hydrogen bonding interactions in the WT compared to the MDR complex. DetMDR1 has 2.3%, 0.8% and 19.2% reduced hydrogen bond formation (compared to the WT) when complexed with DRV, ATV, and LPV respectively. Interestingly, DetMDR2 and DetMDR3 have increased hydrogen bond formation when in complex with DRV and ATV and a 21% and 36% decrease in hydrogen bond formation when in complex with LPV. The increase in hydrogen bonding is likely due to an alteration in the residues involved in an interaction with the PI. While changing the binding pocket may increase hydrogen bond formation, this also alters the conformation of the inhibitor.

5.3 *Discussion*

Three multi-drug resistant HIV-1 protease clinical isolates were selected from patients attending the Wayne State University infectious diseases clinic and failing antiviral treatment therapy on PI based regimens. To explore the structural mechanisms resulting in treatment failure, we performed 40ns MD simulations on the Detroit MDR series. Our results indicate a novel structural role for the I47V, V32I, I54M and L90M resistance mutations.

The V32I and I47V mutations play a structural role in tethering the flaps to the active site. Sequence analysis comparisons of the DetMDR protease isolates showed that DetMDR2 does not contain the V32I and I47V mutation combination. It contains only the I47V mutation and without the V32I mutation there is a loss in vdW contact volume between these two residues. Analysis of this apo protease trajectory displayed a marked increase in RMSD when compared to the WT, and a visual inspection of the trajectory as well as per-residue RMSD analysis confirmed that this is due to the flaps opening. Therefore, we postulate that I47V is responsible for flap opening. However, V32I is a compensatory mutation that may be responsible for tethering the flaps to the active site through its contacts with I47V. Therefore, if I47V is the only mutation it may be responsible for promoting flap opening through the disruption of the vdW interactions between residues 32 and 47.

I54M and L90M may be responsible for asymmetric movement of the protease flaps. This mutation combination is only present in DetMDR1 which is the only protease in the series in where the flaps asymmetrically open and then close after 4 ns. The mechanism explaining asymmetric flap opening will be explored further in future characterization studies of the HIV-1 protease isolates.

5.4 *Author's Contribution*

The author worked with Dr. Dewdney to train Dr. Chordia in the setup, performance, and analysis of MD experiments. He performed setup and analysis of approximately half of the MD experiments. He also wrote and edited much of the conclusions section of the paper.

Chapter 6: Future Directions

6.1 *HIV-1*

Drug resistance remains a major problem for HIV-1 treatment, although many strides have been made in the last decade. Patients who develop drug resistance mutations to multiple inhibitors have very poor prognoses, and the development of new inhibitors that are effective against these variants is still a pressing need. Additionally, even with proper adherence to treatment, it is becoming clear that patients with otherwise undetectable viral plasma titers still maintain reservoirs of virus in privileged areas, including the brain and CNS [65]. PIs and other antiretroviral drugs are unable to gain access to these privileged compartments, allowing the virus to replicate freely and eventually cause significant damage. A majority of patients who otherwise have undetectable viral loads eventually experience some degree of neurological complication, years after initial infection [66]. Our lab has begun work on the development of a small molecule inhibitor targeting the P-glycoprotein ATPase, a promiscuous transporter responsible for pumping compounds out of the Brain/CNS. Effective inhibition of this enzyme may allow for current treatments to act on the otherwise unreachable HIV reservoirs, improving treatment outcomes for many patients.

APPENDIX

Parameters	MDR769 L33F apo	MDR 769 L33F-DRV Complex
PDB Entry	4YOB	4YOA
Data Collection:		
Space Group	$P4_12_12$	$P4_12_12$
Wavelength (Å)	0.979	0.97856
Cell Constants (Å)	a = 45.75 b = 45.75 c = 102.22	a = 45.47 b = 45.47 c = 102.22
Resolution Range (Å)	102.22-1.50 (1.68-1.50)	102.22-1.70 (1.90-1.70)
Number of unique reflections	17,812	12,545
Completeness (%)	99.1 (98.2)	99.9 (100)
Redundancy	14.1 (14.5)	13.7 (14.2)
Mean I/σ (I)	34.4 (3.4)	26.9 (4.2)
R_{merge}	0.046 (0.752)	0.049 (0.615)
Refinement:		
R_{work} (%)	17.60	19.01
R_{free} (%)	19.43	22.46
Number of Atoms:		
Ligand		38
Protease	782	784
Solvent	134	71
Average isotropic B factor (Å^2):		
Ligand		93.94
Protease	23.36	35.11
Solvent	37.49	46.03
RMSD bond length (Å)	0.007	0.014
RMSD bond angle (°)	0.984	1.543
Ramachandran plot:		
Allowed/generous/disallowed (%)	98/2/0	98/2/0

Table A1: Crystallographic data for 4YOA and 4YOB.

REFERENCES

1. Mandell, G.L., J.E. Bennett, and R. Dolin, *Mandell, Douglas, and Bennett's principles and practice of infectious diseases*. 7th ed. 2010, Philadelphia, PA: Churchill Livingstone/Elsevier.
2. Turk, B., *Targeting proteases: successes, failures and future prospects.* Nat Rev Drug Discov, 2006. **5**(9): p. 785-99.
3. Gilbert, P.B., et al., *Comparison of HIV-1 and HIV-2 infectivity from a prospective cohort study in Senegal.* Stat Med, 2003. **22**(4): p. 573-93.
4. Reeves, J.D. and R.W. Doms, *Human immunodeficiency virus type 2.* J Gen Virol, 2002. **83**(Pt 6): p. 1253-65.
5. Ratner, L., et al., *Complete nucleotide sequence of the AIDS virus, HTLV-III.* Nature, 1985. **313**(6000): p. 277-84.
6. Pancera, M., et al., *Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility.* Proc Natl Acad Sci U S A, 2010. **107**(3): p. 1166-71.
7. Robinson, H.L., *New hope for an AIDS vaccine.* Nat Rev Immunol, 2002. **2**(4): p. 239-50.
8. Li, G. and E. De Clercq, *HIV Genome-Wide Protein Associations: a Review of 30 Years of Research.* Microbiol Mol Biol Rev, 2016. **80**(3): p. 679-731.

9. Fletcher, T.M., 3rd, et al., *Nuclear import and cell cycle arrest functions of the HIV-1 Vpr protein are encoded by two separate genes in HIV-2/SIV(SM).* EMBO J, 1996. **15**(22): p. 6155-65.
10. Davies, D.R., *The structure and function of the aspartic proteinases.* Annu Rev Biophys Biophys Chem, 1990. **19**: p. 189-215.
11. Pettit, S.C., et al., *Initial cleavage of the human immunodeficiency virus type 1 GagPol precursor by its activated protease occurs by an intramolecular mechanism.* J Virol, 2004. **78**(16): p. 8477-85.
12. Laco, G.S., *HIV-1 protease substrate-groove: Role in substrate recognition and inhibitor resistance.* Biochimie, 2015. **118**: p. 90-103.
13. Brik, A. and C.H. Wong, *HIV-1 protease: mechanism and drug discovery.* Org Biomol Chem, 2003. **1**(1): p. 5-14.
14. Moore, R.D. and R.E. Chaisson, *Natural history of HIV infection in the era of combination antiretroviral therapy.* AIDS, 1999. **13**(14): p. 1933-42.
15. Flexner, C., *HIV drug development: the next 25 years.* Nat Rev Drug Discov, 2007. **6**(12): p. 959-66.
16. Wlodawer, A., *Rational approach to AIDS drug design through structural biology.* Annu Rev Med, 2002. **53**: p. 595-614.
17. Yanchunas, J., Jr., et al., *Molecular basis for increased susceptibility of isolates with atazanavir resistance-conferring substitution I50L to other*

protease inhibitors. Antimicrob Agents Chemother, 2005. **49**(9): p. 3825-32.

18. Doyon, L., et al., *Selection and characterization of HIV-1 showing reduced susceptibility to the non-peptidic protease inhibitor tipranavir.* Antiviral Res, 2005. **68**(1): p. 27-35.
19. McCoy, C., *Darunavir: a nonpeptidic antiretroviral protease inhibitor.* Clin Ther, 2007. **29**(8): p. 1559-76.
20. Lv, Z., Y. Chu, and Y. Wang, *HIV protease inhibitors: a review of molecular selectivity and toxicity.* HIV AIDS (Auckl), 2015. **7**: p. 95-104.
21. Kuiper, B.D.K., B.J.; Dewdney, T.G.; Chordia, P.; Ross, K.; Brunzelle, J.S.; Kovari, I.A.; MacArthur, R.; Salimnia, H.; Kovari, L.C., *The L33F darunavir resistance mutation acts as a molecular anchor reducing the flexibility of the HIV-1 protease 30s and 80s loops.* Biochemistry and Biophysics Reports, 2015. **2**: p. 160-165.
22. Roberts, J.D., K. Bebenek, and T.A. Kunkel, *The accuracy of reverse transcriptase from HIV-1.* Science, 1988. **242**(4882): p. 1171-3.
23. Deeks, S.G.G., S.J.; Moore, R.D., *Trends in Multidrug Treatment Failure and Subsequent Mortality among Antiretroviral Therapy-Experienced Patients with HIV Infection in North America.* Clinical Infectious Diseases, 2009. **49**(10): p. 1582-1590.

24. Rhee, S.Y., et al., *Human immunodeficiency virus reverse transcriptase and protease sequence database.* Nucleic Acids Res, 2003. **31**(1): p. 298-303.

25. Martin, P.V., J.F.; Proteasa, G.; Jimenez, Y.L.; Wawrzak, Z.; Winters, M.A.; Merigan, T.C.; Kovari, L.C., *"Wide-open" 1.3 A structure of a multidrug-resistant HIV-1 protease as a drug target.* Structure, 2005. **13**: p. 1887-1895.

26. McCoy, A.J., et al., *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.

27. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.

28. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.

29. Joosten, R.P., et al., *PDB_REDO: automated re-refinement of X-ray structure models in the PDB.* J Appl Crystallogr, 2009. **42**(Pt 3): p. 376-384.

30. Chen, V.B., et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 12-21.

31. Schrodinger, LLC, *The PyMOL Molecular Graphics System, Version 1.8*. 2015.

32. Laskowski, R.A. and M.B. Swindells, *LigPlot+: multiple ligand-protein interaction diagrams for drug discovery.* J Chem Inf Model, 2011. **51**(10): p. 2778-86.

33. Karplus, M., *Molecular dynamics simulations of biomolecules.* Acc Chem Res, 2002. **35**(6): p. 321-3.

34. Shaw, D.E.G., J.P.; Bank, J.A.; Young, C., *Anton 2: raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer.* Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, 2014: p. 41-53.

35. Brooks, B.R., *CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations.* Journal of Computational Chemistry, 1983. **4**: p. 187-217.

36. Brooks, B.R., *CHARMM: the biomolecular simulation program.* Journal of Computational Chemistry, 2009. **30**(10): p. 1545-1614.

37. Phillips, J.C., *Scalable molecular dynamics with NAMD.* Journal of Computational Chemistry, 2005. **26**(16): p. 1781-1802.

38. Darden, T.Y., D.; Pedersen, L., *Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems.* Journal of Chemical Physics, 1993. **98**(12): p. 10089-10092.

39. Arnold, K., et al., *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling.* Bioinformatics, 2006. **22**(2): p. 195-201.

40. Humphrey, W., A. Dalke, and K. Schulten, *VMD: visual molecular dynamics.* J Mol Graph, 1996. **14**(1): p. 33-8, 27-8.

41. Grosdidier, A., V. Zoete, and O. Michielin, *SwissDock, a protein-small molecule docking web service based on EADock DSS.* Nucleic Acids Res, 2011. **39**(Web Server issue): p. W270-7.

42. Grosdidier, A., V. Zoete, and O. Michielin, *Fast docking using the CHARMM force field with EADock DSS.* J Comput Chem, 2011. **32**(10): p. 2149-59.

43. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis.* J Comput Chem, 2004. **25**(13): p. 1605-12.

44. Tang, M.W. and R.W. Shafer, *HIV-1 antiretroviral resistance: scientific principles and clinical applications.* Drugs, 2012. **72**(9): p. e1-25.

45. Lambert-Niclot, S., et al., *Factors associated with the selection of mutations conferring resistance to protease inhibitors (PIs) in PI-experienced patients displaying treatment failure on darunavir.* Antimicrob Agents Chemother, 2008. **52**(2): p. 491-6.

46. Chordia, P.D., T.G.; Keusch, B.J.; Kuiper, B.D.; Ross, K.; Kovari, I.A.; MacArthur, R.; Salimnia, H.; Kovari, L.C., *The role of mutations at codons 32, 47, 54, and 90 in HIV-1 protease flap dynamics.* Discoveries, 2014. **2**(4).

47. Maguire, M., et al., *Emergence of resistance to protease inhibitor amprenavir in human immunodeficiency virus type 1-infected patients: selection of four alternative viral protease genotypes and influence of viral susceptibility to coadministered reverse transcriptase nucleoside inhibitors.* Antimicrob Agents Chemother, 2002. **46**(3): p. 731-8.

48. Wensing, A.M., et al., *2014 Update of the drug resistance mutations in HIV-1.* Top Antivir Med, 2014. **22**(3): p. 642-50.

49. Agniswamy, J., et al., *HIV-1 protease with 20 mutations exhibits extreme resistance to clinical inhibitors through coordinated structural rearrangements.* Biochemistry, 2012. **51**(13): p. 2819-28.

50. Yedidi, R.S., et al., *Contribution of the 80s loop of HIV-1 protease to the multidrug-resistance mechanism: crystallographic study of MDR769 HIV-*

1 protease variants. Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 6): p. 524-32.

51. Liu, F., et al., *Effect of flap mutations on structure of HIV-1 protease and inhibition by saquinavir and darunavir.* J Mol Biol, 2008. **381**(1): p. 102-15.

52. Liu, Z., et al., *Conserved hydrogen bonds and water molecules in MDR HIV-1 protease substrate complexes.* Biochem Biophys Res Commun, 2013. **430**(3): p. 1022-7.

53. Ragland, D.A., et al., *Drug resistance conferred by mutations outside the active site through alterations in the dynamic and structural ensemble of HIV-1 protease.* J Am Chem Soc, 2014. **136**(34): p. 11956-63.

54. Saskova, K.G., et al., *Molecular characterization of clinical isolates of human immunodeficiency virus resistant to the protease inhibitor darunavir.* J Virol, 2009. **83**(17): p. 8810-8.

55. Ye, Z., et al., *Reprogramming acyl carrier protein interactions of an Acyl-CoA promiscuous trans-acyltransferase.* Chem Biol, 2014. **21**(5): p. 636-46.

56. Qiu, X. and Z.P. Liu, *Recent developments of peptidomimetic HIV-1 protease inhibitors.* Curr Med Chem, 2011. **18**(29): p. 4513-37.

57. Mills, N., *ChemDraw Ultra 10.0.* Journal of the American Chemical Society, 2006. **128**(41): p. 13649-13650.

58. Wainberg, M.A., et al., *Clinical correlates and molecular basis of HIV drug resistance.* J Acquir Immune Defic Syndr, 1993. **6 Suppl 1**: p. S36-46.

59. Smyth, R.P., M.P. Davenport, and J. Mak, *The origin of genetic diversity in HIV-1.* Virus Res, 2012. **169**(2): p. 415-29.

60. Johnson, V.A., et al., *Update of the drug resistance mutations in HIV-1: March 2013.* Top Antivir Med, 2013. **21**(1): p. 6-14.

61. Wang, Y., et al., *The higher barrier of darunavir and tipranavir resistance for HIV-1 protease.* Biochem Biophys Res Commun, 2011. **412**(4): p. 737-42.

62. Rabi, S.A., et al., *Multi-step inhibition explains HIV-1 protease inhibitor pharmacodynamics and resistance.* J Clin Invest, 2013. **123**(9): p. 3848-60.

63. Rhee, S.Y., et al., *HIV-1 protease mutations and protease inhibitor cross-resistance.* Antimicrob Agents Chemother, 2010. **54**(10): p. 4253-61.

64. Ceccherini-Silberstein, F., et al., *Identification of the minimal conserved structure of HIV-1 protease in the presence and absence of drug pressure.* AIDS, 2004. **18**(12): p. F11-9.

65. Sacktor, N., et al., *Prevalence of HIV-associated neurocognitive disorders in the Multicenter AIDS Cohort Study.* Neurology, 2016. **86**(4): p. 334-40.

66. Simioni, S., et al., *Cognitive dysfunction in HIV patients despite long-standing suppression of viremia.* AIDS, 2010. **24**(9): p. 1243-50.

ABSTRACT

STRUCTURE FUNCION STUDIES OF HIV-1 PROTEASE

by

BRADLEY JAMES KEUSCH

August 2016

Advisor: Dr. Ladislau C Kovari

Major: Biochemistry and Molecular Biology

Degree: Master of Science

HIV-1 is the causative agent of the devastating human disease Acquired Immunodeficiency Syndome (AIDS). While much progress has been made over the past two decades, HIV-1 remains a major global health concern. HIV-1 protease is 99-amino acid homodimer aspartyl protease that is essential to the life cycle of HIV. This has rendered it an attractive and very successful drug target. However, due to the high error rate of the HIV -1 reverse transcriptase, drug resistance mutations in the protease can develop very rapidly in some patients, rendering current protease inhibitors (one of the main classes of drug in common antiretroviral therapy) less effective or completely ineffective. In this thesis, we investigate the structural impact of a number of HIV-1 protease drug resistance mutations. These include L33F, which is selected for on darunavir treatment (one of the most prescirbed protease inhibitors), I47V, and V54I (which we identify as compensatory mutations involved in the tethering of the protease flaps and proper formation of the active site). A fuller understanding of the structural impact of these resistance mutations will hopefully facilitate the development of protease inhibitors that can overcome this common drug resistance.

AUTOBIOGRAPHICAL STATEMENT

Education

08/2013-08/2016	M.S. in Biochemistry and Molecular Biology Department of Biochemistry and Molecular Biology Wayne State University School of Medicine
08/2004- 12/2008	B.S. in Zoology Department of Natural Science Michigan State University

Publications

Chordia, P., Dewdney, T.G., **Keusch, B.J.**, Kuiper, B.D., Ross, K., Kovari, I.A., MacArthur, R., Salimnia, H., Kovari, L.C. 2014 The role of mutations at codons 32, 47, 54, and 90 in HIV-1 protease flap dynamics. *Discoveries* 2(4): e27

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