MQP-BC-DSA-9421

NEW STRUCTURE-BASED HIV-1 PROTEASE INHIBITORS BIND DRUG-RESISTANT VARIANTS IN THE SUBNANOMOLAR RANGE

A Major Qualifying Project Report

Submitted to the Faculty of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Bachelor of Science

in

Biochemistry and Biology & Biotechnology

by

_____
Reinhold Gildemeister

April 28, 2011

APPROVED:

_____
Akbar Ali, Ph.D.
Biochemistry and Molecular Pharmacology
UMASS Medical Center
Major Advisor

_____
David Adams, Ph.D.
Biology and Biotechnology
WPI Project Advisor

ABSTRACT

Every year millions of people contract HIV or die from HIV-AIDS related illnesses. Since current drugs generally target viral enzymes, selective drug pressure coupled with the high mutation rate and infidelity of HIV-1 reverse transcriptase lead to drug resistance. It is thus crucial to develop new, tighter-binding inhibitors with more flexible structures that can better inhibit mutant proteases. Two novel inhibitor cores were designed and synthesized, all while attempting to stay within the confines of the substrate-envelope. Many new drugs were tested, which bound to various drug-resistant mutants in the pM range.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

First I would like to thank Dr. Celia Schiffer for allowing me to work in her lab. I would also like to thank Dr. Akbar Ali for guiding me through the project and teaching me organic chemistry along the way. I would also like to thank Dr. Hong Cao for teaching me about the FRET assay and how to analyze the data. I would also like to thank Professor Dave Adams for helping me with the project and in correcting my paper.

BACKGROUND

Human Immunodeficiency Virus Pandemic

The Human Immunodeficiency Virus (HIV) is a single-stranded positive-sense RNA retrovirus that is further classified as a lentivirus. Lentiviruses infect and replicate in non-dividing cells, in which HIV mainly infects $CD4^+$ T cells, but have also been known to infect macrophages and dendritic cells (Weiss, 1993).

HIV has been considered a pandemic for the past 30 years. The time between initial infection and the time it takes for serious symptoms to occur is very long for lentiviruses, which is why many people are unaware that they are infected, and unfortunately spread the disease unknowingly (NIAID's). Infection eventually leads patients to develop Acquired Immunodeficiency Syndrome (AIDS), in which an individual's immune system is so low that any opportunistic infection can potentially lead to death. Since the beginning of the epidemic, the number of AIDS-related deaths has been estimated to be over 25 million, with 1.8 million in 2009 alone. Around 60 million people have been infected with HIV, and there are an estimated 33.3 million people who are currently infected with HIV (UNAIDS).

The HIV-1 Genome and Structure

The HIV virion is a spherical virus with a diameter of 100 nm and its outer coat, or viral envelope is made up of a lipid bilayer that was budded from the host cell, and around 72 HIV Env proteins. The Env proteins consist of surface tri-mers of gp120/SU and tri-mers of transmembrane gp41/TM proteins. Lining the inside of the viral envelope

is the matrix (MA or p17) proteins. Within this matrix there are around 1,056 capsid (CA or p24) proteins that surround two identical full-length copies of HIV RNA (NIAID's; Pornillos et al., 2011). Each RNA strand is around 9-kb long, flanked on both sides by long terminal repeats (LTR's) which control the production of new viruses by either proteins from the virus or the host cell. In order to protect the genomic RNA inside the virion core, there are nucleocapsid (NC or p7) proteins which nonspecifically bind to nucleic acids, thus protecting the RNA from nucleases and also compacts it within the core. There are three Pol proteins which are essential enzymes that are also encapsulated within the capsid; they are protease (PR), reverse transcriptase (RT), and integrase (IN). Along with these eight proteins, there is a p6 protein also located within the virion, as well as six accessory proteins. Three of the six proteins (Vif, Vpr, and Nef) are found within the viral particle, while (Tat, Rev, and Vpu) are not (Frankel and Young, 1998). **Figure 1** shows both the organization of the HIV-1 genome and the general structure of the virion.

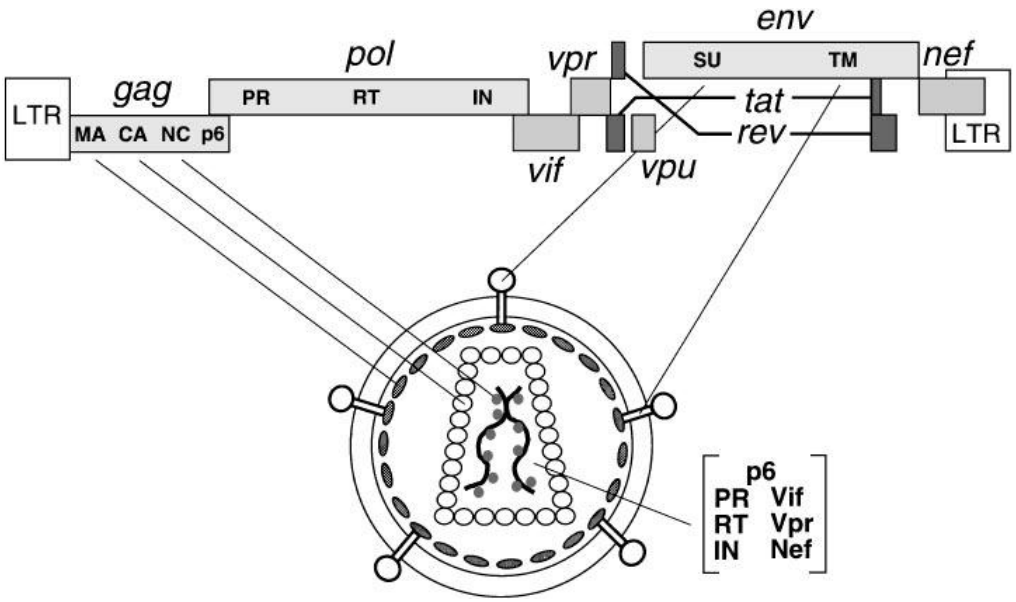


Figure 1: HIV-1 Genome and Structure (Frankel and Young, 1998).

As seen in Figure 1, the HIV-1 genome has nine open reading frames (ORF's) that encode 15 different proteins. Three of these ORF's encode the Gag, Pol, and Env polyproteins, which are the main viral proteins. These polyproteins are then cleaved by the protease as to separate the individual proteins. Gag is cleaved into the four following proteins: MA, CA, NC, and p6; Pol is cleaved into three enzymes: PR, RT, and IN, and Env is cleaved into SU and TM proteins (Frankel and Young, 1998). There is a small spacer peptide (p1) between NC and p6, and between CA and NC there is another small spacer peptide (p2). These spacers help regulate the rate and location of PR cleavage (Freed, 1998). The six other proteins are essentially accessory proteins and are encoded individually; they are: Vif, Vpr, Nef, Tat, Rev, and Vpu (Frankel and Young, 1998).

The Lifecycle of HIV-1

Cells become infected with HIV when the glycoprotein gp120 of a mature virion binds CD4 and interacts with one of two chemokine co-receptors, CCR-5 or CXCR-4. These co-receptors are crucial, because both virus adsorption and cellular infectivity increase for cells which have low amounts of CD4 but high amounts of one of the co-receptors (Kozak et al., 1997). The transmembrane glycoprotein gp41 undergoes a conformational change, which promotes fusion between the virion and the host cell, ultimately releasing the viral capsid within the host cell. The viral capsid is then uncoated and the viral ss-RNA is converted to ds-DNA by the viral RT enzyme. The viral DNA enters the nucleus and is incorporated into the chromosomal DNA by the IN enzyme (Freed, 1998). The viral genes are then expressed, in which Tat greatly increases

the rate of viral transcription. Most of the mRNA's are translated in the cytoplasm, except for the Env mRNA which is translated in the endoplasmic reticulum (ER). Full-length viral RNA's are spliced and transported to the cytoplasm, to be translated or packaged, and is regulated by Rev. Vpu promotes the degradation of CD4 which is coexpressed with Env in the ER, while Nef degrades surface CD4 so that Env can be transported to the cell surface. The viral proteins and RNA are assembled at the cell surface and an immature virion buds from the cell, with its cell surface coated by both gp120 and gp41. The viral PR cleaves the Gag and Gag-Pol polyproteins into their mature active proteins, thus becoming a mature virion (Frankel and Young, 1998). If this final proteolytic step does not occur, then newly budded viral particles are noninfectious (Warnke and Barreto, 2007). For this reason, HIV-1 PR is one of the main therapeutic targets for developing anti-HIV-1 drugs.

Anti-retroviral Drugs Used to Treat HIV

There are currently six different classes of antiretroviral drugs that are FDA-approved for treating HIV infections. These classes are nucleoside reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs), protease inhibitors (PIs), fusion inhibitors (FIs), entry inhibitors/co-receptor inhibitors (CRIs), and integrase inhibitors (INIs), in which the latter three only consist of one drug per class. Each class of drug combats the HIV virus differently, but all do so during the most crucial steps in the viral life-cycle (Mehellou and De Clercq, 2010). **Table 1** shows the six different drug classes and the approved drugs within each class.

Table 1: FDA-Approved Antiretroviral Drugs Used in the Treatment of HIV Infection.

Nucleoside Reverse Transcriptase Inhibitors	Non-Nucleoside Reverse Transcriptase Inhibitors	Protease Inhibitors	Fusion Inhibitors	Entry Inhibitors-CCR5 co-receptor antagonist	HIV Integrase strand transfer Inhibitors
Combivir	Intelence	Amprenavir	Fuzeon	Selzentry	Isentress
Emtriva	Rescriptor	Tipranavir			
Epivir	Sustiva	Indinavir			
Epzicom	Viramune	Saquinavir			
Hivid		Lopinavir			
Retrovir		Ritonavir			
Trizivir		Darunavir			
Truvada		Atazanavir			
Videx		Nelfinavir			
Viread					
Zerit					
Ziagen					

The first approved drug class was the NRTIs, with the first drug being approved in 1987. This drug class inhibits the HIV reverse transcriptase because the molecules get incorporated into the growing DNA strand but lack a 3'-hydroxyl group, thus preventing incoming nucleotides from forming a phosphodiester bond. These drugs actively compete with natural nucleotides so their addition terminates the growing DNA strand and renders it useless. The problem with these drugs is that in order to work, they first must be phosphorylated to their active 5' triphosphate form by using cellular kinases, so they may not become phosphorylated (Warnke and Barreto, 2007).

The second class of anti-HIV drugs are NNRTIs, and work by noncompetitively binding close to the active site of reverse transcriptase which causes a conformational change of the active site (Warnke and Barreto, 2007).

Two drugs inhibit the entry of the virion into the host cell and are classified as an FI and a CRI. These drugs block the virion's glycoproteins from fusing with a host cell. There is only one INI, and it stops the viral DNA from incorporating itself into cellular DNA. Protease inhibitors are competitive inhibitors to the different substrates that HIV-1 protease cleaves. These PIs are non-cleavable molecules; hence the protease cannot cleave the polyproteins into their individual and active proteins (Mehellou and De Clercq, 2010).

HIV-1 Protease

The active HIV-1 protease (PR) is a homo-dimer, in which the active site is formed between two 99-residue monomers, each monomer contributes a catalytically essential aspartic acid (Asp25/Asp25') (Frankel and Young, 1998). HIV PR is similar to other aspartyl proteases, in which a proton from a water molecule is transferred to one of the aspartic acids, followed by the transfer of a proton from the other aspartic acid to the peptide bond which is being cleaved, and briefly creates a tetrahedral non-covalent transition state. Another important feature of the protease is the presence of a conserved water molecule within the binding pocket, which mediates contacts between the amide groups of Ile50/Ile50' and the P2/P1 and P2'/P1' carbonyl oxygen atoms (Ali et al., 2010). In order for PR to bind the substrate, the two mobile flaps of the PR must move away, and once the substrate is bound, they move back over the active site and lock down over the bound substrate (Anderson et al., 2009). The PR cleaves the Gag and Gag-Pol polyproteins at 12 different sites (**Figure 2**) (Perez et al., 2010).

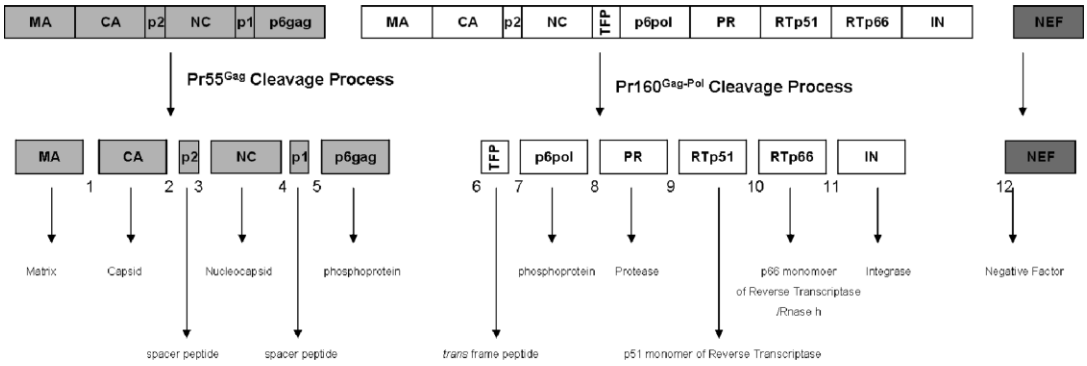


Figure 2. A schematic representation of the 12 substrate cleavage sites within the HIV-1 polyproteins (Perez et al., 2010).

HIV-1 Protease Substrate Cleavage Sites

The PR enzyme recognizes 12 sequences of eight amino acids (P4-P4'), which are quite diverse, and then cleaves the scissile bond between P1 and P1', which is denoted by an asterisk in **Table 2** (Perez et al., 2010). Most of the substrate sites have a branched amino acid residue at the P2 site, a hydrophobic residue at P1, and either an aromatic or proline at P1'. Due to this variance, the enzyme cleaves different substrates with efficiencies that differ by nearly 400-fold (Anderson et al., 2009). One of the fastest substrates to be cleaved is the first one, which is cleaved at the Tyr-Pro peptide bond. This reaction yields a K_m of around 103 µM, and a turnover number (k_{cat}) of 4.9 s^{-1} (Matayoshi et al., 1990).

Table 2. The 12 Sequences Cleaved by HIV-1 Protease.

Substrate Sequences (P4-P4')	Cleavage Domain
SQNY*PIVQ	MA-CA
ARVL*AEAM	CA-p2
ATIM*MQRG	p2-NC
RQAN*FLGK	NC-p1
RQAN*FLRE	NC-TFP

PGNF*LQSR	p1-p6gag
DLAF*LQGK	TFP-p6pol
SFNF*PQVT	p6pol-PR
TLNF*PISP	PR-RTp51
AETF*YVDG	RTp51-RTp66
RKVL*FLDG	RTp66-INT
DCAW*LEAQ	NEF

HIV-1 Protease Inhibitors

There are nine FDA-approved HIV-1 PIs (**Figure 3**), and all are competitive active site inhibitors which bind the wild-type (Q7K) protease from the nanomolar to picomolar range. All but TPV are peptidomimetics which mimic the enzymatic transition state, with several different non-cleavable scaffold cores (Anderson et al., 2009). These PIs are generally hydrophobic and interact with the mainly hydrophobic S2-S2' pockets in the PR active site. Although the PIs are chemically different, their three-dimensional shapes and electrostatic characteristics are quite similar (Nalam and Schiffer, 2008). There are either six or seven bonds between the two critical P2/P1' carbonyl oxygen atoms, which is very similar to the six bonds between the substrates. For the peptidomimetic PIs there also is a secondary hydroxyl group in place of the carbonyl at P1, which interact with the catalytic Asp25/25' protease residues and is necessary for tight PI binding (Ali et al., 2010). Darunavir is the most potent PI (Anderson et al., 2009) and has both six bonds between the two carbonyls along with the secondary non-cleavable hydroxyl group at P1, making DRV a great reference when developing new PIs.

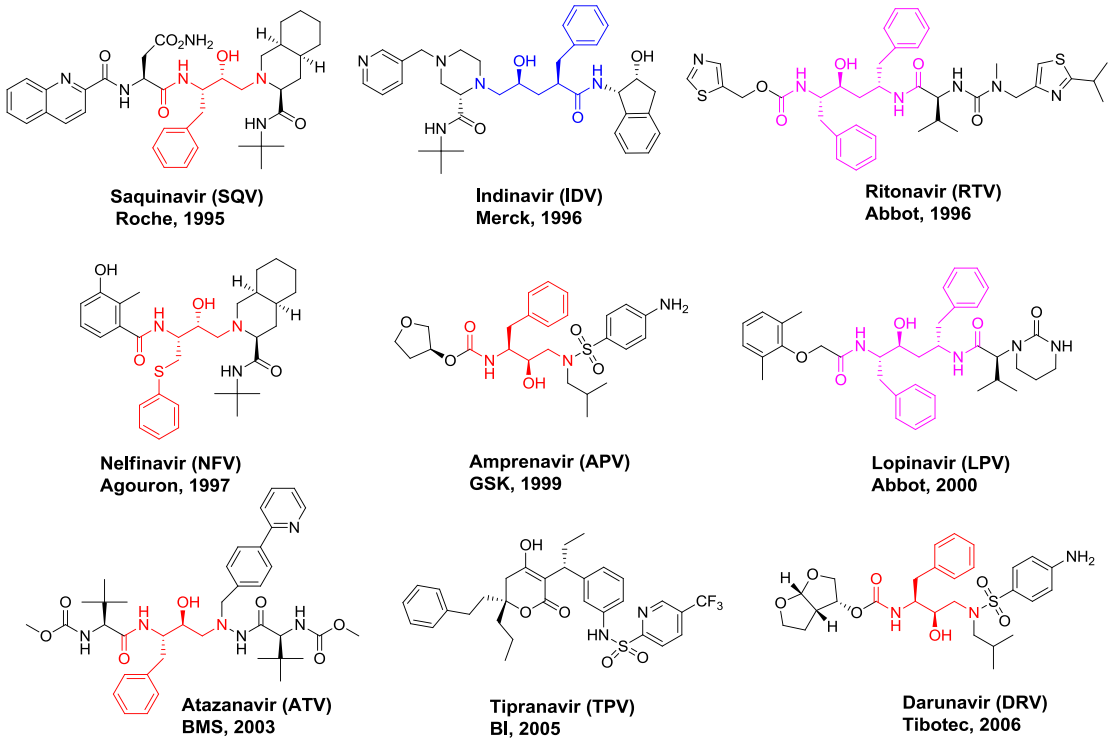


Figure 3. FDA-approved HIV-1 protease inhibitors, with the different scaffold cores shown in color (Ali et al. 2010).

Development of HIV-1 Drug-Resistant Protease Variants

The main reasons for drug-resistant mutations in HIV-1 PR are due to the high mutation rate and infidelity of HIV-1 RT and to selective drug pressure (King et al., 2004). HIV-1 replicates in and infects new cells at an extremely high rate, and since the RT does not have any proofreading ability, the process is error-prone. About one mutation occurs per transcribed viral genome. Most are base substitutions, but others are duplications, insertions, and recombinations (Clavel and Hance, 2004). When patients undergo antiretroviral therapy, observed mutations occur more frequently because the virus must evolve to survive. Once a primary drug-mutation occurs by introducing a PI,

secondary mutations often occur, which help increase the fitness of the protease (Nalam and Schiffer, 2008).

Mutations of the PR occur at residues both inside and outside the active site, and at least 34 of the 99 residues of each monomer show that mutations have clinical significance. The HIV-1 protease genome of most highly advanced patients usually has between five and 15 mutations. Some of the most common active site mutations are D30N, G48V, V82A, I84V, I50V, and I50L, with I84V being the worst of the multi-drug-resistant mutations as it strongly impacts the binding of most PIs (Nalam and Schiffer, 2008; King et al., 2004). The I50V, V82A, and I84V mutations all decrease the residue size, thus losing on average two or more crucial van der Waals contacts between the PR and the PI as compared between the normal wild-type PR and the substrate (King et al., 2004). Mutations outside the active site cause a large conformational change of the PR to process and release its substrate, but PIs are less dynamic because they should stay tightly bound to the PR. Increasing the flexibility of the PR detrimentally affects inhibitor binding by increasing the rate of dissociation between the PR and the PI (Nalam and Schiffer, 2008).

Substrate Envelope Hypothesis

Although the normal PR protein substrates are quite different structurally, various cleavage site peptides adopt a conserved shape/volume. By overlapping the volumes of various bound substrates within the active site, a conserved shape has been identified called the substrate envelope (**Figure 4A**). This method can also be done with the PIs bound to wild-type PR, because they too adopt a conserved shape and contact similar

residues within the active site of the PR, and is called the inhibitor envelope (**Figure 4B**). Overlaying these two conserved envelopes shows that when inhibitor atoms protrude beyond the substrate envelope and contact other PR residues, drug resistance occurs (**Figure 4C**). Most drug resistant mutations in the active site do not contact the substrates, which prove that these primary mutations are caused by the inhibitor protruding beyond the substrate envelope. The better inhibitors fit within this envelope, the less susceptible they are to drug resistance, since a mutation that affects the inhibitor would also affect cleavage of the majority of the substrates (Ali et al., 2010).

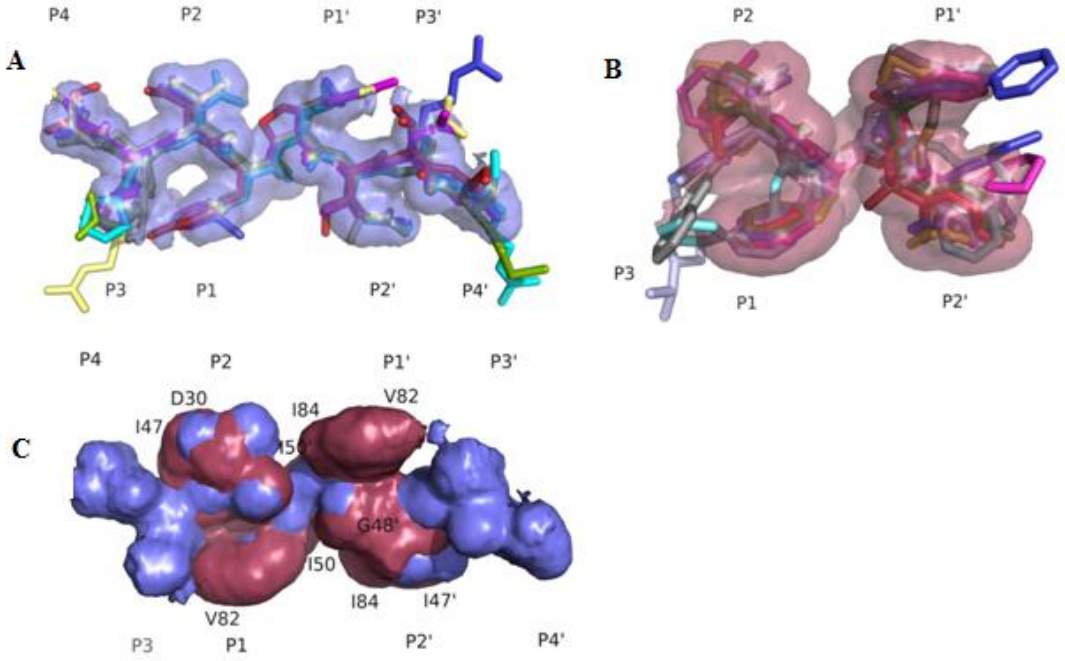


Figure 4. (A) Substrate envelope of HIV-1 protease. (B) The inhibitor envelope within the active site of HIV-1 protease. (C) Overlay of the substrate envelope (blue) with the inhibitor envelope (red). Residues that contact inhibitors due to their protrusion from the substrate envelope and confer drug resistance when mutated are labeled (Ali et al., 2010).

Fluorescent Resonance Energy Transfer (FRET) Assays

In this project, in order to measure the enzymatic inhibitory constants (K_i's) for different inhibitors, FRET assays were used. These assays measure the increase in fluorescence over time. The assay contains a natural substrate of HIV-1 which contains a fluorescent donor (EDANS) at the C-terminus and a quenching acceptor (DABCYL) at the N-terminus. The substrate has the following sequence (Arg-Glu(EDANS)-Ser-Gln-Asn-Tyr-Pro-Ile-Val-Gln Lys(DABCYL)-Arg) (**Figure 5**) where HIV-1 protease cleaves at the Tyr-Pro bond. EDANS is excited at 340 nm and will only fluoresce if it is more than 100 Å away from DABCYL. So fluorescence will only occur if the substrate is cleaved by the protease. Measuring the change of EDANS fluorescence intensity at 490 nm directly correlates with how much substrate has been cleaved (Matayoshi et al., 1990).

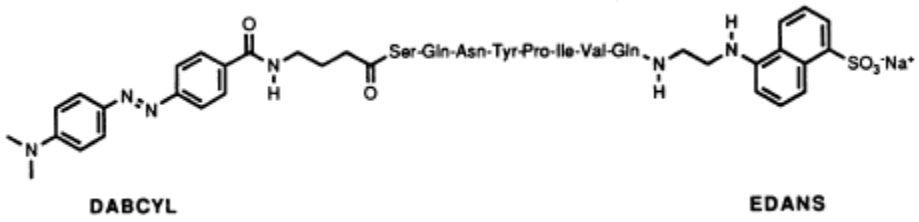


Figure 5. Structure of the fluorogenic HIV-1 substrate (Matayoshi et al., 1990).

From the data collected, initial velocities are plotted against the inhibitor concentrations and are fit by non-linear regression. The K_i value is derived using the Morrison equation (**Figure 6**).

$$v = C\,\frac{\sqrt{(K_i + I - fE)^2 + 4\,k_i\,fE} - (K_i + I - fE)}{2} \qquad \text{Where } (S << K_m)$$

f = fraction of active enzyme, E = 50 nM, S = 1 µM Parameters for non-linear regression fitting for Ki, C and f

$$C = \frac{k_{cat}\,S}{K_m + S}$$

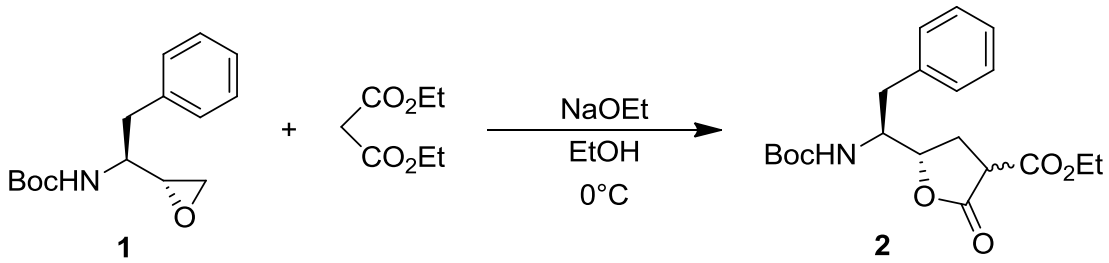
Figure 6. The Morrison equation and its constants.

PROJECT PURPOSE

The purpose of this project was to design novel HIV-1 protease inhibitors (PIs) that fit very well within the conserved substrate envelope, so that none of the functional groups would protrude out of the proposed envelope. By focusing on the conserved envelop the drug should be much more potent, allowing it to bind to drug-resistant proteases. The novel cores had six bonds in between two crucial carbonyls, which is similar to Darunavir and the substrates themselves, and the new cores also did not have an amide group linking the core to the P2' group. This allows the PI to have more degrees of freedom, making it easier to conform to the active site of the protease and many of its mutants. The enzymatic kinetic evaluation on Darunavir and 11 of its analogues was performed to see whether the different functional groups bound tighter to drug-resistant proteases providing lower K_i's.

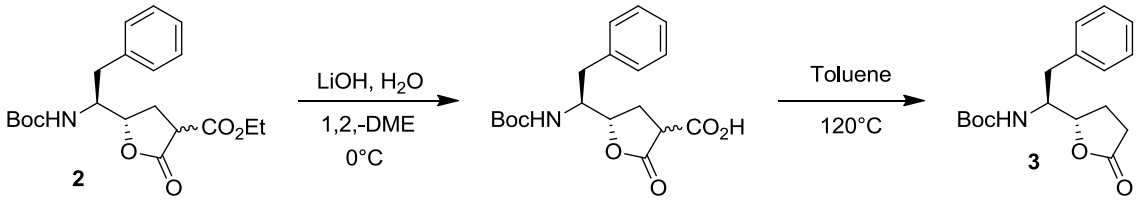
METHODS

Synthesis of (S)-ethyl 5-((S)-1-(tert-butoxycarbonylamino)-2-phenylethyl)-2-oxotetrahydrofuran-3-carboxylate



Epoxide **1** (20 g, 75.94 mmol) was mixed with diethylmalonate (12.68 mL, 83.53 mmol) and dissolved in EtOH (60 mL) at room temperature until the solution became homogeneous. The mixture was cooled to 0 °C and NaOEt (5.68 g, 83.53 mmol) in EtOH (40 mL) was slowly added under dry nitrogen atmosphere, stirred for 30 min, warmed to room temperature, and stirred overnight. Reaction was quenched by cooling the mixture to 5 °C and AcOH (10 mL) in EtOAc (90 ml) was slowly added. EtOAc (200 mL) and a 25% aqueous solution of NaCl (200 mL) were added and layers separated. The organic layer was washed with a 5% aqueous $NaHCO_3$ solution (300 mL), and then by a saturated aqueous NaCl solution (200 mL). The extraction process was done once more with EtOAc (150 mL). The combined organic extract was dried (Na_2SO_4), filtered, and evaporated under reduced pressure. The residue was dissolved in t-butyl methyl ether (MTBE) (100 mL) and was evaporated to dryness once again. The resulting crude off-white crystalline product was dissolved in MTBE (40 mL), filtered, and MTBE (15 mL) and n-heptane (90mL) were added and mixed at 60°C until homogeneous. The product was allowed to crystallize overnight at 4°C. The product was mixed, filtered, and the filter cake was washed with a mixture of MTBE and n-heptane (1:9) (150 mL). The solid was dried providing the pure off-white crystalline solid **2** (20.73 g, 72.3%).

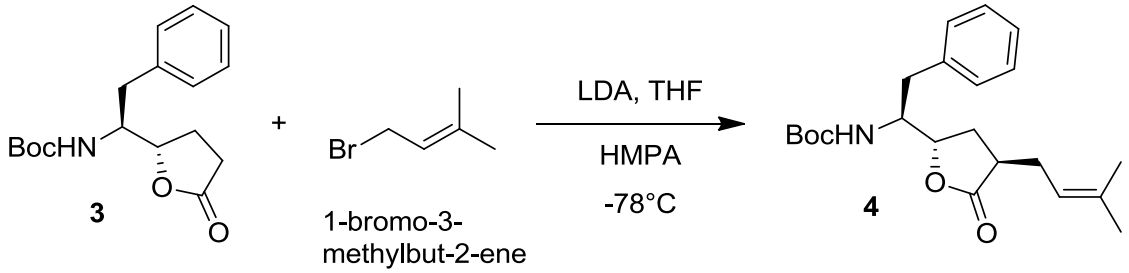
Synthesis of tert-butyl (S)-1-((S)-5-oxotetrahydrofuran-2-yl)-2-phenylethylcarbamate



To a solution of $LiOH \cdot H_2O$ (5.56 g, 132.5 mmol) in water (100 mL) was added a solution of the ester **2** (10 g, 26.5 mmol) in 1, 2-dimethoxyethane (DME) (100 mL). The reaction stirred at room temperature for 4 hours, then cooled to 0 °C, and a 10% aqueous solution

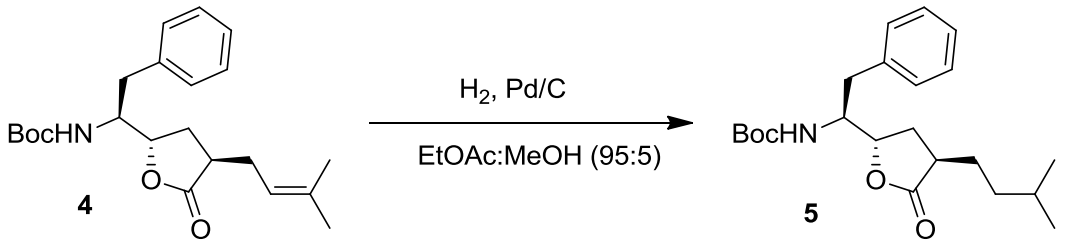
of citric acid (55 mL) was added to lower the pH to 6. The mixture was partially evaporated under reduced pressure at room temperature, and a 2N HCl solution was used to get a pH of 3-4. The resulting solution was stirred at room temperature for 2 hours and was extracted with EtOAc (3 x 300 mL) and once with CH_2Cl_2 (250 mL). Combined organic extract was washed with a saturated aqueous NaCl solution (100 mL), dried (Na_2SO_4), filtered, and evaporated under reduced pressure to yield 10.1g of white foam. The resulting crude acid was dissolved in toluene (90 mL) and heated to 120 °C overnight. The resulting solution was then evaporated under reduced pressure to yield a pale yellow gummy residue. The residue was purified by flash column chromatography on silica gel, using an EtOAc-hexanes (3:7) mixture as eluent to provide the lactone **3** (8.23 g, 87.1%) as an off-white solid.

Synthesis of tert-butyl (S)-1-((2S,4R)-4-(3-methylbut-2-enyl)-5-oxotetrahydrofuran-2-yl)-2-phenylethylcarbamate



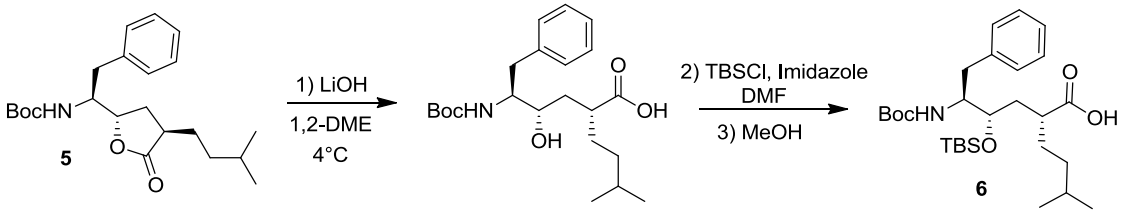
A 1.6 M solution of n-butyl lithium in THF (15 mL, 24.0 mmol) was added dropwise to a −78 °C solution of diisopropyl amine (3.47 mL, 24.53 mmol) in dry THF (24 mL) under dry nitrogen atmosphere, and stirred for 45 min. Lactone **3** (3g, 9.81 mmol) in THF (10 mL) was added drop-wise, and the enolate was allowed to form at this temperature for 45 min. A solution of 1-bromo-3-methylbut-2-ene (1.83 g, 12.26 mmol) in HMPA (10 mL) was added drop-wise to the reaction, which became very difficult to stir. The reaction was left at −78°C for 30 min and then transferred to a −50°C bath. The solution was maintained between −50 and −35°C for 2 hours and the reaction was quenched with a saturated aqueous solution of NH_4Cl (15 mL). After warming to room temperature, the organic layer was extracted with EtOAc (2 x 200 mL), which were then washed with a saturated NaCl aqueous solution (50 mL), dried (Na_2SO_4), filtered, and evaporated under reduced pressure. The yellowish oil was purified by flash column chromatography on silica gel using an EtOAc in hexanes (15-20%) mixture as eluent to provide the ester **4** (1.94 g, 52.8%) as a yellowish oil.

Synthesis of tert-butyl (S)-1-((2S,4R)-4-isopentyl-5-oxotetrahydrofuran-2-yl)-2-phenylethylcarbamate



Ester **4** (1.84 g, 5.18 mmol) was dissolved in an EtOAc-MeOH (95:5, 25 mL) mixture and a 10% Pd/C catalyst (0.400 g) was added and hydrogenated over night using balloons. The reaction mixture was filtered through a pad of celite, the pad was washed with EtOAc, and the filtrate was evaporated under reduced pressure to yield an off-white solid which was purified by flash column chromatography on silica gel using an EtOAc in hexanes (30%) mixture as eluent to provide the lactone **5** (1.82 g, 93.4%) as a white solid.

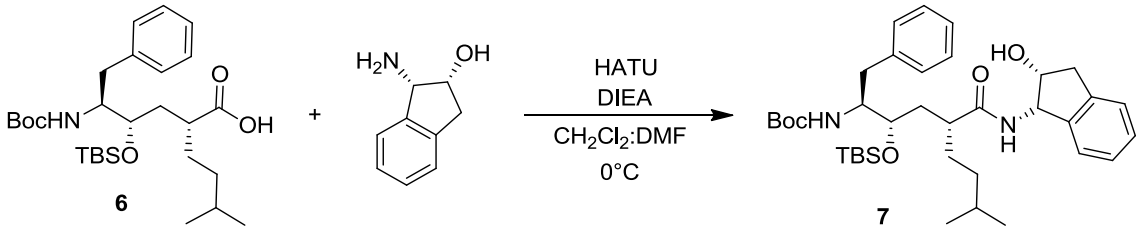
Synthesis of (2R,4S,5S)-5-(tert-butoxycarbonylamino)-4-(tert-butyldimethylsilyloxy)-2-isopentyl-6-phenylhexanoic acid



A 1M aqueous LiOH solution (16 mL) was slowly added at 4 °C to a solution of compound **5** (1.0 g, 2.66 mmol) in 1,2-DME (20 mL), and stirred overnight. Citric acid was then added as to get a pH of 4 and the resulting solution was extracted with EtOAc (3 x 150 mL). The combined organic extract was washed with saturated aqueous NaCl solution (50 mL), dried (Na_2SO_4), filtered, and evaporated under reduced pressure at 20 °C. The crude hydroxyl acid was dissolved in DMF (25 mL) and *tert*-butyldimethylsilyl chloride (4.05g, 26.6 mmol) and imidazole (2.0 g, 29.3 mmol) were added to the solution and stirred for 36 hours. The reaction mixture was treated with methanol (20 mL) and stirred for 8 hours, then evaporated under reduced pressure. The reaction mixture was partitioned between a 10% aqueous citric acid solution (20 mL) and EtOAc (125 mL), and the aqueous layer was further extracted with EtOAc (125 mL). The combined organic extract was washed with a saturated aqueous NaCl solution (20 mL), dried (Na_2SO_4), filtered, and evaporated under reduced pressure. The resulting residue was purified by flash column chromatography on silica gel using a 2-3% methanol in CH_2Cl_2 mixture as an eluent to provide acid **6** (1.24 g, 91.8%) as an almost

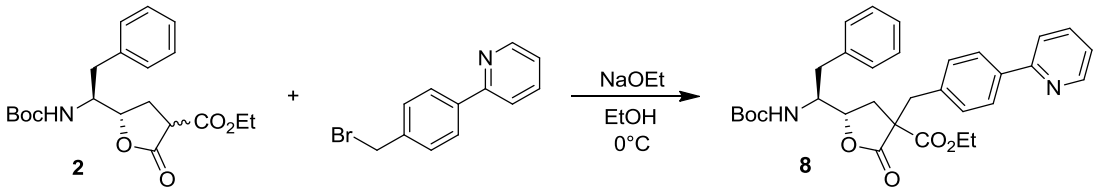
clear gummy substance. Analysis by NMR showed that some racemazation had occurred.

Synthesis of tert-butyl (2S,3S,5R)-3-(tert-butyldimethylsilyloxy)-5-((1S,2R)-2-hydroxy-2,3-dihydro-1H-inden-1-ylcarbamoyl)-8-methyl-1-phenylnonan-2-ylcarbamate



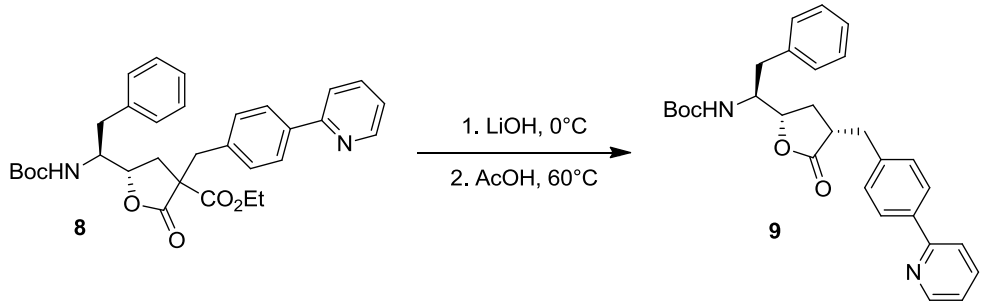
Acid **6** (0.40 g, 0.79 mmol) and (1S,2R)-(−)-cis-1-Amino-2-indanol (0.118 g, 0.79 mmol) were mixed in a 1:1 ratio of CH_2Cl_2:DMF (10 mL), cooled to 0 °C, and N-N-diisopropylethylamine (DIEA) (0.31 g, 2.37 mmol) was slowly added. Solid HATU (0.30 g, 0.79 mmol) was added in one portion after 15 minutes and left at 0 °C for 15 min before moving it to room temperature. Reaction mixture was evaporated under reduced pressure, partitioned between water (20 mL) and EtOAc (125 mL), and the aqueous layer was extracted once more with EtOAc (125 mL). The combined organic extract was washed with a saturated aqueous NaCl solution (20 mL), dried (Na_2SO_4), filtered, and evaporated under reduced pressure.

Synthesis of (5S)-ethyl 5-((R)-1-((tert-butoxycarbonyl)amino)-2-phenylethyl)-2-oxo-3-(4-(pyridin-2-yl)benzyl)tetrahydrofuran-3-carboxylate



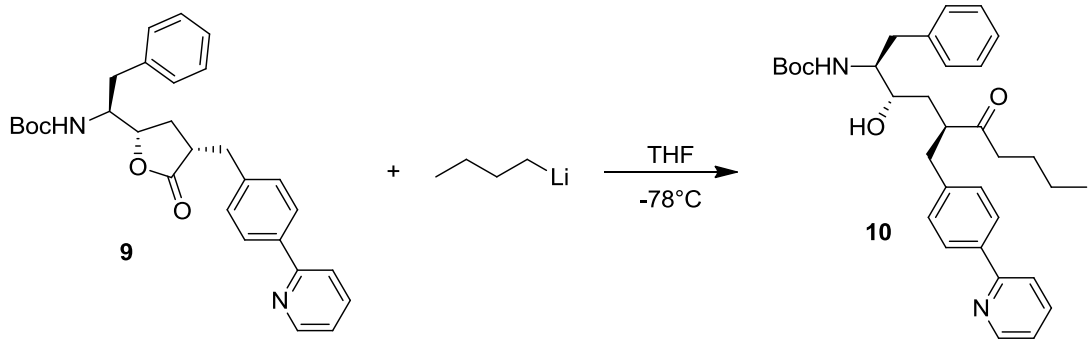
Ester **2** (5.0 g, 13.25 mmol) and 2-(4-(bromomethyl)phenyl)pyridine (3.29 g, 13.25 mmol) were mixed in EtOH (42 mL) for 15 min and cooled to 0 °C. A solution of NaOEt (0.95 g, 13.9 mmol) in EtOH (8.5 mL) was slowly added, and the reaction mixture was left stirring at 4 °C overnight to yield lactone **8**.

Synthesis of tert-butyl (R)-1-((2S,4S)-5-oxo-4-(4-(pyridin-2-yl)benzyl)tetrahydrofuran-2-yl)-2-phenylethylcarbamate



To the crude mixture containing lactone **8**, $LiOH \cdot H_2O$ (2.7 g, 64.29 mmol) was added portionwise over 5 min, and stirred for 2.5 hours at 0 °C. AcOH (3.85 g, 68.57 mmol) was added and the reaction mixture was heated to 60 °C and left stirring overnight. The suspension was cooled to room temperature and extracted with EtOAc (3 x 50 mL). The combined organic extract was washed with a saturated NaCl aqueous solution (20 mL), dried (Na_2SO_4), filtered, and evaporated under reduced pressure. The resulting residue was purified by flash column chromatography on silica gel using an EtOAc-hexanes (2:3) mixture as eluent, evaporated under reduced pressure, and was crystallized by dissolving in a MeCN-H_2O (1:1) (50 mL) mixture at 75 °C for 1 hour and by slowly cooling the mixture to room temperature over a period of 2 hours. The solid was filtered, and the filter cake was washed with a MeCN:H_2O (1:1) (30 mL) mixture, n-heptane (20 mL) to give lactone **9** (2.0 g, 31.8%) as white fluffy crystals.

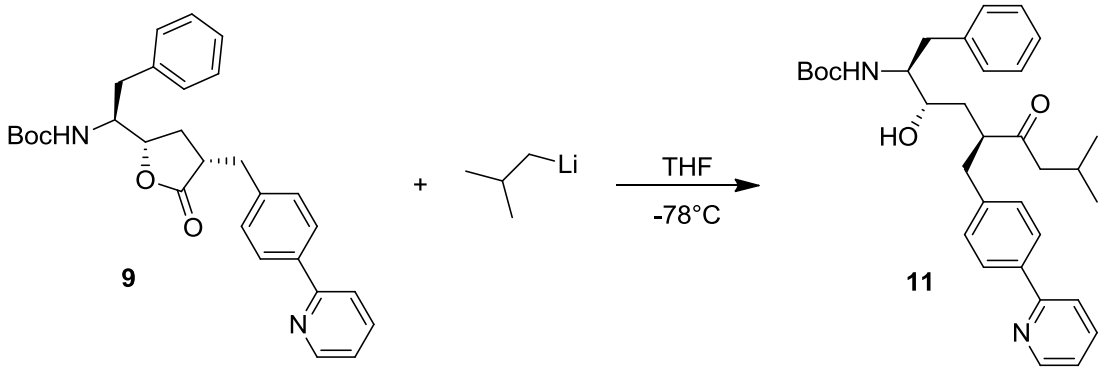
Synthesis of tert-butyl (2S,3S,5S)-3-hydroxy-6-oxo-1-phenyl-5-(4-(pyridin-2-yl)benzyl)decan-2-ylcarbamate



To a solution of lactone **9** (0.166 g, 0.35 mmol) in dry THF (10 mL) at −78 °C and under dry nitrogen atmosphere, a 1.6 M solution of n-butyl Lithium in THF (1.18 mL, 1.89 mmol) was slowly added and left stirring for 6 hours. The reaction mixture was quenched with a saturated aqueous NH_4Cl solution (5 mL) and extracted with EtOAc (2 x 30 mL). The combined organic extract was washed with a saturated aqueous NaCl

solution (15 mL), dried (Na_2SO_4), filtered, and evaporated under reduced pressure. The resulting residue was purified by flash column chromatography on silica gel using an EtOAc-hexanes (3:7) mixture as eluent to provide hydroxyketone **10** (0.120 g, 64.6%) as white crystals. NMR analysis revealed that the product was a 4:1 mixture of the (5S:5R) isomers.

Synthesis of tert-butyl (2S,3S,5S)-3-hydroxy-8-methyl-6-oxo-1-phenyl-5-(4-(pyridin-2-yl)benzyl)nonan-2-ylcarbamate



To a solution of lactone **9** (0.306 g, 0.646 mmol) in dry THF (15 mL) at −78 °C and under dry nitrogen atmosphere, a 1.4 M solution of sec-butyl Lithium in THF (2.3 mL, 3.23 mmol) was slowly added and left stirring for 2 hours. The reaction mixture was quenched with a saturated NH_4Cl aqueous solution (10 mL) and extracted with EtOAc (2 x 40 mL). The combined organic extract was washed with a saturated aqueous NaCl solution (15 mL), dried (Na_2SO_4), filtered, and evaporated under reduced pressure. The resulting residue was purified by flash column chromatography on silica gel using an EtOAc-hexanes (3:7) mixture as eluent to provide hydroxyketone **11** (0.008 g, 2.3%).

HIV-1 Protease Inhibition Assays

All HIV-1 protease inhibitor potencies were determined by fluorescent resonance energy transfer (FRET) assays. The HIV-1 protease substrate, (Arg-Glu(EDANS)-Ser-Gln-Asn-Tyr-Pro-Ile-Val-Gln Lys(DABCYL)-Arg) was purchased from Molecular Probes. The electron donor (EDANS) and electron acceptor (DABCYL) were labeled at the two ends of the substrate. Fluorescence measurements were completed using a fluorescence spectrophotometer (Photon Technology International) at 21 °C. Excitation and emission wavelengths were set at 340 and 490 nm, respectively, and each reaction

was monitored for about 5 minutes. Wild-type HIV-1 protease (Q7K) and its mutant variants were purified by desalted through PD-10 columns (Amersham Biosciences) in which sodium acetate (20 mM, pH 5) was used as the elution buffer. The protease concentrations were around 50 nM as estimated by UV spectrophotometry at 280 nm. All inhibitors were dissolved in dimethylsulfoxide (DMSO) and diluted to their appropriate concentrations. For all experiments, 40 μL of a 2 μM substrate were used in substrate buffer [0.1 M sodium acetate, 1 M sodium chloride, 1 mM ethylene-diaminetetraacetic acid (EDTA), 1 mM dithiothreitol (DTT), 2% DMSO, and 1 mg/mL bovine serum albumin (BSA) with an adjusted pH of 4.7]. Protease (2 μL) and inhibitor (1 μL) or DMSO were mixed with 37 μL of buffer before addition of the substrate. Inhibitor binding dissociation constant (K_i) values were obtained by nonlinear regression fitting (GraFit 5, Erithacus software) to the plot of 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. Each of the inhibitors were measured in triplicate, and their averages were calculated. The 12 drugs were measured against Wild-Type (Q7K), and three multi-drug mutants of HIV-1 protease. The three mutant proteases had the following mutations and were called: M1 (L10I, G48V, I54V, L63P, V82A); M2 (L10I, L63P, A71V, G73S, I84V, L90M); and M3 (I50V and A71V).

RESULTS

The purpose of this project was to design novel HIV-1 protease inhibitors that fit very well within the conserved substrate envelope, thus minimizing their susceptibility to drug resistance. Two novel cores were developed and synthesized. To increase the degrees of freedom of the new inhibitors, they did not include an amide group linking the cores to the P2' groups. Enzymatic kinetic evaluations were performed on the most potent PI (DRV) and 11 new DRV analogues against WT and 3 mutant proteases. After various approaches to create the new cores, a final procedure was finally developed and optimized (**Figure 7**).

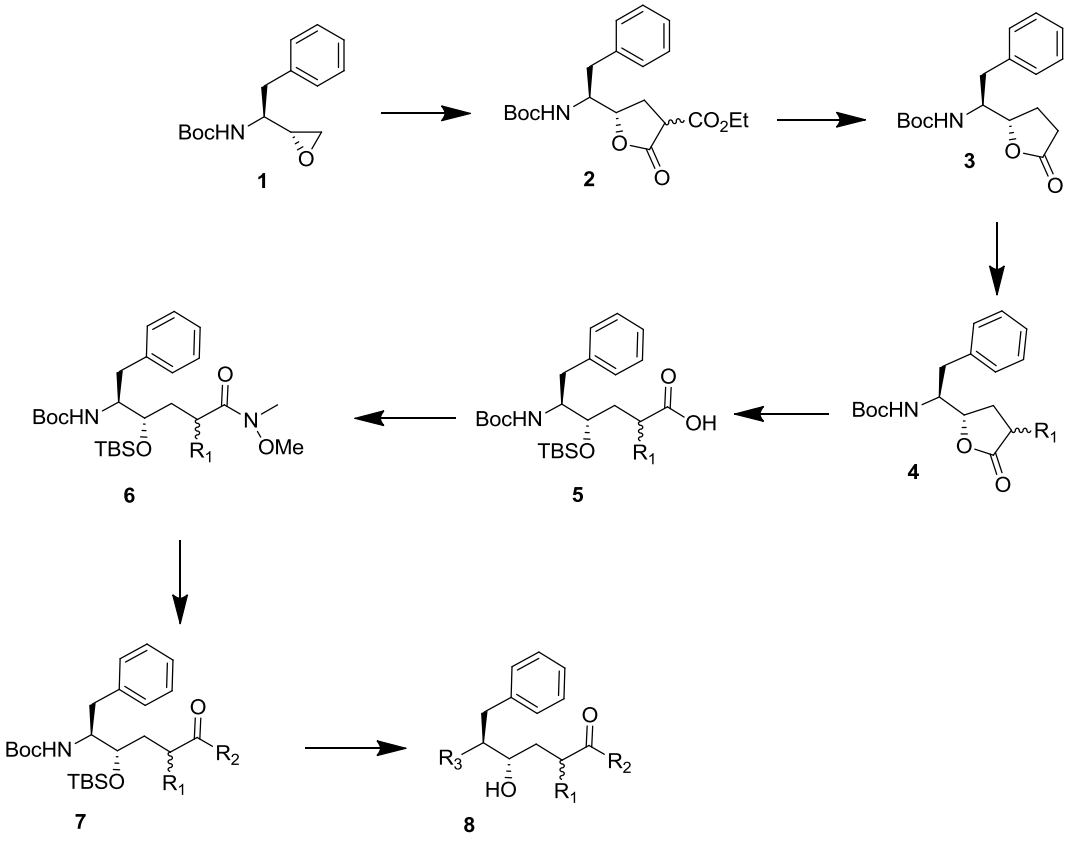


Figure 7. General scheme of the procedure used to develop the new PI cores.

The two novel cores only vary at the third chiral center, with both their chirality and functional group (**Figure 8**).

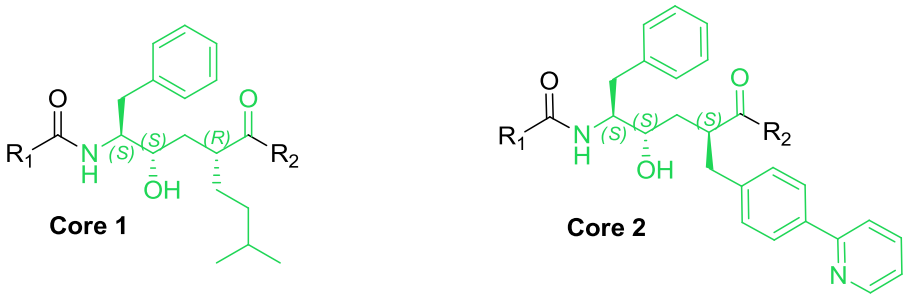


Figure 8. Structures of the two novel HIV-1 protease inhibitor cores, with the main scaffold cores shown in green.

To measure inhibition activities for the 12 compounds, FRET assays were performed, and from the initial velocities for each concentration, a graph was obtained in which the K_i values were found by non-linear fitting. The concentrations of drug used to measure the inhibitors ranged from 0 µM to 8 µM. The 12 drugs were measured against Wild-Type (Q7K), and three multi-drug mutants of HIV-1 protease. The three mutants had the following mutations and were called: M1 (L10I, G48V, I54V, L63P, V82A); M2 (L10I, L63P, A71V, G73S, I84V, L90M); and M3 (I50V and A71V). Four examples of these K_i graphs can be found below (**Figures 9-12**).

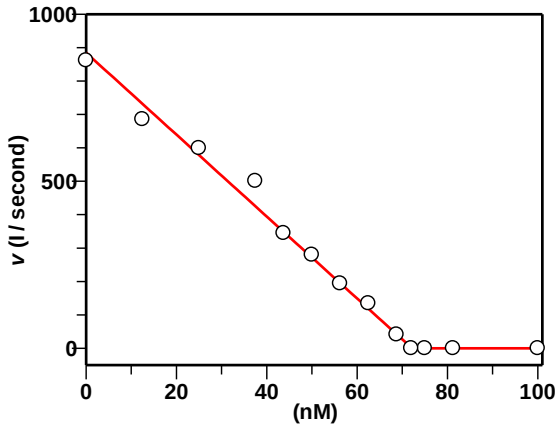


Figure 9. Inhibition curve for compound F with M1 mutant protease, gave a K_i of 0.001 nM.

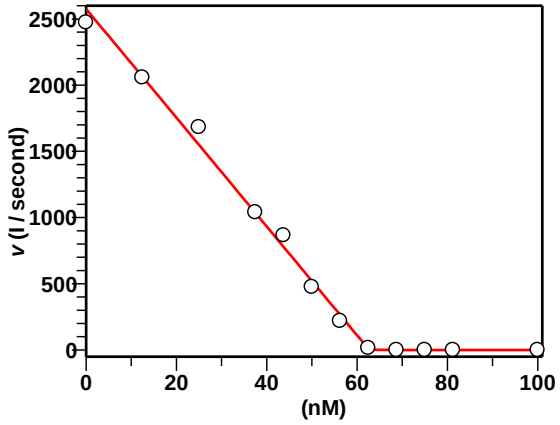


Figure 10. Inhibition curve for compound A with Q7K protease, gave a K_i of 0.001 nM.

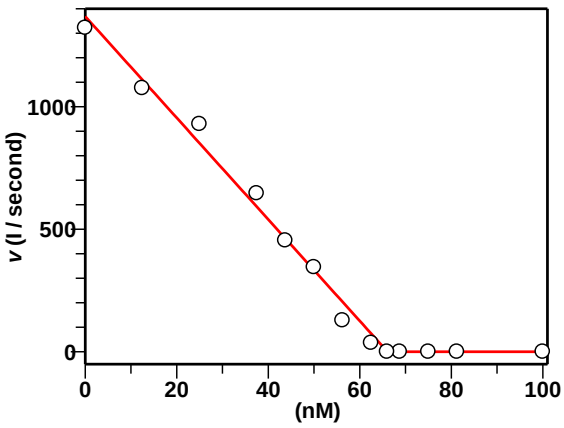


Figure 11. Inhibition curve for compound G with M2 mutant protease, gave a K_i of 0.001 nM.

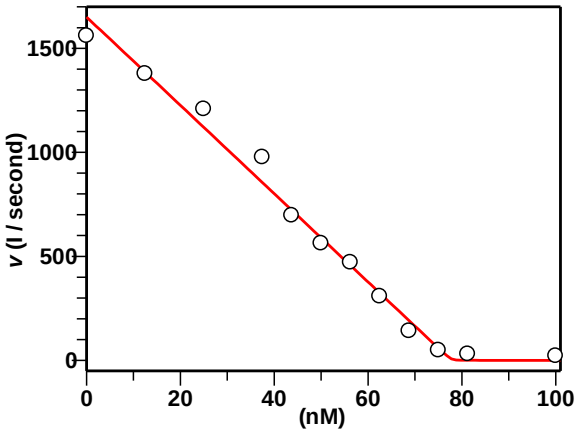


Figure 12. Inhibition curve for compound A with M1 mutant protease, gave a K_i of 0.001 nM.

The inhibition assays were performed in triplicate, and the averages and standard deviations are tabulated in **Table 3**. The table shows that 7 compounds had K_i values that were in the same range as Darunavir, and 4 compounds that were much worse inhibitors. Compounds A, B, C, E, and F showed better inhibition than DRV for both M2 and M3 proteases (right two columns in table). Their values varied, but compounds C and F displayed the best inhibition to M2 protease, with K_i's of 2 pM. Compound C gave an average K_i of 1 pM for M3 protease. None of the compounds had a lower K_i than DRV for Q7K and M1 proteases. Compounds H, I, J, and K gave K_i values that were much lower than DRV for all of the proteases. Only compounds H and K gave K_i's that were close to that of DRV, and were 48 and 55 pM respectively for the Q7K protease. These values are in the pM range, but they are still 20-25 times higher than DRV. Most of the K_i's for compounds H through K were in the nM range, especially for the mutant groups.

Table 3. Average K_i for DRV and 11 new DRV analogues.

Inhibitor	Q7K (nM)	M1 (nM)	M2 (nM)	M3 (nM)
Darunavir	0.002±0.002	0.002±0.002	0.008±0.011	0.006±0.005
Compound A	0.005±0.003	0.004±0.002	0.003±0.001	0.004±0.003
Compound B	0.004±0.001	0.008±0.004	0.006±0.006	0.004±0.005
Compound C	0.009±0.006	0.004±0.003	0.002±0.002	0.001±0.002
Compound D	0.004±0.004	0.004±0.003	0.022±0.034	0.020±0.026
Compound E	0.004±0.002	0.006±0.003	0.007±0.007	0.004±0.004
Compound F	0.004±0.004	0.003±0.002	0.002±0.002	0.004±0.001
Compound G	0.002±0.002	0.012±0.000	0.010±0.013	0.018±0.023
Compound H	0.048±0.014	1.974±0.756	3.256±0.421	1.056±0.193
Compound I	0.303±0.135	4.228±0.595	13.37±1.64	3.238±0.700
Compound J	0.230±0.030	3.986±1.705	3.198±0.817	2.547±0.314
Compound K	0.055±0.020	0.284±0.166	0.540±0.147	0.770±0.123

DISCUSSION

Two new HIV-1 protease inhibitor cores were developed and are expected to fit well within the substrate envelope so that they should be less susceptible to encounter drug resistance. Various Darunavir analogues have K_i values that are in the same range as DRV if not better. So by changing different functional groups, the inhibitors can become better or worse.

In order for a protease inhibitor to tightly bind to the protease, it must fit well within the substrate envelope and not protrude beyond it (Ali et al., 2010). Darunavir fits well within the substrate envelope and is the most potent PI currently available (Anderson et al., 2009). So by creating different DRV analogues, one can greatly enhance the binding affinity of the drug. If the analogue protrudes out of the substrate envelope more than DRV, then the binding affinity will greatly diminish.

Since the synthesis of two novel PI cores was developed, all that needs to be done now is add different functional groups to both sides of the cores. There are many different options to choose from, and the groups that are more likely to fit within the substrate envelope will be the best candidates. Computational analysis can greatly help choose which groups to add. One can predict the most energetically favorable configuration of a compound, and by using a docking program one can predict how well a certain compound will bind to the HIV-1 protease. After many new compounds are synthesized, FRET assays will be done as to see which compounds are the most potent. Then many other assays like solubility, toxicity, etc. are done as to see whether the new compounds have the potential to become new PIs. *In vivo* studies must be conducted and crystallography will be done on any promising compounds.

BIBLIOGRAPHY

Ali A et al. (2010) Molecular basis for drug resistance in HIV-1 protease. *Viruses*. **2**: 2509-2535.

Anderson J et al. (2009) Viral Protease Inhibitors. *Antiviral strategies.* Springer Publishers. p. 85-90.

Clavel F, Hance A (2004) HIV Drug Resistance. *The New England journal of medicine* **350:** 1023-1035.

Frankel AD, Young JA (1998) HIV: Fifteen proteins and an RNA. *Annual Reviews in Biochemistry* **67:** 1-25.

Freed EO (1998) HIV-1 Gag Protein: Diverse Functions in the Virus Life Cycle. *Virology* **251:** 1-15.

King NM et al. (2004) Combating Susceptibility to Drug Resistance: Lessons from HIV-1 Protease. C*hemistry & biology* **11**: 1333-1338.

Kozak SL et al. (1997). CD4, CXCR-4, and CCR-5 dependencies for infections by primary patient and laboratory-adapted isolates of human immunodeficiency virus type 1. *Journal of Virology* **71:** 873-882.

Matayoshi E et al. (1990) Novel fluorogenic substrates for assaying retroviral proteases by resonance energy transfer. *Science* **247:** 954-958.

Mehellou Y, De Clercq, E (2010) Twenty-six years of anti-HIV drug discovery: Where do we stand and where do we go? *Journal of Medicinal Chemistry* **53**: 521-538.

Nalam M, Schiffer C (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.

NIAID's http://www.niaid.nih.gov/topics/HIVAIDS/Understanding/Biology/Pages/biology.aspx

Perez M et al. (2010) Substrate recognition in HIV-1 protease: A computational study. *Journal of Physical Chemistry* **114:** 2525-2532.

Pornillos O et al. (2011) Atomic-level modeling of the HIV capsid. *Nature* **469:** 424-427.

(UNAIDS). Joint United Nations Programme on HIV/AIDS. Report for 2009.

Warnke D, Barreto J (2007) Antiretroviral Drugs. *Journal of Clinical Pharmacology* **47:** 1570-1579.

Weiss RA, (1993) How does HIV cause AIDS? *Science* **260:** 1273-1279.