Assessing Human Immunodeficiency Virus (HIV) Protease Inhibition with Resistant Variants and Novel Inhibitors

A Major Qualifying Project
Submitted to the Faculty of Worcester Polytechnic Institute
in partial fulfillment of the requirements for the Degree in Bachelor of Science
in
Biochemistry

By

_____
Klajdi Kosovrasti

_____
Mina Henes

Date: May 2, 2018

University of Massachusetts Medical School
Department of Biochemistry and Molecular Pharmacology
Schiffer Laboratory
Worcester Polytechnic Institute

_____   _____
Professor Anita Mattson, Advisor   Professor Destin Heilman, Coadvisor

Acknowledgments

Table of Contents

List of Figures

List of Tables

List of Equations

Abstract

Human Immunodeficiency Virus (HIV-1) Protease Inhibitors (PIs) have become one of the most effective anti-viral drugs on the market. Darunavir (DRV), the most potent FDA-approved and clinically prescribed PI has been a cornerstone in Highly Active Antiretroviral Therapy (HAART) and the fight against HIV/Acquired Immunodeficiency Syndrome (AIDS). However, the ability of the HIV protease to mutate, grow resistance against PIs and proliferate rapidly has become a global concern. To address this issue, a new series of PIs were designed using the substrate envelope hypothesis to resist resistant mutations. The first set of PIs, UMass 1-10, are derived from the DRV backbone and have an aniline, methoxy, hydroxymethyl, benzodioxole, or a benzothiazole modification implemented on the P2' site, and 2-methylbutane or isohexyl modification on the P1' site. The second set of compounds, the mono- and di-hydroxyl series, also derived from the DRV backbone and contain a mono- or di-hydroxyl modification on their P2' site, and isobutyl, 2-methylbutane, or isohexyl modification on their P1' site. All of these compounds were kinetically tested against various mutants of the HIV-1 protease, and crystal structures were solved in order to structurally analyze the binding mode. In general, the UMass 1-10 compounds exhibit pico-molar K_i's against protease mutants, comparable to DRV. The mono- and di-hydroxyl series show promising results, as pico-molar potencies are also observed. In combination with crystal structures, these results can be utilized to design new PIs with enhanced inhibitory potencies against a wide range of HIV-1 protease mutants.

Chapter 1: Introduction

Throughout this report, we focus on the Human Immunodeficiency Virus (HIV) protease and how it utilizes evolution via numerous genomic mutations to circumvent the human immune system. The HIV protease is an aspartyl homodimer, with only 99 amino acids in each chain. It is responsible for cleaving along twelve nonhomologous sites which leads to a mature virus. It has the ability to mutate a large portion of its sequence to evade inhibitors while maintaining its normal function. The analysis of novel protease inhibitors specifically designed to resist the rapid mutation rate of the HIV protease is the main focus of this project. Along with the Wild Type (WT) strain, the various protease mutants of interest were I84V, I50V, I50V;A71V, and V82I. Utilizing the backbone of currently prescribed protease inhibitors such as Darunavir (DRV) and Lopinavir (LPV), the P1' and P2' moieties were modified and tested on the mutants previously mentioned. Prior to synthesizing any compounds, modeling simulations can be utilized to estimate how inhibitor modifications will affect binding to the active site and the overall interactions that might incur due to inhibitor modifications.

In order to effectively optimize these new inhibitors, it is crucial to understand how and where natural substrates fit within the enzyme's active site. The specificity of the protease cannot be determined by the sequences of the substrates cleaved. In previous studies conducted by the Schiffer lab, it was determined that there is a consistent consensus volume throughout all HIV-1 protease variants, referred to as the substrate envelope. Based on this, it was determined that the specificity of the enzyme relies on substrate shape rather than sequence (1). Using a substrate envelope guided drug design approach, novel inhibitors can be modeled and synthesized.

To study the effectiveness of these inhibitors, inhibition assays were conducted in order to determine the inhibition constant (K_i) for every respective drug and variant. In addition to K_i, the Michaelis Menten constant (K_m) of WT and numerous mutants were determined via a K_m assay. The K_m parameter was used to analyze the apparent affinity of each mutant to the substrate. Lastly, crystal structures of compounds in complex with mutants of interest were solved. Solved crystal structures were then utilized to visualize the inhibitors inside the substrate envelope. Such analysis allowed us to determine major differences in inhibitor binding and mutant structure compared to DRV. Furthermore, this analysis informed us about the relationship between the calculated K_i's and the observed inhibitor binding inside the substrate envelope. Considering all of the data and

analysis gathered, hypotheses regarding the inhibition capability of each compound can be made. This will help determine which inhibitor(s) are worth undergoing further testing in more complex laboratory procedures, such as viral passaging, computer simulations to determine van der Waals (vdW) interactions, and molecular dynamics simulations to properly visualize the inhibitors in a dynamic state.

1.1 Epidemiology of HIV

Since its first categorized case over 30 years ago, the Centers for Disease Control and Prevention (CDC) estimates that 37 million people worldwide are living with HIV. Of these, the World Health Organization (WHO) claims that 1.1 million die every year due to the virus (2) . In 2015, the United States had 39,513 positive HIV diagnoses, which represents a 4.8% increase from 2014. The three main causes of new HIV infection in order of prevalence are male-to-male sexual contact, heterosexual contact, and injection drug use.

No cure currently exists for HIV, and once a patient is infected, it is necessary for them to remain on therapeutic drug treatments. The virus has a very high risk of remission and drug resistance, which is why it is essential to continue constant treatment. Patients are often treated with a mixture of therapeutic drugs, targeting the virus at various stages of its cycle. No prevention vaccines exist for HIV. However, there are multiple research projects that are exploring the possibility and effectiveness of a preventative HIV vaccine (3) .

HIV-1 specifically attacks and enters CD4 cells, also known as T-helper white blood cells of the immune system. Once the virus enters the host cell, it hijacks the cell's native machinery (e.g., ribosomes, golgi apparatus, etc.) to make copies of its genome. When a patient is infected with HIV, they typically do not express symptoms immediately. Rather, HIV will stay dormant and asymptomatic in the host body, typically 3-12 years and sometimes longer, while building up a viral load (4). Once the viral load reaches a critical value, the host cell lyses and the virions are released into the bloodstream (5). Alternatively, the virions can bud off, leaving the host intact. The released virions mature outside the host cell and bind to uninfected CD4 cells, thus repeating the cycle and conducting further replication. Once a patient's White Blood Cell (WBC) count decreases below 200/mm³ (healthy adults have a WBC count $\geq$ 1000/mm³), they are said to have Acquired Immunodeficiency Syndrome (AIDS) (4).

1.2 HIV Genome Map

      In order to obtain a full understanding of the HIV-1 lifecycle and structure, we must first discuss its genome and its components. At its center, HIV contains two copies of single-stranded, positive sense RNA molecules approximately 7.9 Kb each (6) . The HIV-1 genome can be divided into several genes; gag, pol, and env, all of which are synthesized as a long chain termed the gag-pro-pol polyprotein. These genes are responsible for coding all of the necessary proteins and enzymes needed by the virus to recognize and bind to host cells, enter CD4 cells, replicate, assemble, and finally bud from host cells and mature. The full genome map is shown in Figure 1.
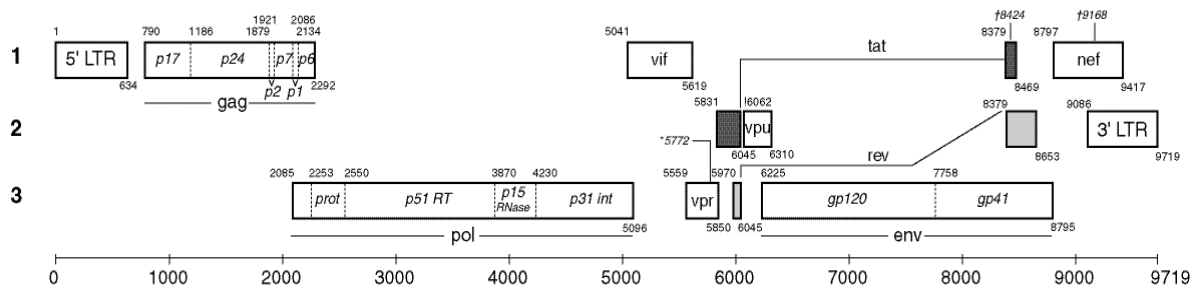
Figure 1: HIV-1 Genome Map

      The gag gene is the first region in the genome. This region encodes for several structural proteins, which are essential for viral assembly and maturation. A summary of the 5' LTR and Gag gene regions, as well as their function, is shown in Table 1.

Table 1: Function Summary of the 5' LTR and Gag Genes

Gene	Region	Function
5' LTR	/	Regulatory regions for transcription initiation and polyadenylation
Gag	P17 (MA)	Codes for the matrix structural proteins
	P24 (CA)	Codes for the capsid structural proteins
	P2	Spacer peptide - regulates conformational changes during maturation
	P7 (NC/p9)	Codes for the nucleocapsid structural proteins
	P1	Spacer peptide - regulates conformational changes during maturation
	P6	Contains binding sites for other proteins and accessory viral proteins

The pol gene region of the HIV genome is responsible for encoding the protease (p10), reverse transcriptase (p51), integrase (p31), and RNase (p15) (see Figure 1). The 3' end of the gag gene includes a cis-acting RNA motif that induces a frameshift when encountered by the ribosomes. This results in the ribosomes continuing uninterrupted translation, resulting in a gag-pol region. The HIV protease gene undergoes autoproteolysis to free itself from the pol gene. After the protease monomer is free, it forms a dimer and goes on to cleave up to 12 different sites on the gag, pol, and env regions (7) .

The env region of the HIV genome, often termed gp160, gives rise to the surface glycoprotein gp120 and the transmembrane glycoprotein gp41 post cleavage by the protease. Together with the CCR5/CXCR4 HIV co-receptors, gp120 and gp41 are vital for recognition and binding to the host CD4 cell (8). Without co-receptor binding, HIV cannot enter the host cell.

The HIV genome also includes regulatory regions. The Tat region produces two forms of Tat (a 72 AA and 101 AA form) that bind to the 5' end of the HIV RNA and initiates transcription (9). The Rev region is an RNA specific binding protein that aids in the transition from early to late HIV gene expression (10). In addition to the regulatory genes, the HIV genome contains four accessory genes; nef, vif, vpr and vpu. These genes aid in facilitating HIV replication in the host cell and enhance virion release post-assembly (11-13).

1.3 HIV-1 Structure and Life Cycle

The general structure of HIV-1 and its components are shown Figure 2. Retroviruses use Ribonucleic Acid (RNA) as their genetic material. The RNA is encapsulated by the nucleocapsid, which in turn is surrounded by the capsid. Enclosed within the capsid are the three main enzymes needed for replication; reverse transcriptase, integrase, and the protease. Reverse transcriptase is necessary to convert RNA into double-stranded DNA (14). Integrase is needed to integrate the produced DNA into the host DNA (15). Finally, the protease is used to cleave the polyprotein, which is an essential step in developing mature infectious virions (16). Glycoproteins on the surface of the membrane allow for recognition of and binding to CD4 cells.
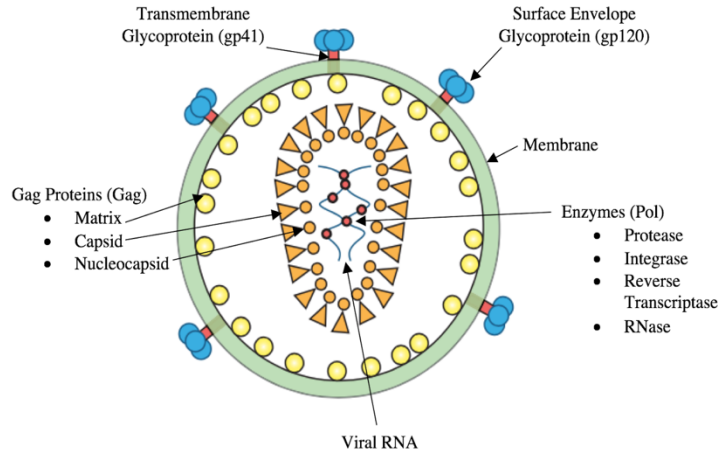
Figure 2: Structure and Components of HIV-1 (17)

Once the virus has identified the CD4 cells, it binds to the CD4 and CCR5 co-receptor on the surface of the cell. The virus then initiates fusion with the cell's membrane. Host cell recognition begins when gp120 binds to the CD4 receptor on the surface of White Blood Cells (WBCs). Binding to the CD4 receptor causes a gp120 conformational change, which exposes the CCR5 co-receptor binding site (18). The double binding results in the fusion peptide of gp41 to insert itself in the membrane of the host cell creating a hairpin loop (see Figure 3). The cell membrane of the virus fuses with the cell membrane of the host cell and the viral capsid enters the cell. Once inside, the viral genetic material is released into the cytoplasm along with the enzymes necessary to aid in the virus' incorporation into the cell's genome (i.e. reverse transcriptase and integrase).
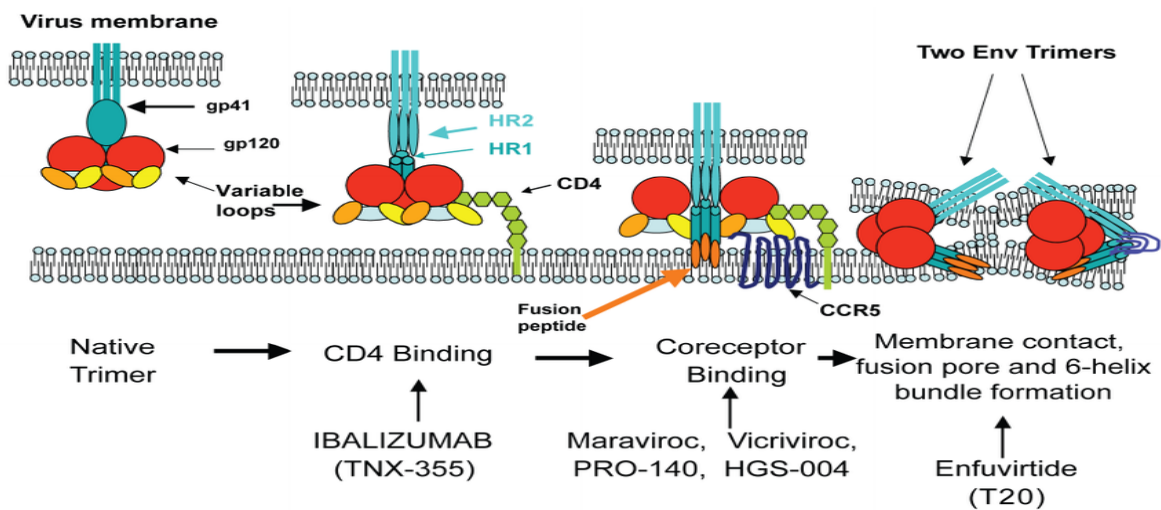
Figure 3: HIV Binding and Entering a Host CD4 Cell (19)

Inside the cell, the viral RNA is first reversely transcripted by reverse transcriptase into double stranded DNA. The double stranded DNA is then integrated into the host chromosome by the enzyme integrase, followed by transcription and translation. This process is depicted in steps 2 through 3 in Figure 4. It is important to note that once viral DNA has been integrated, it remains in the host cell's genome, replicating not just once but as often as the host cell replicates.
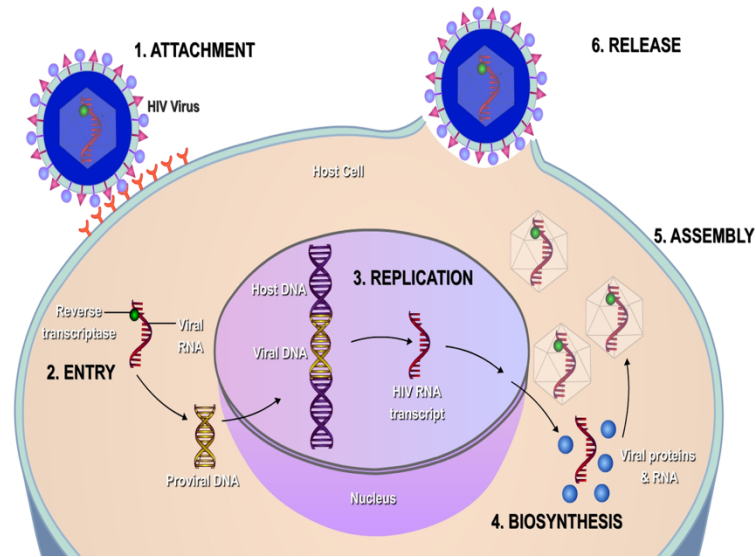
Figure 4: Reverse Transcription and Viral DNA Integration

Once viral DNA is replicated, it is translated back into viral RNA. This process marks the beginning of the virion formation, which are immature viruses. Virions consist of essential initial enzymes and the uncleaved polyprotein, which is transcribed by the host cell's ribosomes. Next, the virions undergo early-stage maturation, which are marked by Env cleavages, giving way to capsid assembly on the cytoplasmic side of the cell membrane.

At this point, the viral RNA along with essential proteins and the polyprotein begin to assemble for budding. The specifics of the budding process are depicted in Figure 5, starting from the bottom of the figure.
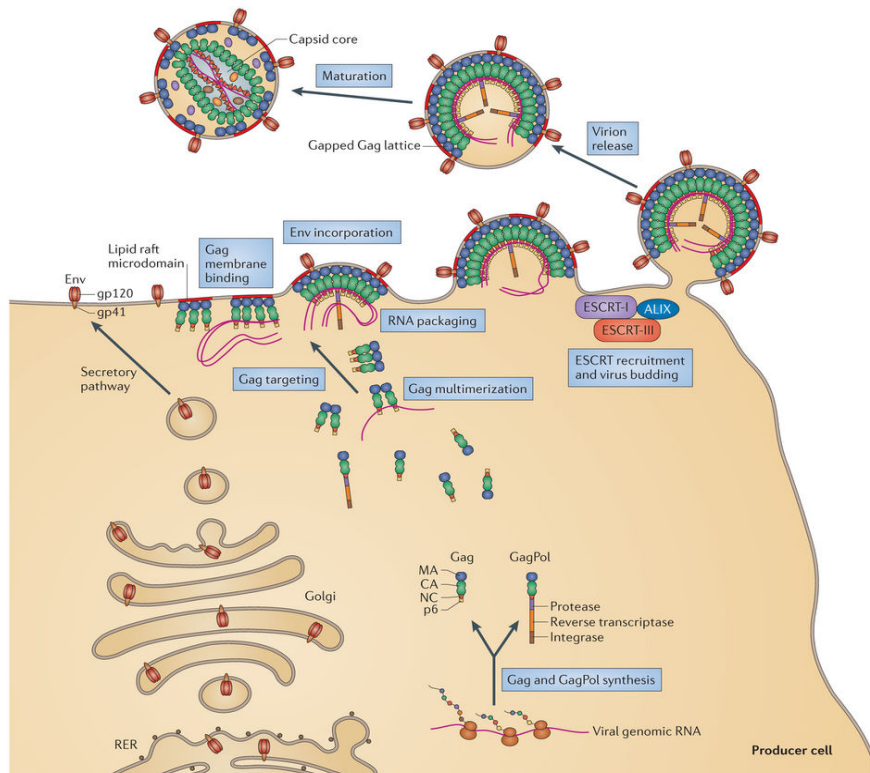
Figure 5: Budding and Maturation Schematic (20)

The viral envelope glycoproteins are trafficked through the cytoplasm from the rough endoplasmic reticulum via the secretory pathway system. The precursors of the Gag polyprotein are synthesized and assembled in the cytoplasm of the cell. The assembled Gag recruits the viral genomic RNA, which was transcribed after DNA integration, and reaches the plasma membrane via a pathway yet to be defined. The Gag-RNA complex latches onto the lipid rafts microdomains via insertion of its amino-terminal myristate into the lipid bilayer and hydrophobic interaction with the phospholipid bilayer. The assembling particle then recruits env as well as endosomal sorting complexes required for completion of the budding (ESCRT-I and ESCRT-III). Lastly, the ESCRT-III collaborates with the Vacuolar Protein Sorting 4 (VPS4) complexes to drive the membrane scission reaction that leads to particle release. In order for the virion to mature, proteolytic cleavage of the Gag and Gag-Pro-Pol polyprotein complexes must occur by the HIV protease. Post budding, the newly-formed virions enter late-stage maturation mediated by the protease, to become infectious, which involves a series of polyprotein cleavages. The mature virions can now go on to infect other CD4 cells and repeat the process.

The HIV life cycle and genome present many opportunities for therapeutic drug targets. These therapies aim to disrupt the normal life cycle of the virus at various stages and halt

replication. In the upcoming sections, we explore the potential therapeutic sites and the classes of medications used against those targets.

1.4 Current FDA Approved HIV-1 Medications and Protease Inhibitors

Considering the life cycle of HIV and its mode of entry into the host cell, researchers have been able to identify four potential pharmacological target sites. The potential target sites include the reverse transcriptase, receptors on the surface of the CD4 cell, HIV integrase, and the HIV protease. Currently, there are 24 FDA approved medications, each falling in one of the following classifications; Nucleoside Reverse Transcriptase Inhibitors (NRTIs), Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs), fusion inhibitors, entry inhibitors, integrase inhibitors, and Protease Inhibitors (PIs). Here we will mainly focus on PIs. Specifically, we will focus on Darunavir (DRV), which is currently the most potent clinically-approved PI prescribed. In addition to DRV, we will also discuss an early generation inhibitor, Lopinavir (LPV). LPV is not as potent at DRV. However, it was utilized to make DRV/LPV hybrid compounds. Structures of DRV and LPV are shown in Figure 6.
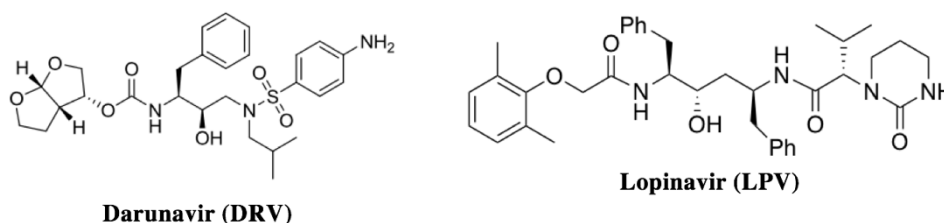
Darunavir (DRV)   **Lopinavir (LPV)**
Figure 6: DRV And LPV Chemical Structures

The HIV-1 protease is highly susceptible to mutations. These mutations allow the virus to evade and develop resistance to targeted drug therapies while maintaining normal function. For that reason, many HIV therapy regimens rely on pharmacokinetic enhancers and combination HIV medications. Pharmacokinetic enhancers strengthen the action of other HIV medications, making them more effective. Combination HIV medications composed of three or more medications are often used as part of the HIV treatment regimen to attack the virus at multiple target sites in its life cycle. The use of combination drugs reduces the possibility of the virus building resistance to a certain medication and thus keeps viral load to a safe minimum.

The protease is perhaps the most enticing therapeutic drug target due to its vital role in virion maturation. Protease Inhibitors (PIs), like DRV and LPV, are competitive inhibitors that bind to the active site of the protease as shown in Figure 7 (7). If the protease cannot bind and cleave its natural substrate, the virus will not have the necessary proteins to replicate and mature.
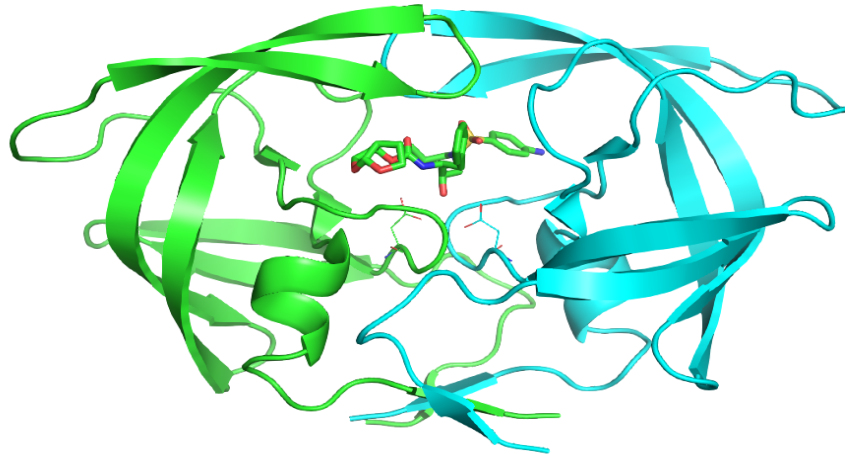

Figure 7: DRV In the Active Site of the HIV-1 Protease

Often, when patients are prescribed PIs, they are usually also prescribed Cobicistat (COBI) (21). COBI is classified as a pharmacokinetic enhancer. When COBI is given in conjunction with PIs, it boosts their action by inhibiting cytochrome P450 3A enzymes (CYP3A). CYP3A is responsible for drug metabolism in the human body. By suppressing the function of CYP3A, the prescribed drugs are able to stay in the bloodstream longer without being metabolized, therefore increasing the time of action against the virus (21).

Combination HIV Medicines

In most HIV patients, physicians opt to follow Highly Active Antiretroviral Therapy (HAART). HAART is a combination of at least three different classes of medications that attack the virus at multiple therapeutic targets simultaneously (22). In a 3-month clinical trial, HAART has shown a decrease in viral load and increased WBC count. However, patients on HAART can show increased drug intolerance and some (~25%) are taken off these medications due to intolerance (22). Physicians can alter therapy based on patient response and tolerance to reach a mixture of mediations that produces the best results for that particular patient.

1.5 HIV-1 Protease and the Substrate Envelope

As mentioned earlier, the HIV-1 protease is an aspartic acid homodimer composed of only 99 residues that cleaves along twelve different sites of the gag-pro polyprotein, giving rise to

structural, enzymatic, and accessory proteins. Analysis of the different cleavage sites shows that these sites are non-homologous and asymmetric in sequence, charge, and size distribution, sharing very little homology (23). The lack of sequence specificity and homology between the cleavage sites has raised an interesting question for the Schiffer laboratory; how does the protease recognize, bind, and cleave along the nonhomologous sites? In order to answer this question, one must first look at the crystal structure of the protease, shown in Figure 8 with the catalytic aspartic acid (D25) and catalytic aspartate (D25') (PDB 1T3R).
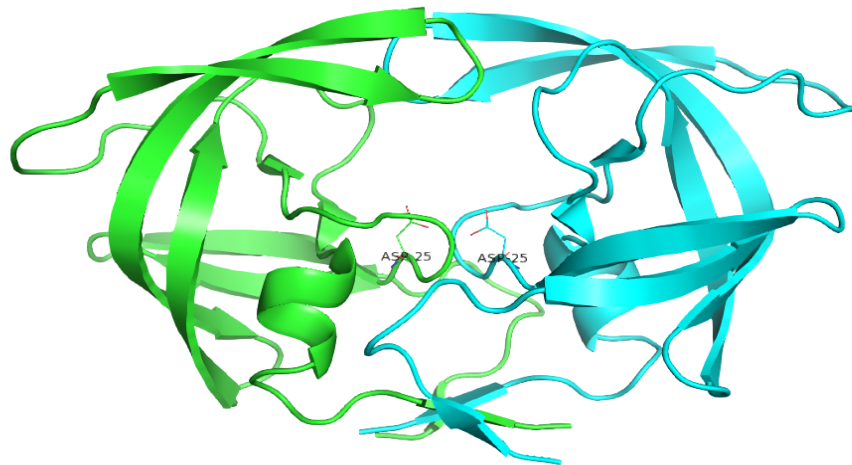

Figure 8: HIV Protease Crystal Structure

The protease possesses a C2 symmetry (180°), but despite the conserved symmetry in its structure and active site, the protease is able to recognize and cleave asymmetric polyproteins. The Schiffer laboratory has conducted numerous studies in an effort to determine what kind of specificity the protease utilizes to bind and cleave its substrate. By analyzing six complexes of HIV-1 protease that correspond to six substrate cleavage sites, the lab was able to superimpose each of the different peptides and come up with a consensus volume that includes all of the peptides. It has been strongly suggested that the protease uses shape specificity rather than sequence specificity. This consensus volume has been referred to as the substrate envelope, which is depicted in Figure 9.
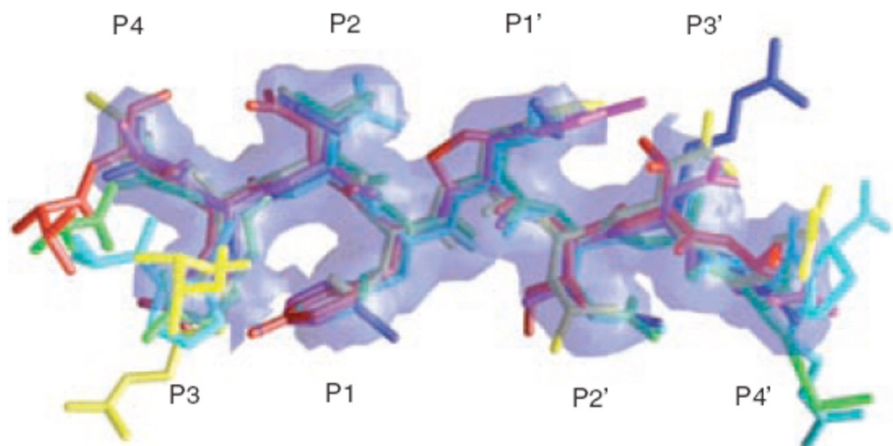
Figure 9: Substrate Envelope with Superimposed Substrates (24)

The six superimposed substrates occupy a relatively similar space inside the active site. By analyzing the relative regions that the substrates occupy, the laboratory has been able to identify and conclude that this is a highly conserved space that the protease is able to recognize and selectively cleave the substrate. By studying the substrate envelope, conclusions are able to be made regarding the specific ways that substrates bind, including where hydrogen bonds and van der Waals (vdW) interaction occur. The Figure 10 depicts the conserved hydrogen bonds that have been observed across all six of the HIV-1 protease complexes tested in the study.
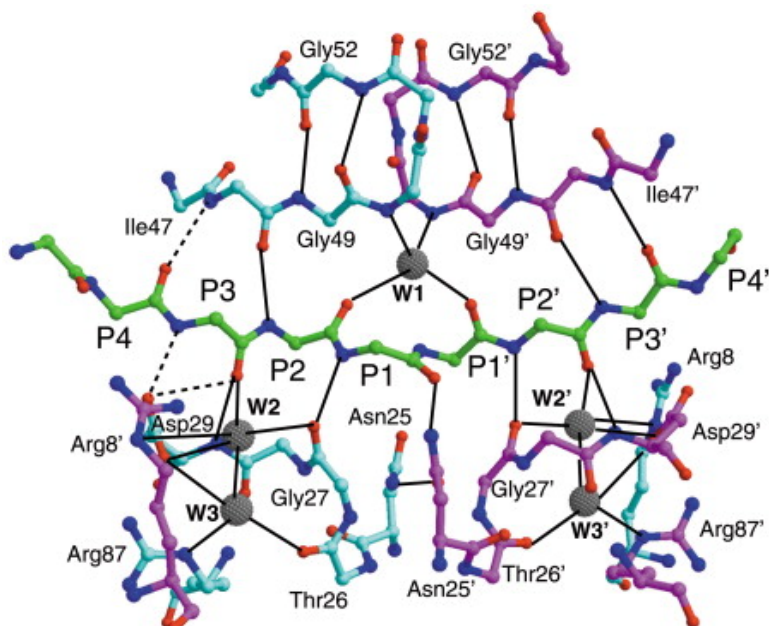

Figure 10: Hydrogen Bonding of Substrate with Active Site of Protease (1)

Most of the hydrogen bonds occur between the backbone of the protease and the backbone of the substrate residues. Eight of these are completely conserved throughout the complexes and

involve the P4-P4' sites of the substrates. The specific protease residues involved in hydrogen bonding are Asn 25', Gly 27/27', Asp 29/29', and Gly 48/48'. Furthermore, there are six partially conserved hydrogen bonds. Four of these involve the backbone atoms and two involve one of the side chains on the substrate and the amino nitrogen atom of Asp 29/29′ and Asp 30/30′. All of these hydrogen bonds are formed between the residues mentioned and the substrate peptide in a nonspecific manner. As the substrate slides through the active site and gets cleaved, new hydrogen bonds are formed with the rest of the uncleaved peptide.

1.5.1 HIV-1 Protease Mode of Action

The HIV-1 protease is a hydrolase; therefore, it requires the presence of a water molecule to facilitate the cleavage of the scissile bond. Within the active site, D25 will be a charged aspartic acid residue while D25' will be in the deprotonated form; aspartate. It is important to note that the protonation states of D25 and D25' are interchangeable. However, both residues cannot be in the aspartate form simultaneously, as this would result in unfavorable repulsions. On the flip side, both residues cannot exist in the aspartic acid form simultaneously, which render the protease unable to cleave its substrate. Figure 11 shows DRV bound to WT; D25 (green) and D25' (blue) are shown in stick form. It is important to note that the hydroxyl of DRV (black circle) orients its hydrogen towards chain B (blue). Thus, showing that D25' is in the aspartate form while D25 is in the aspartic acid form. For the remainder of this paper, we will use green to denote chain A and blue to denote Chain B.
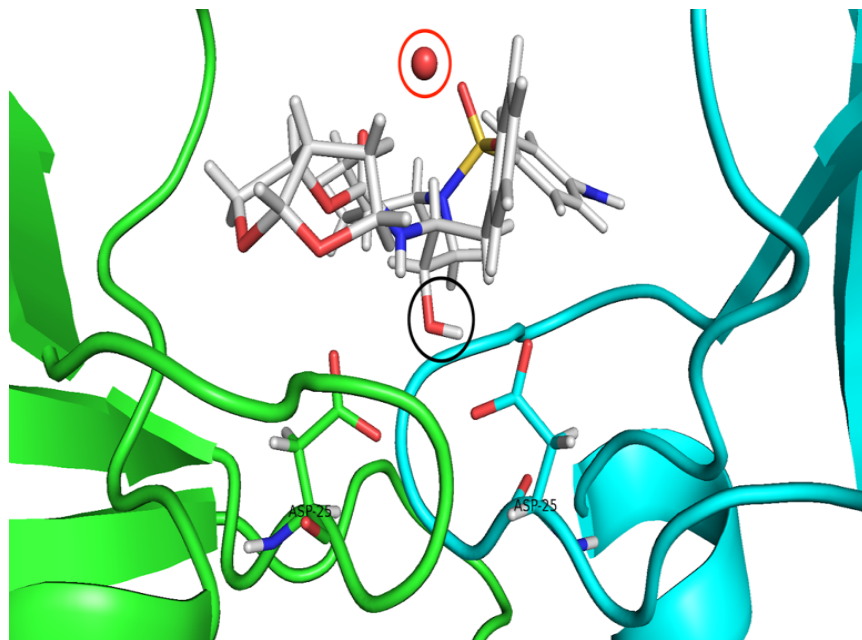

Figure 11: Chain A Versus Chain B

The conserved water molecule highlighted in the red circle shown in Figure 11 is responsible for facilitating the cleavage of the scissile bond. This water molecule is also highly conserved in all solved crystal structures of HIV-1 protease bound to an inhibitor. The proposed protease mechanism is shown in Figure 12.
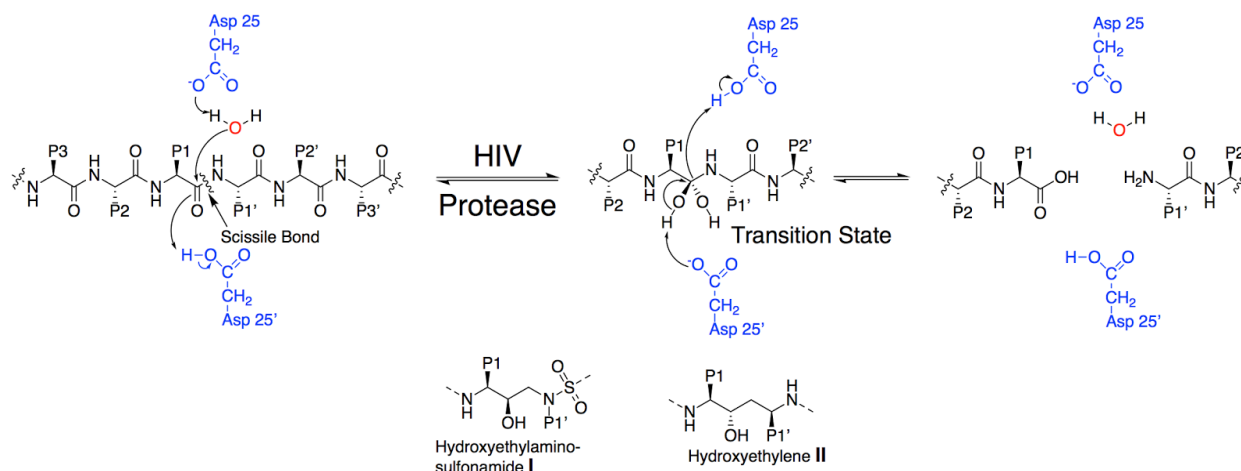

Figure 12: Proposed HIV-1 Cleavage Mechanism

The deprotonated oxygen of D25 attacks the water molecule and becomes neutral. The OH then performs a nucleophilic attack on the carbonyl of the scissile bond, which in turn takes a proton from D25'. This leads to the transition state with two hydroxyls, as shown in the middle of Figure 12. The deprotonated oxygen of D25' then retakes a proton from one of the hydroxyls, which sets up a negatively charged oxygen where one of the lone pairs of electrons collapses on the single bond, creating a double bond. This facilitates the breaking and cleavage of the scissile bond. The lone pair of the scissile nitrogen becomes protonated by taking a proton from D25, bringing D25 and D25' to their starting states, thus resetting the catalytic site for another cleavage. The end product is the C terminus of one protein and the N terminus of another.

Competitive protease inhibitors are peptidomimetics that are designed to mimic the transition state. Transition state analogs bind tightly to the enzyme by converting the short-lived natural transition state to a stable thermodynamic state (25). The inhibitors shown in this study utilize either the hydroxyethylamine-sulfonamide (I) or the hydroxyethylene (II) transition state analogs (see Figure 12). These analogs are modified in four positions; P2, P2', P1, and P1' position. More specifically, compounds are designed to mimic the interactions that occur by the natural transition states prior to cleavage. The hydroxyl group in the peptidomimetics is un-cleavable, which allows the inhibitor to be locked in the active site without getting cleaved.

<u>1.5.2 HIV-1 Protease Wild Type Variants</u>

It is important to note that HIV-1 and HIV-2 are two genetically distinct viruses. HIV-1 is the main topic of this paper. Within the subgroup of HIV-1, there are multiple strains that are genetically distinct. Figure 13 shows the four groups of HIV-1 and their subgroups. Group M is the "Major" group of HIV-1 that is mainly responsible for the current HIV-1 epidemic.
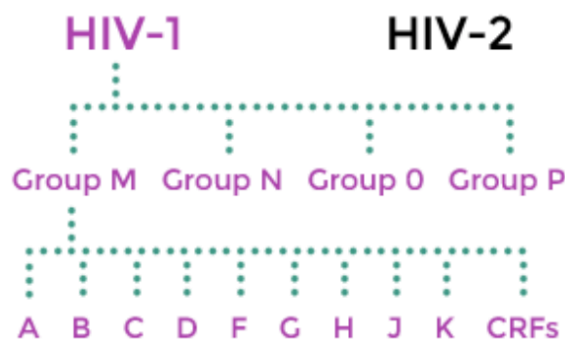

Figure 13: HIV-1 Groups and Subgroups (26)

The presence of multiple groups and subgroups gives rise to many wild type strains of the HIV-1 protease. Here we look at the two WT strains tested in this study. The Celia Schiffer wild type (CS_WT) contains the Q7K mutation, which corresponds to genotype SF2. This mutation avoids autoprotolysis. The Ron Swanstrom wild type (RS_WT) contains Q7K, K41R, P63L, and V64I, which corresponds to genotype NL4-3. The corresponding amino acid sequence and their mutations are shown in Figure 14.

CS_WT:
PQITLWKRPLVTIRIGGQLKEALLDTGADDTVLEEMNLPGKWKPKMIGGIGGFIKVRQYDQIPVEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNF
RS_WT:
PQITLWKRPLVTIKIGGQLKEALLDTGADDTVLEEMNLPGRWKPKMIGGIGGFIKVRQYDQILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNF

Figure 14: CS_WT Vs RS_WT

Throughout this study, we utilize RS constructs that contain the five baseline mutations in addition to the resistant mutation. For example, RS_I84V contains Q7K, K41R, P63L, V64I, and I84V. The CS_WT construct was utilized for comparative analysis between the wild types.

<u>1.5.3 HIV-1 Protease Mutations and Drug Resistance</u>

Designing new inhibitors using the substrate envelope hypothesis is of crucial importance considering how the protease develops drug resistance. Protease mutations will typically occur where inhibitor atoms protrude from the substrate envelope and contact protease residues. Making contact with residues beyond the substrate envelope puts selective pressure on the protease to mutate. When protease residues mutate, inhibitor binding is greatly impaired, as inhibitors are

designed to interact via hydrogen bonds with specific residues. However, substrate recognition, binding, and cleavage is not affected, since hydrogen bonding with the substrate is non-specific.

Early protease inhibitors, such as Ritonavir (RTV), were designed without the substrate envelope hypothesis. Despite this, RTV has been shown to be a potent inhibitor (K_i = 55 pM in WT/Q7K). A key feature of all protease inhibitors is that they are designed to be substrate transition state analogs with an un-cleavable hydroxyl moiety at the P1 position (see Figure 12). Although RTV is a potent inhibitor, it is not confined within the substrate envelope when compared to DRV, as shown in Figure 15 (RTV in blue and DRV in green).
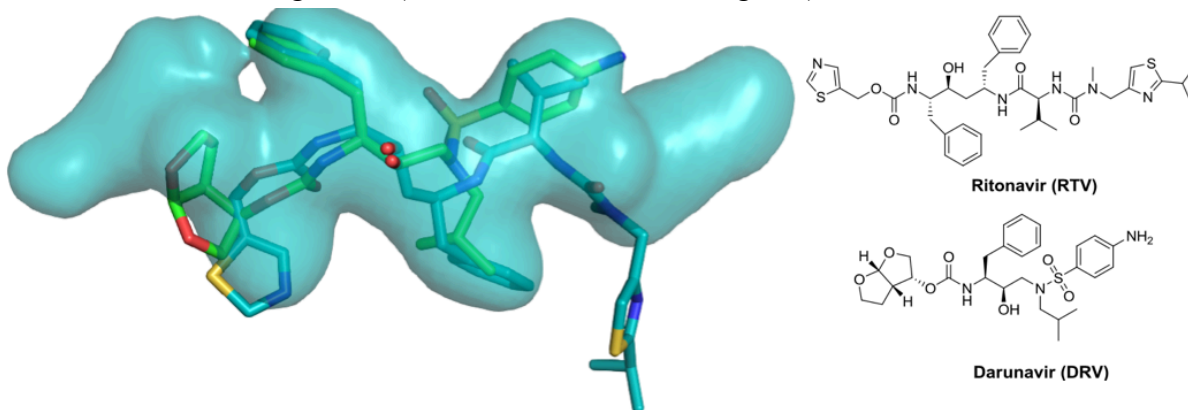

Figure 15: Ritonavir (RTV) And DRV Inside the Substrate Envelope

As a consequence of the P2' moiety protruding from the substrate envelope, patients that were treated with RTV show resistant mutations such as I82V and I84V shortly after treatment, thus failing RTV therapy. Currently, RTV is no longer prescribed as a PI due to its numerous side effects. The same mechanism allows the protease to develop resistance to many of the currently used PIs.

The use of structural-based drug design gave rise to DRV, the most clinically potent PI prescribed today (27). DRV fits fairly well inside the substrate envelope (see Figure 16) and makes many backbone interactions through its novel bis-THF moiety on the P2 position. The use of the substrate envelope hypothesis along with the bis-THF moiety makes DRV extremely potent against WT (K_i = 5 - 10 pM). Despite picomolar inhibition, DRV is still capable of inducing mutations on the protease, such as I82V, I84V, and I50V;A71V. These can greatly impair DRV's ability to competitively bind to the active site.
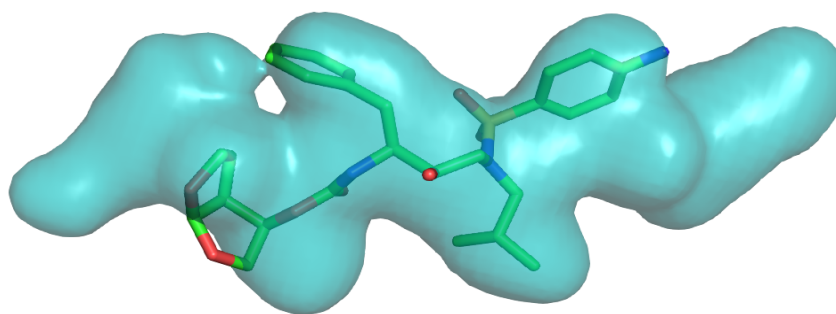
Figure 16: DRV Inside the Substrate Envelope

       The observed electrostatic interactions of DRV inside the active site, in the form of hydrogen bonds, are shown in Figure 17. Starting from the P2' position, three hydrogen bonds are observed between the backbone nitrogen of D29 and D30 with the bis-THF moiety. These bonds are of crucial importance to the measured potential of the DRV. Two carbonyl groups form a coordinated four-way hydrogen bonding network between the backbone nitrogen of I50 and I50' through a conserved water molecule. This water molecule is highly conserved along with those four hydrogen bonds because they play an important role in closing the flaps of the protease. The uncleavable hydroxyl group interacts with the catalytic D25 and D25' residues. The nitrogen atom between the P2 and P1 moieties forms a hydrogen bond with the carbonyl of the G27. Lastly, the P2' moiety forms a hydrogen bond with the backbone carbonyl of D29' and with the side chain of D30' mediated by a water molecule. It is important to note that while the P1 and P1' moieties do not form hydrogen bonds, they are important when considering Van der Waals interactions.
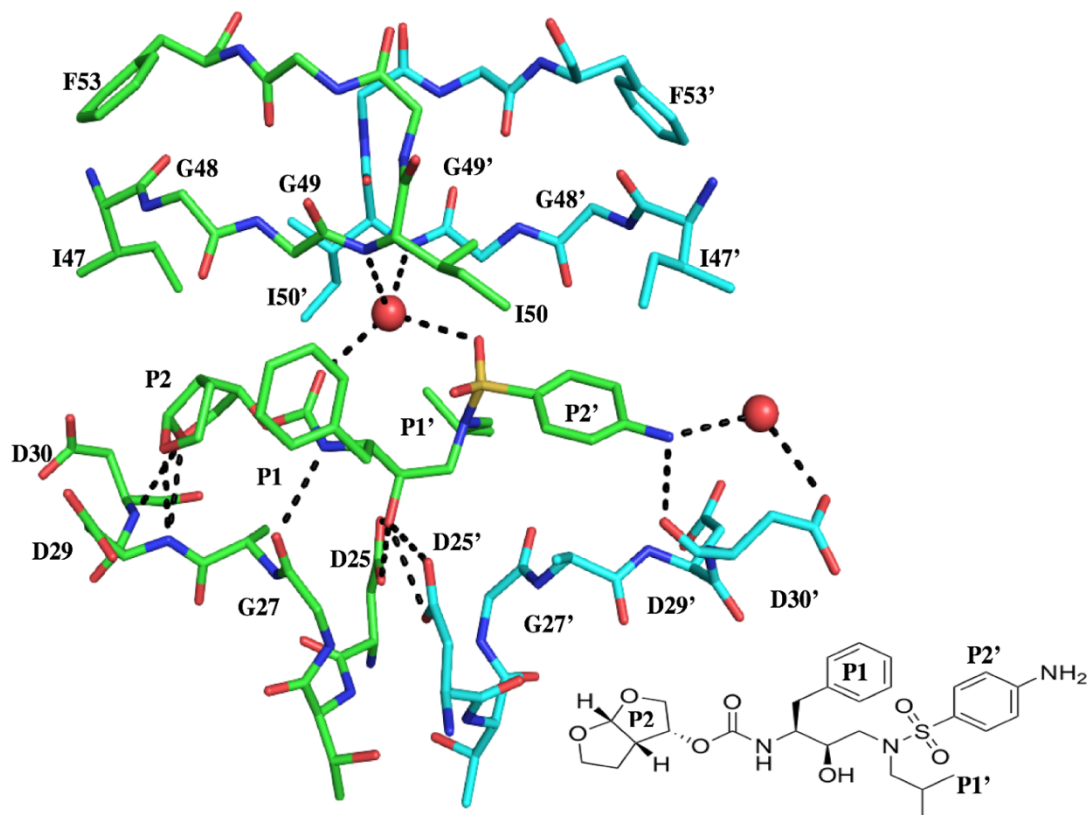
Figure 17: DRV Hydrogen Bonds Within the Active Site

When considering where to modify DRV, three possible locations arise; the P1, P1', and P2' moieties. Several modifications to the P1' and P2' moieties have been made and two series of compounds have been synthesized focused on these two positions. First, the UMass 1 - 10 compounds compared two modifications at the P1' position and five modifications at the P2' position. The UMass compounds were tested through enzymatic assays to obtain K_i data and crystal structures were solved to visualize hydrogen bonds. This analysis gave rise to the LR series. This series attempted to modify the best performing UMass compounds to obtain greater inhibition. A subdivision of the LR series includes six compounds that are a hybrid of DRV and LPV.

1.6 Evolution of Resistant Drug Design and Novel Protease Inhibitors

As mentioned earlier, the UMass compounds were made by modifying the P1' and P2' moieties. In the P1' position, compounds 1 - 5 had a 2-methylbutane group substitution, while compounds 6 - 10 had an isohexyl group substitution. It is important to note that these compounds are grouped according to their P2' moity, such that UMass 1 and 6 are a pair, UMass 2 and 7 are a pair and so on, as shown in Figure 18. UMass 1 and 6 contained the same amine group in the P2' position as DRV. The remainder of the compounds had either a methoxy, hydroxymethyl, benzodioxole, or a benzothiazole group in the P2' position. For various reasons discussed in the discussion section, DRV and UMass 3 were chosen to be further modified and give rise to the LR and LR2 series compounds.

The LR series contains a mono-hydroxyl moiety while the LR2 series contains a di-hydroxyl

Figure 18: UMass 1-10 Compounds

moiety on the P2' position, as shown in Figure 19. These compounds also experimented with the stereochemistry of the hydroxyl groups. Lastly, the P1' position was modified to contain either isobutyl, 2-methylbutane, or isohexyl groups. These modifications were made to study the effects of adding extra methyl groups at the P1' position. These modifications gave rise to twelve compounds shown in Figure 19. Like the UMass compounds, the LR/LR2 series were synthesized by corresponding substitutions, illustrated by the bold lines.
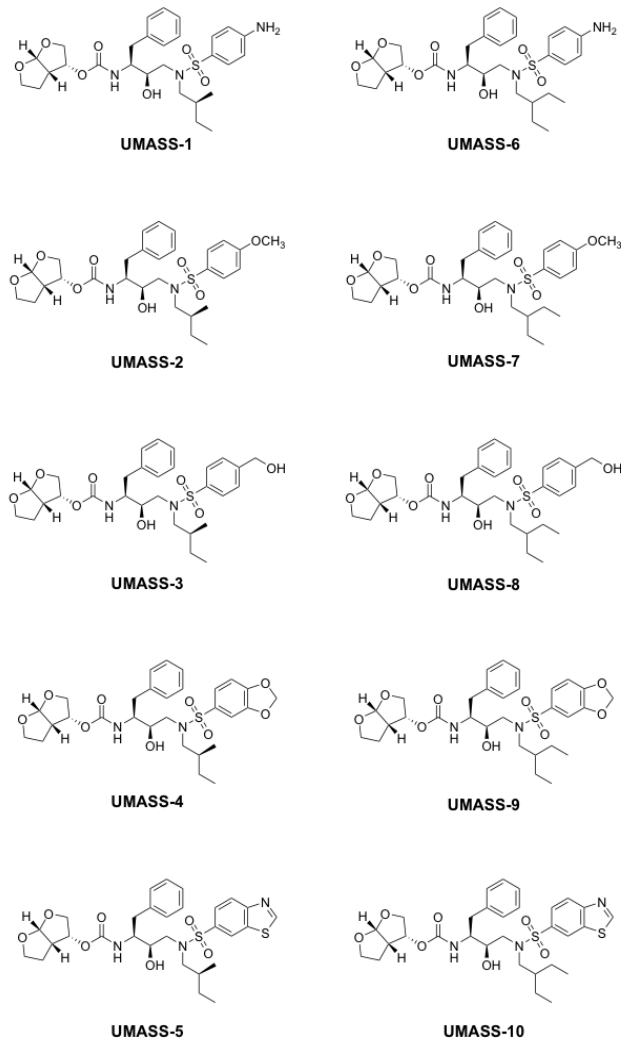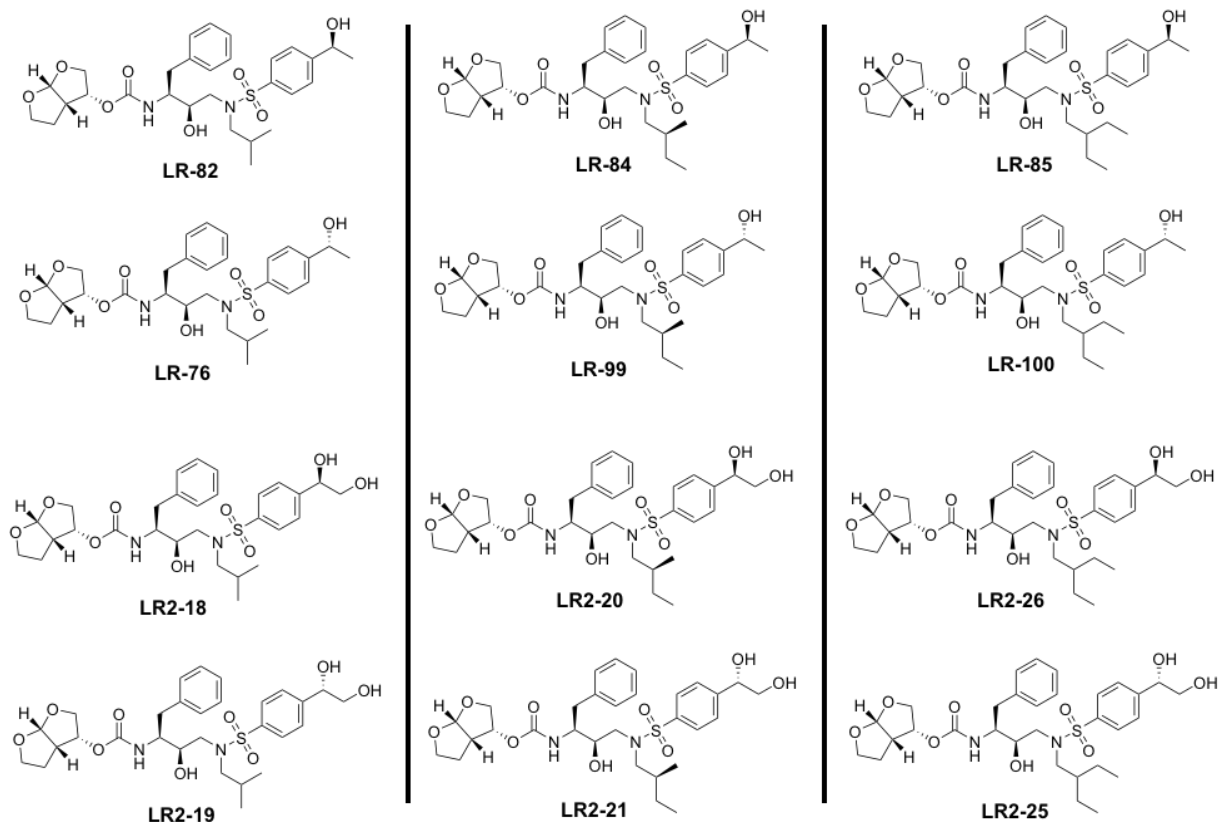
25

Figure 19: LR and LR2 Series

The LR2 series was further expanded to include six new compounds that are hybrids between LPV and DRV. These compounds are shown in Figure 20. Once again, these compounds correspond as pairs, based on their P1' moiety.
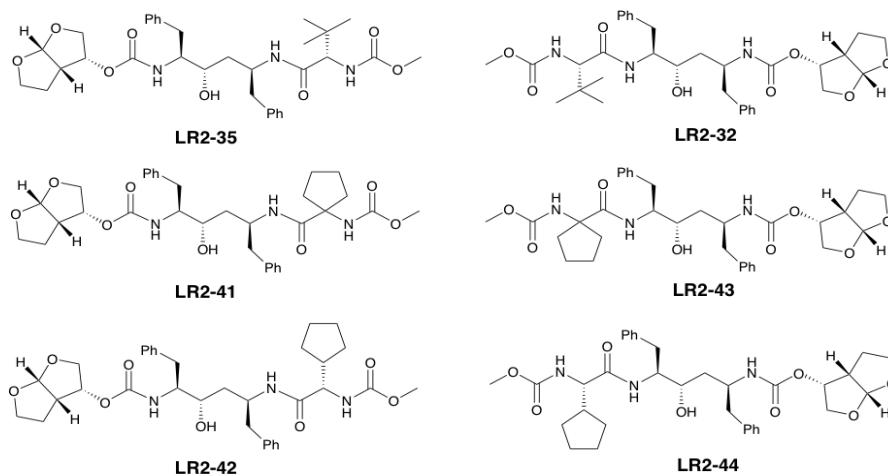

Figure 20: LR2 Series Continued - LPV/DRV Hybrids

The bis-THF moiety of DRV has shown to make highly favorable interactions with the backbone residues of the active site, making for a valid argument and idea to see if substitution of this group into older compounds would be beneficial to their inhibition potency. Thus, the

motivation behind the hybrid compounds was to study if first generation compounds can be modified with the bis-THF moiety of DRV to increase their inhibition.

1.7 Experimental Design

In order to qualitatively determine inhibition potential, computer modelling and chemical properties for each compound, such as ClogP and Molecular Weight (MW), were determined. Computer programs such as ChemDraw and Maestro were utilized to carry out initial analysis. By using already solved crystal structures, the P1' and P2' positions of already bound ligands can be modified into compounds of interest in Maestro and the type of interactions made can be determined if they are favorable or unfavorable. Furthermore, chemical properties can be calculated in ChemDraw and used to analyze how well a compound would be able to enter the cell (ClogP) in a natural environment and how much a compound weighs (MW).

All of the above compounds presented promising results, so the next step in determining their potency was carrying out kinetic inhibition assays in order to determine the inhibition constant (K_i) for each compound. Once the K_i's were determined, they were compared to the K_i of DRV and those closest to the potency of DRV were selected for further studying. Crystal structures of compounds of interest in multiple protease constructs were solved. Furthermore, Van der Waals (vdW) interactions for each inhibitor-protease construct were calculated to learn about hydrophobic interactions between the inhibitor and the protease. Lastly, Molecular Dynamics (MD) simulations were carried out to understand the inhibitor-protease interaction in a dynamic state versus a static state (i.e. crystal structures). The MD simulations were utilized to supplement vdW, K_i, and structural data. All of this analysis combined can be studied and investigated in order to understand how and why specific compounds inhibit the protease the way they do and how protease structure varies between compounds and between mutants.

As previously mentioned, protease mutations greatly affect inhibitor binding, yet substrate binding and cleavage is able to occur without any visible effects. An important baseline kinetic parameter is the apparent affinity constant (K_m), which is a measure of the affinity the enzyme has for its substrate. The K_m values for each variant tested had not been determined previously. Determining K_m values of the mutants and comparing those values with the K_m of WT protease would inform us if the affinity for substrate for each mutant is affected by the specific mutations. In order to determine these K_ms, a K_m assay was carried out for each variant. All of the data obtained for each compound and mutant was crucial in characterizing the ability of the designed

compounds to inhibit protease activity, how each mutant behaves under inhibitor pressure and what unique behavioral and structural features each mutant possesses.

Chapter 2: Materials and Methods

<u>2.1 K$_m$ Assay</u>

      K$_m$ assays were carried out on all protease variants of interest to calculate the apparent affinity constant, K$_m$. The fluorogenic substrate shown in Figure 21, contains the same FRET pair as the K$_i$ assay substrate. K$_m$ assay substrate is a natural sequence (MA/CA), however it is not optimized like the K$_i$ assay substrate. This substrate does not need to be optimized and must be a natural sequence to calculate a biologically relevant K$_m$ value.
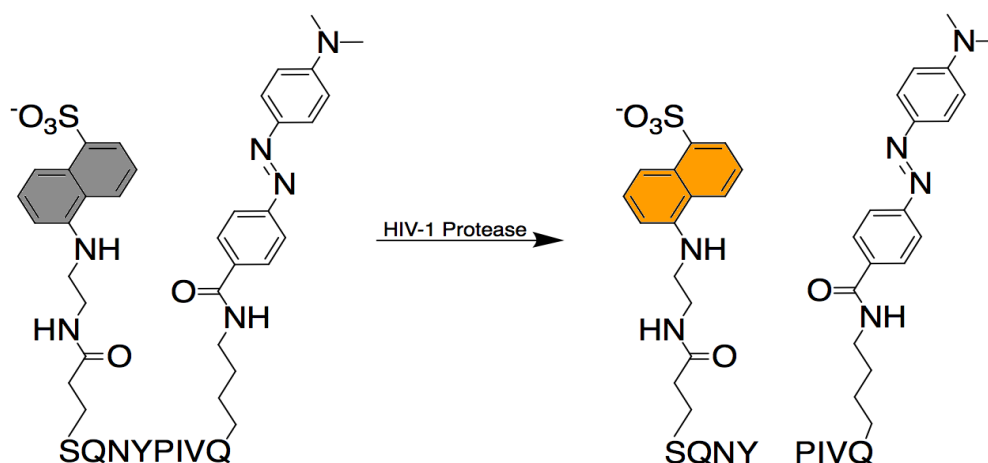
Figure 21: K$_m$ Assay Substrate

<u>2.1.1 Determining Km Values</u>

      K$_m$ assays were done in nonbinding surface 96-well black half-area plates. All assays were conducted in 6% DMSO for wells 1-11 and 8% DMSO for well 12 with a total reaction volume of 60 μL. Each plate was used to test one protease construct in triplicates. Plate setup schematic is shown in Figure 22. Increasing substrate concentration (0-40 μM) in 2X assay buffer [100 mM sodium Acetate and 200mM sodium chloride] and the appropriate DMSO concentrations were centrifuged [1000 x g for one minute] using a plate centrifuge at 20°C. This assay did not have an incubation period. The reaction was initiated by the addition of 5 μL of 10 nM of HIV-1 protease (RS_WT, CS_WT, RS_I82V, RS_I84V, RS_I50V, and RS_I50V;A71V). Fluorescence was monitored using a PerkinElmer EnVision plate reader (excitation at 340 nm, emission at 492 nm). Substrate concentration points were globally fitted to the Michaelis-Menten equation to obtain the K$_m$ value of the protease constructs.
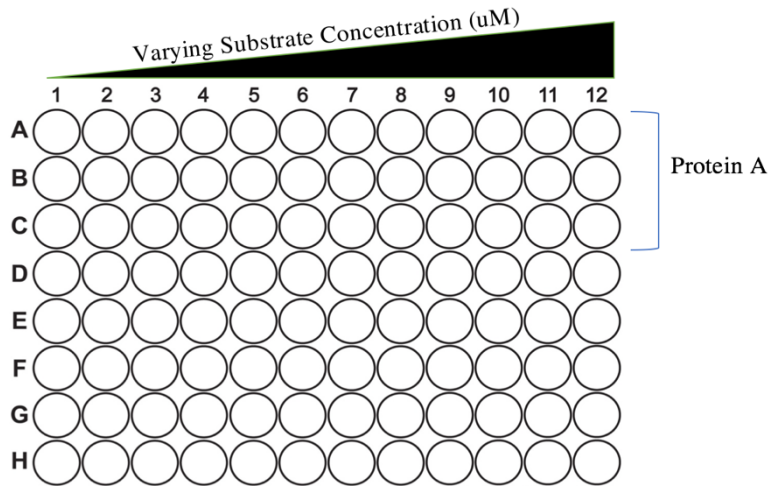
Figure 22: K$_m$ Assay 96-Well Plate Setup

Progression curves were generated for each triplicate across the twelve substrate concentrations. The progression curves were inputted into a log equation to obtain the initial velocity of the reaction. The initial velocity was then used to calculate the K$_m$ using the Michaelis-Menten equation shown below.

$$\frac{V_i}{V_0} = 1 - \frac{\left([E]_T + [I]_T + K_i^{app}\right) - \sqrt{\left([E]_T + [I]_T + K_i^{app}\right)^2 - 4[E]_T[I]_T}}{2[E]_T} \qquad \textit{(Equation 1)}$$

An example of progression curves and a K$_m$ graph is shown in Figure 23. The calculations and following graphs were generated using Prism 7 software.
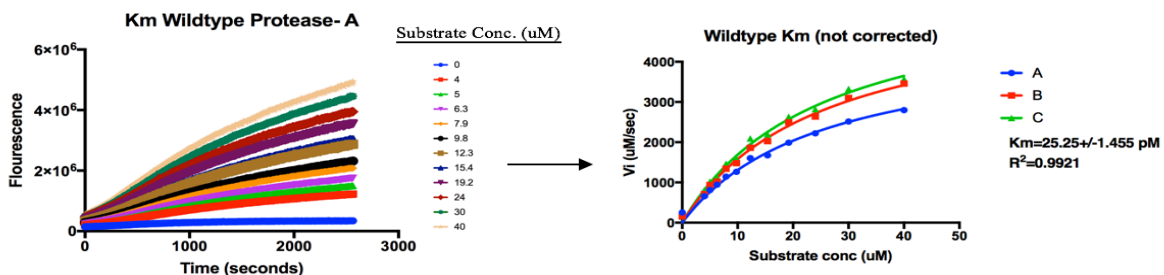
Figure 23: Processing K$_m$ Data

2.1.2 Correcting for the Inner Filter Effect

Utilizing fluorescence change for kinematic assays is convenient for monitoring enzyme kinetics. However, fluorescence loses linearity and therefore, accuracy, with high substrate concentrations. Referred to as "quenching", the fluorophore is overcrowded by the free-floating quencher once cleave occurs, leading to significantly reduced emissions for substrate

concentrations over 20 μM (28). This is known as the inner filter effect and can alter enzymatic assay results. As seen in the resulting graph in Figure 23 above, the reaction curve seems to level off quite early in the reaction. The value of K_m is biochemically determined to be half of V_{max}, and this is not represented by the graph. A correction assay for higher concentrations of substrate was carried out in order to provide a ratio of the correction value needed to be applied to the obtained results for the K_m assays. This ratio increases the values of the observed K_ms, leading to an accurate estimate of the K_ms of various mutants.

A solution of 6% DMSO was prepared. Three rows of a 96-well plate, each containing 12 wells, were utilized for this assay. 27.5 uL of 6% DMSO were added to each of the wells. Increasing substrate concentration (0-40 μM) in 2X assay buffer [100 mM sodium Acetate and 200mM sodium chloride] were serially diluted into each of the rows. A popular donor for developing FRET pairs called EDANS was used for this assay. 5 uL of 6 uM EDANS in 2% DMSO was suspended into each well and the fluorescence reading of the plate was recorded five times. Fluorescence was monitored using a PerkinElmer EnVision plate reader (excitation at 340 nm, emission at 492 nm). After a series of calculations, the obtained correction ratios were determined and are depicted in the results section. These ratios were then applied to the obtained initial velocities from each K_m assay for each substrate concentration. The results of the corrected K_m values are shown in the results section.

2.2 HIV-1 Enzyme Inhibition Assays (K_i)

Enzyme inhibition assays were carried out on all inhibitors of interest to calculate the inhibition constant, K_i. Inhibitors such as DRV and all of the tested inhibitors are tight binding inhibitors. For that reason, an assay of high sensitivity is needed to accurately calculate a K_i. The fluorogenic substrate, shown in Figure 24, is a highly optimized substrate that was used in all K_i assays presented in this paper. The Fluorescence Resonance Energy Transfer (FRET) pair of the substrate consists of the fluorophore (EDANS) and the quencher (DABCYL). In un-cleaved substrate molecules, the DABCYL is in close proximity to the EDANS, thus "quenching" EDANS fluorescence. When the substrate is cleaved by the HIV-1 protease between the phenylalanine and the leucine, EDANS and DABCYL separate. As a result, EDANS fluorescence can be measured as a function of time. The arginine residues on either end of the substrate increase its solubility in solution.
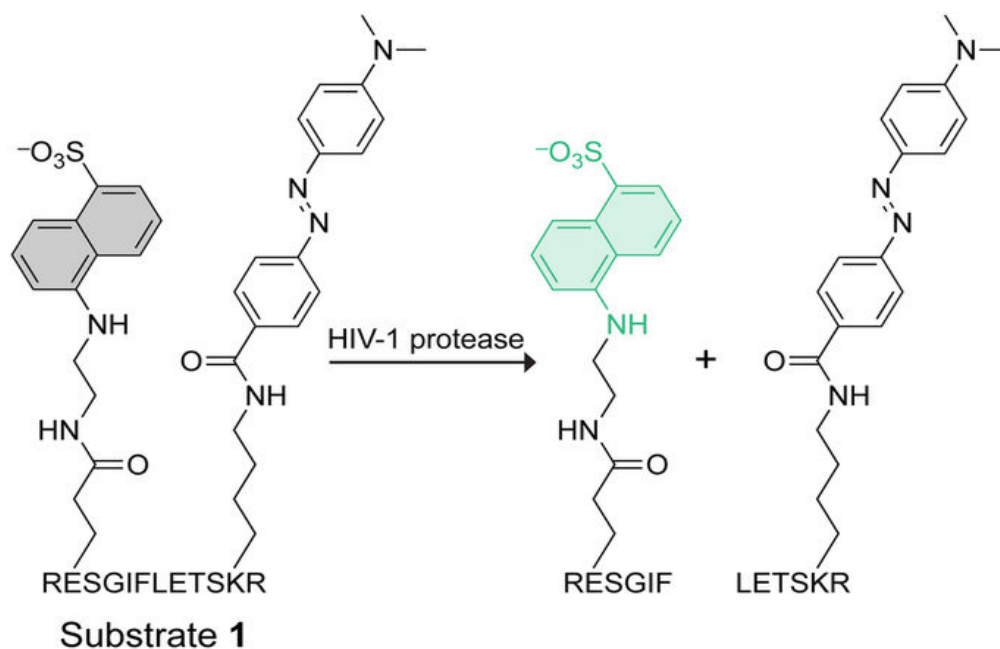
Figure 24: Fluorogenic Assay Substrate (29)

K_i assays were done in nonbinding surface 96-well black half-area plates. All assays were conducted in 4% DMSO with a total reaction volume of 60 μL. Each plate was used to test two inhibitors in triplicates. The last two rows were reserved for a DRV control, which was ran as a replicate. Inhibitor concentration were varied in each lane of the well, ranging from zero inhibitor concentration to the desired maximum inhibitor concentration in either ⅔ or ½ dilutions. Concentrations of inhibitor were optimized for the specific inhibitors and mutants being tested. Plate setup schematic is shown in Figure 25. For each assay, 5 mL of a 0.77 nM HIV-1 protease (WT, V82I, I84V, I50V, and I50V;A71V) was prepared by a series of two dilutions. First, the concentration of the stock protein was measured using the A_{280} and Beer's Law. Then, the stock protein was diluted down to 500 nM by adding enough 2X assay buffer [100 mM sodium Acetate and 200mM sodium chloride] to yield 500 μL. The protein was diluted once more to yield 5 ml with a final concentration of 0.77 mM. 27.5 μL of 0.77 nM protein were added to each well. After all the components have been added to the plate, the plate was centrifuged at 1000 x g at 20°C for one minute using a plate centrifuge. The plate was then preincubated at room temperature for 1 hour. After the incubation period, the reaction was initiated by the addition of 5 μL of the HIV-1 optimized substrate to a final concentration of 10 μM. Fluorescence was monitored using a PerkinElmer EnVision plate reader (excitation at 340 nm, emission at 492 nm). The triplicates for each tested inhibitor contained 12 inhibitor concentration point. The concentration points were

globally fitted to the Morrison K_i equation for tight binding inhibitors to obtain the K_i value of the inhibitor for each triplicate to a specific protease construct.
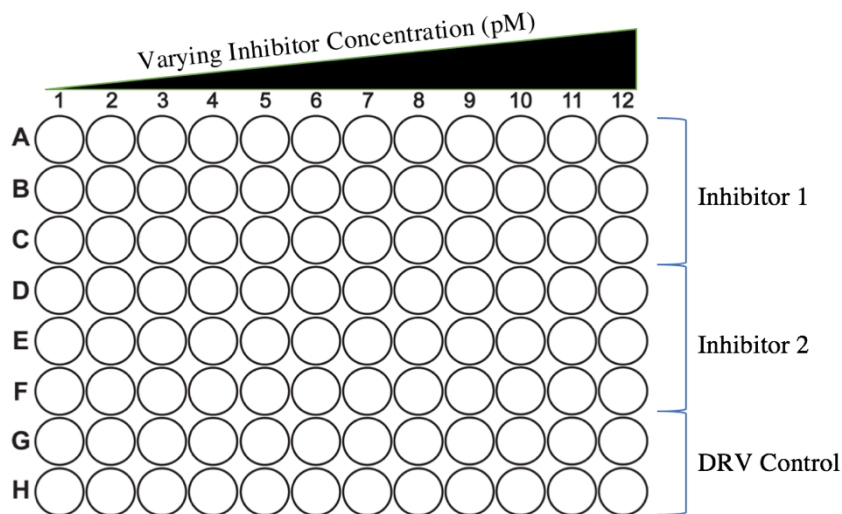

Figure 25: K_i Assay 96-Well Plate Setup

Progression curves were generated for each triplicate across the twelve inhibitor concentrations. The progression curves were inserted into a log equation to obtain the initial velocity of the reaction. The initial velocity was then used to calculate the K_i using the Morrison K_i equation for tight fitting inhibitors. An example of progression curves and a K_i graph is shown in Figure 26. The calculations and following graphs were generated using GraphPad Prism 7 software. For a more detailed schematic of the K_i assay calculation procedure, see Appendix B.
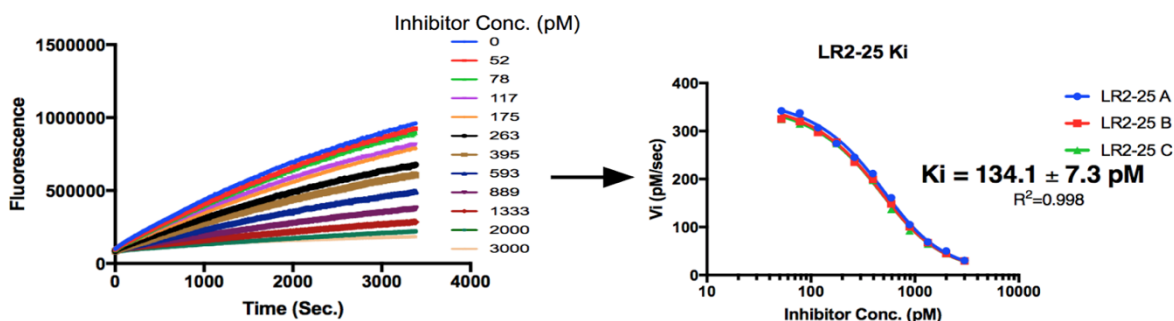
Figure 26: Processing K_i Data

Why Do We Use the Morrison K_i Equation?

Regular inhibitors require high concentrations relative to the total enzyme concentration in order to inhibit enzyme activity. Tight-binding inhibitors, on the other hand, require concentrations relatively similar to the concentration of total enzyme in order to exhibit high inhibition of enzyme activity. Competitive inhibitors facilitate rapid formation of the Enzyme-Inhibitor (EI) complex.

33

This complex is maintained considerably longer by tight-binding inhibitors compared to regular inhibitors, so the concentration of the EI complex in solution is no longer negligible compared to the overall concentration of I. This adds another variable to the system as a whole and to the determination of K_i. The process of enzyme-substrate and enzyme-inhibitor interactions with tight binding inhibitors is detailed in the schematic below.
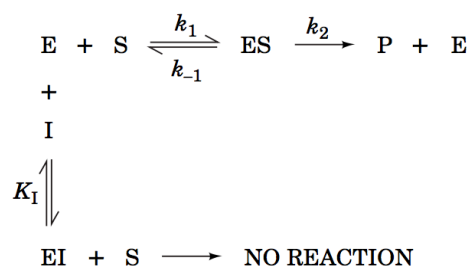
$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} P + E$$
$$+$$
$$I$$
$$K_I \updownarrow$$
$$EI + S \longrightarrow NO\ REACTION$$

Figure 27: Schematic of Enzyme Kinetics

When a competitive inhibitor binds an enzyme, it blocks the substrate from binding to the enzyme's active site and undergoing catalysis. For regular inhibitors, high inhibitor concentrations are required in order to observe a considerable EI concentration. Due to the high inhibitor concentration relative to total enzyme concentration, the change in total inhibitor concentration [I] during EI complex formation is negligible, so it is assumed that total inhibitor concentration does not change. This is demonstrated by the equation below.

$$v = \frac{V_{max}[S]}{K_m\left(1 + \dfrac{[I]}{K_{i_a}}\right) + [S]}$$

(Equation 2) (30)

This equation does not consider a change in [I] and it neglects the EI complex. With tight binding inhibitors, constant [I] cannot be assumed since total enzyme and inhibitor concentrations are relatively similar. Free inhibitor concentration decreases with the formation of the EI complex, indicating a direct correlation. For this reason, Morrison and colleagues determined a kinetic equation that accounted for a change in [I] when the EI complex is formed (Equation 2). Modern softwares, like GraphPad Prism, utilize this equation for rapid K_i calculations given specific constraints. These constraints include a known E_t, K_m, and substrate concentration, [S]. This

equation takes into account the experimentally determined initial speed of the reaction in the absence of the inhibitor (v_o), the determined speed of the reaction under the influence of the inhibitor (v), the level of fluorescence (Y) measured by the Envision, and inhibitor concentration (X). Using all of these specifications, the program is able to extrapolate an accurate K_i value.

$$Y = \frac{V_o*(1 - (E_T + X + Q)\text{-}\sqrt{(E_T + X + Q)^2 - (4 * E_T * X))}}{2 * E_T}$$

$$Q = K_i * \left(1 + \frac{S}{K_m}\right) \qquad\qquad \textit{(Equation 3) (29)}$$

2.3 Crystallography

Crystals were obtained through the hanging drop method. In order to co-crystallize the protein, bound to an inhibitor of interest, three- to five-fold molar excess of inhibitor to protease were initially incubated at 4°C overnight, then incubated at room temperature until rod-like crystals were observed. The final concentration of protease was between 1-2 mg/mL. A 2:1 ratio of inhibitor–protein volumes were combined to set up hanging drops of 5 µL. The reservoir solution consisted of 23-24% (w/v) Ammonium Sulfate with 0.1M Bis-Tris-Methane Buffer at pH 5.5. Crystals were grown at room temperature and were evident within 24–72 hours. The crystals used for data collection were transferred into a cryoprotectant containing 25% glycerol, mounted in the Mitegen Micromounts and flash-frozen over a nitrogen stream. Intensity data for 13c wild-type protease complex were collected at −80 °C on an in- house Rigaku X-ray generator equipped with an R-axis IV image plate. Then 360 frames were collected per crystal with an angular separation of 0.5° and no overlap between frames. Crystals of all complexes were of the $P2_12_12_1$ space group, with one dimer per asymmetric unit.

2.4 Van der Waals

Van der Waals (vdW) interactions were determined by running a computer script that is part of the Schrodinger family of programs. The output of the script is an excel file that displays the Lennard Jones Potential, in Kcal/mol, for each residue in the protease, calculated by the following equation:

$$V_{(r)} = 4\epsilon[(\sigma/r)^{12}-(\sigma/r)^6 \qquad\qquad \textit{(Equation 4) (31)}$$

where
- V = the intermolecular potential between the two atoms or molecules
- ϵ = the well depth and a measure of how strongly the two particles attract each other.

- σ = the distance at which the intermolecular potential between the two particles is zero. σ gives a measurement of how close two nonbonding particles can get and is thus referred to as the vdW radius. σ is equal to one-half of the internuclear distance between nonbonding particles
- r is the distance of separation between both particles (measured from the center of one particle to the center of the other particle)
- $A = 4\epsilon\sigma^{12}$, $B = 4\epsilon\sigma^{6}$

Equation 4 is often expressed as:

$$V_{(r)} = (A/r^{12}) - (B/r^{6})$$

(Equation 5) (31)

The script used by the lab can calculate the Lennard Jones Potential for all residues and ignores residues with a vdW value under a certain cutoff value. The output can then be sorted by chain and residue number and plotted using the Prism 7 software. The script can be run for all constructs of interest and the vdW values for multiple constructs can be compared using Prism 7.

Chapter 3: Results

The HIV-1 protease is arguably the most important therapeutic drug target in the fight against HIV/AIDS. The protease's importance to viral maturation means that effective and efficient inhibition would essentially stop the production of new virions and the maturation of immature ones. Despite its incredibly sensitive task of recognizing, binding, and cleaving 12 non-homologous substrates, the HIV-1 protease is capable of mutating up to half its amino acid sequence and still recognize and cleave its natural substrate. To understand how mutations effect natural substrate cleavage, K_m values were obtained using a natural MA/CA substrate with a FRET pair attached. The K_m values for several HIV-1 protease variants are shown in Table 2. The results of the K_m assays show that drug resistant mutations such as I84V, V82I, and I50V;A71V retain the same affinity to substrate as the WT. The assays helped demonstrate the compensatory nature of the A71V mutation; the I50V mutant alone shows a K_m of 2620 µM while the I50V;A71V mutant has a K_m of 73.2 µM. The K_m values are only suggestive of how the variants bind to the natural substrate. To understand how these inhibitors bind to the variants, the K_i value for each individual inhibitor-variant combination must be calculated. For a more details on the K_m results, including raw data, see Appendix A.

Table 2: HIV-1 Protease Variants K_m Values

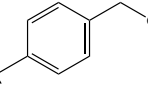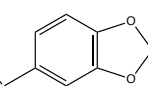| Variant	Corrected K_m (µM)
RS_WT	62.4 ± 4.9
CS_WT	55.9 ± 6.09
RS_I84V	66.4 ± 4.343
RS_V82I	61.7 ± 4.39
RS_I50V	2620 ± 6566
RS_I50V;A71V	73.2 ± 9.08
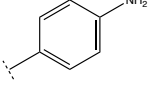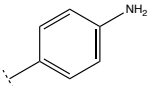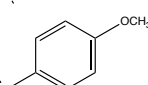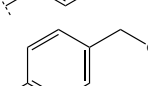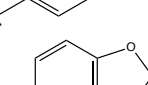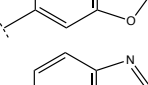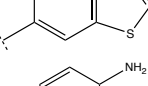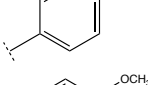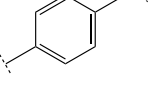
The high potency of DRV and its pico-Molar K_i in HIV-1 WT protease and multiple drug resistant mutants was a revolutionary step in protease inhibitor design. However, development of drug resistant mutants to even the most potent drug on the market still remains a concern. The group at the Schiffer Laboratory is focused on optimizing the DRV backbone at the P1' and P2' positions in hopes to discover an even more potent protease inhibitor. The main design criteria for the UMass 1-10 compounds, LR/LR2 series, and DRV/LPV hybrids were: (1) fits well inside the substrate envelope; (2) based on the backbone of an already known tight fitting inhibitor; (3)

optimized P1' and P2' to maintain contacts and increase binding potentials with evolutionarily favorable residues of the protease, especially in multi-drug resistant mutants; and (4) the compound binds at sub-picomolar affinity and presents favorable viral passing data. At the beginning of this study, premade compounds were tested for their respective K_i values in various HIV-1 protease constructs and crystal structures were solved.

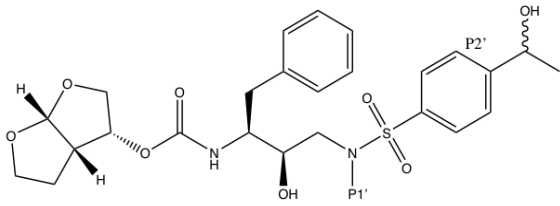K$_i$ Calculations and Structural Data for UMass 1-10 and LR/LR2 Series

A series of compounds premade by the Schiffer laboratory were tested against mutants of interest for their K_i values. Table 3 shows the calculated K_i values of compounds UMass 1-10 in RS_I84V and RS_I50V;A71V.
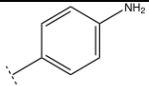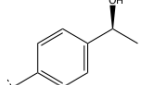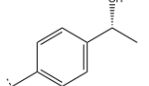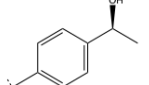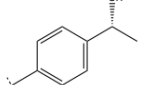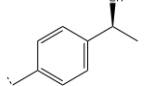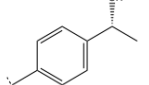
Table 3: UMass 1-10 K$_i$ Values in RS_I84V and RS_I50V;A71V

Inhibitor	P1'	P2'	Ki – I84V (pM) (fold Change)	Ki – I50V;A71V (pM) (fold Change)
DRV			25.0 ± 3.6 (5)	67.5 ± 4.9 (14)
UMass 1			26.1 ± 3.7 (5)	110.3 ± 5.9 (22)
UMass 2			< 5	15.1 ± 2.7 (3)
UMass 3			9.9 ± 2.7 (2)	79.9 ± 5.9 (16)
UMass 4			10.5 ± 1.8 (2)	32.9 ± 3.0 (6.7)
UMass 5			7.0 ± 1.7 (1.4)	7.8 ± 0.9 (1.6)
UMass 6			12.8 ± 3.1)2.6)	100.0 ± 9.9 (20)
UMass 7			12.1 ± 4.5 (2.4)	18.2 ± 3.0 (3.6)
UMass 8			< 5	55.4 ± 4.0 (11)
UMass 9			7.6 ± 1.6 (1.5)	42.3 ± 2.6 (8.5)
UMass 10			14.3 ± 9.3 (3)	5.8 ± 1.1 (1.2)

The K_i results, in combination with the solved crystal structures of the UMass compounds, led the Schiffer Laboratory to further modify the hydroxyl P2' moiety of UMass 3 and 8. The modification gave rise to the LR series which contains a mono-hydroxyl P2' moiety as well as the LR2 series which contains a di-hydroxyl P2' moiety (see Figure 19). It is important to note that the P1' moieties of UMass 3 and 8 were utilized as well as the DRV P1' moiety. Table 4 shows the calculated K_i values of the LR series in RS_I84V, while Table 5 shows the calculated K_i values of the LR2 series in RS_I84V.

Table 4: LR Series K_i Values in RS_I84V

Inhibitor	P1'	P2'	Ki – I84V (pM)
DRV			25.0 ± 3.6
LR-82			48.5 ± 3.8
LR-76			108.9 ± 7.0
LR-84			110.6 ± 7.7
LR-99			94.9 ± 4.9
LR-85			19.9 ± 1.8
LR-100			28.3 ± 2.9

The determined K_i's for these compounds showed promising results for LR-85 and LR-100, both ranging right around DRV's affinity (19.9 and 28.3 pM, respectively). For this reason, crystal structures of these compounds in complex with RS_WT were solved to visualize how they bind within the active site. It is important to note that while crystal structures are usually solved in the RS_WT construct, K_i data is usually shown for the RS_I84V construct. This is due to the fact

that in RS_WT, all inhibitors are extremely potent and observing fold differences between inhibitors is sometimes not possible. Also, the interactions made and maintained in the two constructs are virtually identical between the two mutants, as the mutation on residue 84 does not affect active site binding. For these reasons, K_i calculations in RS_I84V are primarily carried out during initial testing to observe fold differences between inhibitors.
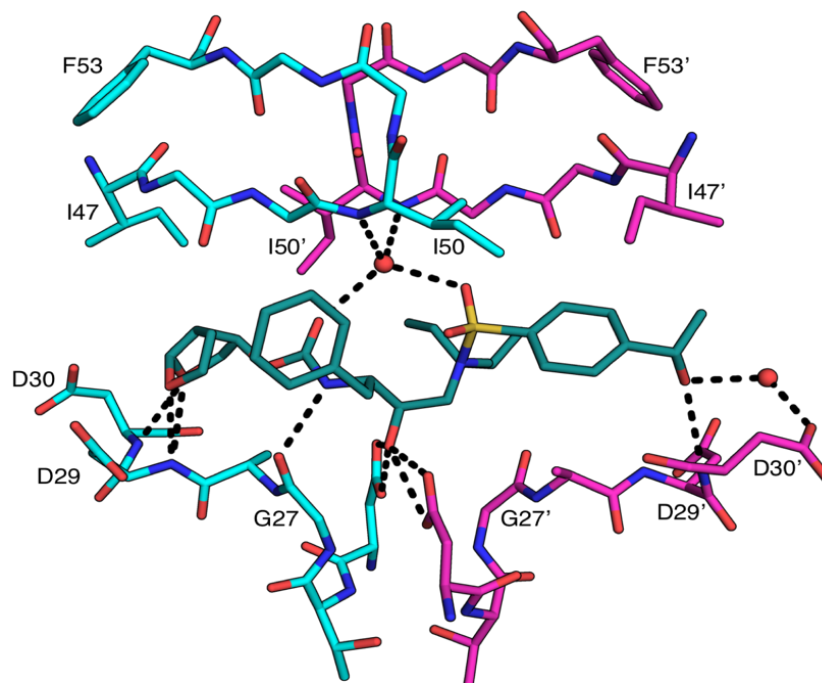
Figure 28: LR-85 In Complex with RS_WT

Overall, LR-85 shows similar binding as DRV. The same hydrogen bonds that DRV creates in the active site are made and maintained by LR-85. The hydroxyl substitution at the P2' position seems to maintain the same backbone and water-mediated bonds that are seen by the amino group on DRV with residues D29' and D30'. The un-cleavable hydroxyl group is nicely contacting D25 and D25', and the P2-bis-THF group is not disturbed, hovering right over and contacting D29 and D30.
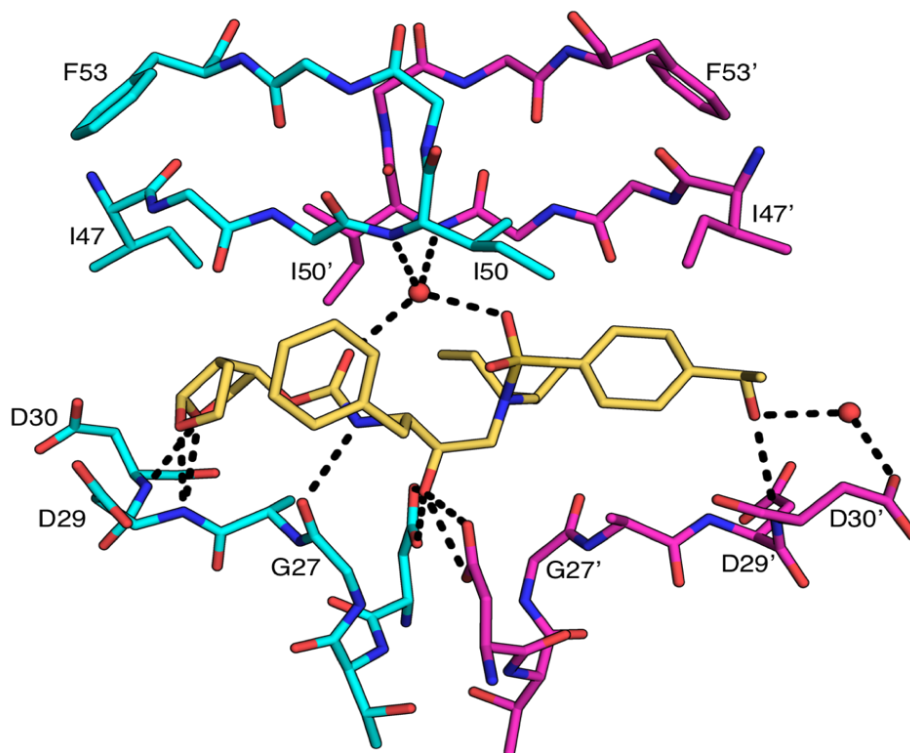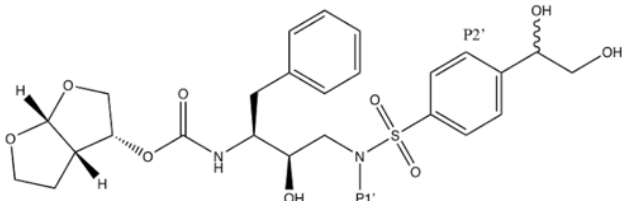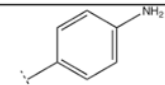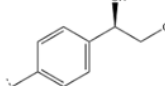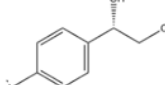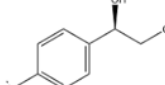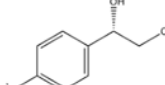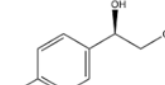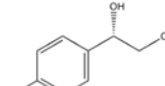
Figure 29: LR-100 In Complex With RS_WT

  The same results observed with the LR-85 crystal structure are observed with LR-100. Visibly identical interactions are made and maintained in the active site, even though the stereochemistry of the P2' hydroxyl group is reversed in LR-100. Replacing the P2' amino moiety with a hydroxyl extension does not seem to disturb any of the interactions made by DRV, but also does not contribute to more interactions within the active site.

  Based on the results obtained from the LR series, and in an attempt to increase interaction within the active site, the group designed the LR-2 series where a second hydroxyl group was added to the P2' extension and the stereochemistry of this group was varied. A total of six compounds maintaining the same P1' variables as the last set were synthesized and tested for their K_i's. The results are shown in Table 5 below.

Table 5: LR2 Series K$_i$ Values in RS_I84V

Inhibitor	P1'	P2'	Ki – I84V (pM)
DRV			25.0 ± 3.6
LR2-18			98.5 ± 8.2
LR2-19			283.1 ± 24.2
LR2-20			216.2 ± 11.0
LR2-21			504.0 ± 39.4
LR2-26			57.4 ± 4.0
LR2-25			134.1 ± 7.3

Of these compounds, LR2-26 showed equipotent inhibition to DRV, resulting in a K$_i$ of 57.4 ± 4.0 pM, a two-fold increase from DRV. Based on the LR compounds, the lab had hypothesized that adding an additional hydroxyl group to the P2' moiety would increase interactions within the active site, leading to a better experimental K$_i$. In order to analyze why this was not the case, a crystal structure of LR2-26 in the RS_WT construct was solved (see Figure 30).
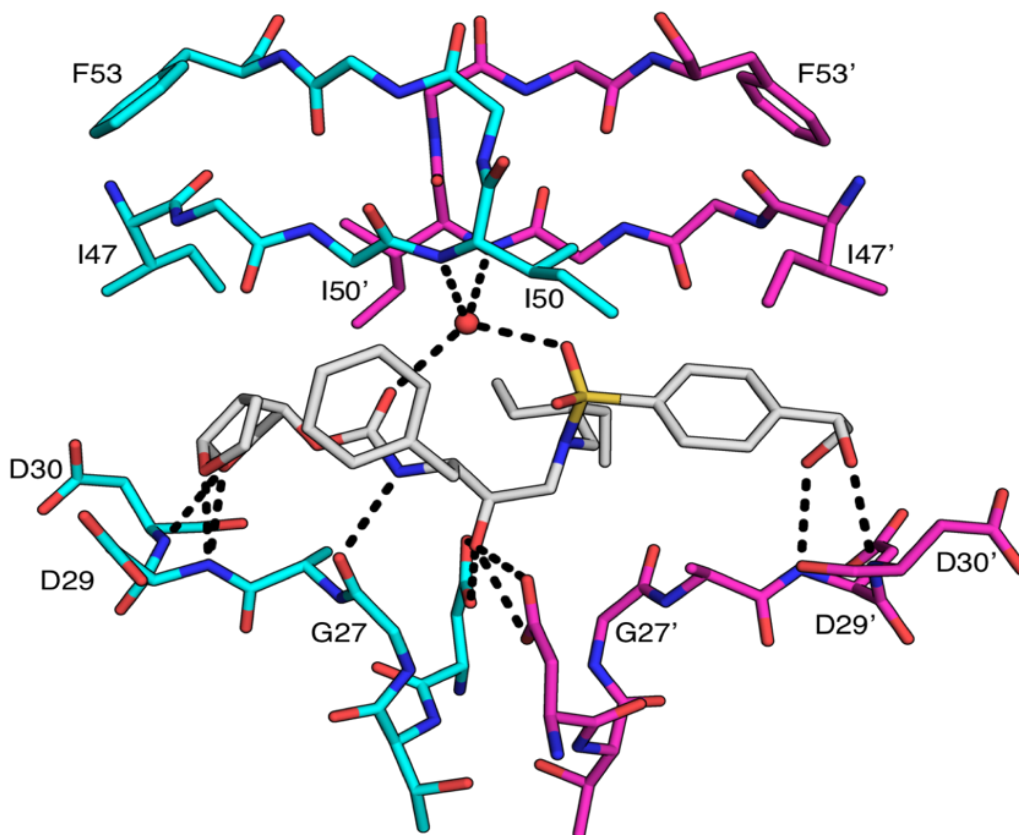
Figure 30: LR2-26 In Complex With RS_WT
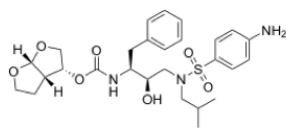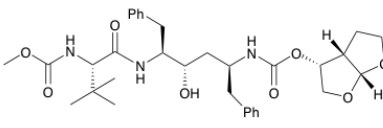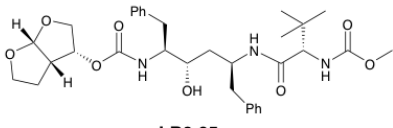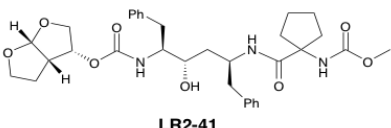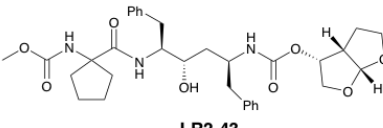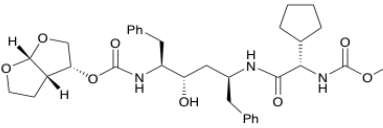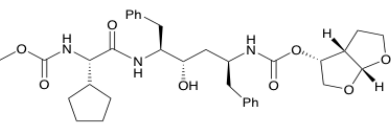
Compared to LR-85 and LR-100 series, the P2' group of LR2-26 interacts with residues D29' and D30' in a very different manner. The water mediated bond is lost with LR2-26, and instead, two hydrogen bonds are formed with the backbone of D29' and D30'. Traditionally, gaining backbone interactions with conserved residues is deemed beneficial to the potency of an inhibitor. Although a backbone interaction was gained from the addition of a hydroxyl, the obtained experimental K_i provides an insight of the importance of a water mediated bond in that area of the active site.

Ki Calculations and Structural Data for LPV/DRV Series

In addition to the mono- and di-hydroxyl compounds, another set of compounds were designed as hybrids between two highly potent drugs, LPV and DRV. These hybrids included the DRV P1 and P2 moieties and the LPV P1' and P2' moieties. Table 6 shows the calculated K_i values of the LR2 hybrid series in RS_WT, RS_I84V, and RS_I50V;A71V. As seen in Table 6, LR2-35 and LR2-32 showed near DRV inhibition against RS_WT. LR2-42 and LR2-44 were about 30 and 80 orders of magnitude less potent than DRV in WT, respectively. LR2-41 and LR2-

43 showed extremely poor inhibition against RS_WT with K_i values equal to 10.6 nM and 55.5 nM, respectively. This extreme loss of potency is believed to be caused by poor compound solubility in DMSO. Furthermore, when tested against major drug resistant mutants, such as RS_I84V and RS_I50V;A71V, we can see that the LPV/DRV hybrids do not inhibit nearly as tightly as DRV. Lastly, it is important to note that when all six LPV/DRV inhibitors were tested against RS_I50V;A71V, only LR2-35 and LR2-42 had a measurable K_i. For the rest of the LPV/DRV hybrids, the raw data against this construct was ambiguous and could not be fitted to the one phase association equation normally utilized.

Table 6: LPV/DRV Hybrids K_i Values in RS_WT, RS_I84V, and RS_I50V;A71V

Compound	Ki - Wild Type (nM)	Ki – I84V (nM) (fold change)	Ki – I50V;A71V (nM) (fold change)
DRV	< 0.005	0.026 ± 0.0056 (7)	0.035 ± 0.0041 (9)
LR2-32	0.018 ± 0.0025	0.21 ± 0.014 (11)	0.27 ± 0.043 (15)
LR2-35	0.013 ± 0.0013	0.37 ± 0.021 (29)	0.074 ± 0.0049 (6)
LR2-41	7.3 ± 0.5	191.0 ± 29.6 (26)	39.5 ± 2.5 (5)
LR2-43	46.4 ± 3.2	297.7 ± 17.4 (6)	484.4 ± 69.9 (10)
LR2-42	0.16 ± 0.011	1.18 ± 0.26 (7)	1.4 ± 0.1 (9)
LR2-44	0.24 ± 0.016	0.71 ± 0.10 (3)	2.52 ± 0.42 (11)

Following the results of the K_i assays, crystal structures of LR2-32 and LR2-35 in complex with RS_WT and RS_I84V were solved to understand how these inhibitors bind to the active site and how they fit within the substrate envelope space. Crystal structures of LR2-41, LR2-43, LR2-42, and LR2-44 in complex with RS_WT were also solved but are not shown here due to their poor inhibitory activity. Figure 31 shows the binding of LR2-32 in complex with RS_WT and RS_I84V, while Figure 32 shows the binding of LR2-35 in complex with RS_WT and RS_I84V.

LR2-32

Ki (RS_WT) = 18.2 ± 2.5 pM   Ki (RS_I84V) = 207.6 ± 14.1 pM
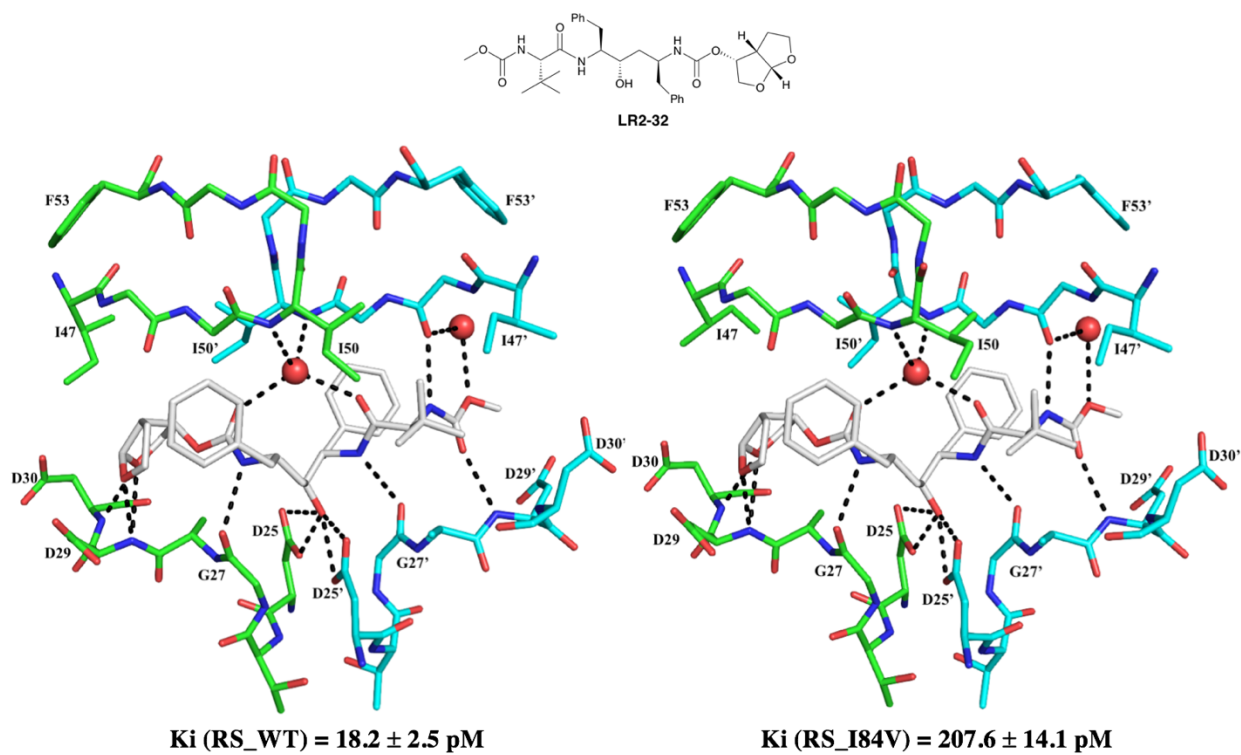
Figure 31: LR2-32 In Complex with RS_WT and RS_I84V

The solved crystal structures of both LR2-32 and LR2-35 showed very similar binding. For both inhibitors, the P2 bis-THF moiety maintained the three hydrogen bonds normally observed with DRV. As with DRV, G27 and G27' both maintained their hydrogen bonds with the nitrogen atoms on the backbone of the inhibitor. The conserved water molecule and its network of four hydrogen bonds is also maintained. Furthermore, D25 and D25' maintain the hydrogen bonds with the un-cleavable hydroxyl group. The P2' carbamate moiety contacts the backbone carbonyl of G49' directly and through a water mediated bond. Lastly, the carbamate also contacts the backbone nitrogen of D29'. Despite having similar binding, it is important to note that LR2-35 orients its P2 bis-THF moiety on chain B versus chain A, unlike DRV and LR2-32.

LR2-35

Ki (RS_WT) = 12.7 ± 1.3 pM   **Ki (RS_I84V) = 371.7 ± 21.1 pM**
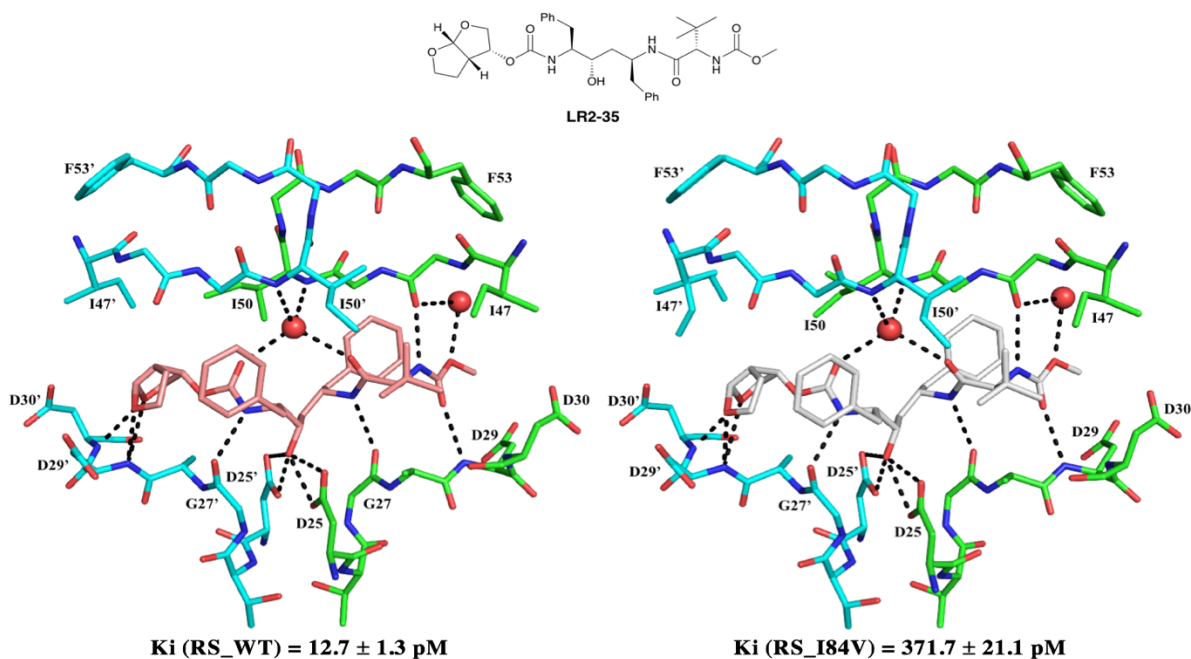
Figure 32: LR2-35 In Complex With RS_WT and RS_I84V

The solved crystal structures were then overlaid in the substrate envelope to see how well these inhibitors fit in that chemical space. Figure 33 shows LR2-32 in the substrate envelope while Figure 34 shows LR2-35 in the substrate.
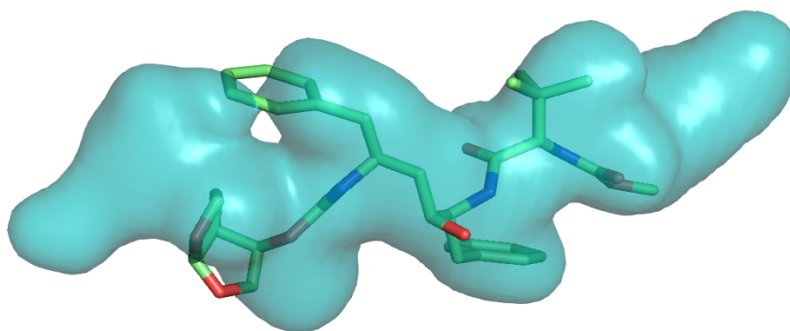
Figure 33: LR2-32 Inside the Substrate Envelope

LR2-32 fits in the substrate envelope in a similar fashion as DRV. The P2 bis-THF moiety as well as the P1-phenyl group extend slightly beyond the envelope. The rest of the inhibitor fits into the envelope nicely. The binding of the P2 bis-THF moiety of LR2-35 does not affect the observed hydrogen bonds. However, this binding mode means that LR2-35 fits into the envelope drastically different than LR2-32. As seen in Figure 34, the P2 bis-THF and the P1 phenyl group fit entirely inside the substrate envelope. While the P1' phenyl group of LR2-32 fits inside the envelope nicely, the P1' phenyl group of LR2-35 does not. Lastly, like LR2-32, the P2' carbamate group of LR2-35 tucks in entirely inside the envelope.
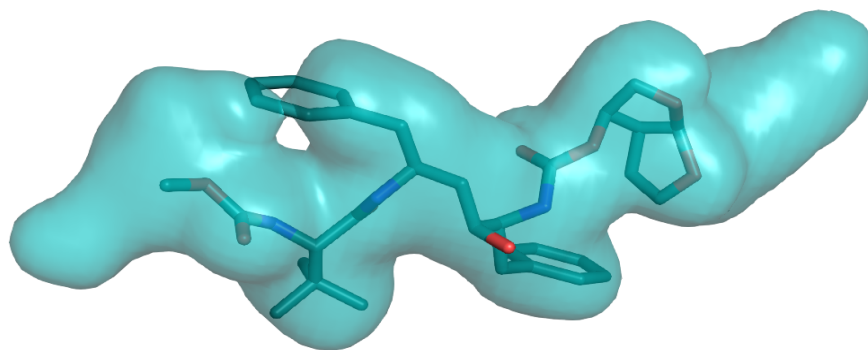
Figure 34: LR2-35 Inside the Substrate Envelope

Van der Waals energy was also calculated for LR2-32 and LR2-35 in complex with both RS_WT and RS_I84V. Table 7 summarizes the calculated vdW potentials. DRV vdW potential is also shown as a control.

Table 7: DRV, LR2-32, and LR2-35 vdW Potential in RS_WT and RS_I84V

Construct/Inhibitor	vdW (Kcal/mol)
RS_WT DRV	-80.1
RS_WT LR2-32	-89.6
RS_WT LR2-35	-89.3
RS_I84V DRV	-79.6
RS_I84V LR2-32	-88.2
RS_I84V LR2-35	-89.1

The overall vdW potential for LR2-32 in complex with RS_WT and RS_I84V is -89.6 Kcal/mol and -88.2 Kcal/mol, respectively. The vdW potential for DRV in complex with RS_WT and RS_I84V is -80.1 Kcal/mol and -79.3 Kcal/mol, respectively. In both RS_WT and RS_I84V, this represents at least a 9 Kcal/mol difference compared to DRV.

The overall vdW potential for LR2-35 in complex with RS_WT and RS_I84V is -89.3 Kcal/mol and -89.3 Kcal/mol, respectively. The vdW potential for DRV in complex with RS_WT and RS_I84V is -80.3 Kcal/mol and -79.6 Kcal/mol, respectively. In both RS_WT and RS_I84V, this represents at least a 9 Kcal/mol difference compared to DRV.

Chapter 4: Discussion and Future Direction

4.1 The Role of Adaptive Resistance in HIV-1 Protease

Throughout this study, it has been evident how elegant and intelligent of a protein the HIV-1 protease is. DRV, as well as other potent inhibitors created in the last decade, are slowly losing their potency against mutated variants of the protease. Mutants such as I84V and I50V;A71V are able to withstand the potency of DRV by as much as 10-fold compared to WT, increasing the protease's chances for survival and therefore, leading to more successful HIV proliferation. Patient isolates continue to show new mutations developed by the protease, with some isolates showing as many as 26 new mutations within its genome. Despite all of these mutations, the protease is able to maintain its substrate specificity with a K_m comparable to that of WT protease. The rate at which the protease mutates is directly dependent on the selective pressure exerted on specific residues by inhibitors used in therapies, such as DRV and LPV.

Kinetic data in conjunction with crystal structures and other molecular dynamics simulations have helped create an idea of favorable and unfavorable inhibitor-protease interactions in the active site. Identifying highly variable residues, such as V32, I50, V82 and I84, has helped in the inhibitor design process such that new inhibitor designs avoid making direct interactions with these residues. Furthermore, focusing on increasing interactions with the backbone of invariant residues, such as D29 has shifted focus towards designing inhibitors that interact with these residues. The UMass series are a perfect example of this, modifying the P2' moiety so more interactions are made with the backbone of D29 and D30. The UMass series show promising results, as their potency is comparable to DRV amongst WT and some DRV resistant mutants, such as I84V.

The mono-hydroxyl and di-hydroxyl compounds also show inhibition comparable to DRV in WT and the I84V mutant, suggesting that such substitutions in the P2' moiety have no drastic impact on the potency of the inhibitor. Crystal structure data also shows that the mono-hydroxyl substitutions bind in a similar fashion as the DRV P2' amine moiety, interacting directly with the backbone of D29' and creating a water-mediated bond with the backbone of D30'. On the other hand, both of the hydroxyl chains on the di-hydroxyl compounds interact directly with D29' and D30' without the formation of a water-mediated bond. In terms of potency, the mono-hydroxyl compounds resulted in lower K_is than the di-hydroxyl compounds, suggesting that the water-mediated bond to the backbone of D30' is essential to the tight binding of the inhibitor. This result

is quite interesting, as it was originally thought that interacting in a direct hydrogen bond with both D29' and D30' would increase the potency of the compounds.

Other inhibitor designs, such as the DRV/LPV hybrids, focused on combining the scaffolds of two of the most potent inhibitors in the hopes of developing an even more potent series. The surprising results showed that the combined scaffolds led to a drastic decrease in potency and molecular simulations highlighted the instability of the new scaffold. In particular, the bis-THF moiety proved to be extremely unstable, sampling space away from the backbone of D29 and D30 thus failing to make direct contact with them.

Although the UMass and mono- and di-hydroxyl series show promising potency results, viral passaging data is necessary to verify how effective they are in avoiding new induced mutations. However, these compounds fail to deliver more potent inhibition than current FDA approved inhibitors, proving that successful P2' modifications are extremely difficult to carry out. This may be due to the limited invariant residues present in that region of the active site. Taking the results of recent compound data into account, the laboratory continues to explore different substitutions in order to optimize the DRV scaffold in the hopes of developing a more potent inhibitor that induces little to no selective pressure on the HIV-1 protease.

4.2 Optimizing the DRV scaffold: P2' and P1' Moiety Substitutions

The K_i and crystallographic data of the UMass 1-10 compounds, the LR/LR-2 compounds, and the LPV/DRV hybrids were instrumental in helping us develop new compounds. Despite the potency of DRV and recent attempts to further optimize inhibition of HIV-1 protease, the virus still remains one of the most infectious in the world. The virus is capable of mutating up to half of its genome, allowing the HIV-1 protease to develop resistance to even the most potent inhibitors (32). Localizing specific residues susceptible to mutations and analyzing solved crystal structures are crucial methods to understanding and developing optimized next-generation inhibitors.

Optimizing the structure of the already potent DRV has proven to be a challenge, as no inhibitor has been able to top its low pico-Molar binding affinity and resistance profile. When optimizing the structure of DRV, there are several key features of the backbone that have proven to be robust and are usually conserved. The un-cleavable hydroxyl group makes two hydrogen bonds with the catalytic aspartic acid/aspartate (D25/25'). Perhaps the next robust feature of DRV is the highly conserved hydrogen bonding network between the P2-bis-Tetrahydrofuran (P2-bis-THF) group and residues D29 and D30 of chain A in the active site. No other inhibitor groups

have been able to make and maintain such tight hydrogen bonding in the active site pocket than the bis-THF group. Lastly, another conserved moiety is the phenyl group in the P1 position. This moiety contributes to the observed VDW interactions and for that reason, is usually conserved. The P1' position is usually variable and various alkyl chains have been placed on the amide of the P1'. It is usually preferred to place small groups in the P1' position to avoid sticking out of the substrate envelope and making unfavorable contacts with drug resistant residues such as V82 and I84. Although the isobutyl P1' moiety of DRV is small, I84V is still a major drug resistant mutation for DRV. This suggests that despite the isobutyl group not protruding from the envelope, I84 is still capable of mutating to resist the binding of DRV. On the P2' position, the aniline group makes a water mediated hydrogen bond with the side chain of residue D30' and another hydrogen bond with the carbonyl backbone of D30'. However, these bonds are not nearly as conserved and stable as those of the P2-bis-THF moiety.

Although the P1 moiety of DRV is usually conserved in most all DRV analogs, several compounds have attempted to modify this group. One such example is GRL-10413 developed by by the Amano laboratory in conjunction with the Ghosh laboratory, attempts to modify the P1 moiety to make contact with the invariant residues R8' and G49 (see Figure 35A, PDB 5KAO). In the solved crystal structure, the 1-chloro-2-methoxybenzene P1 group was observed to have two conformations in the crystal structure. The first shows the chlorine making a halogen bond with the carbonyl oxygen of G49. The second show the chlorine making two halogen bonds with the nitrogen(s) of R8' (see Figure 35B).
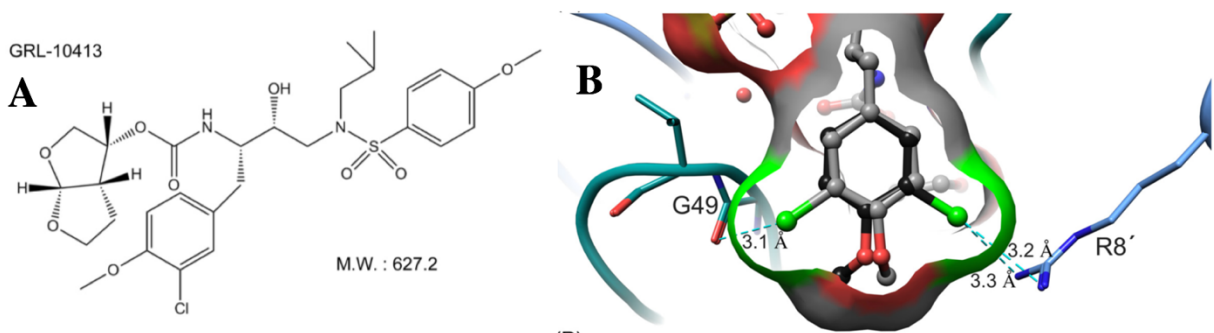
Figure 35: Molecular Structure and Bonds of GRL-10413

Based on the solved crystal structure of GRL-10413, nine di-halogenated-P1 compounds were designed with the main goal of maintaining both halogen bonds with G49 and R8' simultaneously (see Figure). The compounds utilize chlorine, fluorine, and bromine in the P1 position. In the P1' position, these compounds either contain the DRV isobutyl, the UMass 1 2-

methylbutane, or the UMass 6 isohexyl. The DRV like P2 bis-THF and P2' amine groups were unchanged. Prior to synthesis, these compounds were tested through varies computational methods that can predict the potency of each inhibitor.
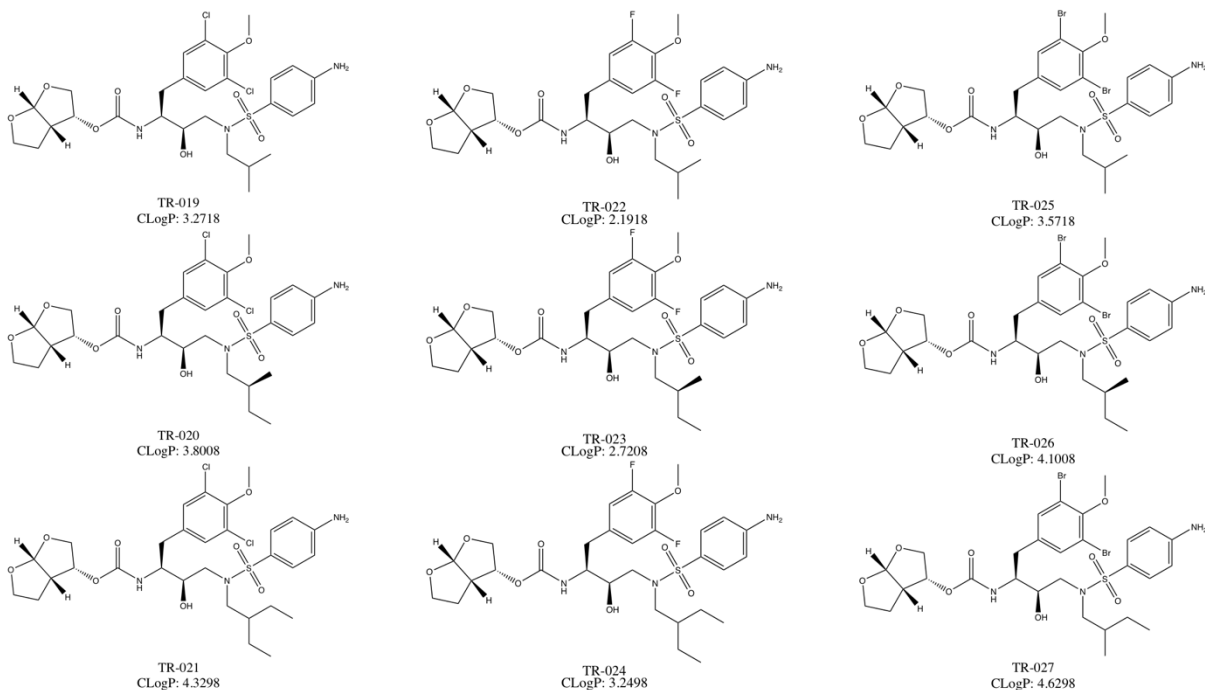
Figure 36: Di-Halogenated P1 Moiety Compounds

Further exploration and analysis of the LR2 compounds led to various conclusions regarding the P2' moiety and the stability of compounds with different linkers. LR2-32 and LR2-35 were of particular interest due to their observed K_i's in the RS_WT scaffold. Further exploration of their crystal structures provided a better understanding of how these compounds bind in the active site. The carbamate group extending from the carbonyl linker of both compounds was observed to make favorable interactions with residues D29' and D30'. Based on this analysis, another set of compounds were designed using the DRV backbone and the sulfonamide linker, KK-01 and KK-03, respectively (Figure37). These compounds contained a carbamate moiety on the P2' position extending from the phenyl ring in the hopes to increase interactions with the D29' and D30' residues (Figure 37). These compounds were sketched in ChemDraw and modeled in Maestro in order to analyze their chemical properties as well as electrostatic interactions within the active site of the HIV-1 protease.

KK-01
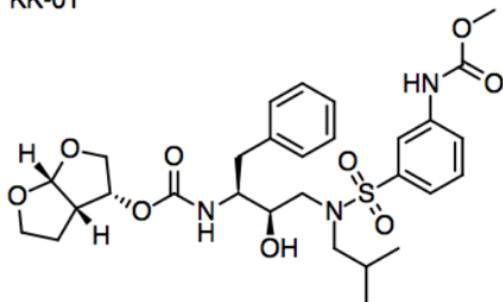
KK-03

Chemical Formula: $C_{29}H_{39}N_3O_9S$
Molecular Weight: 605.70
CLogP: 3.0828
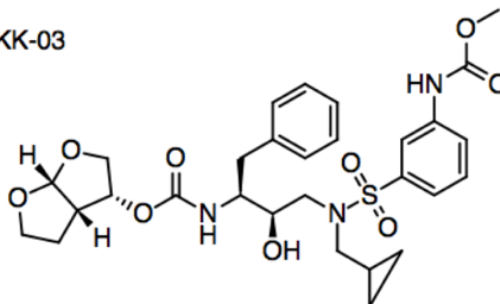
Chemical Formula: $C_{29}H_{37}N_3O_9S$
Molecular Weight: 603.69
CLogP: 2.5988

Figure 37: Carbamate compounds KK-01 and KK-03

Appendix

Appendix A: K[m] Data

Variant	Uncorrected K$_m$ (µM)	Corrected K$_m$ (µM)
RS_WT	24.9 ± 1.5	62.4 ± 4.9
RS_2Mut[1]	17.3 ± 1.6	55.1 ± 12.41
RS_4Mut[2]	12.9 ± 1.27	21.1 ± 2.71
RS_4GMut[3]	60.6 ±	71.0 ± 4.61
RS_8Mut[4]	38.3 ± 3.8	123.8 ± 30.1
RS_10Mut[5]	31.3 ±	66.4 ± 4.3
L76V	45.8 ±	299.9 ± 102.3
L33F	20.4 ± 1.2	34.7 ± 2.2
V32I	17.6 ± 1.5	30.6 ± 3.0
V32I;L33F	34.4 ± 3.9	306.6 ± 130.8
ATA 21[6]	45.5 ±	691.8 ± 103.4
KY 26[7]	25.9 ±	74.4 ± 13.4
SLK 19[8]	6.7 ±	11.0 ± 1.0
VEG 23[9]	18.5 ± 1.7	9.9 ± 0.8
VSL 23[10]	41.9 ± 5.3	219.7 ± 77.4

[1] V82F, I84V

[2] V32I, M46I, V82F, I84V

[3] K45I, M46I, V82F, I84V

[4] I13V, G16E, V32I, L33F, K45I, M46I, V82F, I84V

[5] I13V, G16E, V32I, L33F, K45I, M46I, A71V, L76V, V82F, I84V

[6] L10F, **K20M**, **V32I**, L33F, M36I, **M46I**, **I47V**, **I54M**, I62V, L63P, **G73T**, **I84V**, L89V, **L90M**

[7] L10V, **K20M**, **V32I**, L33F, K43T, **M46I**, **I47V**, **I54M**, I62V, L63P, A71I, I72L, **G73S**, V77I, **V82A**, L89V, **L90M**

[8] L10I, M36I, **G48M**, F53L, **I54V**, I62V, L63P, A71V, **V82A**, **I84V**, **I85V**, L89V

[9] L10I, **K20V**, **V32I**, L33F, M36I, **M46I**, **I47V**, **I54M**, Q58E, L63P, A71V, **G73S**, **I84V**, **L90M**

[10] L10I, **K20R**, L33F, M36I, K43T, **G48V**, **I50V**, **I54S**, I62V, L63P, A71V, I72V, **G73T**, V77I, **V82A**, **I85V**

Appendix B: K_i Sample Calculation/Processing the K_i Data

1. Prepare 100 mL 2X Assay Buffer [100 mM Sodium Acetate – 200 mM Sodium Chloride]
   - Start with 3M stock Sodium Acetate
     $C_1V_1 = C_2V_2$
     3M (X) = 0.1M (100mL)
     X = 3.3 mL of 3M Sodium Acetate
   - 200 mM Sodium Chloride (M * MW * V = g)
     $0.2M * 58.44 \frac{g}{mol} * 0.1 \text{ L} = 1.17g$ of NaCl
   - Add H_2O to the 100 mL

2. Prepare 5 mL of 4% DMSO
   - 200 μL of 100% DMSO + 4800 μL of H_2O

3. Inhibitor Prep for 2/3 serial dilution and 60 μL final well volume
   - $\frac{Well\ Volume}{Inhibitor\ Volume} = \text{Fold Dilution} \rightarrow \frac{60\ \mu L}{27.5\ \mu L} = 2.2\ Fold\ Dilution$
   - To well 12, 82.5 μL of inhibitor will be added, then 55 μL will be serially diluted to well 11 leaving behind 27.5 μL. The serial dilution process will be repeated until well 2. Well 1 is a control well and does not get any inhibitor; 55 μL will be taken about of well 2 and discarded.
   - Well 12 gets 82.5 μL * 3 replicate = 247.5 μL → prepare 275 μL
   - Final Well (well 12) inhibitor concentration = 1500 pM
     $$1500\ pM * 2.2\ fold\ dilution = 3,300\ pM$$
   - 275 μL * 3,300 pM = 11 μL (4% of 275 μL because the assay is in 4% DMSO) * X
     X = 82.5nM Inhibitor Concentration

     32 μL (82.5nM Inhibitor) = 500 nM Stock * X [prepare 32 μL to avoid pipetting volumes under 5 μL]
     X = 5.28 μL of 500 nM stock Inhibitor

     32.0 μL – 5.28 μL = 26.72 μL of 100% DMSO

     Scheme: 5.28 μL of 500 nM Inhibitor + 26.72 μL DMSO → 11 μL of that + 264 μL H_2O → gives us 275 μL needed to start the dilution at well 12

4. Protein Prep at 0.35nM (A_{280} will differ for each construct, this is just an example calculation)
   - 0.35 nM * 2.2-Fold Dilution = 0.77 nM
   - A_{280} 10-Fold = 0.172 (A_{280} is measured in 90 μL refolding buffer + 10 μL protein = 10-fold dilution)
     1.72 / 24,980 = 69 μM * 40% Active Protein (we assume only 40% activity because the protein is not entirely pure) = 27.6 μM
     500 μL * 500 nM = 27600 nM * X
     X = 9.06 μL + 491 2X Assay Buffer
   - 0.77nM Protein Prep
     5 mL * 0.77 nM = 500 nM * X
     X = 7.7 μL + 492.3 μL 2X Assay Buffer

5. Substrate Prep
   - Total Assay Volume * [S] = Concentration added to assay * Volume needed
     60 μL * 10 μM = 5 μL * X
     X = 120 μM
   - Minimum volume needed for machine * X from previous = 30 μL * Concentration
     1500 μL * 120 μM = 30 μL * X
     X = 6 mM
   - 750 μL 2X Assay Buffer PLUS 720 μL H_2O PLUS 30 μL of 6 mM substrate

6. Plate Setup
    - 27.5 μL of 4% DMSO are added to wells 1 – 11
    - 82.5 μL of inhibitor are added to well 12
    - Serially dilute 55 μL from well 12 to 11 and so on until well 2. Well 1 does not get inhibitor. Take 55 μL from well 2 and discard
    - Add 27.5 μL of the 0.77 nM protein to all wells
    - Spin plate at 1000 x g for 1 minute at room temperature
    - Pre-incubate for 1 hour are room temperature
    - Reaction starts when 5 μL of substrate is inserted by the Envision
    - Run Envision program to collect 200 reads per well

References

1. Prabu-Jeyabalan, M., Nalivaika, E., and Schiffer, C.A. (2002) Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. *Structure (London, England : 1993).* **10,** 369

2. Anonymous (2017) Global Health Observatory (GHO) Data HIV/AIDS**2017,**

3. Anonymous (2017) What Are Vaccines and What Do They Do?

4. Al-Jabri, A.A. (2007) Mechanisms of Host Resistance Against HIV Infection and Progression to AIDS. *Sultan Qaboos University medical journal.* **7,** 82

5. Otto O. Yang, Annie-Chen Tran, Spyros A. Kalams, R. Paul Johnson, Margo R. Roberts, and Bruce D. Walker (1997) Lysis of HIV-1-Infected Cells and Inhibition of Viral Replication by Universal Receptor T Cells. *Proceedings of the National Academy of Sciences of the United States of America.* **94,** 11478-11483

6. Ganser-Pornillos, B.K., Yeager, M., and Pornillos, O. (2012) Assembly and Architecture of HIV**726,** 441-465

7. Hélène C. F. Côté, Zabrina L. Brumme, and P. Richard Harrigan (2001) Human Immunodeficiency Virus Type 1 Protease Cleavage Site Mutations Associated with Protease Inhibitor Cross-Resistance Selected by Indinavir, Ritonavir, and/or Saquinavir. *Journal of Virology.* **75,** 589-594

8. Tamamis, P., and Floudas, C.A. (2014) Molecular recognition of CCR5 by an HIV-1 gp120 V3 loop. *PloS one.* **9,** e95767

9. Karn, J., and Stoltzfus, C.M. (2012) Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harbor perspectives in medicine.* **2,** a006916

10. Burch, C.L., Leonard, C.W., Weeks, K.M., Swanstrom, R., Dang, K.K., Bess Jr, J.W., Watts, J.M., and Gorelick, R.J. (2009) Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature.* **460,** 711-716

11. Lin Li, Hai Shan Li, C David Pauza, Michael Bukrinsky, and Richard Y Zhao (2005) Roles of HIV-1 auxiliary proteins in viral pathogenesis and host-pathogen interactions. *Cell Research.* **15,** 923-934

12. Heigele, A., Sauter, D., Münch, J., and Kirchhoff, F. (2014) HIV-1 accessory proteins: Nef. *Methods in molecular biology (Clifton, N.J.).* **1087,** 115

13. Andrew, A., and Strebel, K. (2014) HIV-1 accessory proteins: Vpu and Vif. *Methods in molecular biology (Clifton, N.J.).* **1087,** 135

14. Goodsell, D. (2002) HIV Reverse Transcriptase**2017,**

15. Thang Chiu, and David Davies (2004) Structure and Function of HIV-1 Integrase. *Current Topics in Medicinal Chemistry.* **4,** 965-977

16. Sarafianos, S.G., Marchand, B., Das, K., Himmel, D.M., Parniak, M.A., Hughes, S.H., and Arnold, E. (2009) Structure and Function of HIV-1 Reverse Transcriptase: Molecular Mechanisms of Polymerization and Inhibition. *Journal of Molecular Biology.* **385,** 693-713

17. Sundquist, W.I., and Kräusslich, H. (2012) HIV-1 assembly, budding, and maturation. *Cold Spring Harbor perspectives in medicine.* **2,** a006924

18. Delhalle, S., Schmit, J., and Chevigne, A. (2012) Phages and HIV-1: From Display to Interplay. *International Journal of Molecular Sciences.* **13,** 4727-4794

19. Olga Latinovic, Janaki Kuruppu, Charles Davis, Nhut Le, and Alonso Heredia (2009) Pharmacotherapy of HIV-1 Infection: Focus on CCR5 Antagonist Maraviroc. *Clinical Medicine. Therapeutics.* **1,** 1497

20. Eric O Freed (2015) HIV-1 assembly, release and maturation. *Nature Reviews. Microbiology.* **13,** 484-496

21. Deeks, E. (2014) Cobicistat: A Review of Its Use as a Pharmacokinetic Enhancer of Atazanavir and Darunavir in Patients with HIV-1 Infection. *Drugs.* **74,** 195-206

22. Brechtl, J.R., Breitbart, W., Galietta, M., Krivo, S., and Rosenfeld, B. (2001) The use of highly active antiretroviral therapy (HAART) in patients with advanced HIV infection: impact on medical, palliative care, and quality of life outcomes. *Journal of pain and symptom management.* **21,** 41

23. Akbar Ali, Rajintha M Bandaranayake, Yufeng Cai, Nancy M King, Madhavi Kolli, Seema Mittal, Jennifer F Murzycki, Madhavi NL Nalam, Ellen A Nalivaika, Aysegül Özen, Moses M Prabu-Jeyabalan, Kelly Thayer, and Celia A Schiffer (2010) Molecular Basis for Drug Resistance in HIV-1 Protease. *Viruses.* **2,** 2509-2535

24. King, N.M., Prabu-Jeyabalan, M., Nalivaika, E.A., and Schiffer, C.A. (2004) Combating Susceptibility to Drug Resistance. *Chemistry & Biology.* **11,** 1333-1338

25. Schramm, V.L. (2013) Transition States, analogues, and drug development. *ACS chemical biology.* **8,** 71

26. Anonymous (2017) HIV STRAINS AND TYPES

27. Ghosh, A.K., Dawson, Z.L., and Mitsuya, H. (2007) Darunavir, a conceptually new HIV-1 protease inhibitor for the treatment of drug-resistant HIV. *Bioorganic & Medicinal Chemistry.* **15,** 7576-7580

28. Palmier, M.O., and Van Doren, S.R. (2007) Rapid determination of enzyme kinetics from fluorescence: Overcoming the inner filter effect. *Analytical Biochemistry.* **371,** 43-51

29. Windsor, I.W., and Raines, R.T. (2015) Fluorogenic Assay for Inhibitors of HIV-1 Protease with Sub-picomolar Affinity. *Scientific reports.* **5,** 11286

30. Mark Brandt (2016) Inhibition kinetics

31. Rabia Naeem (2017) Lennard-Jones Potential

32. Shafer, R.W., and Schapiro, J.M. (2008) HIV-1 Drug Resistance Mutations: an Updated Framework for the Second Decade of HAART. *AIDS Reviews.* **10,** 67-84