Substrate Envelope Based Design of New HIV-1 Protease Inhibitors Active Against Drug-Resistant HIV-1

A Major Qualifying Project

Submitted to the Faculty of the
WORCESTER POLYTECHNIC INSTITUTE

In Partial Fulfillment of the Requirements for the
Degree of Bachelor of Science
By

_____

Kevin Yee

Date: April 29, 2010

Approved:

_____ _____

Professor Destin Heilman (WPI), Primary Advisor

Dr. Celia Schiffer (UMass Medical)

Table of Contents

Table of Figures

Abstract

With over 30 different Human Immunodeficiency Virus (HIV) drugs approved by the FDA, drug resistance is still a major problem. This experiment suggested a new general structure-based design for drug discovery based on the substrate envelope hypothesis in order to successfully design and synthesize a series of novel HIV-1 protease inhibitors. The design inhibitors utilize the 2-ethylbutyl group as new high affinity P1′ ligands. The designed compounds showed highly potent inhibitory activity against wild-type and a panel of MDR HIV-1 Protease variants. Further development of these protease inhibitors may lead to more effective treatments against drug-resistant HIV-1.

Introduction

The Human Immunodeficiency Virus Pandemic (HIV)

Since the United States declared Human Immunodeficiency Virus (HIV) a pandemic in 1981, over 30 HIV drugs have been approved by the FDA(*1*). Despite the progress of these drugs, 56,000 new infections occur each year and over 30 million people are currently affected worldwide(*2*). By 2007, HIV had contributed to over 2 million deaths with 565,927 deaths in the United States(*2*). Many of these deaths occurred in developing countries, where major societal problems such like healthcare, nutrition, and poverty are frequently encountered(*1*). With the growing deaths and drug resistance encountered during therapy, further research for new HIV drugs is essential.

The current HIV drugs used to treat HIV-1 belong to four classes: nucleoside and nucleotide analogues, reverse transcriptase inhibitors, fusion inhibitors, and protease inhibitors(*3*). A common treatment of HIV used is Highly Active Antiretroviral Treatment (HAART), which is a combination of antiretroviral drugs(*3*). While HAART has been effective against HIV, its effectiveness is decreasing due to the increase of resistant mutations within the virus(*2*). Although current drug discovery protocols approach this problem by testing numerous analogues, this approach does not account for virus mutations and drug resistance(*2*). In order to combat drug resistance, new inhibitors must inhibit both the wild-type virus and mutant forms of HIV(*2*). As seen in darunavir, which is the most potent protease inhibitor, one way of combating resistance is by designing drugs based on the substrate envelope. With HIV recognizing various substrate sequences, the substrate envelope is a conserved uniform region of substrates binding within the binding site(*4*). By using a protease inhibitor that fits within the substrate envelope, viral resistance can be evaded (*4*).

The HIV Structure

HIV is characterized as a retrovirus, which are viruses that contain RNA as its genetic material and use reverse transcriptase to produce DNA from its genome and integrates the DNA into the host genome (*1*). HIV is subgrouped with lentiviruses, which replicate in non-dividing cells (*5*). With lentiviruses, serious symptoms tend to occur after a long period from the initial infection and can deliver a significant amount of genetic information to the DNA of the host cell(*1*).

The HIV envelope has a 100 nm diameter and is created from the membrane of a human cell during the budding process(*1*). Proteins are fixed within the viral envelope and include the Env protein, which protrudes from the surface of the virus(*1*). Env consists of glycoprotein 120 (gp120) and glycoprotein 41 (gp41), which assists in anchoring the structure to the viral envelope(*1*). Within the viral envelope is a capsid that surrounds two strands of HIV RNA(*1*). The capsid is made up of 2,000 copies of p24, a viral protein(*1*). At the end of each RNA are long terminal repeats that control production of new viruses triggered by proteins from HIV or the host cell(*1*). Additionally, the HIV core is comprised of p17, an HIV matrix protein, and p7, a HIV nucleocapsid protein used in later development within the virus life cycle(*1*).

The HIV-1 Genome and Life Cycle

HIV-1 is a complex retrovirus that encodes 15 distinct proteins and express nine open reading frames(*6*). The gag, pol, and env genes serve to produce the core proteins necessary for the formation of the virus. The tat, rev, nef, vif, vpr, and vpu genes serve as regulatory genes to control reproduction of the virus(*1*).

Three of the open reading frames encode the Gag, Pol, and Env polyproteins. Each polyprotein is proteolyzed into individual proteins(*6*). The four Gag proteins (matrix (MA), capsid (CA), nucleocapsid

(NC), and p6) and the two Env proteins (surface or gp120 (SU) and transmembrane or gp41 (TM)) are structural components that form the core of the virion and outer membrane(*6*). The three Pol proteins (protease (PR), reverse transcriptase (RT), and integrase (IN)) provide the essential enzymatic functions and are encapsulated within the particle(*6*). HIV-1 also encodes six additional accessory proteins(*6*). Vif, Vpr, and Nef and are found within the viral particle, while Tat and Rev provide gene regulatory functions, while Vpu indirectly aids in assembling the virion(*6*). The retroviral genome is encoded by a ~9 kb RNA and two genomic length RNA molecules that are packaged in the particle(*6*).

Knowledge of the replication cycle of the virus is essential in order to understand the mechanism of action for antiviral drugs and to find potential drug targets (Figure 2). During the initial step of HIV replication, there is an interaction between the envelope proteins of the virus (SU), the cell surface receptors (CD4 receptors), and the chemokine coreceptors (CXCR4 and CCR5) of the host(*7*). This initial interaction causes a conformational change in the envelope protein and promotes a fusion between the host cytoplasmic membrane and the viral envelope(*7*). This fusion process is promoted by TM and allows the viral capsids to enter the cell through the membrane(*7*).

Once inside the membrane, the virion core is then uncoated to expose a viral nucleoprotein complex which contains MA, RT, IN, Vpr, and RNA(*6*). The viral DNA is then inserted and integrated into the chromosomal DNA of the host cell by the integrase(*7*). The IN protein catalyzes the intergration of the viral DNA into a host chromosome and the DNA is repaired(*6*). The expression of the infected DNA allows for production of precursor viral proteins(*7*). The complex is then transported to the nucleus, where the RNA is reverse transcribed by RT into a partially duplex linear DNA(*6*). The viral transcripts are expressed from the promoter located in the 5' LTR with Tat increasing the rate of transcription(*6*). Spliced RNAs are then transported from the nucleus to the cytoplasm in order for Rev to regulate the translation(*6*). Once the viral mRNAs are translated in the cytoplasm, the Gag and Gag-Pol polyproteins

become localized to the cell membrane and the Env mRNA is translated at the endoplasmic reticulum(*6*). The Gag and Gag-Pol, Vir, Vpr, Nef and genomic RNA then assemble the core particles and the virion begins budding at the surface(*6*). The expression of the infected DNA allows for production of precursor viral proteins(*7*). In order to provide the SU and TM proteins for the outer membrane during budding, the Env polyprotein must first be released with CD4, which is the cell surface HIV-1 receptor in the ER(*6*). Vpu then promotes the CD4 degradation and Env is transported to the cell surface where it is prevented from binding CD4(*6*). Nef facilitates the routing of CD4 from the cell surface and golgi apparatus to the lysosomes and results in receptor degradation and preventing interactions with Env (*6*).

A particle is released from the cell surface with SU and TM on the surface of the cell and the virion undergoes maturation and requires Vir to regulate the proteolytic processing of Gag and Gag-Pol polyproteins by the protease(*6*). The precursor viral proteins which are assembled at the host cell surface, form new viral particles, and leave the host cell through the process of budding(*7*). During budding, an outer layer and an envelope are acquired from the host cell and the protease enzyme cleaves the precursor viral proteins into mature products(*7*). If the protease does not cleave the precursor proteins then the viral particles cannot initiate the replication cycle in other cells(*7*).

Comparison of Potential Drug Targets

HIV contains many enzymes and receptors which are crucial to the lifecycle of HIV and can make attractive drug targets including integrase(*8*). By disrupting the integrase, HIV cannot integrate its DNA with the host. Although many integrase inhibitors have been published over the years, no integrase inhibitor has been approved by the FDA(*9*). Some of these drugs have made clinical trials and most of which have shown severe liver and kidney toxicity(*9*). The past compounds L-870,810 (Merck) and S-1360 (Shionogi) in the past have looked promising, but most integrase inhibitors have shown severe liver

and kidney cell toxicity(*9*). Despite these toxicity issues, integrase may serve as a future potential drug target.

While there are no integrase inhibitors approved by the FDA, protease inhibitors are the most common FDA approved inhibitors and the reverse transcriptase inhibitors are the second most common FDA approved inhibitors. Reverse transcriptase inhibitors were first marketed in 1997 and revolutionized HIV therapy with good antiviral activity and convenient regimens. While protease inhibitors and nonnucleoside reverse transcriptase inhibitors have similar potency against the virus, protease inhibitors are more effective than nonnucleoside reverse transcriptase inhibitors with patients infected with high viral loads (*10*). This increase in effectiveness is due to the high genetic barrier for resistance in protease inhibitor based regimens(*11*). Protease inhibitors have been effective for treating patients with an interruption in their treatment(*12*). Current protease inhibitors have not been developed using the substrate envelope. By design protease inhibitors that fit within the substrate envelope will be essential for future HAART and evading viral resistance(*4*).

Nucleoside and Nucleotide Analogues

Nucleoside and nucleotide analogues act as DNA chain terminators and were the first antiviral drugs to be approved for treatment of HIV(*7*). This class of HIV drugs seeks to inhibit reverse transcription of the viral RNA genome into DNA which is an important event in the early stage of the viral life cycle(*3*). The nucleoside analogs refer to the similarity of nucleic acids, but differ by the replacement of a hydroxyl group in the 3' position and by adding another substituent incapable of forming the 5' to 3' phosphodiester linkage(*3*). DNA elongation is hindered by the incorporation of the analogue(*3*). The reverse transcriptase incorporates these nucleoside and nucleotide analogues competitively into the viral DNA after the phosphorylation by cellular kinases(*3*). The synthesis of the viral DNA is then stopped by these analogues and no additional nucleotides can be added(*3*).

Despite the effectiveness of these drugs, several mutations occur in the reverse transcriptase and impair the ability of the reverse transcriptase to incorporate an analogue in DNA(*3*). The most common mutations that occur that lead to resistance are the M184V, Q151M, and the K65R mutation(*3*). Methionine 184 is located at the active site of the catalytic region of the reverse transcriptase and is replaced by a valine(*3*). The M184V mutation is usually the first mutation to overtake the wild-type virus within a few weeks and affects the drug lamivudine(*3*). The different side chain of valine interferes with the positioning of lamivudine triphosphate analogues within the catalytic site(*3*). The group of mutations at the Q151M complex are mutations that demonstrate resistance against stavudine and didanosine(*3*). The Q151M mutation affects a residue located near the nucleotide binding site of the reverse transcriptase(*3*). After this initial mutation, many secondary mutations accumulate and increase the resistance activity within the enzyme(*3*). Although this resistance occurs in less than 5% of all HIV strains with resistance to nucleoside analogues, this mutation complex provides high level resistance to most analogues(*3*). The K65R mutation confers resistance to most analogues except zidovudine and is a frequent mutation(*3*). Despite the high drug resistance encountered against HIV, nucleoside and nucleotide analogues are essential to HAART therapy.

Nonnucleoside Reverse Transcriptase Inhibitors

Nonnucleoside reverse transcriptase inhibitors bind noncompetitively and inhibit reverse transcriptase(*3*). These inhibitors bind to the reverse transcriptase enzyme and block activity at the DNA polymerase by causing a conformational change and disrupting the catalytic site(*7*). Nonnucleoside reverse transcriptase inhibitors do not require phosphorylation to become active and are not incorporated into the DNA(*7*). By affecting the flexibility of the enzyme, the binding of the inhibitors is blocked and the reverse transcriptase is unable to synthesize DNA(*3*).

Most of the mutations that occur with nonnucleoside reverse transcriptase inhibitors are located in the pocket target, which reduce the affinity of the drug(*3*). The subtle differences in the

interactions between the various nonnucleoside reverse transcriptase inhibitors and the hydrophobic pocket often cause mutations at Y181C, Y188C, K103N, G190A, and V106A(*3*).

Fusion Inhibitors

The HIV membrane contains gp120, and gp41(*7*). In the envelope structure, gp120 molecules make up the cap, gp41 forms the stalk and it is anchored by the viral lipid bilayer(*7*). By utilizing a sequence of interactions between the glycoprotein complex (gp120-gp41) and specific cell surface receptors, HIV-1 can enter target cells through a conformational change in gp41(*3*). During the early stages of this process, two motifs from gp41 interact with the heptad repeat(HR) 1 and 2 and unite to form a 6-helix hairpin bundle structure in order to attach the virus to the target(*7*). The membrane of the virus and the target are brought close together and fused by rearranging gp41(*3*). The hydrophobic region of gp41, HR2, folds onto a proximal hydrophobic region, HR2, which shortens the molecule(*3*). The only FDA approved fusion inhibitor is Enfuvirtide, which is a 36 amino acid peptide derived from HR2. Enfuviritide binds to the HR1 region of gp41 and blocks the formation of the 6 helix bundle necessary for fusion. Destabilizing HR1 binding prevents HIV from affecting the target(*3*).

Viral resistance usually occur from mutations that are located in a stretch of ten amino acids within HR1(*3*). Changes to the amino acids in gp41 and gp120 outside HR1 are associated with significant differences in the effectiveness of Enfuvirtide(*3*).

Highly Active Antiretroviral Treatment (HAART)

In 1996, Highly Active Antiretroviral Treatment (HAART) became available and was effective because it would combine multiple mechanisms and drugs in order to avoid resistance(*3*). The development of HAART transformed a fatal disease into a chronic and manageable disease(*8*). The strategy used by HAART involves multiple mechanisms for virus inhibition(*3*). The probability that that these mutants would be able to resist all of the drugs in the regimen would be much lower than any

single drug alone(*3*). Currently there are four classes of inhibitors: nucleoside and nucleotide analogues, reverse transcriptase inhibitors, fusion inhibitors, and protease inhibitors(*3*).

The HIV-1 Protease

The HIV-1 protease is encoded in the viral Pro gene, which is downstream of the gag gene and upstream from the pol gene and is expressed as the polyprotein precursor Gag-Pro-Pol(*8*). In order for HIV to become active, the Gag-Pro-Pol precursor must dimerize in order for HIV protease to become active(*8*). Although little is known about how the protease excises itself from the polyprotein precursor, the initial cleavage occurs in cis at a novel cleavage site(*8*).

HIV-1 protease is a homodimeric aspartyl protease with its active site at the dimer interface with two aspartic acids located at the base of the active site(*8*). The enzymatic mechanism is a push pull mechanism with an acid-base catalysis with a water molecule(*8*). The water molecule transfers a proton to the carboxyl groups of the aspartic acids and then is transferred to the targeted cleaved peptide bond(*8*). During this process, a non-covalent tetrahedral intermediate is briefly formed (Figure 3,Figure 4)(*8, 13*). A mobile beta turn serves as a flap to cover the active site within each subunit(*8*). In order for the substrate to have access to the active site, the flaps must be able to move away (Figure 5) (*8*). Once the substrate is recognized and bound, the flaps must move over the active site and lock down over the bound substrate. This step completes the active site cavity and permits substrate cleavage(*8*). The turnover number (K_{cat}) for the HIV protease dimer, which is the number of enzymatic reactions per unit time, is 4.9 s^{-1}(*14*). The Michaelis constant, K_m, is the approximate affinity for enzyme to substrate(*14*). A small K_m means V_{max} has reached a low concentration of substrate and approximately indicates increased binding affinity to the substrate(*14*). The measured K_m for the HIV protease from another experiment was 103 µM and the V_{max} was 164 nM min^{-1}.

HIV protease cleaves the Gag and Gag-Pro-Pol polyproteins at ten sites and the cleavage site recognize eight different substrates throughout the P4' and P4 pockets(*8*). These ten sites are quite diverse and are cleaved with different efficiencies(*8*). Despite these differences, most of these substrate sites contain many common characteristics such as a branched amino acid residue at the P2 site, a hydrophobic residue at the P1 site, and an aromatic or proline at the P1' site(*8*). However, the difference in substrate sequences result in variation of cleavage rates and contribute to the order the substrates are cleaved(*8*).

Development of Resistance

The high levels of virus production along with the low fidelity rate of the HIV's reverse transcriptase results in a development of diverse viral mutations which can infect target cells at an extremely rapid rate (*3*). During transcription, the reverse transcriptase is error prone with mutations averaging one mutation per viral genome transcribed(*3*). The reverse trasnscriptase is responsible for making a double stranded DNA copy from a single stranded RNA template molecule; however, reverse transcriptase has no proofreading ability(*15*). While substitution is the most common error, duplications, insertions, and recombination also frequently occur(*3*). When under selective pressure, even a single substitution of an amino acid can produce high levels of resistance within a matter of weeks(*3*). In fact, nearly 70% of 99 residues in HIV-1 protease are known to mutate(*8*).

The Substrate Envelope Hypothesis

The conserved shape substrates that adopt within the active site of the protease is called the substrate envelope (Figure 6) (*8*). The protease recognizes different substrates based on shape, rather than a specific amino acid sequence(*4*). Key points of resistance tend to occur in clusters when the inhibitors protrude outside of the substrate envelope(*4*). This can be seen when superimposing the substrate envelope and the inhibitor envelope, which is the volume occupied by the overlapped inhibitors (Figure 7) (*16*). As seen in Figure 7, the inhibitor envelope protrudes outside the substrate

envelope perhaps explaining why current drugs face drug resistance. However, when superimposing the substrate envelope with darunavir, there is little resistance encountered (Figure 8). By developing an inhibitor that fits within this uniform region, the inhibitor would be more likely to evade viral resistance (*4*). A mutation that would affect inhibitor binding would also reduce the affinity for the substrate, preventing HIV's propagation(*4*). By using the substrate envelope as a basis for drug design, drug resistance for mutations should be less likely, since these enzymes must allow the substrate to bind in order to propagate(*4*).

Protease Inhibitors

HIV-1 protease is a complex enzyme with 2 identical halves and its active site is located at the base of the cleft(*7*). The HIV-1 protease cleaves the large viral precursor polypeptide chains into small and functional proteins, allowing HIV to propagate(*7*). However, by creating low molecular weight molecules that tightly fit into the HIV protease, HIV replication is hindered(*17*).

Currently there are nine FDA approved protease inhibitors: saquinavir (1995), ritonavir (1996), indinavir (1996), nelfinavir (1997), amprenavir (1999), lopinavir (2000), atazanavir (2003), fosamprenavir (2003), tipranavir (2005), and darunavir (2006)(*8*). The inhibitors are competitive active site that bind to the protease from the nanomolar to picomolar range(*8*). Despite the many protease inhibitors over the last decade, there is a need for more potent inhibitors with improved pharmacokinetics, decreased side effects, and increased effectiveness against mutant HIV proteases(*8*). Many of the protease inhibitors have closely overlapping structures and interactions(*8*).

Seven of the nine inhibitors are peptidomimetics, which are small proteins designed to mimic a peptide and to mimic the enzymatic transition state(*8*). The enzymatic transition state contains several noncleavable dipeptide isosteres to mimic the transition state of substrate cleavage(*8*). These inhibitors fit into the P2-P2' region of the active site with a hydrophobic cyclical side chain at P1 and a bulk

functional group at P1' region(*8*). Many of the drugs contain large hydrophobic moieties that interact with the hydrophobic P2-P2' pocket in the active site(*17*). However, Tipranavir is an exception and has demonstrated potent inhibition against clinical isolates resistant to multiple inhibitors due to its molecular flexibility and allows tipranavir to fit in the active site of the protease that has become to resistant to other protease in inhibitors(*8*). Darunavir is the second nonpeptidic protease inhibitor, replacing the tetrohydrofuranyl (THF) with a bis-THF component(*8*).

Resistance to Protease Inhibitors

While protease inhibitors have strong affinity for the active site, resistance can emerge within the substrate binding site of the protease (Figure 9) (*3*). Mutations allow the protease variant to maintain its function by cleaving 10 natural substrates in the Gag and Gag-Pro-Pol polyproteins(*17*). Amino acid changes however can reduce the affinity for most binding inhibitors because they can modify the number and points of contact between the inhibitors and protease(*3*). When comparing protease inhibitors to the natural substrates, inhibitors tend to fit tighter and occupy more space within the active site(*3*). However, inhibitors tend to be smaller than substrates in order to maintain bioavailability and have a different shape than the substrates(*17*). Natural substrates are less tight and more variable and allows for cleavage of the polyproteins required for proper assembly of the viral particle(*3*). Substitutions of amino acids near the cleavage sites of the gag polyprotein increase the level of resistance and explicative capacity of HIV(*3*). However, only 34 of 99 residues of HIV-1 protease have had clinical significance(*17*).

The most common multidrug resistant variants are M1 (L10I, G48V, I54V, L63P, V82A) (Figure 10), M2 (D30N, L63P, N88D), M3 (L10I, L63P, A71V, G73S, 184V, L90M) (Figure 11), and M4(I50V, A71V) (Figure 12) (*18*). The M1 variant is resistant to nelfinavir and the M4 variant is resistant to amprenavir and darunavir(*18*). However, the mutations D30N, G48V, V82A, I84V, I50V, and I50L affect inhibitor binding by altering specific points within the active site(*17*). D30N is a common mutation associated

with nelfinavir; G48V is a common mutation for saquinavir; I50V is a common mutation associated with amprenavir and darunavir; I50L is a common mutation with atazanavir; V82A and I84V affect the binding of all inhibitors(*17*).

There are also mutations that alter the balance between substrate recognition and inhibitor binding(*17*). By increasing the flexibility of the HIV-1 protease, it may affect the inhibitor by increasing the rate of dissociation between the protease and the inhibitor(*17*). When looking at HIV-1 patients, only five to fifteen mutations are found on the protease gene with mutations both inside and outside the active site(*17*). Common outside mutations include L10I, I54V, I54T, A71V, A71T, V77I, and L90M and these mutations not only affect inhibitor binding, but also compensate for viability and fitness of the enzyme(*17*). When overlaying most protease inhibitors over the substrate envelope, many of protease inhibitors protrude beyond the substrate envelope at the P3 and P2' subsites(*16*).

Optimizing the Protease Inhibitors

Two characteristics for an effective protease inhibitor are having high potency on both the wild-types and having a broad range of drug resistant strains(*19*). Protease inhibitors also must have favorable pharmacokinetic properties such as oral administration(*19*). Prior to darunavir, most discovery teams focused on optimizing for wild-type potency rather than compounds effective on resistant strains(*19*). The binding affinity of Darunavir was found to be much higher than other analyzed protease inhibitors and has been shown to fit well within the substrate envelope(*8*).

Due to the high binding affinity and effectiveness against resistant strains, structure activity relationship should be analyzed in order to optimize analogues of darunavir. Darunavir contains the bis-THF moiety for improved interaction in the P2'-pocket(*19*). The additional hydrogen bonding between bis-THF ring and protease inhibitor backbone results in activity against a broad range of clinical isolates with an IC_{50} in the nanomolar range including multidrug resistant HIV-1(*8*). When introducing the *p*-

nitro, *p*-acetyl, and p-iodo groups, the withdrawing substituent groups on darunavir did not influence the potency on the wild-type virus. However, changing the substituent groups has demonstrated increased activity in the mutant viruses(*19*).

When looking at the most effective protease inhibitors, the NH interaction with Asp30 was found in the compounds with the highest activity on the wild-type enzyme(*19*). The backbone NH of Asp30' also forms a hydrogen bond with the bond accepting groups substituted on the benzene ring(*19*). Since the scaffold was anchored by hydrogen bonds to the aspartyl dyad, the positions of a variety of substituent groups can be predicted in the target subsites(*20*). This was due to the nature of the arrangement of lone pairs in the substituent group and leading to a stronger interaction and is crucial against the wild-type protease(*19*). The position of the oxygen atoms might also influence the presence of van der Waals interactions of the methyl group side chains present in the P2' pocket(*19*). Other essential substituent groups are the *p*-NH_2, *p*-OH, and *p*-CH_2NH_2, which form a hydrogen bond with the backbone of the carboxylic acid of Asp30(*19*). Another water mediated hydrogen bond was also forms with the side chain of Asp30 in the P2' pocket(*19*).

While the first generation of protease inhibitors binding to the HIV protease is mainly entropy driven, new protease inhibitors are both entropy and enthalpy driven(*19*). The enthalpic driven binding contributes to the tight interactions between the inhibitor and the enzyme(*19*). The affinity to backbone interactions allow the protease inhibitor to retain their wild-type potency(*19*). Unlike the newer protease inhibitors, the first generation protease inhibitors do not fit well in the substrate envelope and resistance is encountered at positions that protrude beyond the substrate envelope(*19*).

Testing For Drug Resistance

The HIV protease allows for correct processing of viral polyproteins and maturation of HIV by cleaving the polyproteins(*14*). In order to identify the effectiveness of protease inhibitors, an assay

based on intramolecular fluorescence resonance energy transfer (FRET) can allow for a quick and practical method for screening large number of inhibitors(*14*). The natural substrate of HIV protease can be covalently labeled with the donor and acceptor(*21*). The energy transfer between the donor and acceptor occurs due to long range dipole-dipole interactions(*21*).

The specific assay used fluorgenic substrates 4-(4-dimethyaminophenylazo)benzoic acid (DABCYL) and Ser Gln Asn Tyr Pro Ile Val Gln-5[(2-aminoethyl)amino]naphthalene-1 sulfonic acid (EDANS) (Figure 13) (*14*). EDANS, which attaches at the C terminus, is a fluorescent donor and DABCYL, which attaches to the N terminus, is a non-fluorescent quenching acceptor (Figure 14) (*14*). Prior to cleavage from the protease, the fluorescence in EDANS is reduced significantly due to intramolecular resonance energy transfer (RET) with DABCYL(*14*). Since RET is no longer significant after 100 Å, fluorescent activity of EDANS is restored after cleavage from the protease and liberation from the DABCYL peptide fragment(*14*). By measuring the initial velocity of the fluorescence activity over time, the effectiveness of protease inhibitors can be analyzed; an decrease in fluorescence suggests a decrease in protease activity(*14*). After measuring the initial velocity, the K_i values were calculated using the Morrison equation(*18*). Since these HIV protease inhibitors have picomolar K_i values, the Morrison allows for precise and accurate calculation of inhibitor binding.

Drug Discovery and Future Protease Inhibitors

Many of the initial leads for the HIV protease inhibitors were found within the pharmaceutical companies libraries which were originally designed for aspartyl protease renin(*8*). Structure based drug design was then used to optimize inhibitor design and many laboratories would cocrystallize patented compounds(*8*). With these patented compounds, other laboratories would determine find out alternative scaffolds that would preserve the same contacts but with better pharmacokinetics and bioavaibility(*8*). Protease inhibitors that fit within the substrate envelope can also be discovered via computational design and by using quantitative structure activity relationship (QSAR)(*2*). Using QSAR

and protein crystallography is a simple and practical way of using binding affinity data for the design of new HIV protease inhibitors(*2*).

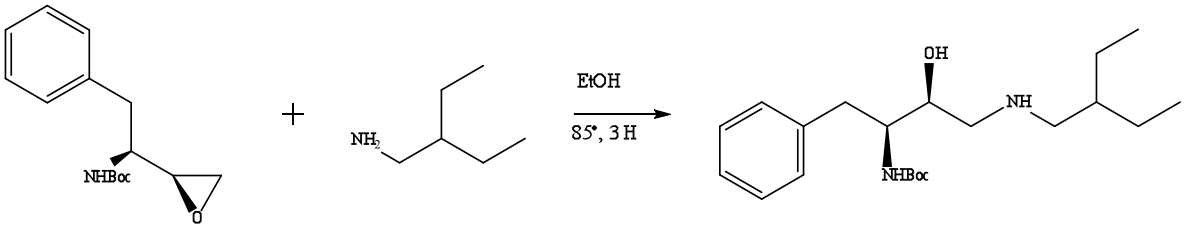
When compared to previous protease inhibitors, darunavir fits better into the substrate envelope, encounters less drug resistance than previous inhibitors, and shows higher binding affinity(*17*). While current strategies focus on analyzing failure modes of existing drugs and designing new compounds with high efficiency against known mutants, many of these drugs encounter new and unanticipated modes of resistance(*18*). By using structure activity relationship (SAR) and computational methods, tight binding inhibitors that are not susceptible to escape mutations even when specific mutations are unknown need to be designed(*17*).

Materials and Methods

Synthesis of the HIV-1 Protease Inhibitors

Synthesis of *tert*-butyl (2S,3S)-4-(2-ethylbutylamino)-3-hydroxy-1-phenylbutan-2-ylcarbamate (KY-15)

Ring Opening of Epoxide



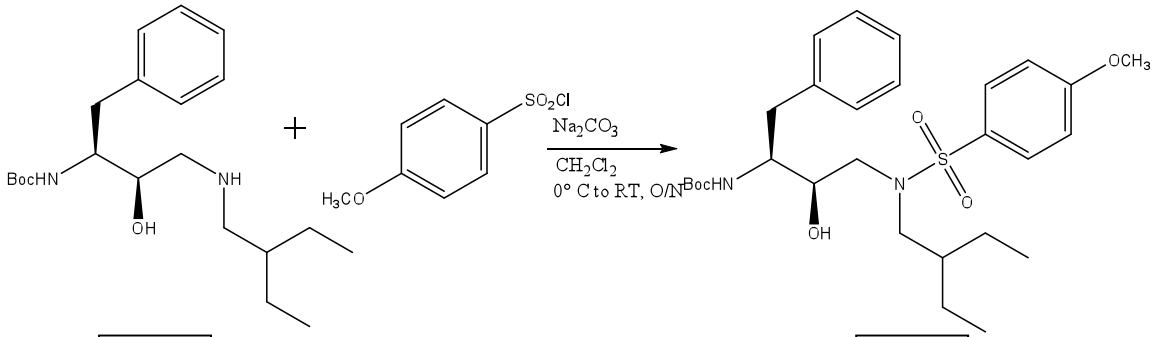
KY-15

Procedure:

To a solution of the chiral epoxide, (1S, 2S)-(1-oxiranyl-2-phenylethyl) carbamic acid tert-butyl ester (7.89, 30 mmol), in EtOH (100 ml) was added to 2-ethylbutanol-1-amine (3.04 g, 30 mmol) and the mixture was heated at 85° C for 3 hours. After the solution was cooled to room temperature, the solvent was removed under pressure. The product was then purified by recrystallization from an EtOAC-hexanes (1:10) mixture and provided the product as a white solid (5.09, 46.5%).

Synthesis of *tert*-butyl (2S,3R)-4-(N-(2-ethylbutyl)-4-methoxyphenylsulfonamido)-3-hydroxy-1-phenylbutan-2-ylcarbamate (KY-17)

Synthesis of (R)-(hydroxyethylamino)sulfonamides



KY-15 KY-17

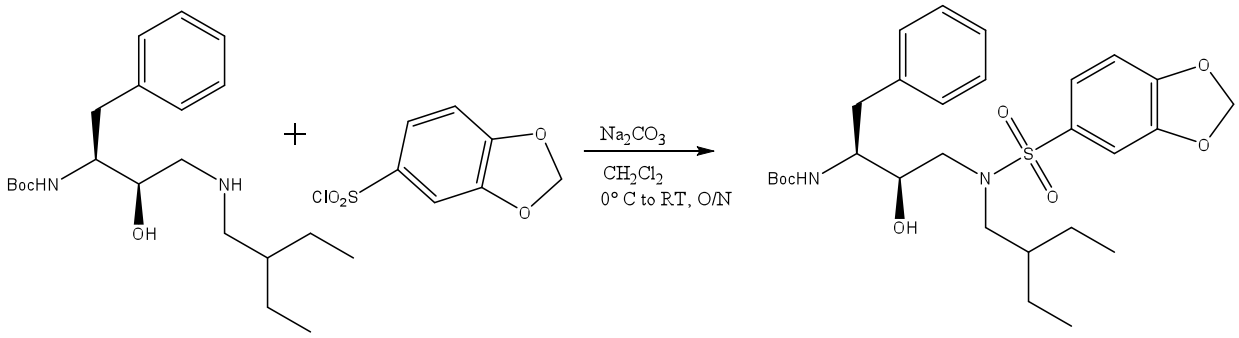
Procedure:

To an ice cooled solution of the above secondary amine, KY-15 (1.82 g, 5 mmol), in CH_2Cl_2 (50 mL) was added to an aqueous solution of Na_2CO_3 (0.848 g, 8 mmol in 20 mL of H_2O) followed by the slow addition of 4-methylphenylsulfonylchloride solid (1.07 g, 5.20 mmol). After 15 minutes, the reaction mixture was warmed to room temperature and stirred. No starting material was detected by TLC and the reaction mixture was diluted with CH_2Cl_2 and the organic layer extract was collected. After collecting the organic layer, the extract was dried with Na_2SO_4, filtered, and evaporated under reduced pressure. Using an EtOAc-hexanes (1:4) mixture as an eluent, the residue was purified by flash chromatography on silica gel. This provided a product (2.35 g, 87.9%) as a white foamy solid. After the H^1 NMR was taken, the residue was again purified by flash chromatography on silica gel using an EtOAc-hexanes mixture (1:4) as an eluent. This provided a pure product, KY-17 (1.94 g, 72.6%), as a foamy white solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.72-7.69 (m, 2H), 7.32-7.19 (m, 5H), 6.99-6.96 (m, 2H), 3.97 (br s, 1H), 3.90 (s, 3H), 3.75 (s, 3H), 3.11-2.98 (m, 4H), 2.93-2.81 (m,2H), 1.46-1.38 (m, 2H), 1.33 (s, 9H), 1.29-1.24 (m, 3H), 0.84-0.79 (m, 6H).

Synthesis of *tert*-butyl (2S,3R)-4-(N-(2-ethylbutyl)benzo[d][1,3]dioxole-5-sulfonamido)-3-hydroxy-1-phenylbutan-2-ylcarbamate (KY-19)

Synthesis of (R)-(hydroxyethylamino)sulfonamides



KY-15 KY-19

Procedure:

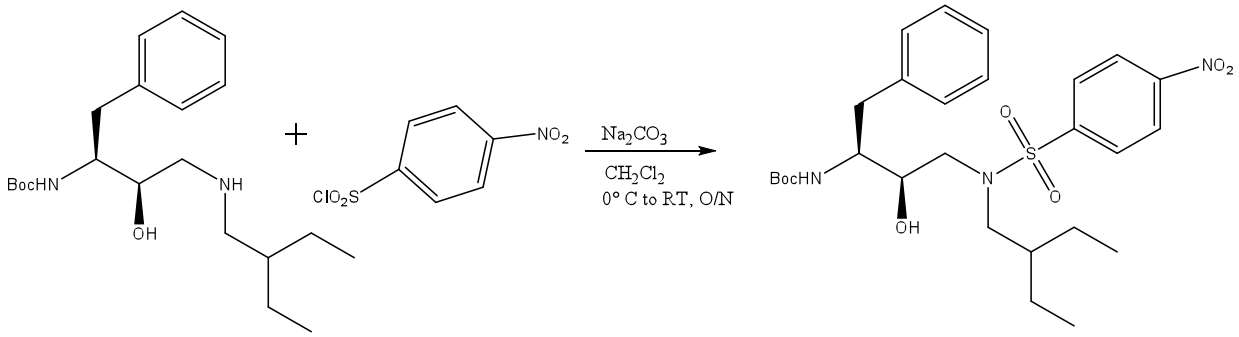
To an ice cooled solution of the above secondary amine, KY-15 (2.19 g, 6 mmol) in CH_2Cl_2 (50 mL) was added to an aqueous solution of Na_2CO_3 (1.4 g, 13.2 mmol in 20 mL of H_2O) followed by the slow addition of 1,3-benzodioxole-5-sulfonylchloride solid (1.40 g, 6.24 mmol). After 15 minutes, the reaction mixture was warmed to room temperature and stirred. However starting material was detected by TLC and the reaction mixture was cooled down. Na_2CO_3 (0.5 g in 5 mL H_2O) and 1,3-benzodioxole-5-sulfonylchloride (0.3 g) were added to the reaction mixture. After 15 minutes, the reaction mixture was warmed to room temperature and stirred. No starting material was detected by TLC. The reaction mixture was diluted with CH_2Cl_2 and the organic layer extract was collected. After collecting the organic layer, the extract was dried with Na_2SO_4, filtered, and evaporated under reduced pressure. Using an EtOAc-hexanes (1:5) mixture as an eluent, the residue was purified by flash chromatography on silica gel. This provided a product (2.35 g, 87.9%) as a white foamy solid. After the NMR was taken, the residue was purified by flash chromatography on silica gel. This provided a pure product, KY-19 (2.70 g, 82.0%), as a foamy white solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.33 (dd, *J* = 8.4, 2.0 Hz, 1H), 7.30-7.28 (m, 1H), 7.27-7.20 (m, 3H), 7.18 (d, *J* = 1.6 Hz, 1H), 6.88 (d, *J* = 8.4 Hz, 1H), 6.09 (s, 2H), 4.59 (d br, *J* = 6 Hz, 1H), 3.93 (s, 1H), 3.75 (s, 1H), 3.14-2.98 (m, 3H), 2.93-2.82 (m, 2H), 1.51-1.37 (m, 1H), 1.34 (s, 12H), 1.31-1.24 (m, 2H), 0.86-0.79 (m, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 156.23, 151.68, 148.53, 137.94, 131.64, 129.80 (2C), 128.72 (2C), 126.69,

123.37, 108.58, 107.86, 102.55, 79.94, 76.92, 73.19, 55.09, 54.74, 53.95, 39.20, 35.89, 28.48 (2C), 23.28, 22.96, 10.80, 10.52.

Synthesis of *tert*-butyl (2S,3R)-4-(N-(2-ethylbutyl)-4-nitrophenylsulfonamido)-3-hydroxy-1-phenylbutan-2-ylcarbamate (KY-21)

Synthesis of (R)-(hydroxyethylamino)sulfonamides



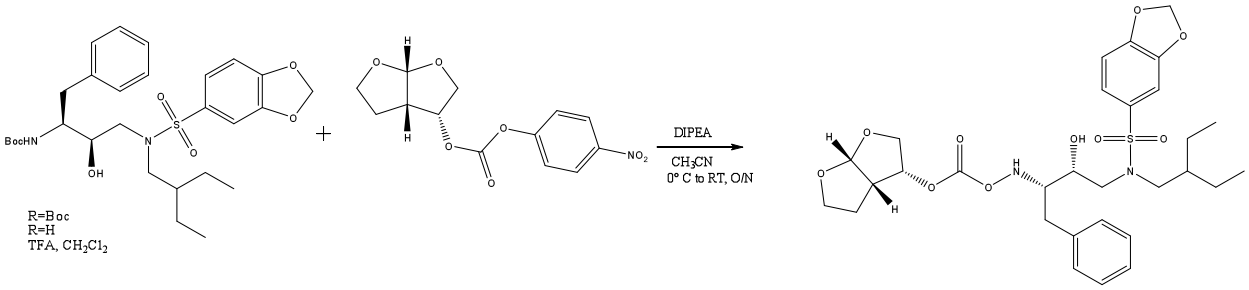
KY-15 KY-21

Procedure:

To an ice cooled solution of the above secondary amine, KY-15 (2.19 g, 6 mmol) in CH_2Cl_2 (50 mL) was added to an aqueous solution of Na_2CO_3 (1.59 g, 15 mmol in 20 mL of H_2O) followed by the slow addition of 4-nitrobenzenesulfonylchloride solid (1.53 g, 6.29 mmol). After 15 minutes, the reaction mixture was warmed to room temperature and stirred. The reaction mixture was diluted with CH_2Cl_2 and the organic layer extract was collected. After collecting the organic layer, the extract was dried with Na_2SO_4, filtered, and evaporated under reduced pressure. This product was purified by recrystallization from EtOH and provided the product (1.91 g, 57.9%) as a white solid.

Synthesis of *N*-(2-ethylbutyl)-*N*-((2*R*,3*S*)-3-(((3*R*,3a*S*,6a*R*)-hexahydrofuro[2,3-b]furan-3-yloxy)carbonyloxyamino)-2-hydroxy-4-phenylbutyl)benzo[d][1,3]dioxole-5-sulfonamide (KY-23)

Boc Deprotection and Coupling



KY-19 KY-23

Procedure:

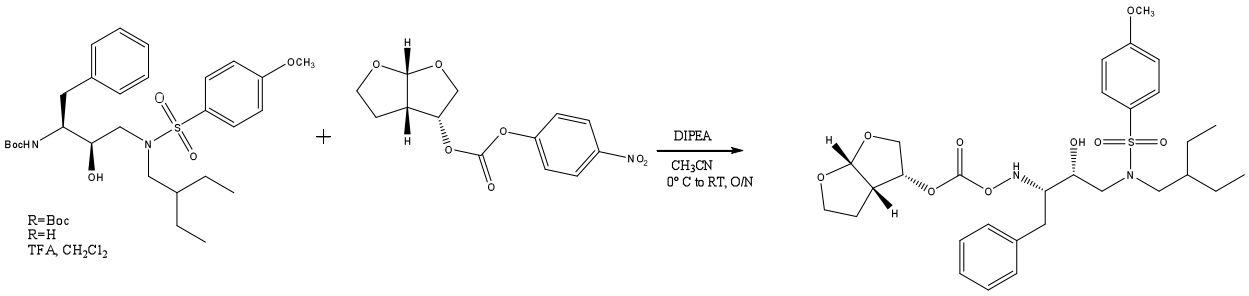
A solution of the above (R)-(hydroxyethylamino)sulfonamide compound, KY-19 (0.5 g, 0.91 mmol), in trifluoroacetic acid (3 mL) and CH_2Cl_2 (10 mL) was stirred at room temperature for 1 hour. After this period the reaction mixture was concentrated under reduced pressure and the residue was dissolved in toluene (5 mL) and the solvent again was evaporated under reduced pressure. The residue was dissolved in acetonitrile (10 mL) and cooled to 0° C. Diisopropylethylamine (DIPEA) (0.45 mL, 2.73 mmol) and THF alcohol mixed carbonate (0.26 g, 0.91 mmol) were added to the solution. The mixture was stirred at room temperature overnight. However, starting material was detected in the TLC and additional THF alcohol mixed carbonate (0.20 g) was added to the solution and was stirred for an additional 3 hours. The reaction mixture was diluted with EtOAc (70 mL) and was washed with water (15 mL) and saturated sodium chloride solution (15 mL). The organic portion was extracted, dried with Na_2SO_4, filtered, and concentrated under reduced pressure. The residue was purified by flash chromatography on silica gel using an EtOAc-hexanes (1:1) mixture as an eluent to provide the pure product (0.50 g, 60.2%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.34 (dd, *J* = 8.4, 2.0 Hz, 1H), 7.30-7.26 (m, 2H), 7.22 (br d, *J* = 7.2 Hz, 3H), 7.17 (d, *J* = 3 Hz, 1H), 6.90 (d, *J* = 8.0 Hz, 1H), 6.10 (s, 2H), 5.64 (d, *J* = 4.8 Hz, 1H), 5.01 (dd, *J* = 14.4, 6.4 Hz, 1H), 4.88 (d, *J* = 9.2 Hz, 1H), 3.95 (dd, *J* = 9.6, 6.4 Hz, 1H), 3.89-3.78 (m, 3H), 3.71-3.65 (m, 3H), 3.14-

3.07 (m, 2H), 3.05 (dd, *J* = 14.7, 9.2 Hz, 1H), 2.94-2.87 (m, 1H), 2.86-2.77 (m,2H), 1.48-1.40 (m, 3H), 1.36-1.25 (m, 3H), 0.85-0.81 (m, 6H).

Synthesis of *N*-(2-ethylbutyl)-*N*-((2*R*,3*S*)-3-(((3*R*,3a*S*,6a*R*)-hexahydrofuro[2,3-b]furan-3-yloxy)carbonyloxyamino)-2-hydroxy-4-phenylbutyl)-4-methoxybenzenesulfonamide (KY-25)

Boc Deprotection and Coupling



KY-17 KY-25

Procedure:

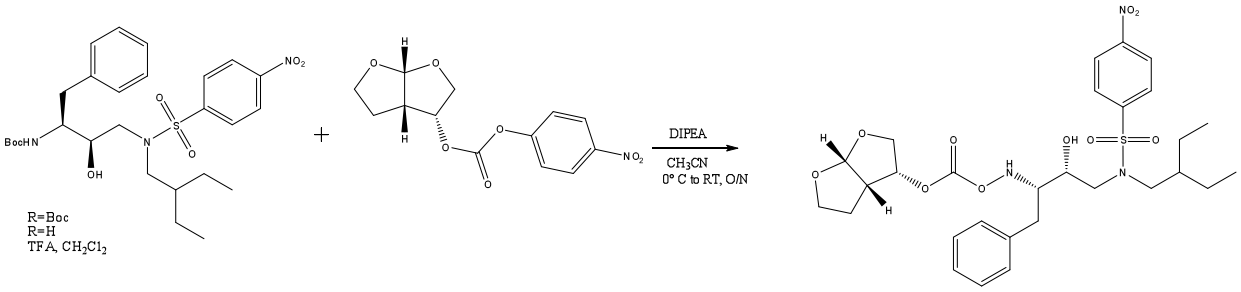
A solution of the above (R)-(hydroxyethylamino)sulfonamide compound, KY-17 (0.5 g, 0.94 mmol), in trifluoroacetic acid (3 mL) and CH_2Cl_2 (10 mL) was stirred at room temperature for 1 hour. After this period the reaction mixture was concentrated under reduced pressure and the residue was dissolved in toluene (5 mL). The solvent was then evaporated under reduced pressure. The residue was dissolved in acetonitrile (10 mL) and cooled to 0° C. Diisopropylethylamine (DIPEA) (0.47 mL, 2.82 mmol) and THF alcohol mixed carbonate (0.29 g, 0.94 mmol) were added to the solution. The mixture was stirred at room temperature overnight. No starting material was detected in the TLC. The reaction mixture was diluted with EtOAc (100 mL) and was washed with water (15 mL) and saturated sodium chloride solution (15 mL). The organic portion was dried with Na_2SO_4, filtered, and concentrated under reduced pressure. The residue was purified by flash chromatography on silica gel using an EtOAc-hexanes (1:1) mixture as

an eluent to provide the product (0.49 g). After detecting impurities in the NMR, the residue was again then purified by flash chromatography on silica gel using an EtOAc-hexanes (2:5) mixture as an eluent.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.73-7.69 (m, 2H), 7.30-7.21 (m, 6H), 7.01-7.69 (m, 2H), 5.64 (d, *J* = 5.2 Hz, 1H), 5.01 (dd, *J* = 6, 4.4 Hz, 1H), 4.88 (d, *J* = 9.6, 1H), 3.96-3.92 (m, 1H), 3.88-3.79 (m, 2H), 3.71-3.65 (m, 2H), 3.16-3.03 (m, 3H), 2.98-2.87 (m, 2H), 2.84-2.78 (m, 2H), 1.66-1.58 (m, 3H), 1.48-1.42 (m, 3H), 1.34-1.25 (m, 4H), 0.845-0.799 (m, 6H).

Synthesis of *N*-(2-ethylbutyl)-*N*-((2*R*,3*S*)-3-(((3*R*,3a*S*,6a*R*)-hexahydrofuro[2,3-b]furan-3-yloxy)carbonyloxyamino)-2-hydroxy-4-phenylbutyl)-4-nitrobenzenesulfonamide (KY-27)

Boc Deprotection and Coupling



KY-21 KY-27

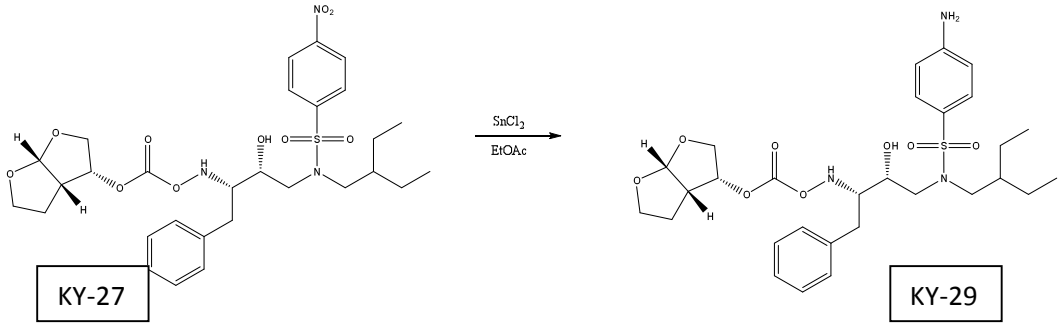
Procedure:

A solution of the above (R)-(hydroxyethylamino)sulfonamide compound, KY-21 (0.6 g, 1.092 mmol), in trifluoroacetic acid (3 mL) and CH_2Cl_2 (10 mL) was stirred at room temperature for 2 hours. After this period the reaction mixture was concentrated under reduced pressure and the residue was dissolved in toluene (5 mL) and the solvent again was evaporated under reduced pressure. The residue was dissolved in acetonitrile (10 mL) and cooled to 0° C. Diisopropylethylamine (DIPEA) (0.529 mL, 3.276 mmol) and THF alcohol mixed carbonate (0.323 g, 1.092 mmol) were added to the solution. The resulting mixture was stirred at room temperature overnight. However, starting material was detected

in the TLC and $C_{13}H_{13}NO_7$ (0.25 g) and DIPEA (0.2 mL) were added to the solution and was stirred overnight. The reaction mixture was diluted with EtOAc (100 mL) and was washed with water (15 mL) and saturated sodium chloride solution (15 mL). The organic portion was dried with Na_2SO_4, filtered, and concentrated under reduced pressure. The residue was purified by flash chromatography on silica gel using an EtOAc-hexanes (1:1) mixture as an eluent to provide the pure product (0.55 g, 83.0%) as a foamy yellow solid.

Synthesis of 4-amino-*N*-(2-ethylbutyl)-*N*-((2*R*,3*S*)-3-(((3*R*,3a*S*,6a*R*)-hexahydrofuro[2,3-b]furan-3-yloxy)carbonyloxyamino)-2-hydroxy-4-phenylbutyl)benzenesulfonamide (KY-29)

General Procedure for Reduction of the Nitro Group



Procedure:

A mixture of the nitro compound, KY-27 (0.55 g, 0.45 mmol), and $SnCl_2 \cdot 2H_2O$ (1.86 g, 8.25 mmol) in EtOAc (20 mL) was heated at 80°C for 3 hours. The reaction mixture was allowed to cool to room temperature and was treated with saturated aqueous $NaHCO_3$ solution (15 mL). It was diluted with EtOAc (100 mL) and the organic layer was collected. The aqueous layer was further extracted with EtOAc (100 mL, 2 x). The combined organic extract was washed with saturated NaCl solution, dried with Na_2SO_4, filtered, and evaporated to yield a yellowish solid. Flash chromatography on silica gel, using an EtOAc-hexanes (1:1) mixture as an eluent provided the product (0.280 g, 29.5%) as a white solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.56-7.53 (m, 2H), 7.38-7.24 (m, 2H), 7.25-7.19 (m,3H), 6.71-6.67 (m, 2H), 5.64 (d, *J* = 5.2 Hz, 1H), 5.00 (dd, *J* = 14.4, 6.0 Hz, 1H), 4.88 (d, *J* = 9.2 Hz, 1H), 4.16 (s, 2H), 3.96-3.92 (m, 1H), 3.89-3.80 (m,3H), 3.71-3.65 (m, 2H), 3.15-3.02 (m, 3H), 2.95-2.87 (m, 2H), 2.83-2.77 (m, 2H), 1.64-1.57 (m, 1H), 1.58-1.40 (m, 3H), 1.33-1.25 (m, 4H), 0.84-0.80 (m, 6H).

HIV-1 Protease Inhibition Assays

The HIV-1 protease inhibitors activities were determined by fluorescence resonance energy transfer (FRET) method. The protease substrate used was Arg-Glu(EDANS)-Ser-Gln-Asn-Tyr-Pro-Ile-Val-Gln-Lys(DABCYL)-Arg, which was purchased from Molecular Probes. The energy donor (EDANS and acceptor (DABCYL) dyes were labeled at both ends of the peptide. Fluorescence measurements were carried out on a fluorescence spectrophotometer (Photon Technology International) at 30° C and excitation and emission wavelengths were set at 340 and 490 nm. Each reaction was recorded for 10 minutes.

Wild-type HIV-1 protease (Q7K) was desalted through PD-19 columns (Amersham Biosciences) and sodium acetate (20 mM, pH 5) were used as an elution buffer. The protease concentrations were around 50 nM estimated by UV spectrophotometry at 280 nm. The protease inhibitors were dissolved in dimethylsulfoxide) DMSO and diluted to appropriate concentrations. Protease (2 µL) and inhibitor (2 µL) or DMSO were mixed and incubated for 20-30 minutes at room temperature before initializing the substrate cleavage action. In all experiments, 150 µL of 1 µM substrate were used in substrate buffer [0.1 M sodium acetate, 1 M sodium chloride, 1 mM ethylenediaminetetraacetic acid (EDTA), 1 mM dithiothreitol (DTT), 2% DMSO and 1 mg/mL bovine serum albumin (BSA) with an adjusted pH 4.7]. Inhibitor binding constant (K_i) values were obtained by nonlinear regression fitting (GraFit 5, Erithacus software) to plot the initial velocity as a function of inhibitor concentrations based on the Morrison equation. The initial velocities were derived from the linear range of reaction curves.

Results

The substrate envelope hypothesis has suggested the mechanism of darunavir's potency and effectiveness against drug resistant HIV-1. By basing future protease inhibitor design on the substrate envelope, the inhibitor can have tighter binding with the protease. Along with the tighter binding, mutations that decrease inhibitor binding would also decrease substrate affinity. The compounds KY-23, KY-25, and KY-29 were analogs of darunavir and synthesized with the addition of 2-ethylbutyl group and various sulfonyl chloride groups. In order to determine the inhibitor binding to the protease, a fluorescence resonance energy transfer (FRET) assay was performed by using various inhibitor concentrations. The initial velocities were measured and used with the Morrison equation in order to calculate the K_i values.

The compounds, KY-23, KY-25, and KY-29, were selected based on computational analysis of prospective compounds that would fit within the substrate envelope. The general mechanism can be seen in Figure 15. The initial step required the opening of the epoxide ring with the primary amine, forming an amino alcohol with the 2-ethylbutyl ligand. The second step required the reaction of the amino alcohol with various sulfonyl chloride groups, yielding a sulfonamide with selected substituent groups. The third step required deprotection of the Boc group and the addition of bis-THF. Bis-THF has been shown to be effective for interaction in the P2' pocket during the initial discovery of darunavir. The addition of the bis-THF, yielded the final compounds KY-23, KY-25, and KY-29.

The HIV-1 protease inhibitors activities were determined by fluorescence resonance energy transfer (FRET) method. The inhibitor binding constant (K_i) values were obtained by nonlinear regression fitting (GraFit 5, Erithacus software) in order to plot the initial velocity as a function of inhibitor concentrations based on the Morrison equation. The initial velocities were derived from the linear range of reaction curves. When compared to a Dixon and Lineweaver-Burk plot, the Morrison equation provided extreme precision and accuracy for the K_i values.

A donating group (EDANS) and accepting group (DABCYL) is attached was attached to a natural substrate of HIV protease. Prior to protease cleavage, DABCYL would quench EDANS, resulting in no detection of fluorescence form EDANS. After the cleavage of the substrate from protease, EDANS was no longer quenched by DABCYL, allowing EDANS to fluoresce. The activity of the protease can be monitored by analyzing the change of EDANS fluorescence intensity.

In Figure 16, Figure 17, Figure 18, and Figure 19, the x axis represented time in seconds and the y axis represented fluorescence intensity. The legend showed increasing concentrations of protease inhibitor and controls of DMSO, which had no protease inhibitor present. The initial velocity was taken and was analyzed with the Morrison equation seen in Figure 20. As seen in all the figures, the DMSO sample had the highest slope and the slopes were gradually lower more inhibitor was introduced. The lowest slopes seen in Figure 16 (wild-type), Figure 17 (M1), Figure 18 (M3), and Figure 19 (M4) were inhibitor concentrations 3.8 μM, 4.5 μM, 3.5 μM, and 3.2 μM, respectively. As seen with all the compounds, the increase of inhibitor concentration resulted in a lower slope, suggesting decreased protease activity. While only the results from KY-29 were shown, FRET assays were done for all compounds.

The calculated K_i values were tabulated into Table 1 and listed the K_i values in nM for compounds KY-23, KY-25, and KY-29 with the wild-type and mutant multidrug resistant HIV proteases, including the darunavir resistant mutant, I50V/A71V. The new darunavir analogs synthesized contained a 2-ethylbutyl ligand, unlike darunavir with had and ethyl ligand. When comparing KY-29 and darunavir, the structures contain the same primary amine and differ with the ligand. The simple addition of this ligand increased inhibitor binding affinity for most compounds, except the M1 and wild-type protease. When comparing the substituent groups of KY-23 and KY-25 with the other compounds, both KY-23 and KY-25 contained bulky polar groups. However, KY-25 contained a 4-OCH_3 group which is allowed to

rotate freely unlike KY-23. The 4-OCH_3 substituent was extremely effective against the M1 and wild-type proteases. Although KY-23 was not as effective with the M1 and wild-type protease, the polar cyclic structure of KY-23 was more potent against the M4 protease than any other compound.

The inhibitors synthesized showed increased activity against both the wild-type and mutant proteases when compared to darunavir. When comparing the compounds with darunavir, these compounds showed a significant increase in inhibitor binding affinity for at least one of the mutants and the wild-type protease. The results indicated that the addition of the 2-ethylbutyl ligand and substituent group increased inhibitor binding and suggests that these analogs fit better into the substrate envelope than darunavir.

Discussion

The darunavir analogs, KY-23, KY-25, and KY-29, incorporated the 2-ethylbutyl ligand and various substituent groups in order to improve the fit within P1 pocket. By developing drugs with the substrate envelope hypothesis, mutations that decrease inhibitor binding would also decrease substrate affinity. In order to test the potency of these analogs, a FRET assay was performed. By monitoring the fluorescence of EDANS over time, the protease activity could be monitored. The initial velocity of fluorescence over time graph was measured and the Morrison equation was used to calculate the K_i values.

When comparing darunavir and synthesized compounds with the wild-type protease, both KY-23 and KY-25 had a 10 fold increase for inhibitor binding affinity. As seen in KY-29 and darunavir, the addition of the 2-ethylbutyl did not change the K_i value with the wild-type protease. Although the 2-ethylbutyl ligand did not affect the K_i value, this suggested that the change in the substituent group might increase the binding affinity of these protease inhibitors. Since KY-23 and KY-25 contain bulky polar substituent groups, the results suggested that these groups might provide a better fit in the substrate envelope.

KY-25 had a 10 fold tighter inhibitor binding affinity than darunavir when comparing the compounds and darunavir with the M1 mutant. After looking at the crystal structure, the V82A mutation would most likely affect the binding of the 2-ethylbutyl ligand and the substituent group. The valine to alanine mutation would change the isopropyl group into a methyl group. This change to a smaller amino acid structure might negatively affect binding with these inhibitors. As seen in Table 1, darunavir had the tighter inhibitor binding affinity than KY-29. Since the ethyl group in darunavir was less bulky than the 2-ethylbutyl group, the 2-ethybutyl ligand was too big and protruded outside the

substrate envelope. However, KY-25 contains both the 2-ethylbutyl ligand and a 4-OCH_3 group. Since 4-OCH_3 is a polar bulky group and can rotate freely, there might be stronger hydrogen bonding interactions with the alanine change. The interaction of 4-OCH_3 and alanine might cause a conformational change, allowing the 2-ethylbutyl ligand to bind tightly into the active site. Due to the few hydrogens and the cyclic nature of KY-29, the inability to rotate might affect the hydrogen bonding with the alanine.

When comparing the compounds and darunavir with the M3 mutant, KY-29 had the tightest inhibitor binding affinity and had a slight lower K_i value than darunavir by 0.002 nM. After looking at the crystal structure, the I84V mutation would most likely affect binding of the 2-ethylbutyl ligand and the substituent group. This mutation changed an isoleucine to valine, substituting a bulky sec-butyl group with a smaller isopropyl group. The addition of the 2-ethylbutyl ligand increased the inhibitor binding affinity when comparing KY-29 and darunavir. Due to the larger size and bulk of the 2-ethylbutyl group, this ligand might be able interact with the smaller valine. Unlike the wild-type protease, both KY-23 and KY-25 had higher K_i values, suggesting a lower inhibitor binding affinity. The polar bulky substituent groups of KY-23 and KY-25 might protrude out of the substrate envelope, affecting the binding of the inhibitor. When comparing these polar bulky groups with the 4-NH_2 groups in KY-29 and darunavir, the smaller nonpolar primary amines might better interact with the valine residues.

The M4 mutant protease is the classic drug resistant variant of darunavir. KY-23 had a tighter inhibitor binding affinity than darunavir by a 100 fold when comparing the compounds and darunavir with the M4 mutant. After looking at the crystal structure, the I50V mutation would most likely affect binding of the 2-ethylbutyl ligand and the substituent group. This mutation substituted an isoleucine for a valine, changing the bulky sec-butyl group to smaller isopropyl group. The addition of the 2-ethylbutyl ligand significantly increased the inhibitor binding affinity when comparing KY-29 and darunavir. Since

the 2-ethylbutyl ligand is much larger than the ethyl ligand, the larger group can compensate for the change in valine. However, the smaller 4-NH_2 group might not be large enough to interact with the valine. However KY-23 might interact better than KY-25 because KY-23 contains a cyclic bulky polar group. The cyclic shape of the substituent group might prevent the cyclic group from rotating while the 4-OCH_3 group is free to rotate. The rotation seen in the 4-OCH_3 group might cause some steric hindrance with the valine.

The method of determining these compounds employed a new general strategy for drug discovery. Rather than randomly testing various darunavir analogs, the substrate envelope allowed for structured-based design. When comparing the compounds with darunavir, these compounds showed a significant increase in inhibitor binding affinity for at least one of the mutants and the wild-type protease. Due to the increased potency of these drugs, these compounds are currently being tested on animals for pharmacokinetics and pharmacodynamics studies. Although these compounds were showed an increased binding affinity compared with darunavir, there was still some drug resistance with the mutant proteases. Future studies can focus on further analyzing crystal structures to locate positions which protrude from the substrate envelope and new analogs can be synthesized to improve these problems. Although none of the compounds outperformed darunavir for all the protease mutants, a combination of these compounds might be useful for HAART therapy.

A major problem with HAART is that it requires many pills for treatment of HIV. By developing more potent drugs, less medication is required to achieve the same effect. Introducing less amount of drug into the body might also solve some of the toxicity problems encountered with HAART as well. Along with the toxicity and potency, the high yielding synthesis for these analogs might prove to be more economical to the pharmaceutical companies.

These compounds were significantly more potent against the wild-type and MDR HIV-1 protease variants than any protease inhibitor. By designing drugs based off of the substrate envelope hypothesis, mutations that alter inhibitor binding would also affect substrate binding as well. Further development of these protease inhibitors may lead to more effective treatments against drug-resistant HIV-1. While this study specifically looked at the designing drugs against the HIV-1 protease, the substrate envelope can be used to design drugs against other drug targets and other diseases.

Figures and Tables

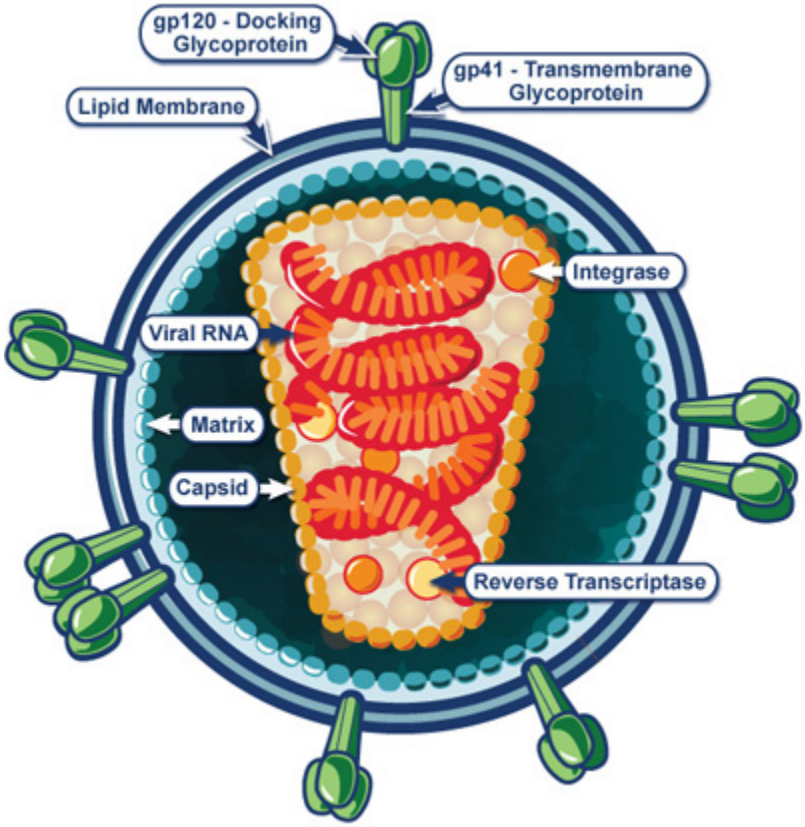


Figure 1: The HIV Viron

Above is the HIV-1 viral envelope structure of both the viral envelope and inner core(*1*).

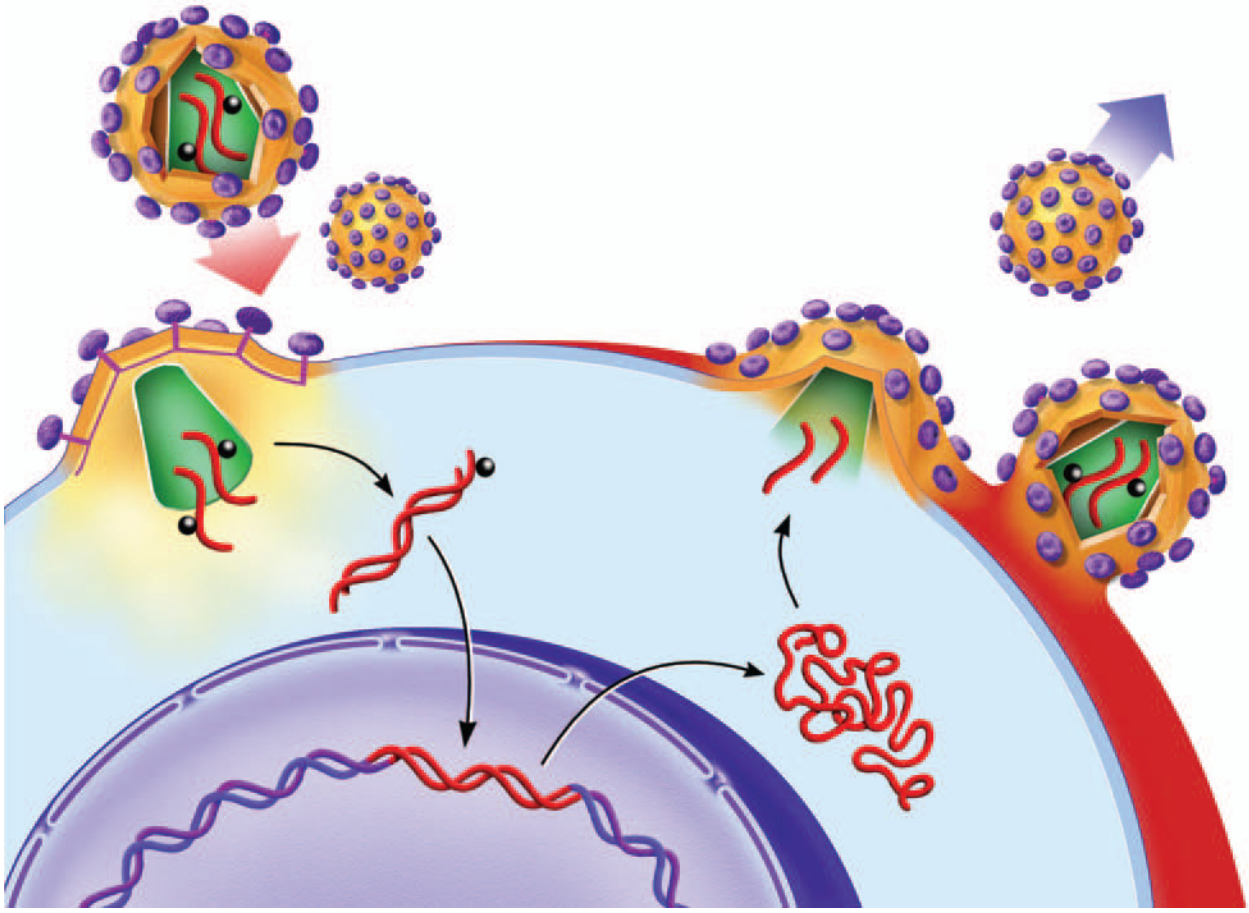


Figure 2: The HIV-1 lifecycle

This figure shows a simple scheme of the HIV-1 life cycle and shows the following interactions: the envelope proteins of the virus and the CD4 receptor and coreceptors of the host cell which bind the viral envelope to the host membrane. The viral DNA is then enters the nucleus and integrates with the host chromosomal DNA and is catalyzed by the viral enzyme integrase. Expression of the genes leads to production of viral proteins and RNA. The HIV protease then cleaves the gag and gag-pol proteins into mature components. Viral proteins and RNA are assembled at the cell surface into new viral particles and leave the cell through budding. While budding, the viral acquires an outer layer and an envelope(*7*).

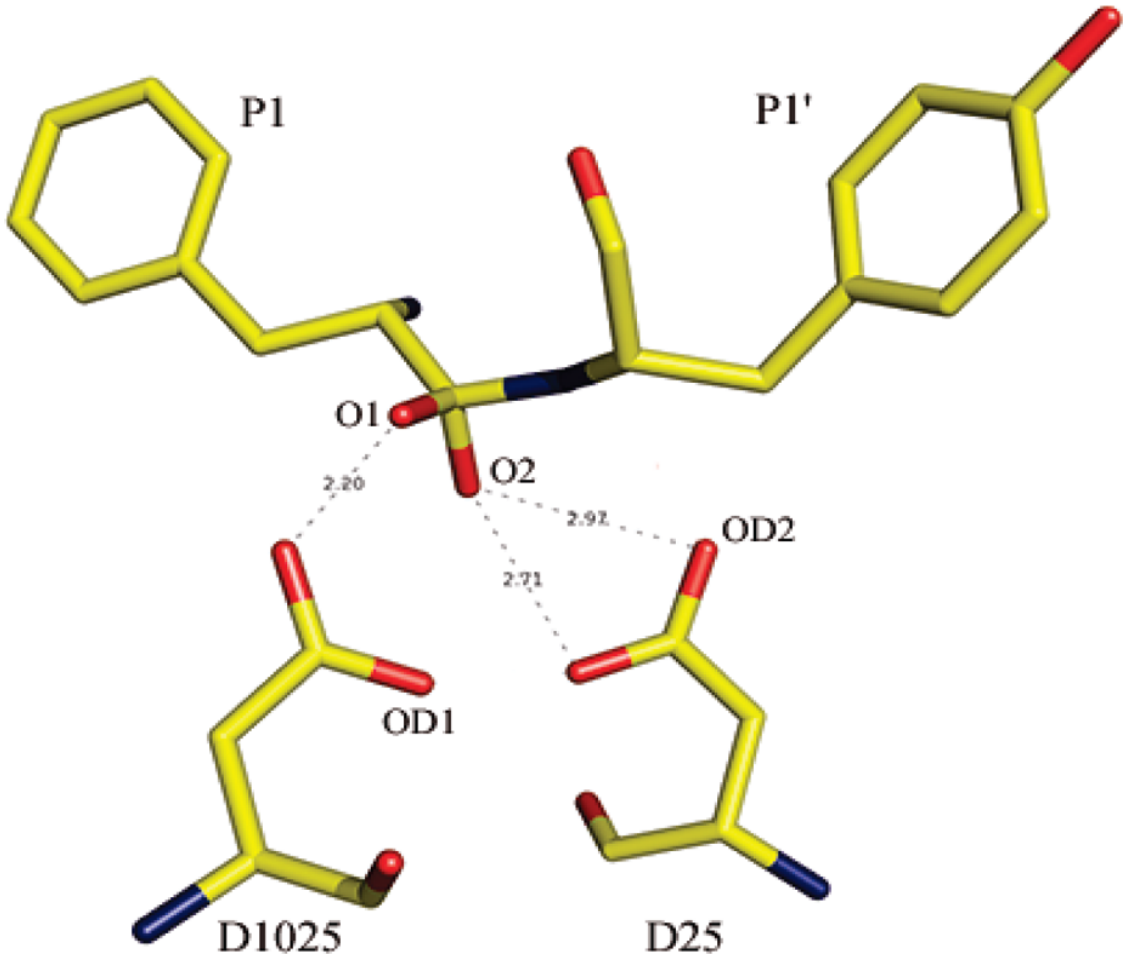


Figure 3: Hydrogen bonding between the aspartates and *gem*-diol of the tetrahedral intermediate at the catalytic center (*13*).

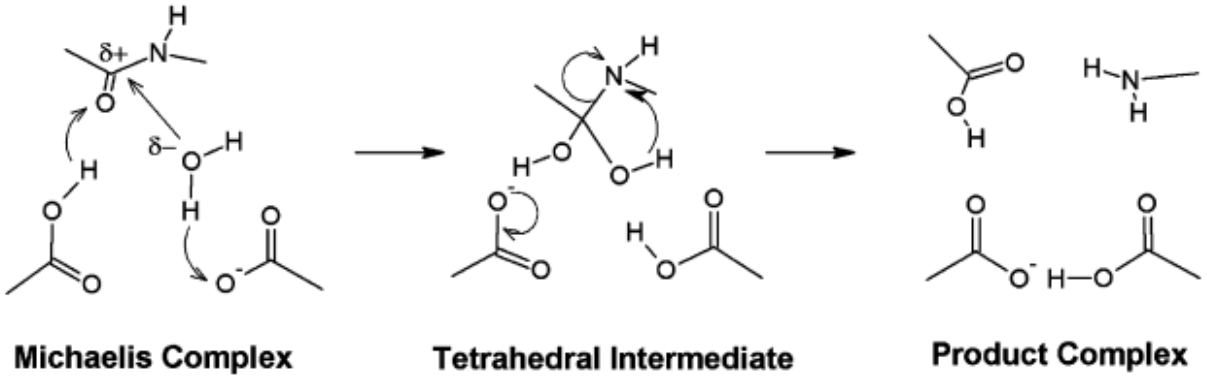


Figure 4: Schematic diagram of the reaction mechanism based on the non-covalent tetrahedral intermediate

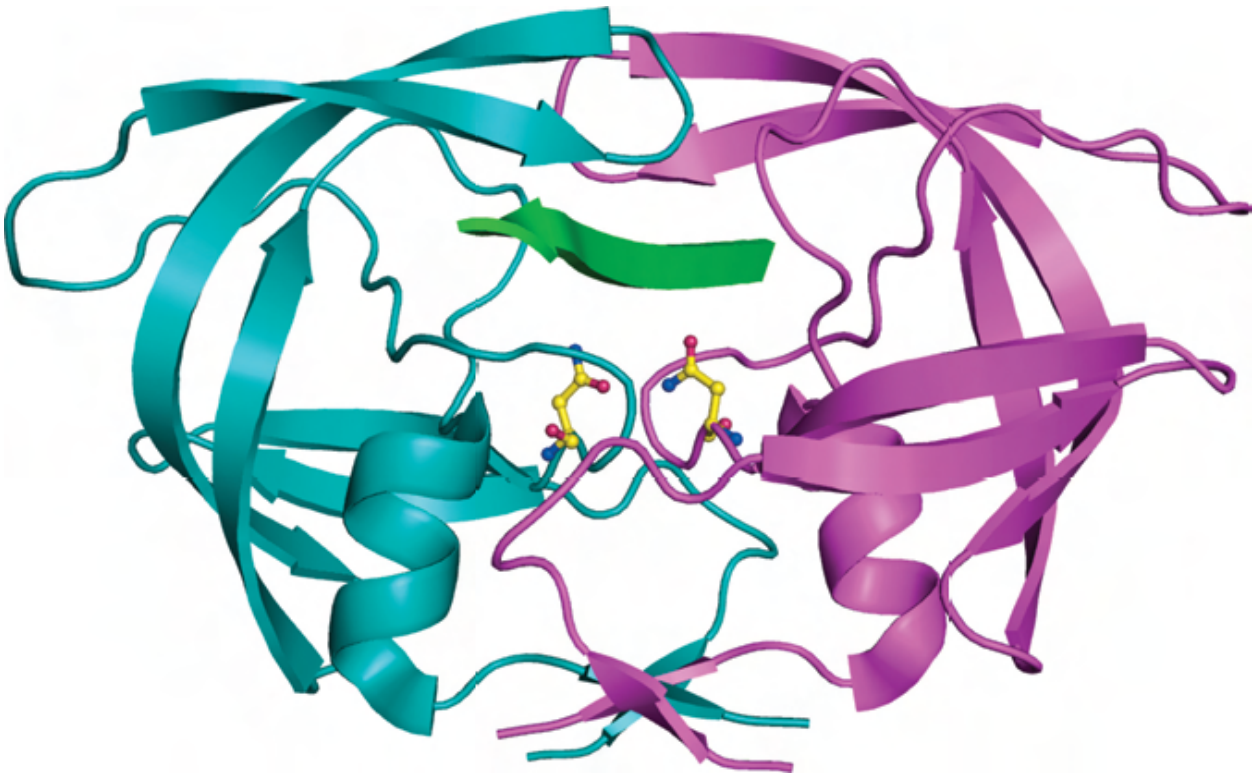


Figure 5: Ribbon diagram of the crystal structure of the HIV-1 Protease

A ribbon diagram of crystal structure of a substrate complex of the homo-dimer HIV-1 protease. The substrate is shown in green; the catalytic aspartic acids are shown in yellow; and each monomer is shown in cyan and pink(*8*).

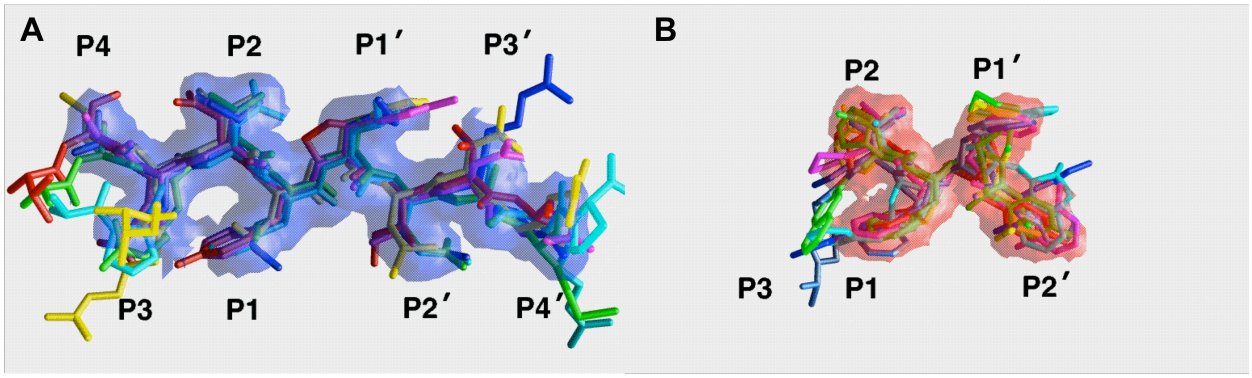


Figure 6: The substrate envelope (A) and the inhibitor envelope (B) (*16*)

The substrate envelope is the conserved area of the natural substrates of HIV protease, while the inhibitor envelope is the conserved area of all the current HIV protease inhibitors.

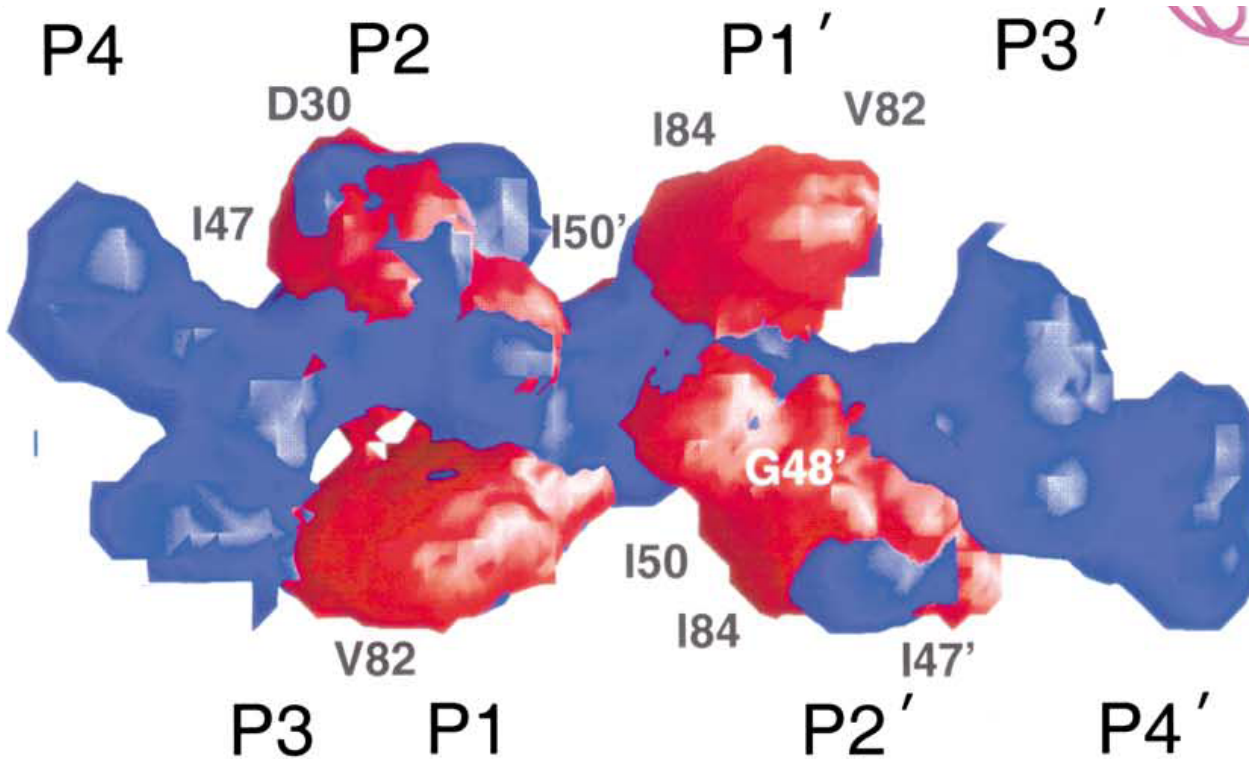


Figure 7: Superposition of the substrate envelope and the inhibitor envelope(*16*)

The superposition of the substrate envelope (blue) and inhibitor envelope (red). Areas that the inhibitor envelope protrudes the substrate envelope confer drug resistance when they are labeled.

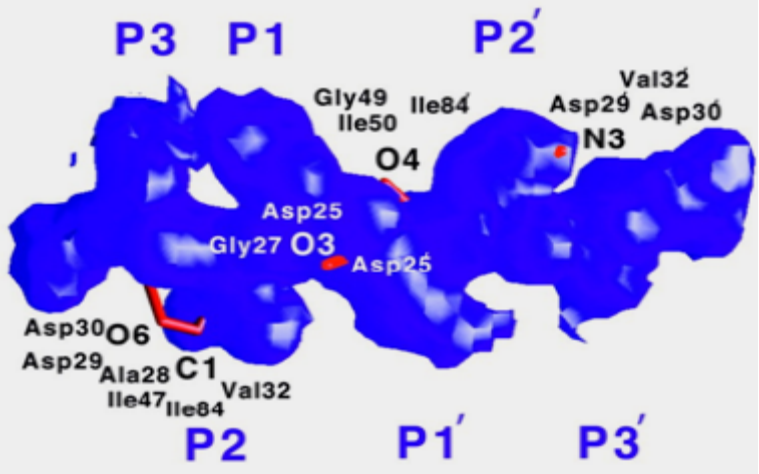


Figure 8: Superposition of the substrate envelope (blue) with darunavir (red); the places where darunavir protrudes out from the substrate envelope encounter viral resistance(*22*).

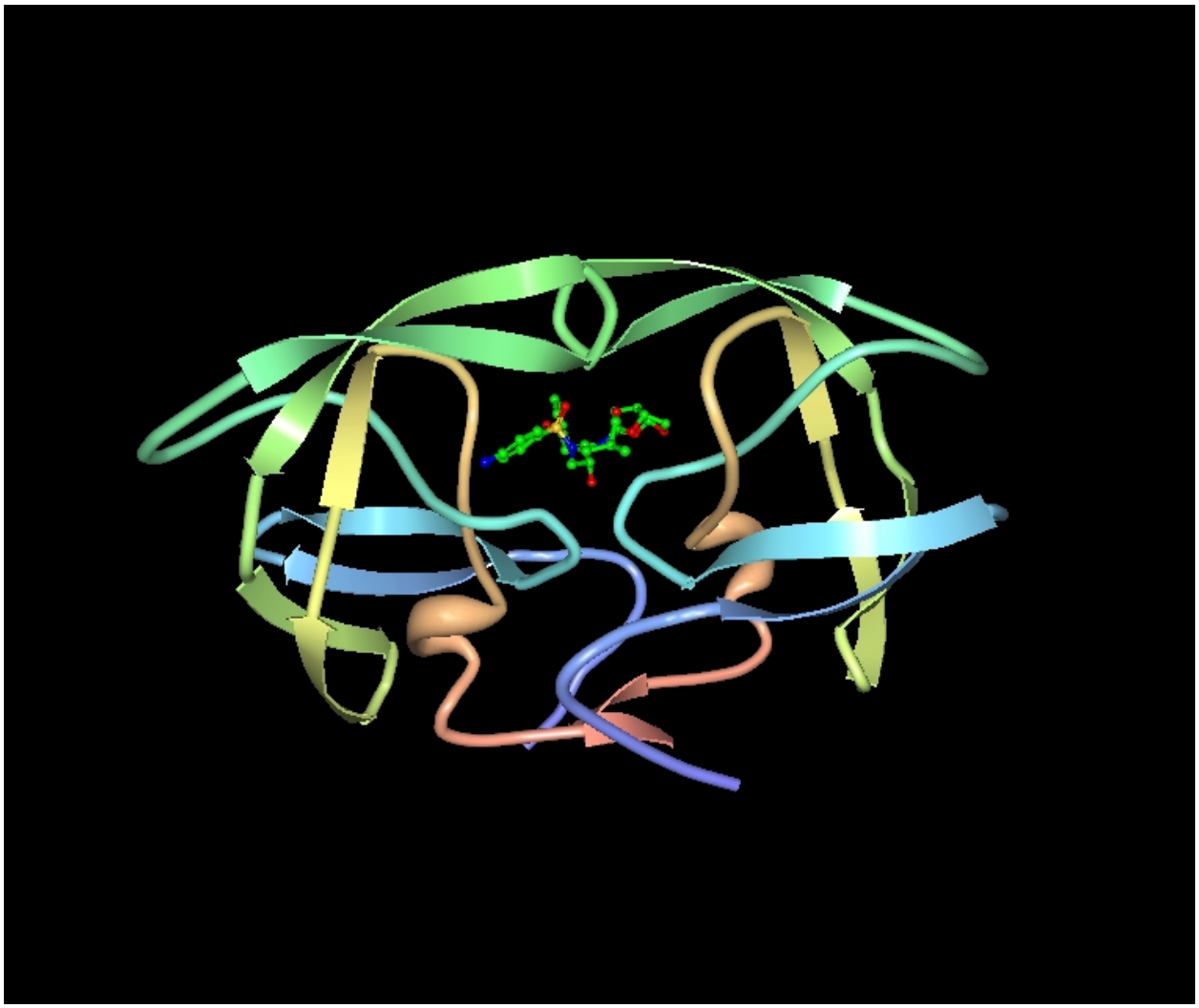


Figure 9: The wild-type HIV protease (3EM6) with darunavir binding in the active site

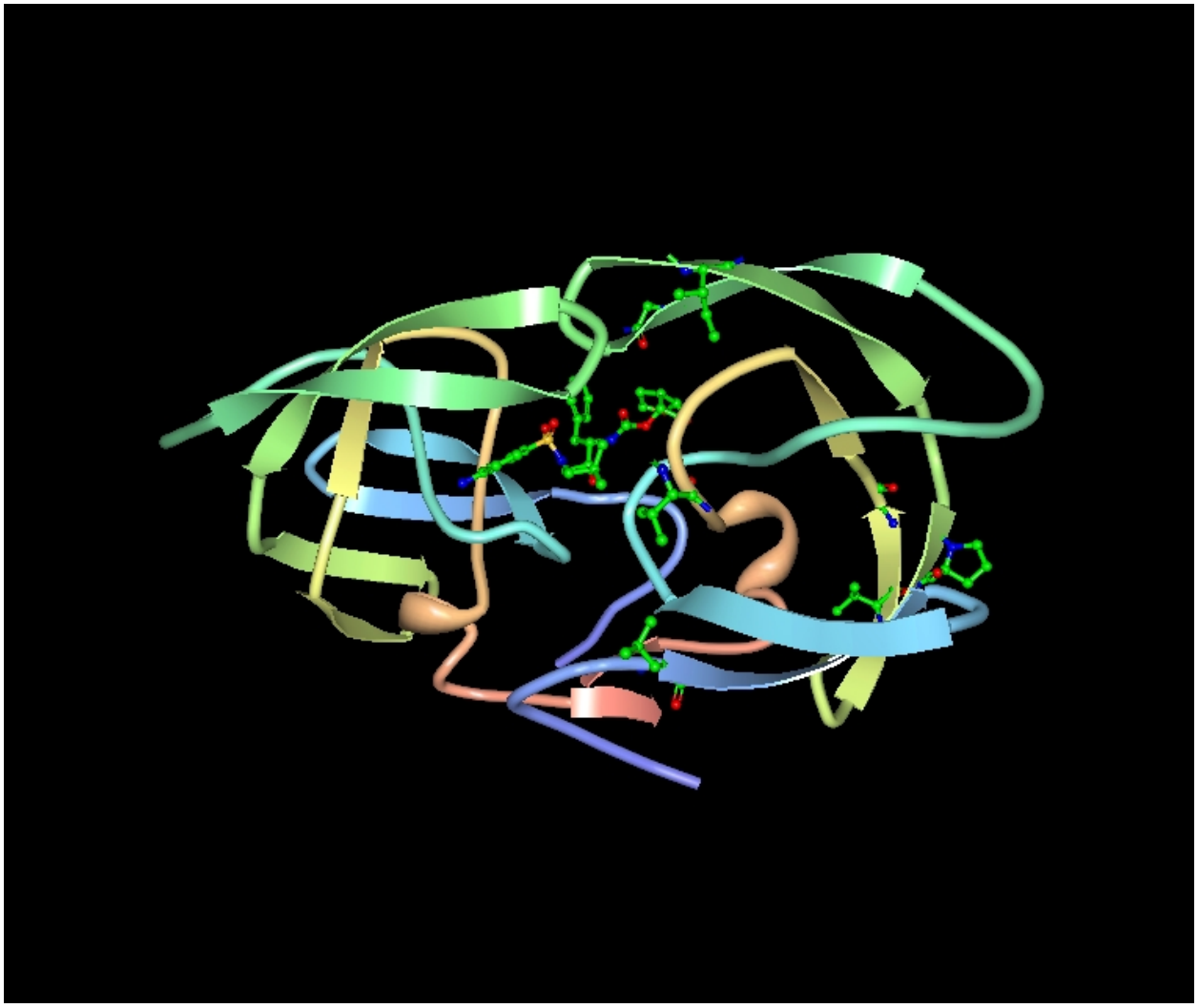


Figure 10: The common M1 mutations in HIV protease (3EM6) with darunavir binding in the active site. As seen the V82A mutation can interefere with darunavir binding to the site.

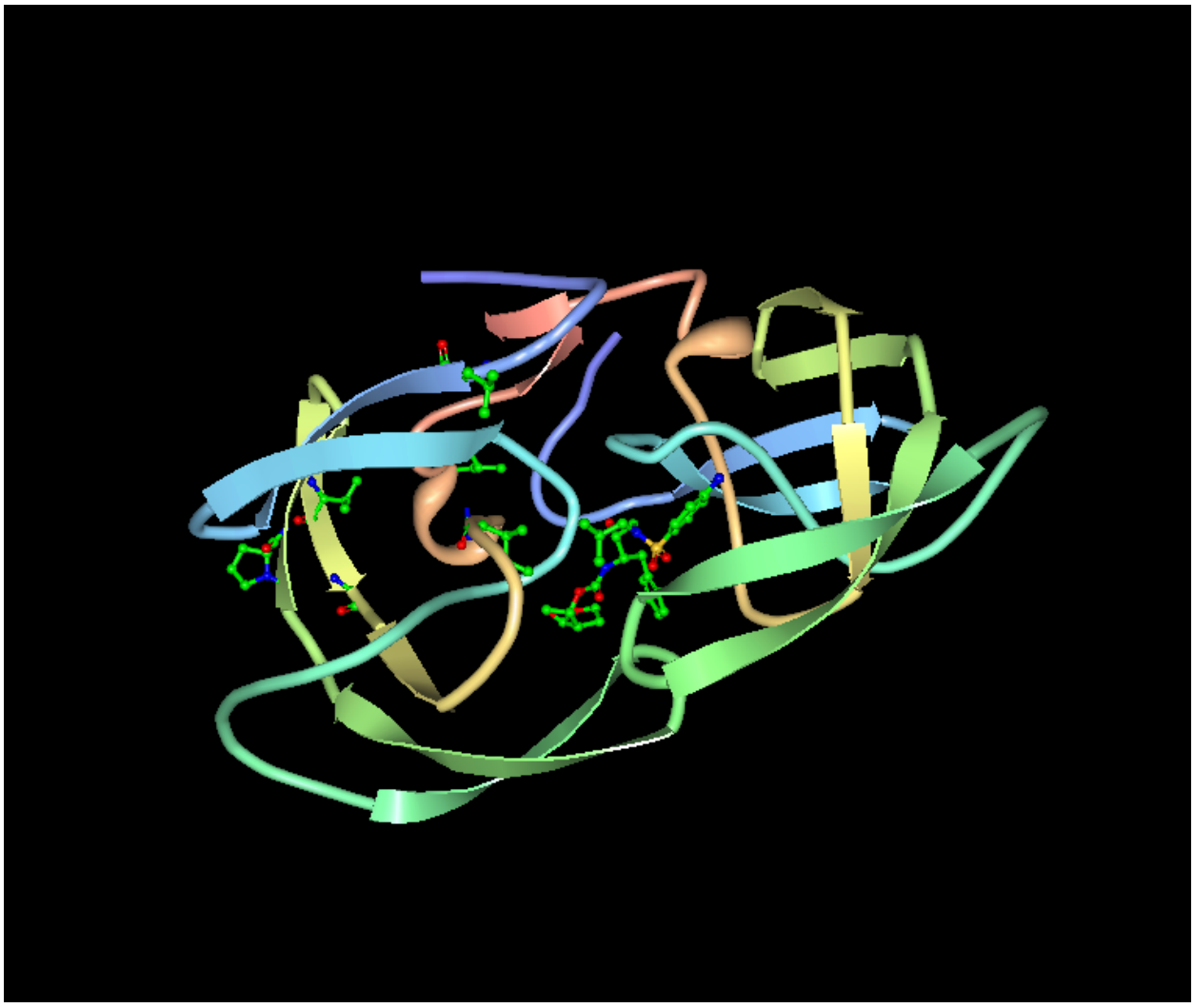


Figure 11: The common M3 mutations in HIV protease (3EM6) with darunavir binding in the active site. As seen the I84V mutation can interfere with darunavir binding to the site.

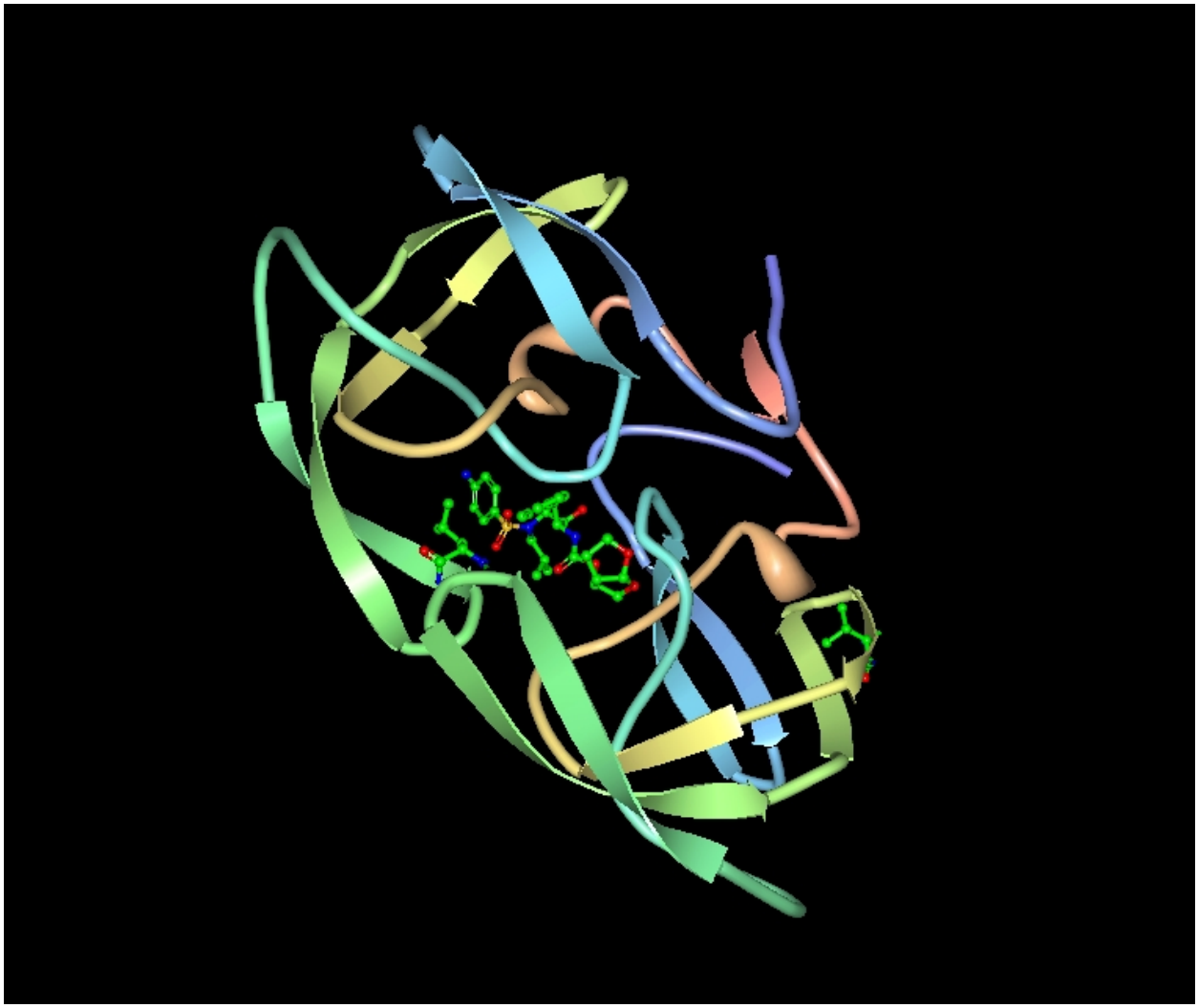


Figure 12: The common M4 mutations in HIV protease (3EM6) with darunavir binding in the active site. As seen the I50V mutation can interfere with darunavir binding to the site.

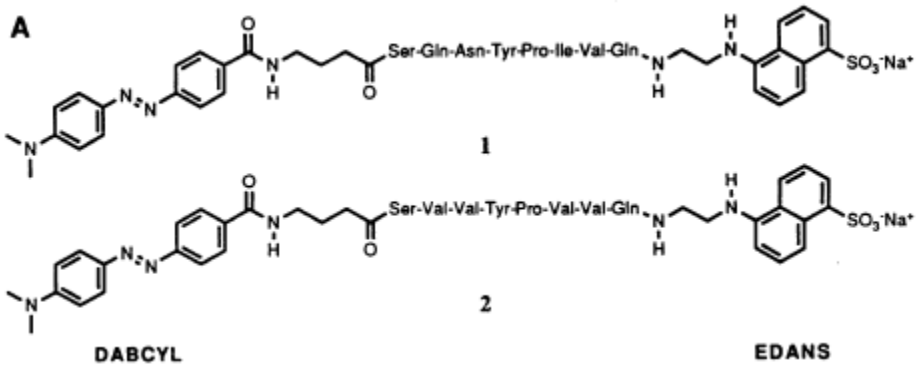


Figure 13: Structures of fluorogenic substrates DABCYL and EDANS (*14*)

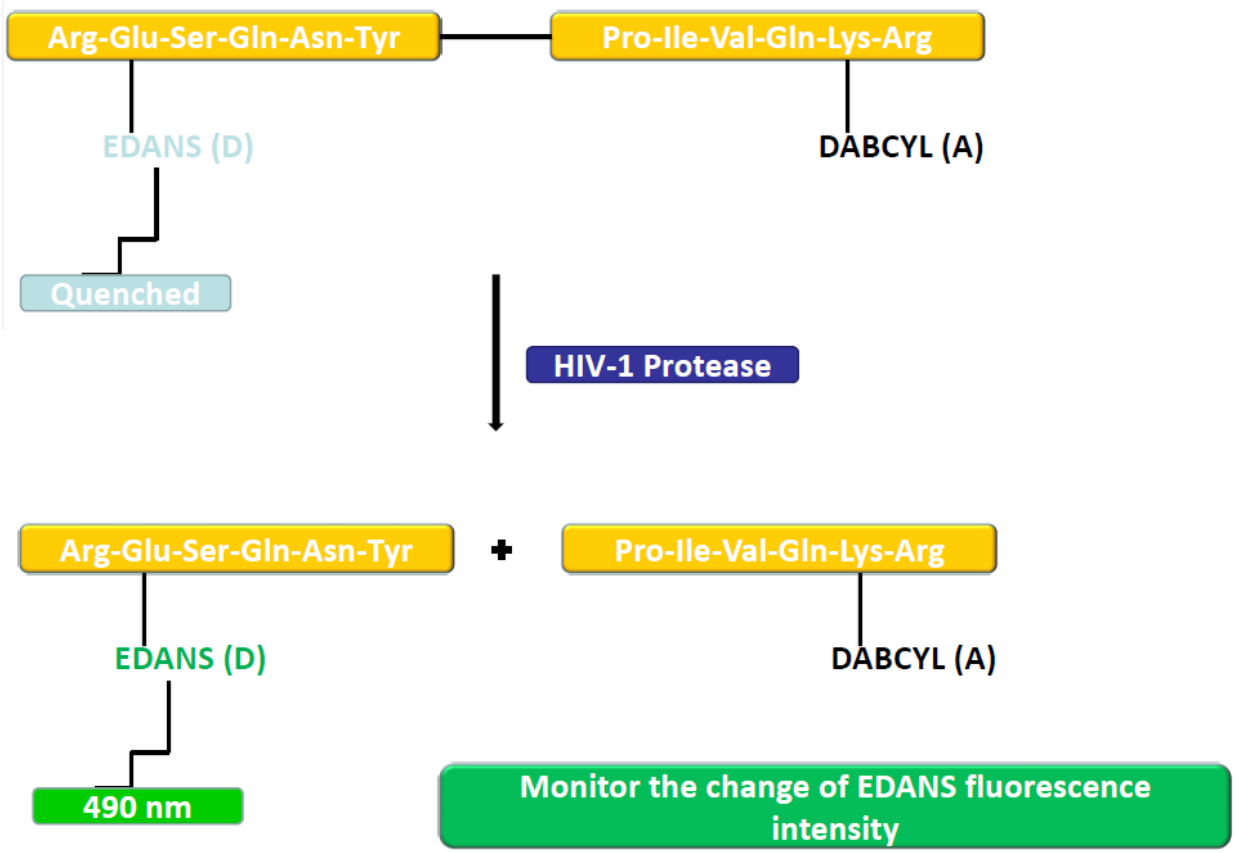


Figure 14: Overview of the FRET assay

A donating group (EDANS) and accepting group (DABCYL) is attached to a natural substrate of HIV protease. When the substrate is uncleaved, DABCYL quenches EDANS and there is no detection of fluorescence. After HIV-1 protease cleaves the substrate, EDANS is not longer being quenched by DABCYL and EDANS fluorescence can be detected. The effectiveness of the protease inhibitor can be monitored by the change of EDANS fluorescence intensity.

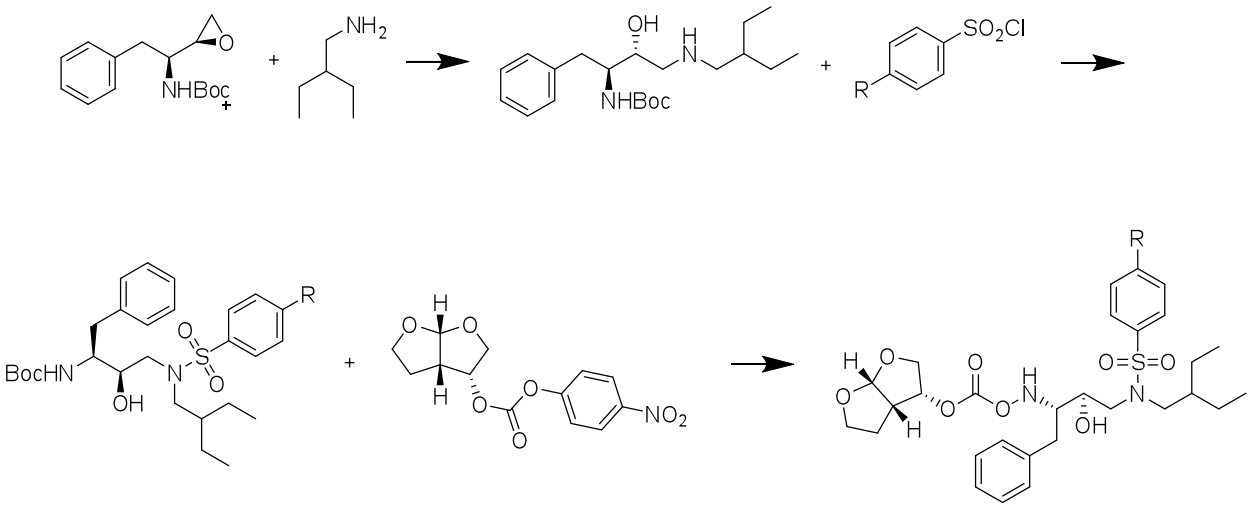


Figure 15: General reaction for the HIV protease inhibitors

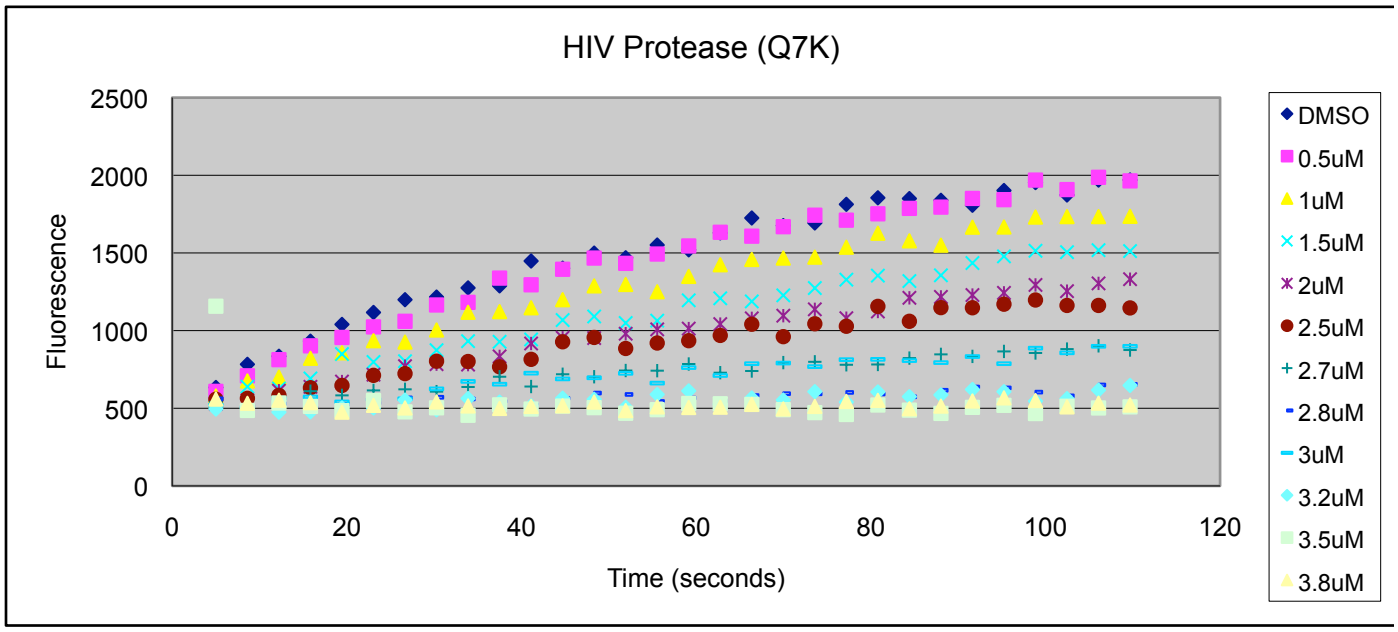


Figure 16: KY-29 Fluorescence Resonance Energy Transfer for HIV Protease (Q7K)

The x axis represented time in seconds and the y axis represented fluorescence intensity. The legend showed increasing concentrations of protease inhibitor and a control of DMSO, which had no protease inhibitor present. The initial velocity was taken and was analyzed with the Morrison equation. As the inhibitor concentration increased the lower fluorescent the slope, suggesting decreased HIV protease activity.

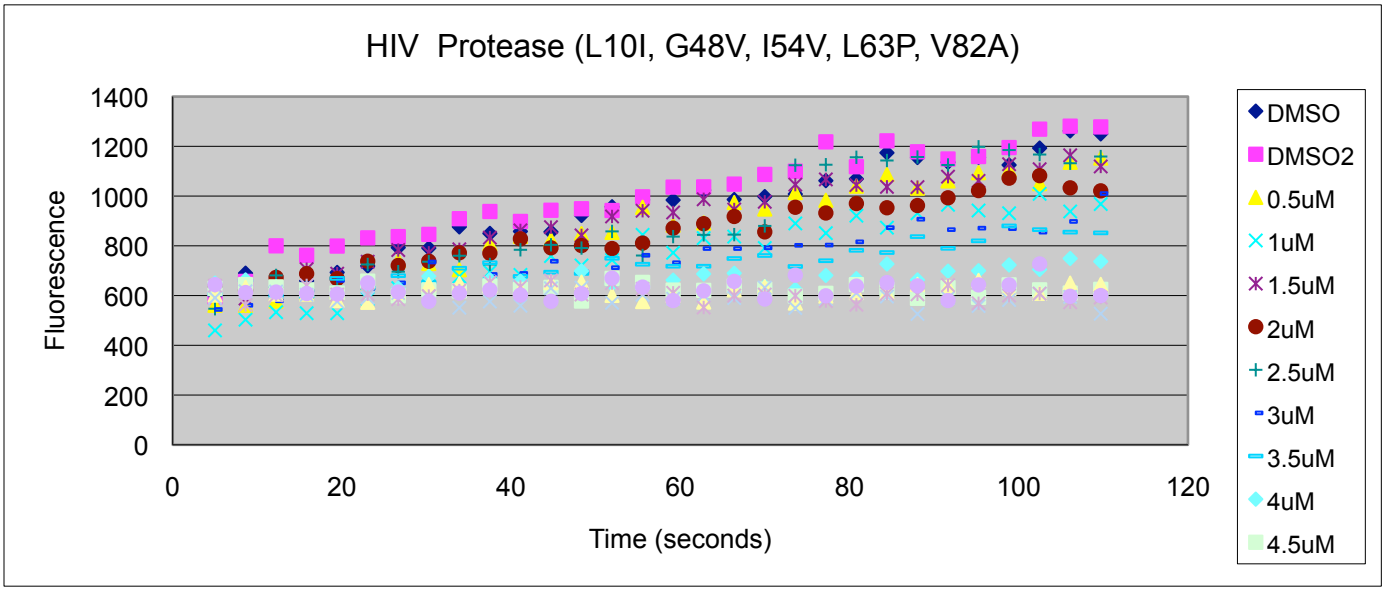


Figure 17: KY-29 Fluorescence Resonance Energy Transfer for HIV Protease (L10I, G48V, 154V, L63P, V82A)

Figure 17 showed the results of the FRET assay with the mutant HIV protease (L10I, G48V, I54V, L63P, V82A) against inhibitor KY-29. The x axis represented time in seconds and the y axis represented fluorescence intensity. The legend showed increasing concentrations of protease inhibitor and two controls of DMSO, which had no protease inhibitor present. The initial velocity was taken and was analyzed with the Morrison equation. As the inhibitor concentration increased the lower fluorescent the slope, suggesting increased inhibition of the HIV protease.

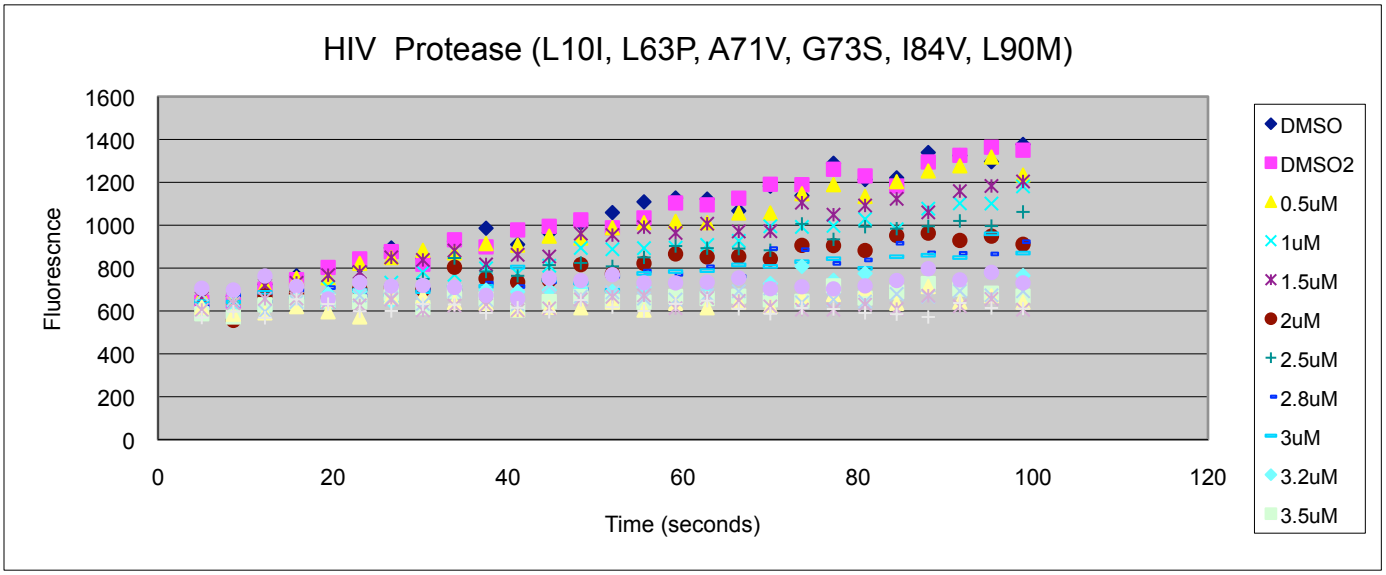


Figure 18: KY-29 Fluorescence Resonance Energy Transfer for HIV Protease (L10I, L63P, A71V, G73S, I84V, L90M)

Figure 18 showed the results of the FRET assay with the mutant HIV protease (L10I, L63P, A71V, G73S, I84V, L90M) against inhibitor KY-29. The x axis represented time in seconds and the y axis represented fluorescence intensity. The legend showed increasing concentrations of protease inhibitor and two controls of DMSO, which had no protease inhibitor present. The initial velocity was taken and was analyzed with the Morrison equation. As the inhibitor concentration increased the lower fluorescent the slope, suggesting increased inhibition of the HIV protease.

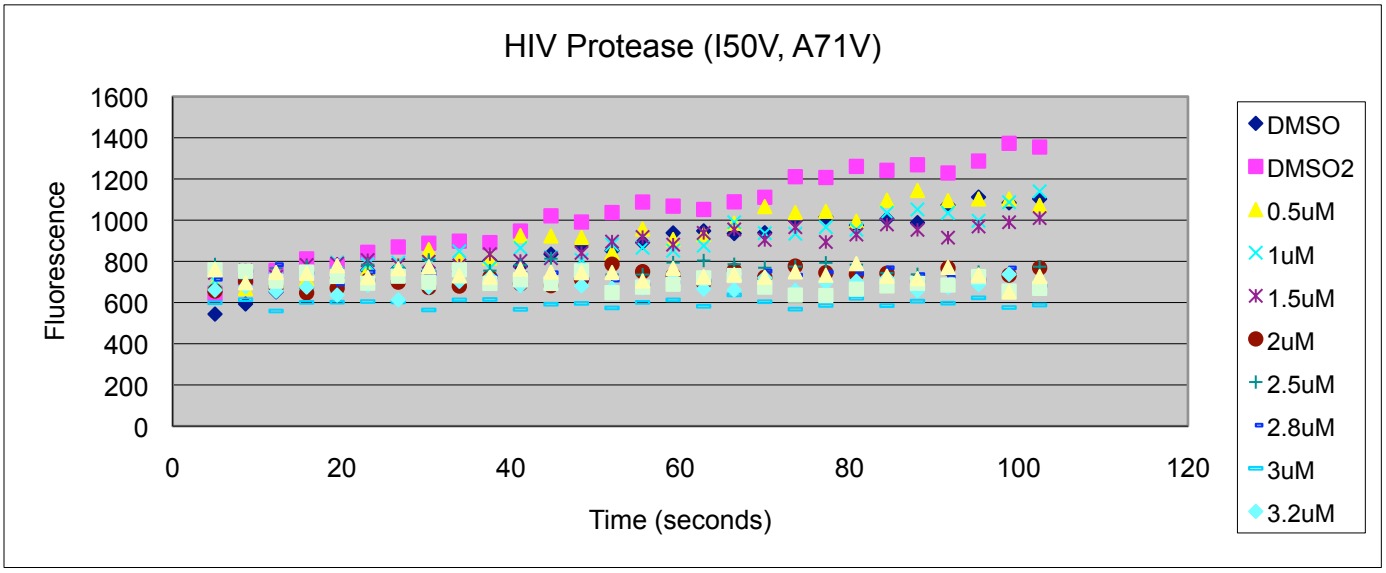


Figure 19: KY-29 Fluorescence Resonance Energy Transfer for HIV Protease (I50V, A71V)

Figure 19 showed the results of the FRET assay with the mutant HIV protease (L10I, L63P, A71V, G73S, I84V, L90M) against inhibitor KY-29. The x axis represented time in seconds and the y axis represented fluorescence intensity. The legend showed increasing concentrations of protease inhibitor and two controls of DMSO, which had no protease inhibitor present. The initial velocity was taken and was analyzed with the Morrison equation. As the inhibitor concentration increased the lower fluorescent the slope, suggesting increased inhibition of the HIV protease.

$$v = C\frac{\sqrt{(K_i+[I]-f[E])^2+4K_i f[E]}-(K_i+[I]-f[E])}{2} \quad C = \frac{k_{cat}[S]}{K_m+[S]}$$

Figure 20: The Morrison Equation was used to determine the K_i values.
K_m= 103 µM K_{cat} = 4.9 s^{-1}

The initial velocities were measured from these graphs and the Morrison equation was used to calculate the inhibitor binding affinity or K_i values seen in Figure 20. K_i is the binding affinity; [I] is the protease inhibitor concentration; [E] is the protease enzyme concentration; [S] is the substrate concentration; *f* is one of the fitting parameters; K_{cat} is the turnover number or the catalytic rate of the enzyme; K_m is the Michaelis constant and characterizes an enzyme's approximate affinity for substrate.

		Ki (nM)			
Compound	**R'**	Wt(Q7k)	L10I,G48V,I54V,L63P,V82A (M1)	L10I,L63P,A71V,G73S,184V,L90M (M3)	I50V,A71V (M4)
KY-23	3,4-OCH_2O-	0.0008	0.034	0.104	0.006
KY-25	4-OCH_3	0.0002	0.001	0.134	0.010
KY-29	4-NH_2	0.0050	0.034	0.022	0.134
DRV	4-NH_2	0.0050	0.025	0.024	0.240

Table 1: K_i values against various HIV protease mutations

Acknowledgements

I would like to thank Dr. Celia Schiffer for help eagerness to help and allowing me to work in her laboratory. I would also like to thank Dr. Akbar Ali and Dr. Hong Cao for their time and helping with the organic synthesis and fluorescence assays. Finally, I would like to thank Professor Destin Heilman for guiding me along the way.

1. (2009) HIV/AIDS, In *NIAID's HIV/AIDS Research Program*, National Institute of Allergy and Infectious Diseases.
2. Jorissen, R. N., Reddy, G. S. K. K., Ali, A., Altman, M. D., Chellappan, S., Anjum, S. G., Tidor, B., Schiffer, C. A., Rana, T. M., and Gilson, M. K. (2009) Additivity in the Analysis and Design of HIV Protease Inhibitors, *Journal of Medicinal Chemistry 52*, 737-754.
3. Clavel, F., and Hance, A. J. (2004) HIV Drug Resistance, *N Engl J Med 350*, 1023-1035.
4. Sripriya Chellappan, Visvaldas Kairys, Miguel X. Fernandes, Celia Schiffer, and Michael K. Gilson. (2007) Evaluation of the substrate envelope hypothesis for inhibitors of HIV-1 protease, *Proteins: Structure, Function, and Bioinformatics 68*, 561-567.
5. Desport, M. (2010) *Lentiviruses and Macrophages: Molecular and Cellular Interactions*, Caister Academic Press.
6. Frankel, A. D., and Young, J. A. T. (1998) HIV-1: Fifteen Proteins and an RNA, *67*, 1-25.
7. Warnke, D., Barreto, J., and Temesgen, Z. (2007) Antiretroviral Drugs, *47*, 1570-1579.
8. Anderson, J., Schiffer, C., Lee, S.-K., and Swanstrom, R. (2009) Viral Protease Inhibitors, In *Antiviral Strategies*, pp 85-110.
9. Liao, C., Karki, R. G., Marchand, C., Pommier, Y., and Nicklaus, M. C. (2007) Virtual screening application of a model of full-length HIV-1 integrase complexed with viral DNA, *Bioorganic & Medicinal Chemistry Letters 17*, 5361-5365.
10. Raboud, J. M., Rae, S., Vella, S., Harrigan, P. R., Bucciardini, R., Fragola, V., Ricciardulli, D., Montaner, J. S. G., and Team, I. S. (1999) Meta-analysis of Two Randomized Controlled Trials Comparing Combined Zidovudine and Didanosine Therapy With Combined Zidovudine, Didanosine, and Nevirapine Therapy in Patients With HIV, *22*, 260.
11. Havlir, D., Eastman, S., Gamst, A., and Richman, D. (1996) Nevirapine-resistant human immunodeficiency virus: kinetics of replication and estimated prevalence in untreated patients, *70*, 7894-7899.
12. Choudhary, S. K., Rezk, N. L., Ince, W. L., Cheema, M., Zhang, L., Su, L., Swanstrom, R., Kashuba, A. D. M., and Margolis, D. M. (2009) Suppression of Human Immunodeficiency Virus Type 1 (HIV-1) Viremia with Reverse Transcriptase and Integrase Inhibitors, CD4+ T-Cell Recovery, and Viral Rebound upon Interruption of Therapy in a New Model for HIV Treatment in the Humanized Rag2-/-{gamma}c-/- Mouse, *83*, 8254-8258.
13. Das, A., Mahale, S., Prashar, V., Bihani, S., Ferrer, J. L., and Hosur, M. V. (2010) X-ray Snapshot of HIV-1 Protease in Action: Observation of Tetrahedral Intermediate and Short Ionic Hydrogen Bond SIHB with Catalytic Aspartate, *Journal of the American Chemical Society*.
14. Matayoshi, E., Wang, G., Krafft, G., and Erickson, J. (1990) Novel fluorogenic substrates for assaying retroviral proteases by resonance energy transfer, *Science 247*, 954-958.
15. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2002) *Molecular Biology of the Cell, Fourth Edition*, Garland.
16. King, N. M., Prabu-Jeyabalan, M., Nalivaika, E. A., and Schiffer, C. A. (2004) Combating Susceptibility to Drug Resistance: Lessons from HIV-1 Protease, *Chemistry & Biology 11*, 1333-1338.
17. Nalam, M. N., and Schiffer, C. A. (2008) New approaches to HIV protease inhibitor drug design II: testing the substrate envelope hypothesis to avoid drug resistance and discover robust inhibitors, *Current Opinion in HIV and AIDS 3*, 642-646 610.1097/COH.1090b1013e3283136cee.
18. Altman, M. D., Ali, A., Kumar Reddy, G. S. K., Nalam, M. N. L., Anjum, S. G., Cao, H., Chellappan, S., Kairys, V., Fernandes, M. X., Gilson, M. K., Schiffer, C. A., Rana, T. M., and Tidor, B. (2008) HIV-

1 Protease Inhibitors from Inverse Design in the Substrate Envelope Exhibit Subnanomolar Binding to Drug-Resistant Variants, *Journal of the American Chemical Society 130*, 6099-6113.

19. Surleraux, D. L. N. G., Tahri, A., Verschueren, W. G., Pille, G. M. E., de Kock, H. A., Jonckers, T. H. M., Peeters, A., De Meyer, S., Azijn, H., Pauwels, R., de Bethune, M.-P., King, N. M., Prabu-Jeyabalan, M., Schiffer, C. A., and Wigerinck, P. B. T. P. (2004) Discovery and Selection of TMC114, a Next Generation HIV-1 Protease Inhibitors, *Journal of Medicinal Chemistry 48*, 1813-1822.
20. Sripriya Chellappan, G. S. Kiran Kumar Reddy, Akbar Ali, Madhavi N. L. Nalam, Saima Ghafoor Anjum, Hong Cao, Visvaldas Kairys, Miguel X. Fernandes, Michael D. Altman, Bruce Tidor, Tariq M. Rana, Celia A. Schiffer, and Michael K. Gilson. (2007) Design of Mutation-resistant HIV Protease Inhibitors with the Substrate Envelope Hypothesis, *Chemical Biology & Drug Design 69*, 298-313.
21. Lakowicz, J. (1999) *Principles of Fluorescence Spectroscopy*, 2nd ed., Kluwer Academic/ Plenum Publishers, New York.
22. King, N. M., Prabu-Jeyabalan, M., Nalivaika, E. A., Wigerinck, P., de Bethune, M.-P., and Schiffer, C. A. (2004) Structural and Thermodynamic Basis for the Binding of TMC114, a Next-Generation Human Immunodeficiency Virus Type 1 Protease Inhibitor, *J. Virol. 78*, 12012-12021.