1-1-2012

Drug resistance mechanisms and drug design strategies for human immunodeficiency virus and hepatitis c virus proteases

Yong Wang
Wayne State University,

DRUG RESISTANCE MECHANISMS AND DRUG DESIGN STRATEGIES FOR HUMAN IMMUNODEFICIENCY VIRUS AND HEPATITIS C VIRUS PROTEASES

by

YONG WANG

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2012

MAJOR: BIOCHEMISTRY AND
MOLECULAR BIOLOGY

Approved by:

_____
Advisor Date

_____

_____

_____

_____

ACKNOWLEDGMENTS

First of all, I would like to thank my adviser, Dr. Ladislau C. Kovari, and doctoral dissertation committee members, Dr. Brian F. P. Edwards, Dr. Zhe Yang, and Dr. Philip E. Pellett for the great academic guidance and suggestions.

The lab members including Iulia A. Kovari, Zhigang Liu, Tamaria G. Dewdney, and Samuel J. Reiter are good colleagues to work with as well as great friends to share the experience of life. Furthermore, I would like to express my gratitude to all the supporting staff of the Department of Biochemistry and Molecular Biology, Wayne State University School of Medicine and our departmental chair, Dr. Bharati Mitra.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

1.1 HIV epidemiology

According to the report of the Joint United Nations Program on HIV/AIDS (UNAIDS), about 2.6 million people became infected with AIDS in 2009 and an estimated 33.3 million people were living with the disease by the end of the year (*1*). Though the number of new infections and AIDS-related deaths have been steadily decreasing over the past few years (*1*), AIDS is still a serious global public health problem and a heavy socioeconomic burden.

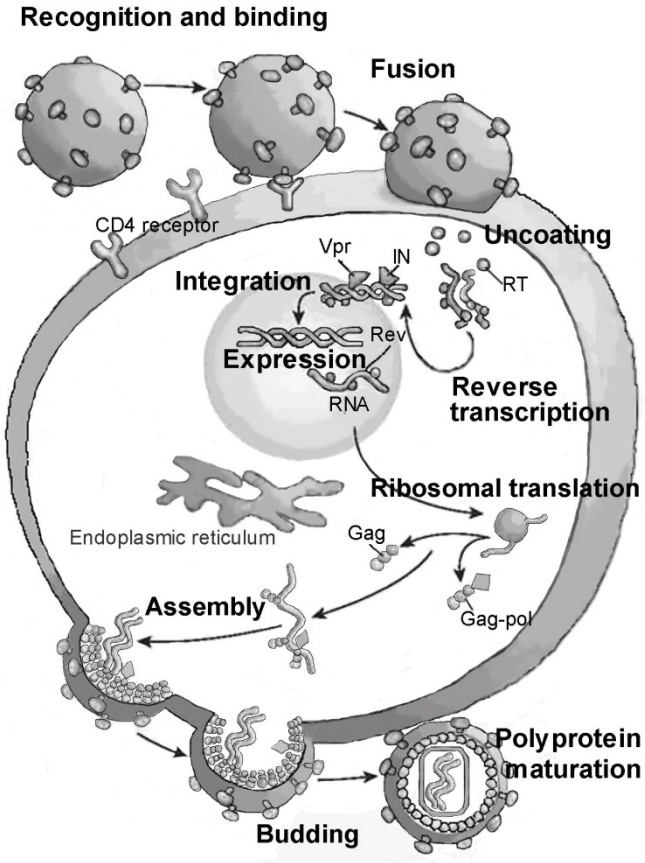


Figure 1.1 HIV life cycle. Adapted based on the figure produced by Pomerantz RJ and Horn DL (*2*).

1.2 HIV Life cycle

The human immunodeficiency virus (HIV) infects CD4 positive T cells through the interaction between viral envelope glycoprotein (gp120 and gp41) and CD4 receptor as well as co-receptors (CCR5 or CXCR4). After the virus anchors to the CD4 positive T cell, the HIV virion fuses with the target cell. The nucleocapsid containing the viral genome, viral proteins, and reverse transcriptase are released to the cytoplasm of the target cell. The viral genome encodes three essential enzymes, the reverse transcriptase, integrase, and protease. The first viral encoded enzyme, reverse transcriptase, reverse-transcribes viral RNA to DNA. As reverse transcriptase degrades the RNA template, a complementary DNA strand is synthesized to form a double stranded duplex called proviral DNA. Integrase then inserts the proviral DNA into the host genome. The inserted proviral DNA is first transcribed into a messenger RNA and then translated into a polyprotein. At this step, the third viral enzyme, the HIV protease, is required to cleave the polyprotein into eight separate proteins and two small spacer peptides. The mature viral structural proteins assemble around the viral genome into new functional viruses. As the new viron buds from its host cell membrane, it is coated with viral envelope proteins and departs to begin a new infectious cycle.

1.3 The HIV-1 genome

The HIV genome consists of a long single stranded RNA molecule. The viral genome contains three main open reading frames, the group-specific antigen gene (*gag*), the polymerase gene (*pol*), and the envelope gene (*env*). The *gag* gene encodes the p55 precursor polyprotein which is processed to generate the matrix protein (MA, p17), capsid

protein (CA, p24), nucleocapsid protein (NC, p7), spacer peptide 1 (SP1, p2), spacer peptide 2 (SP2, p1), and p6. The *pol* gene encodes three viral enzymes, which are the reverse transcriptase, integrase, and protease. The reading frames of *gag* and *pol* differ by one nucleotide. During translation, ribosomal frame-shifting produces the Gag-Pol polyprotein. The frequency of this frame-shifting is about 5 to 10% (*3*). Therefore, the three viral enzymes are expressed in the Gag-Pol precursor. The *env* gene encodes a precursor, gp160, which is processed by the host protease to generate the external glycoprotein gp120 and the transmembrane glycoprotein gp41. The Tat protein encoded by the transactivator gene (*tat*) is an essential viral regulatory factor interacting with an RNA element to activate transcription. Another essential regulatory factor is coded by the regulator of virion gene (*rev*), which promotes the nuclear export of the viral RNA. The product of virion infectivity factor gene (*vif*) preserves viral infectivity by inducing the degradation of an antiretroviral cytidine deaminase, APOBEC3G (apoplipoprotein B mRNA-editing, enzyme-catalytic, polypeptide-like 3G) (*4*). The viral protein R encoded by *vpr* gene regulates nuclear import of the HIV-1 pre-integration complex (*5*), arrests cell growth (*6*), and induces host cell apoptosis (*7*). The viral protein U gene (*vpu*) only exists in HIV-1. Its gene product induces CD4 degradation in the endoplasmic reticulum and enhances virion release from HIV-1-infected cells (*8*). The negative regulator protein gene (*nef*) encodes a multifunctional protein, which actually has a positive role in maintaining high viral load (*9*). The long terminal repeats (LTRs) are located at each end of the HIV-1 genome.

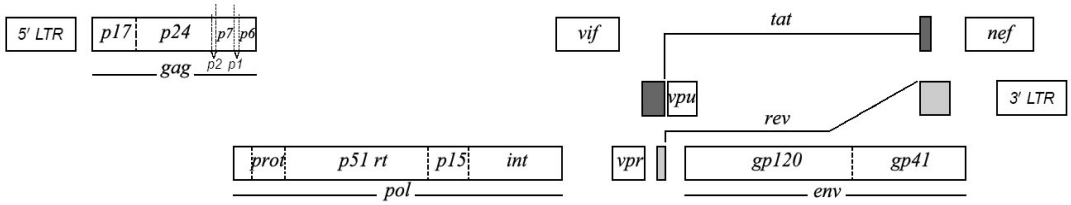


Figure 1.2 The HIV-1 genome. Open reading frames are shown as rectangles.

1.4 The HIV-1 protease

In the HIV life cycle, the protease is an essential element for viral maturation. The HIV protease is a homodimeric aspartyl protease, and each monomer is composed of 99 amino acid residues with a catalytic aspartate at position 25. HIV-1 protease cleaves Gag and Gag-Pol polyprotein precursor encoded by the HIV-1 virus genome at nine processing sites to produce mature active proteins. The Pol polyproteins is first cleaved off from the Gag-Pol polyproteins and then further digested into protease, reverse transcriptase (p51), RNase H (p15), and integrase. The active site is not fully exposed, being covered by a two flexible β-hairpin flaps. The flaps need to open to allow the substrates to access the active site. The HIV-1 protease enzyme activity can be inhibited by blocking the active site of the protease.

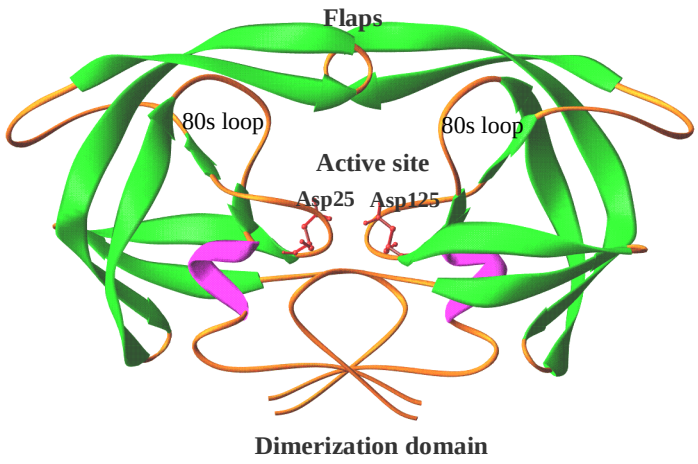


Figure 1.3 Structure of HIV-1 protease.

1.5 Anti-HIV drugs

Since HIV was discovered in 1983 (*10*), 26 anti-HIV compounds were approved by the U.S. FDA within almost three decades. The latest HIV drug, rilpivirine from Johnson & Johnson Pharmaceuticals, was approved in May 2011. Anti-HIV drugs have been developed to target the various steps in the HIV life cycle. Based on the targeted step of HIV life cycle, anti-HIV drugs can be classified into five different categories: fusion inhibitors, entry inhibitors, transcriptase inhibitors, integrase inhibitors, and protease inhibitors. The cell entry inhibitors, enfuvirtide and maraviroc, block the entry of the viral particle into the host CD4+ lymphocytes. Enfuvirtide interacts with the glycoprotein gp41 on the viral membrane to prevent the fusion of the viral particle and the host cell membrane (*11*). Maraviroc is an antagonist of the CCR5 co-receptor, which only acts on the macrophage-tropic (R5) HIV-1 strains (*12*). The inhibition of reverse transcriptase stops the production of proviral DNA. There are two classes of reverse transcriptase inhibitors, (i) Nucleoside/Nucleotide Reverse Transcriptase Inhibitor (NRTI) and (ii) Non-nucleoside Reverse Transcriptase Inhibitor (NNRTI). In the case of NRTIs, both nucleotide analog reverse-transcriptase inhibitors, such as tenofovir and adefovir, and nucleoside analog reverse-transcriptase inhibitors, such as zidovudine (AZT), didanosine (ddI), zalcitabine (ddC), stavudine (d4T), lamivudine (3TC), abacavir (ABC), emtricitabine (FTC), entecavir (INN), and apricitabine (ATC), get incorporated into the elongating DNA copy of HIV viral genome and cause premature termination of reverse transcription. The NNRTIs, such as efavirenz, nevirapine, delavirdine, etravirine, and rilpivirine bind to pockets other than the active site of reverse transcriptase and cause an

allosteric effect that inhibits reverse transcription. Integrase inhibitors are the newest class of antiretroviral agents. Raltegravir is the first integrase inhibitor approved by the U.S. FDA. Elvitegravir is undergoing Phase 3 clinical trials. Both compounds inhibit the strand transfer activity of integrase and halt the process of viral DNA integration (*13, 14*). The proteolytic cleavage of the polyprotein precursor is also a critical step for viral maturation. The inhibition of HIV protease effectively blocks the viral life cycle. The HIV-1 protease and its inhibitors are introduced in the following sections.

1.6 Current HIV-1 protease inhibitors

Based on the 2010 global report from UNAIDS, there are more than five million people receiving antiretroviral therapy, and this number is 13-fold higher than the number for 2004 (*1*). The combination therapy of HIV-1 protease inhibitors and reverse transcriptase inhibitors referred to as Highly Active Antiretroviral Therapy (HAART) is the current AIDS therapy that is most effective. The AIDS-related mortality has dropped sharply, and AIDS has gradually become a controllable, chronic disease. However, the extended length of time that patients take the medication and the error-prone replication of HIV genomes facilitate the evolution of drug resistant strains, which are the greatest obstacle in the treatment of HIV-1/AIDS patients. Only a few mutations in HIV-1 protease are needed to confer drug resistance. There have been numerous studies published regarding multi-drug resistance with a focus on HIV-1 protease (*15*).

The indispensable role of HIV-1 proteases in viral maturation makes it a popular target for drug design. The numerous structures of HIV-1 protease have facilitated the design of new and improved inhibitors. There are nine HIV-1 protease inhibitors

approved by the U.S. Food and Drug Administration (FDA); those inhibitors include: saquinavir, indinavir, ritonavir, nelfinavir, amprenavir, lopinavir, atazanavir, tipranavir, and darunavir. Unfortunately, resistance has been observed during the treatment with all of the approved HIV-1 protease inhibitors, and resistance cannot be overcome by increasing drug concentration due to the toxicity of protease inhibitors. Consequently, resistance has caused protease inhibitor treatment failure in a great number of patients (*16*). In the following section, the FDA-approved HIV-1 protease inhibitors are introduced. There are structural similarities among those inhibitors, and they share a similar binding pattern that might be the cause of the cross resistance characteristic of HIV-1 protease inhibitors.

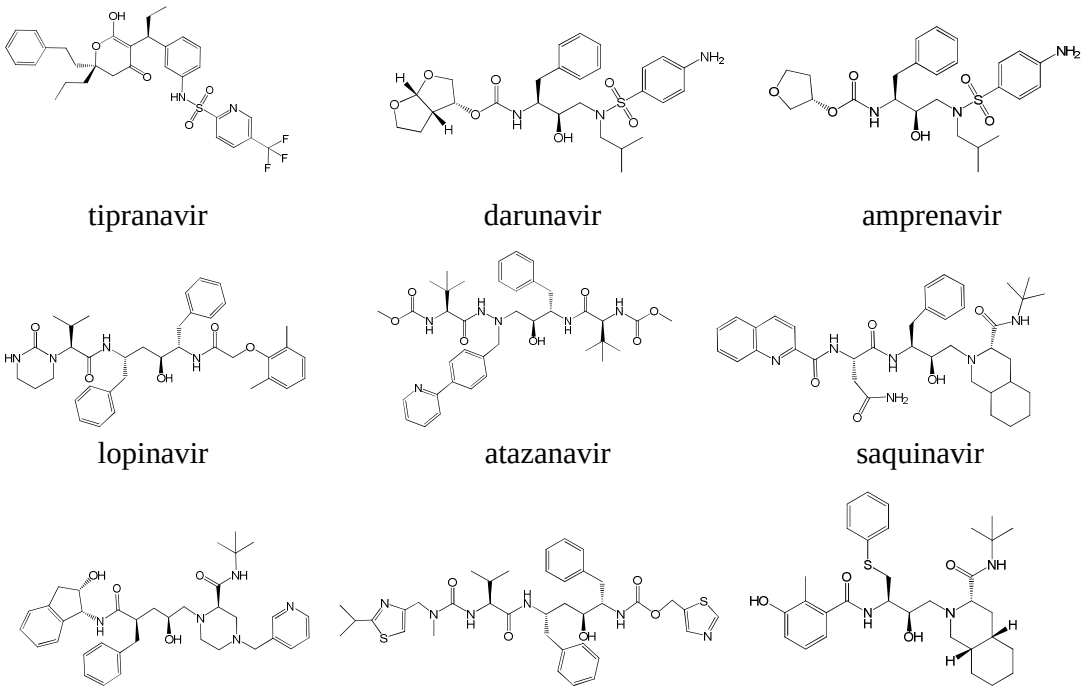


Figure 1.4 Chemical structures of the HIV-1 protease inhibitors approved by the US Food and Drug Administration.

1.6.1 Saquinavir

Saquinavir (brand name: Invirase), the first FDA-approved HIV-1 protease inhibitor, was invented by Hoffmann-La Roche Inc. The original design for the precursor of saquinavir comprised a proline at the P1’ site and a phenylalanine at the P1 site. The rationale is that HIV-1 protease cleaves the substrate between a phenylalanine and a proline while mammalian proteases do not cleave substrates containing proline at the P1’ site. In the final structure of saquinavir, the proline was replaced by a (S,S,S)-decahydro-isoquinoline- 3-carbonyl (DIQ) group to enhance the inhibitory potency (Figure 1.2). The carbonyl of the DIQ group contacts the bridging water molecule that interacts with the inhibitor and the flaps of HIV-1 protease (*17*).

1.6.2 Ritonavir

Ritonavir (brand name: Norvir), invented by Abbott Laboratories, was originally designed as an HIV-1 protease inhibitor, but it was found later that ritonavir boosts the circulating concentration of other HIV-1 protease inhibitors by inhibiting the cytochrome P450 (*18*). Due to this function of ritonavir, it is still used with the second generation of HIV-1 protease inhibitors. Boosted HIV-1 protease inhibitors elevate the drug resistance barrier. With the isopropyl thiazolyl P3 group, ritonavir is longer than other FDA-approved HIV-1 protease inhibitors (Figure 1.2).

1.6.3 Lopinavir

Lopinavir (brand name: Kaletra, containing lopinavir and ritonavir), from Abbott Laboratories, was developed on the basis of ritonavir. The core region of lopinavir, a hydroxyethylene dipeptide isostere, is the same as that of ritonavir (Figure 1.2). The P2

and P2’ group are altered in lopinavir relative to ritonavir. The 5-thiazolyl P2 group of ritonavir is changed to a phenoxyacetyl group, and the 2-isopropylthiazolyl P2’ group of ritonavir is replaced by a six-member cyclic urea. In general, the new P2 and P2’ groups are smaller, which is designed to decrease the contact with highly variable residues at the 82 site of HIV-1 protease (*19*). The substitution of the P2 and P2’ groups improves the inhibitory potency of lopinavir against the drug-resistant variants of HIV-1 protease (*19*).

1.6.4 Indinavir

Indinavir (brand name: Crixivan) was developed by Merck & Co., Inc. An advantage of indinavir is that it effectively inhibits both HIV-1 and HIV-2 while the disadvantage is that the circulating indinavir concentration decreases quickly. The low plasma concentration of indinavir usually leads to treatment failures (*20*). In addition, indinavir has a low sulibility that may result in the patients developing kidney stones. For this reason, indinavir has been replaced by second generation protease inhibitors.

1.6.5 Nelfinavir

Nelfinavir (brand name: Viracept) was made by Agouron Pharmaceuticals. One end of the nelfinavir molecule has the same DIQ group as saquinavir (Figure 1.2). The other end of nelfinavir contains a 2-methyl-3-hydroxybenzamide group. Another novel aspect of the nelfinavir design is that the S-phenyl group at the P1 site magnifies the potency of this inhibitor.

1.6.6 Amprenavir

Amprenavir (brand name: Agenerase), developed by Vertex Pharmaceuticals, has a

benzyl group at the P1 site and an isobutyl group at the P1' site (Figure 1.2). The P1' group and the phenyl amide P2' group are connected by a sulfonamide. The asymmetry of the P1 and P1' groups may favor the internal pseudosymmetry of HIV-1 protease. Amprenavir contains fewer chiral centers than previous HIV-1 protease inhibitors. This improvement simplifies the chemical synthesis and increases the oral availability (*21*).

1.6.7 Tipranavir

Tipranavir (brand name: Aptivus) is the only non-peptidomimetic HIV-1 protease inhibitor, which received approval from the FDA in 2005. Due to dissimilarities with other HIV-1 protease inhibitors, tipranavir encounters a relatively different HIV-1 protease resistance profile (*22, 23*). An innovation in the tipranavir design is the functional substitution of the bridging water molecule connecting the inhibitor and protease flaps. The lactone oxygen atom of the dihydropyrone ring of tipranavir interacts directly with the Ile 50 residues in the flap region of the HIV-1 protease. The direct interaction stabilizes the protease-inhibitor complex. Though the structure of tipranavir is different from previous inhibitors, its contact residues of HIV-1 protease share many similarities with other HIV-1 protease inhibitors. Furthermore, the adverse effects of tipranavir are considerable, with one of the worst being hemorrhage (*24*).

1.6.8 Atazanavir

Atazanavir (brand name: Reyataz), approved in 2003, is an aza-dipeptide analogue, which exhibits potent anti-HIV activity. The 50% effective concentration (EC_{50}) of atazanavir in cell culture is 2.6 to 5.3 nM (*25*). A unique character of atazanavir is the presence of a large phenylpyridyl P1 group that is asymmetric relative to its benzyl P1'

group.

1.6.9 Darunavir

Darunavir, approved in 2006, is the latest HIV-1 protease inhibitor on the market. The hydrogen bonds that darunavir forms with the backbone of the HIV-1 protease slow down the development of drug resistance (*26*). The structure of darunavir is very similar to that of amprenavir. The only difference is that the P2 group of darunavir is a double-ringed bis-tetrahydrofuran that replaces the single-ringed tetrahydrofuran of amprenavir. This change allows darunavir to have more hydrogen bonds with the Asp 29 residues of HIV-1 protease (*27*). Darunavir or atazanavir coadministrated with ritonavir is a current first-line antiretroviral regimen (*28*).

1.6.10 The common mechanism of inhibitor binding

Most current HIV-1 protease inhibitors were designed to mimic the substrate transition state. The hydroxyl group of the inhibitor interacts with the carboxyl group of the active site residues, Asp 25 and Asp 25´, by hydrogen bonds. Moreover, the inhibitor-contacting residues of HIV-1 protease are relatively conserved, such as Gly 27, Asp 29, Asp 30, and Gly 48. Despite the fact that diverse chemical groups have been incorporated in HIV-1 protease inhibitors, the conserved contact profile of the chemical groups still makes the protease inhibitors vulnerable to prevent drug resistance.

Most inhibitor development efforts are to modify the available HIV-1 protease inhibitors, which inevitably causes the new inhibitors to inherit structural similarity with previous inhibitors. There have been many attempts to further modify current inhibitors; some have succeeded, while others have not shown significant benefits. A series of

compounds has been tested to optimize the P1 group of lopinavir, but none of them has shown significant benefits over the original lopinavir (*29*). The successful examples of modified inhibitors are darunavir and lopinavir, which were modified compounds of amprenavir and ritonavir, respectively.

1.6.11 New design approaches for HIV-1 protease inhibitors

Resistance has been observed with all the current HIV-1 protease inhibitors due to their structural similarities. Therefore, novel scaffolds of HIV-1 protease inhibitors are considered to be the key to overcoming drug resistance. With the combination therapy of structurally diverse HIV-1 protease inhibitors, HIV-1 protease is less likely to develop resistance to the inhibitors in the combination group.

Designing inhibitors based on the substrate envelope of HIV-1 protease is a recent strategy to develop HIV-1 protease inhibitors. A substrate envelope is the overlapping volume occupied by multiple substrates. Based on this theory, a substrate-based HIV-1 protease inhibitor, RO1, has been designed and synthesized (*30*). The results demonstrated that RO1 has 5 to 10 fold higher inhibitory efficacy than current HIV-1 protease inhibitors (*30*). Among the current FDA-approved HIV-1 protease inhibitors, only tipranavir forms hydrogen bonds with the flap residues of the protease. Inhibitor RO1 forms more hydrogen bonds with HIV-1 protease, which enhances the binding energy and is also a step towards the diversity of binding patterns. However, the contact residues of RO1 are still the old pattern, and no novel contact has been established. The difficulty of complete alteration of the binding pattern might be due to the inhibitor design concept of substrate mimics. To develop new scaffolds of HIV-1 protease

inhibitors, it is necessary to adopt drug discovery strategies other than substrate mimics.

New scaffolds of HIV-1 protease inhibitors might be the key to overcoming drug resistance. It is very risky to explore new structures of inhibitors. One reason is that the new scaffolds are less likely to be superior to the well-established old scaffolds. Mozenavir's structure is highly dissimilar from other HIV-1 protease inhibitors, which was developed based on the cyclic urea scaffold (*31*). However, the clinical studies showed that mozenavir did not show significant advantages compared to other protease inhibitors on the market, and Gilead Sciences, Inc. stopped the development of mozenavir (*32*). The other reason is that new scaffolds with fewer study backgrounds may cause severe adverse effects. The relatively different structure of tipranavir may be not only the reason for generating altered drug-resistance profiles, but it is also the cause of the unpredictable side effects, such as hemorrhage. Based on the above summary, the non-peptidomimetic HIV-1 protease inhibitors seem to have a higher tendency to cause adverse effects.

The discovery of new HIV-1 protease scaffolds may rely on high-throughput screening techniques. By screening large chemical libraries with diverse compound structures, the leads of new HIV protease inhibitor scaffolds could be found. After extensive modifications and tests, the leads of new scaffolds might become the next generation of HIV-1 protease inhibitors possessing novel chemical structures and alternative binding patterns to HIV-1 protease. A set of HIV-1 protease inhibitors that bind and function by various mechanisms can greatly decrease the probability for the protease to simultaneously develop resistance to all inhibitors.

1.7 Protein-ligand interactions

The flexibility of the protein is important in the affinity and specificity of ligand binding as well as product release. Optimal interactions between the enzyme and the substrate are formed through induced conformational changes (*33*). The major interactions between the protein and the ligand are electrostatic interactions, hydrogen bonding, hydrophobic interactions, and van der Waals interactions. An electrostatic interaction is formed when oppositely charged groups interact at a distance less than 4 Å. (*34*). The charged state of some amino acid residues determines the local electrostatic potential of the protein, which is an important component of molecular recognition. The hydrogen bonds are interactions between a partially electronegative atom and a hydrogen atom covalently bound with another partially electronegative atom. Both the main chain of a protein and the side chain of polar amino acid residues can serve as hydrogen bond donors and acceptors. In contrast with hydrophilic interactions, such as electrostatic interaction and hydrogen bonding, which are usually on the surface of a protein, hydrophobic interactions assemble the core of a protein during folding. Van der Waals interactions represent attractive or repulsive force between proximate atoms. Despite the fact that van der Waals interactions are weak, the considerable number of van der Waals interactions determines the important role of this interaction in macromolecules, such as protein folding, protein-protein interactions, and protein-ligand interactions.

1.7.1 Substrate binding to the HIV-1 protease

Understanding substrate binding is an important step in drug design. Based on the MA/CA substrate cleavage site (Ser-Gln-Asn-Tyr-Pro-Ile-Val), the P1’ Pro substitution

with 20 different residues has been tested to identify the preferred P1’ residues of HIV-1 protease, and hydrophobic residues, such as Leu, Ile, Val, Phe, and Tyr, are favorable at the P1’ site (*35*).

The substrate envelope represents the conserved volume occupied by the nine HIV-1 protease substrates. Designing inhibitors that fit into the overlapping volume of substrates can avoid the contacts with mutation hotspots of the protease (*36*). Based on this theory, the inhibitors designed to fit inside the consensus volume defined by the substrates exhibited potent inhibition against HIV-1 protease and low levels of resistance to a panel of drug-resistant HIV-1 protease variants (*37*).

The most widely accepted mechanism for the hydrolysis of HIV-1 protease substrates is the general acid and base mechanism (*38*) (Figure 1.3). In the active site of the HIV-1 protease, one aspartate residue is protonated while the other aspartate residue is deprotonated (*39*). During cleavage, the protonated aspartate residue transfers the hydrogen atom to the scissile carbonyl oxygen (*40*). A water molecule in the active site attacks the carbonyl group of the scissile peptide bond of the substrate and the cleavage site of the substrate forms a tetrahedral oxyanion intermediate (*41*).

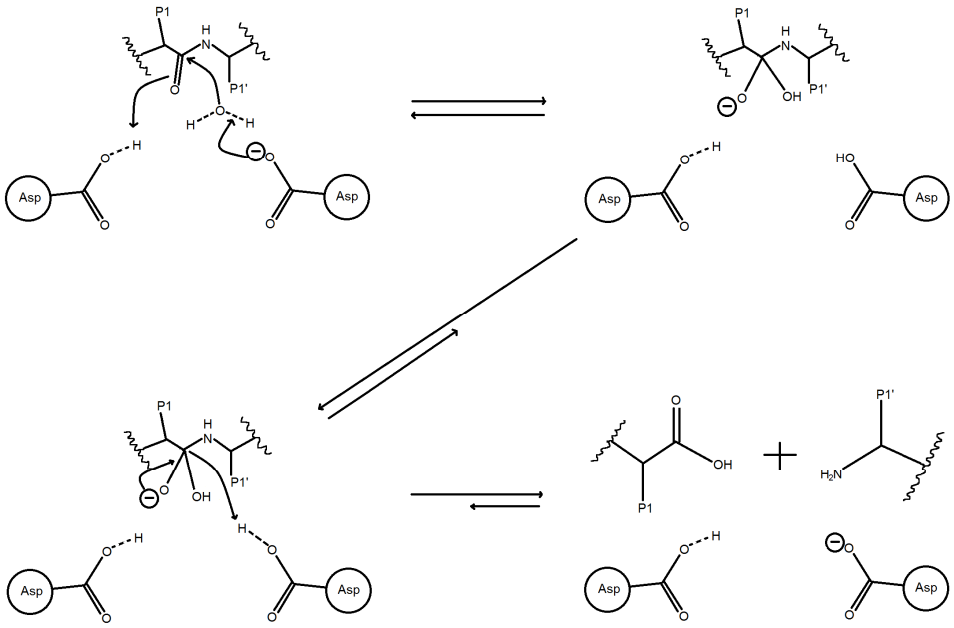


Figure 1.5 HIV-1 protease substrate cleavage mechanism.

The Gly-rich flaps of HIV-1 protease alter the conformation from an “open” form to a “closed” form after ligand binding. NMR studies have demonstrated the flexibility of the flap region of the HIV-1 protease (*42*). The flexibility of the flaps is considered to be essential for substrate binding and product release. After the substrate is incorporated into the protease actives site, substrate binding stabilizes the interactions between the flaps of HIV-1 protease. A water molecule bridges the flap residue Ile50 and Ile50’ with hydrogen bonds, and this water molecule is not involved in catalysis (*40*).

1.8 Summary of the HIV-1 protease structures in the dissertation

The HIV protease structures mentioned in the dissertation are listed in the following table. The residues involved in crystal packing were identified using PDBePISA (Protein Interfaces, Surfaces and Assemblies) server.

Table 1.1 Summary of the HIV-1 protease structures

PDB code	Description	Space Group	Unit cell		Residues involved in crystal packing
			Length (Å)	Angle (°)	
2HB4	Apo WT HIV-1 protease isolate NL4-3	$P\ 4_12_12$	a=48.99 b=48.99 c=105.62	α=90 β=90 γ=90	4, 6-7, 35, 37, 41-42, 44, 46, 49-57, 61, 63, 71-73, 77-79, 81, 88, 91-94
2HB2	Apo MDR HIV-1 protease variants with mutations 24I, 46I, 53L, 63P, 77I, and 82A	$P\ 4_12_12$	a=49.67 b=49.67 c=109.61	α=90 β=90 γ=90	4, 6-7, 35, 42, 50-57, 61, 71-73, 77-79, 81, 88, 91-94
1HHP	Apo WT HIV-1 protease isolate BRU	$P\ 4_12_12$	a=50.2 b=50.2 c=106.8	α=90 β=90 γ=90	4, 6-7, 35, 37, 42, 49-57, 61, 63, 71-73, 77-79, 81, 88, 91-94
2PC0	Apo WT HIV-1 protease isolate BH10	$P\ 4_12_12$	a=46.42 b=46.42 c=101.37	α=90 β=90 γ=90	4, 6-7, 39-41, 46, 48-55, 60-61, 63, 71-74, 79, 81, 88, 91-94
1TW7	Apo MDR HIV-1 protease 769 with mutations 10I, 36V, 46L, 54V, 62V, 63P, 71V, 82A, 84V, and 90M	$P\ 4_1$	a=45.04 b=45.04 c=105.74	α=90 β=90 γ=90	4, 6-7, 34, 37-44, 48-61, 63, 71-74, 77-81, 88, 91-94
2Q5K	WT HIV-1 protease in complex with lopinavir	$P\ 2_12_12_1$	a=51.34 b=58.04 c=61.13	α=90 β=90 γ=90	1-4, 6-7, 12-19, 34-46, 48-53, 54-55, 57-58, 59, 61-63, 65, 67-72, 77-81, 88, 91-94, 96, 98-99
2O4P	WT HIV-1 protease in complex with tipranavir	$P\ 2_12_12$	a=58.61 b=86.26 c=45.89	α=90 β=90 γ=90	1-2, 4, 6-7, 12, 14, 17-21, 34-46, 48-49, 51-53, 55, 57-58, 61, 63, 65, 67-68, 70-73, 79-83, 88, 91-92, 94, 99
3SPK	MDR HIV-1 protease 769 in complex with tipranavir	$P\ 6_1$	a=63.12 b=63.12 c=83.54	α=90 β=90 γ=120	1-4, 6-8, 10, 12-14, 16-21, 34-35, 37-41, 43-48, 51-53, 58, 65, 67-68, 74, 79-81, 94, 96, 98
3BVB	WT HIV-1 protease in complex with darunavir	$P\ 2_12_12$	a=58.26 b=85.91 c=46.05	α=90 β=90 γ=90	2, 4, 6-7, 12, 14, 17-20, 34-37, 40-46, 48-53, 55, 57-58, 61, 63, 65, 67-68, 70-73, 79-83, 88, 91-92, 94, 96
3SO9	MDR HIV-1 protease 769 in complex with darunavir	$P\ 2_12_12_1$	a=28.9 b=66.38 c=90.84	α=90 β=90 γ=90	1-2, 4, 6, 14-18, 21, 34-46, 48-58, 60-65, 70, 72, 77-82, 87, 92-94, 96, 98-99
1A8K	WT HIV-1 protease in complex with CA/p2 P1'F peptide	$P\ 112_1$	a=59.6 b=52.2 c=61.7	α=90 β=90 γ=90	2, 4, 6, 10-12, 14, 19, 21, 34-35, 41-46, 52-55, 57, 63, 65, 67-68, 70-72, 79-81, 92
N/A	MDR HIV-1 protease 769 in complex with CA/p2 P1'F peptide	$P\ 2_12_12_1$	a=28.77 b=65.39 c=92.82	α=90 β=90 γ=90	1-2, 4, 6, 14-19, 21, 29, 34-37, 39-46, 49-58, 60-61, 63-65, 67-68, 70-72, 77-82, 87, 91-94, 96, 98-99
1KJ7	WT HIV-1 protease in complex with p2/NC peptide	$P\ 2_12_12_1$	a=51.51 b=59.43 c=61.74	α=90 β=90 γ=90	1-4, 6-7, 12-14, 17, 19, 34-35, 37, 41-46, 48-55, 61, 63, 65, 67-72, 77-81, 88, 91-94, 96
N/A	MDR HIV-1 protease 769 in complex with p2/NC peptide	$P\ 2_12_12_1$	a=28.62 b=63.85 c=91.11	α=90 β=90 γ=90	1-2, 4, 6, 14-17, 19, 21, 34-44, 46-58, 59-61, 63-65, 67-72, 74, 77-82, 87, 91-94, 96, 98-99

1.9 Hepatitis C virus and treatment

The hepatitis C virus (HCV), a member of the *Flaviviridae* family, is a single-stranded positive sense RNA virus. It is the major risk factor for liver cirrhosis and hepatocellular caricinoma and the virus infects approximately 170 million people worldwide, four million of which are in the United States (*43, 44*). The classic treatment for HCV infection is a combination of pegylated interferon and ribavirin. Both antiviral drugs function indirectly since neither of them targets a specific HCV protein or RNA. The pegylated interferon and ribavirin combination treatment requires a weekly injection for 24 or 48 weeks, yet only less than 50% patients achieve a sustained virological response. Furthermore, there are a number of serious adverse effects associated with interferon treatment, such as leucopenia, thrombocytopenia, neutropenia, depression, fatigue, and "flu-like" symptoms, that makes withstanding the long treatment period difficult for patients (*45*). Consequently, there is a need for safer and more effective drugs for the treatment of HCV.

1.10 HCV life cycle

The HCV infection starts with the virus binding to the cell surface receptor. Several putative receptor molecules have been proposed, such as the cluster of differentiation 81 (CD81) (*46*), scavenger receptor B type I (SR-BI) (*47*), dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin (DC-SIGN or CD209) (*48*), liver/lymph node-specific intercellular adhesion molecule-3 (ICAM-3)-grabbing integrin (L-SIGN or CD209L) (*49*), low-density lipoprotein receptor (LDL-R) (*50*), and asialoglycoprotein receptor (ASGP-R) (*51*). Due to the narrow host specificity, HCV only

infects humans and chimpanzees. Genetically humanized mice expressing human CD81 and occludin are able to establish HCV infection (*52*). L-SIGN molecules abundantly expressed on liver sinusoidal endothelial cells play a crucial role in HCV tissue tropism (*49*).

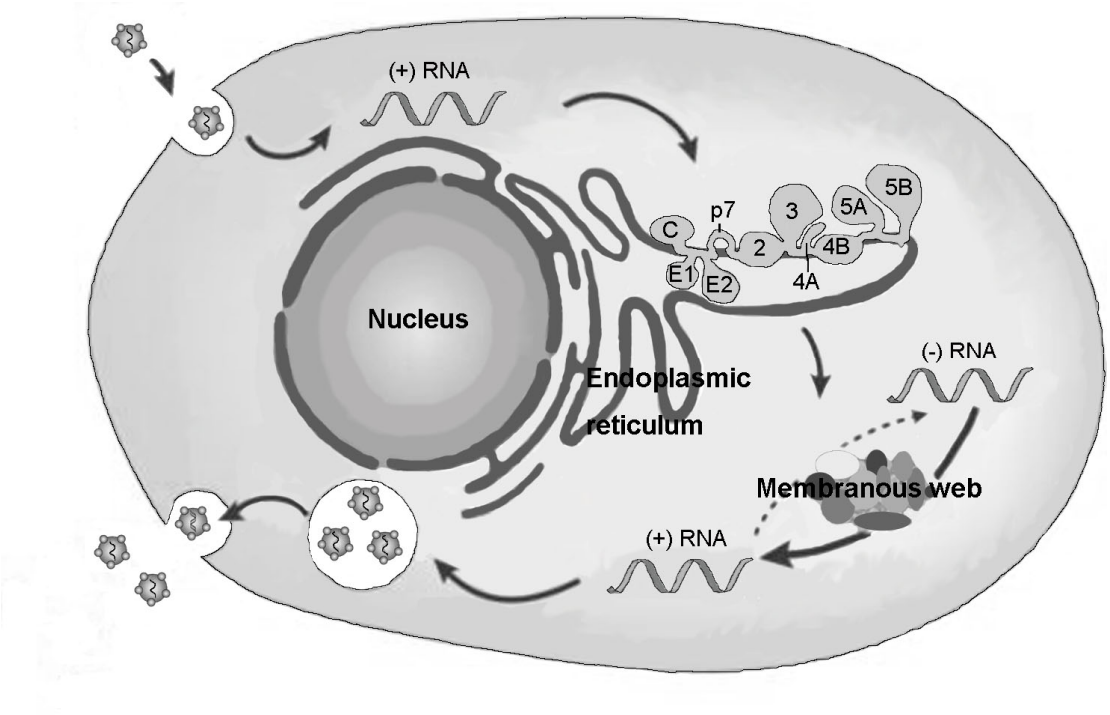


Figure 1.6 HCV life cycle. Adapted from the figure produced by Moradpour D, *et al* (*53*)

After attaching to a host receptor, the viral surface glycoproteins trigger receptor-mediated endocytosis. By uncoating the capsid shell of the virion, positive-strand genomic RNAs are released from the virus, and the RNAs are attached to the endoplasmic reticulum. A complementary negative-strand RNA is synthesized based on the viral RNA. The negative-strand RNA is then utilized as a template to produce thousands of positive-strand viral RNA copies by the viral polymerase NS5B. The positive-strand RNA is used for polyprotein translation. The internal ribosome entry site (IRES) of HCV RNA regulates the assembly of translation initiation complexes by

recruiting both cellular proteins and viral proteins (*54*). Then, a large precursor polyprotein with approximately 3,000 amino acid residues is synthesized using the viral genomic RNA. A total of ten proteins are produced by the cleavage of the polyprotein. The structural proteins are released by host enzymes while the non-structural proteins are cleaved off by the viral enzyme. The zinc-dependent NS2-3 autoprotease liberates the NS3 by self-cleavage. The efficient NS2-3 proteinase activity requires the last 130 residues of NS2 and the first 180 amino acids of NS3 (*55*). The NS3 assembles with its cofactor NS4A and performs self-cleavage at the NS3-NS4A junction to liberate NS4A. The NS4A-NS4B, NS4B-NS5A, and NS5A-NS5B junctions are cleaved by NS3/4A protease (*56*). A host peptidase cleaves the E1-E2 junction. The envelope proteins, E1 and E2, form a heterodimer after N-glycosylation (*57*).

The core proteins are self-assembled and bind the viral RNA to generate the nucleocapsid. The glycoproteins E1 and E2 associated with the double lipid layers form an envelope to cover the nucleocapsid. The newly formed viral particles are secreted from the host cell to initiate a new infectious cycle.

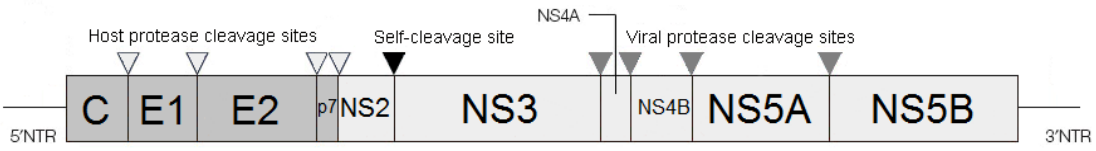


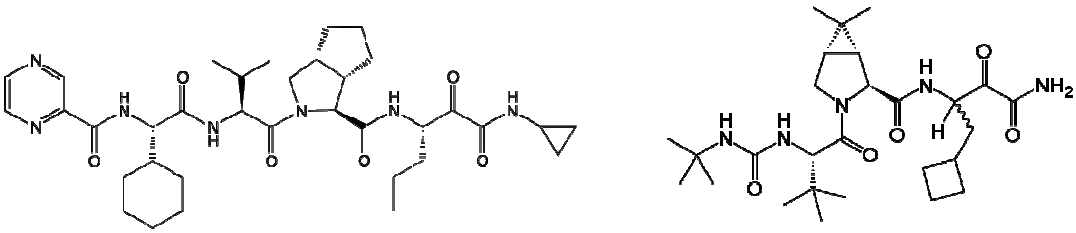
Figure 1.7 The HCV genome. Adapted from the figure produced by Tan SL, *et al*.(*58*)

1.11 The HCV protease

The HCV protease emerges as a desirable drug target, because it is essential for viral maturation. The single polyprotein encoded by the HCV genome is proteolytically

processed by the HCV NS3/4A serine protease which is responsible for the cleavage at four junctions of the HCV polyprotein including NS3/4A (self-cleavage), NS4A/4B, NS4B/5A, and NS5A/5B (*59-61*).

The extensive functional studies of NS3/4A protease and the solved crystal structures have improved our understanding of this protease. NS3 is a bifunctional enzyme; the protease domain is located at the N-terminus and a helicase domain at the C-terminus. The protease and helicase domains function independently (*62, 63*). NS4A, a 54-residue peptide, is required as a co-factor to increase the activity of the NS3 protease (*64, 65*); this may be due to the incorporation of NS4A into the β sheet at the N-terminus of the NS3 protease with tight non-covalent bonds (*66*). Another possible function of NS4A is to anchor NS3 to the membrane (*67*).



telaprevir boceprevir

Figure 1.8 Chemical structures of HCV protease inhibitors.

1.12 The HCV protease inhibitor, telaprevir

Currently, two HCV protease inhibitors have been approved by the US FDA, telaprevir (VX-950) and boceprevir. Telaprevir is an oral drug approved in May 2011 which is a potent inhibitor of the genotype 1 HCV protease. Telaprevir was developed by Vertex Pharmaceuticals as a peptidomimetic inhibitor, which is derived from the N-terminal product of the NS4B/5A. Telaprevir inhibits the NS3/4A protease by binding

to the active site with its α-ketoamide moiety. The crystal structure of the NS3 protease with the NS4A peptide or NS3/4A fusion from HCV genotypes 1a and 1b have been solved (*66, 68*). The shallow substrate-binding groove of the protease increases the difficulty of inhibitor binding, but α-ketoamide inhibitors exhibited an unusual binding mechanism in the catalytic site of the protease. By optimization of α-ketoamide as a serine trap group to covalently anchor inhibitors and to increase their affinity to the protease, telaprevir, a potent, covalent, reversible, and selective inhibitor of HCV NS3/4A protease, was developed (*69*). The covalent bond between telaprevir and the protease is hydrolysable, which makes telaprevir a reversible inhibitor.

After a treatment with telaprevir for 2 to 3 days, the viral load in patients with genotype 1 chronic hepatitis C was reduced by hundred to thousand folds (*70*). At the end of a 14-day treatment with Telaprevir, up to a 10,000-fold reduction in HCV viral load was observed, while in some patients the virus became undetectable (*70*). Telaprevir dramatically shortened the duration of therapy, and fewer adverse effects were noted (*70, 71*). Therefore, with such a drug, the efficacy and tolerability of HCV treatment could be greatly enhanced. Furthermore, telaprevir cured the genotype 1 hepatitis C patients who had failed to achieve a sustained virological response with interferon-based therapy.

At least six major HCV genotypes have been identified with varied geographic distributions. Each genotype is further divided into subtypes. Genotype 1a HCV is prevalent in North and South America; genotype 1b, in Europe and Asia; genotype 3, in Australia and South Asia; genotype 4, in Egypt, France and Central Africa. Genotype 2 exists in Japan, China, the United States, and Europe. Genotype 5 is observed in South

Africa and Saudi Arabia. Genotype 6 is restricted to Hong Kong, Macau, and Vietnam.

The various HCV genotypes respond differently to treatments. For example, patients infected with genotype 2 HCV showed better response to the interferon therapy relative to the other genotypes (*72*). Although telaprevir has demonstrated strong antiviral activity in patients with chronic genotype 1 HCV infection in the clinical trial, its efficacies with the HCV genotypes 4, 5, and 6 have not been fully investigated. The sequence similarities between NS3 protease of genotype 1 and NS3 protease of genotype 4, 5, and 6 vary approximately 60% to 80%. Due to the relatively low sequence similarity between the HCV genotype 1 protease and proteases from other HCV genotypes, the inhibitory efficacies of non-genotype 1 proteases is likely affected. Enzyme assays demonstrate that the potency of telaprevir against the HCV NS3/4A proteases of genotype 2 is similar as that of telaprevir against the genotype 1 HCV proteases, but telaprevir exhibited reduced activity against the genotype 3 HCV proteases (*73*). Interestingly, the protein sequence similarities between the genotype 1 and genotype 2 proteases are less similar compared to the protease sequences between genotypes 1 and 3. Therefore, structural information is critical to explain the varied potency of telaprevir against non-genotype 1 HCV proteases. A previous study showed the mutations V36A/G/L/M and T54A/S in NS3/4A protease from HCV genotype 1 strain were associated with low to medium levels of telaprevir resistance (*74*). The mutation V36L exists in genotype 4a, 4d, 5a, and 6f HCV protease sequences. Thus, the comparison of telaprevir binding between the HCV genotype 1 protease and proteases from other genotypes will provide an insight of how to modify the inhibitor to restore the enzyme inhibition.

CHAPTER 2 METHODS

2.1 Synthesis of lopinavir analog

The synthesis of the lopinavir analog (Figure 2.1) was modified from the lopinavir synthesis developed by Abbott Laboratories (*75*). The amount of reagents described in the following sections was optimized for lopinavir analog synthesis.

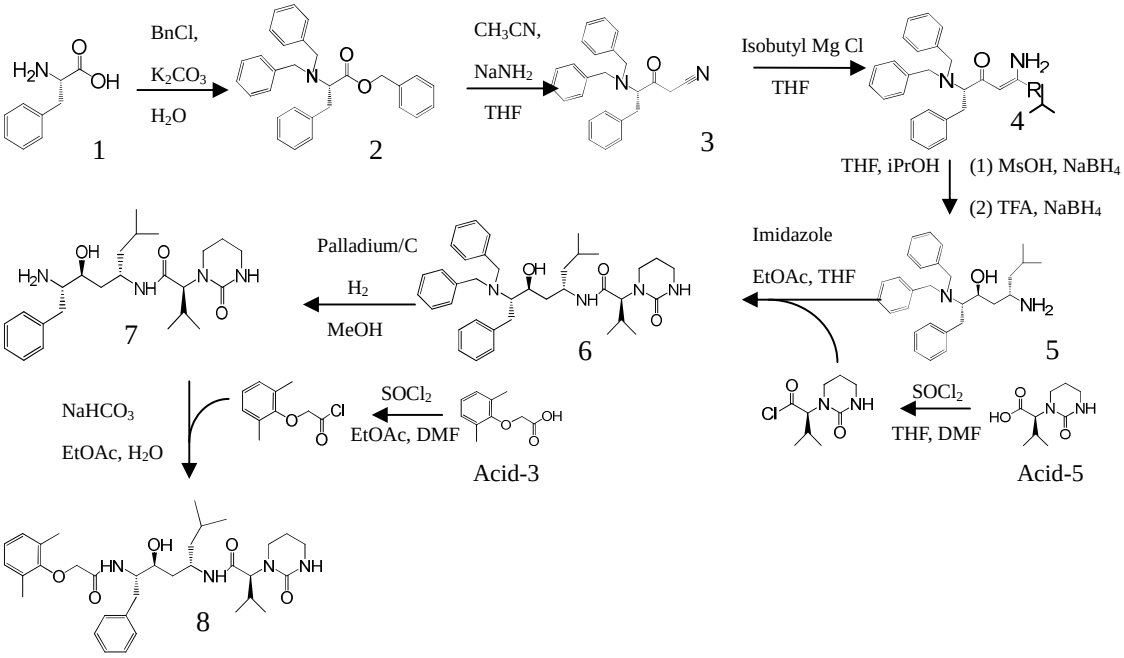


Figure 2.1 Synthetic scheme for the lopinavir analog.

2.1.1 Synthesis of (L)-N,N-dibenzylphenylalanine benzyl ester (compound 2)

In a round flask, 15 g L-phenylalanine (compound 1) and 49.8 g K_2CO_3 were dissolved in water, and 44 ml benzyl bromide was then added to the aqueous solution. The solution in the flask was boiled overnight. The reaction was stopped by cooling down the flask to room temperature. Ethyl acetate and water were used to extract the organic

phase from the product mixture. The benzyl ester in the organic phase was further washed with saturated sodium chloride solution. The organic phase was dried *in vacuo* to yield a thick yellow oil. The product was purified by a silica gel column chromatography (ethyl acetate:n-hexane = 1:4 v/v) to yield 7.1 g benzyl ester (compound 2, 18% yield).

2.1.2 Synthesis of (4S)-4-dibenzylamino-3-oxo-5-phenyl-pentanonitrile (compound 3)

7.1 g benzyl ester dissolved in 19 ml anhydrous tetrahydrofuran (THF) was cooled to -45°C with an acetone solution mixed with dry ice. 4.26 g sodium amide was dissolved in 47 ml anhydrous THF in a separate flask and cooled to -45 °C. The air in both sealed flasks was replaced by nitrogen. 6.14 ml acetonitrile was added to the sodium amide solution and reacted for 30 min to generate an anion solution. The cold anion solution was transferred to the benzyl ester solution by syringe. The mixture was stirred at -45 °C for 7 hr. The progress of the reaction was monitored by sampling and examination using thin layer chromatography (TLC). The Rf is 0.55 (ethyl acetate:n-hexane = 1:3 v/v). The reaction was quenched with 50 ml 25% (w/v) aqueous citric acid. The organic phase was extracted, washed with saturated sodium chloride solution, and dried with magnesium sulfate. The crude product was dissolved in boiling ethanol. As the solution was cooled down resulting in 4.5 g white product (74% yield per reaction).

2.1.3 Grignard reaction

3.8 g of magnesium turnings were cut and weighted in a dry two-funnel flask that was connected to a condensation tower. The reaction system was dried thoroughly by heating and filled with argon. After cooling the reaction system with ice, 17.7 ml of 1-chloro-2-methyl propane was added to the flask. Iodine powder (10 mg) and a few

drops of 1,2-dibromoethane were added to the solution mixture as an activating agent. The reaction was stirred at room temperature for 30 min. The boiling of this solution indicated the generation of the Grignard reagent, isobutyl magnesium chloride. In a separate flask, 4.5 g of compound 3 was dissolved with anhydrous THF and transferred to the flask containing Grignard reagent. The reaction mixture was stirred over night at room temperature. The reaction was stopped by addition of 10% (w/v) citric acid. The product in the organic phase was extracted and purified by silica gel column chromatography with (ethyl acetate:n-hexane = 1:20 v/v, Rf 0.39), which yielded 2.5 g of enaminone (compound 4, 47% yield).

2.1.4 Reduction reaction

A two-step reduction reaction generated the largest percentage of desired enantiomer. 1 g sodium borohydride in 54 ml anhydrous THF was cooled to 0°C before 4.3 ml methansulfonic acid was slowly added to the suspension. The slow addition was to ensure that the temperature of the suspension remained below 5°C. A solution of enaminone (2 g) dissolved in THF and 5 ml isopropanol was added the previous suspension. The mixture was stirred at 4°C for 28 hr. In the second step of reduction, 5 ml trifluoroacetic acid was slowly added to 2 g sodium borohydride in 26 ml THF to ensure a temperature below 15°C. The suspension was stirred at room temperature for 4 hr. The reaction was quenched with 3 N sodium hydroxide. By adding tert-butyl methyl ether, the organic layer was separated and washed with 0.5 N sodium hydroxide, 20% (w/v) ammonium chloride, and 6% (w/v) sodium chloride.

2.1.5 Coupling of Acid-5

In a flask filled with nitrogen gas, 0.7 g Acid-5, purchased from Biofine International Inc. (Vancouver, Canada), was dissolved in 9.7 ml THF and cooled to 4°C. 0.35 ml thionyl chloride was added dropwise to the Acid-5. The mixture was stirred for 5.5 hr until it became a thick slurry. The slurry was dried *in vacuo*, dissolved in heptane, and further dried to thoroughly evaporate the residual thionyl chloride. The dried product, acyl chloride, was dissolved in 3.45 ml dimethylformamide. 1.5 g of product 5 was added to 0.96 ml ethyl acetate containing 0.72 g dissolved imidazole. The solution was rapidly transferred to the acyl chloride and stirred for 1 hr at 0°C followed by overnight stirring at 30°C. The reaction was quenched with an aqueous hydrochloric acid solution by adjusting the aqueous layer to pH 3. The aqueous layer contained unreacted Acid-5, and the organic layer was extracted for further clean up. The crude product was loaded on a silica column. The impurity was eluted with a mobile phase of 90% hexane, 10% isopropyl alcohol, and 0.1% ammonium hydroxide. The product was eluted with 89% dichloromethane, 10% methanol, and 1% ammonium hydroxide.

2.1.6 Deprotection and Acid-3 coupling

Two benzyl protective groups from compound 6 were removed by catalytic hydrogenation with palladium on carbon. The reaction was under 65~75 psi hydrogen pressure for two days. 0.076 ml thionyl chloride was added to 0.167 g Acid-3 (Biofine International Inc.,Vancouver, Canada) dissolved in 0.32 ml ethyl acetate and 0.1 ml dimethylformamide and stirred for 5 hr at 50°C. The activation of Acid-3 was completed under nitrogen atmosphere. 0.25 g deprotected compound 7 and 0.28 g sodium

bicarbonate were dissolved in ethyl acetate/water (1:1, v/v). The activated Acid-3 was transferred to the compound 7 by syringe and stirred at room temperature for 1 hr. The final compound 8 in the organic layer was washed with 5% (w/v) sodium bicarbonate solution and water sequentially. The dry final product (compound 8) was 0.2 g. The products of each reaction were confirmed by mass spectrometry.

2.2 Protein expression and purification

Table 6.1 lists the protein sequences of MDR HIV-1 proteases. Active and inactive MDR HIV-1 protease genes were codon optimized for *E. coli* expression with the software DNA 2.0 (*76*), synthesized by GENEART, Inc. (Regensburg, Germany), and inserted into the pET21b plasmid. The inactive MDR 769 82T protease had an active site mutation, D25N, to eliminate catalytic activity. To prevent auto-proteolyses, the Q7K mutation was introduced into the active MDR HIV-1 protease gene. All MDR HIV-1 proteases were expressed in the *E. coli* strain BL21 (DE3) pLysE. *E. coli* cells were lysed by sonication. The HIV-1 protease expressed in inclusion bodies was separated by centrifugation and dissolved in a buffer of 50 mM Tris, 25 mM sodium chloride, 0.2% β-mercaptoethanol, and 6 M urea. The dissolved HIV-1 protease variants were purified using a Q-Sepharose ion exchange column (Amersham Biosciences, Piscataway, NJ) pre-equilibrated with the buffer containing 6 M urea. The optimal buffer pH varied from 7.8-8.6 for each of the protease variants. The flow-through containing the pure HIV-1 protease was refolded in the dialysis buffer (20 mM sodium phosphate, 0.2% β-mercaptoethanol, and 10% glycerol) with step-wise decreases in urea concentration. The final protein buffer was 20 mM sodium acetate, pH 5.0, 1 mM dithiothreitol (DTT),

and 10% (v/v) glycerol. The proteases prepared for crystallization were concentrated to 1.5 mg/ml using Amicon concentrators with a 5-kDa molecular mass cut-off (Millipore Corporate, Billerica, MA).

2.3 Protease inhibition assays

HIV protease Forster Resonance Energy Transfer (FRET) substrate I, purchased from AnaSpec, Inc. (Fremont, CA, USA), was used in the half-maximal inhibitory concentration (IC_{50}) determination experiments. The FDA-approved HIV-1 protease inhibitors were kindly provided by the NIH AIDS Research and Reference Reagent Program (www.aidsreagent.org). The CA/p2 pseudo peptide with a reduced scissile peptide bond [-ψ(CH2NH)-] was synthesized in the Department of Chemistry, Wayne State University. The fluorescence emitted by substrate cleavage was monitored by a microplate reader (SpectraMax M5, Molecular Devices, Sunnyvale, CA) at a 340 nm excitation wavelength with an emission wavelength of 490 nm. The HIV-1 protease reaction buffer was adjusted to pH 4.7 [0.1 M sodium acetate, 1.0 M sodium chloride, 1.0 mM ethylenediaminetetraacetic acid (EDTA), 1.0 mM DTT, 10% dimethylsulfoxide (DMSO), and 1mg/ml bovine serum albumin (BSA)]. In the reaction buffer containing 5 μM FRET substrates, the concentration of all the HIV-1 proteases used in the enzyme assays was adjusted to a substrate cleavage velocity of 0.08 Relative Fluorescence Units (RFU)/sec. The final HIV-1 protease concentration was approximately 7 nM. The protease inhibitor was serially diluted in DMSO from 10 μM to 0.013 nM. The protease variants and inhibitors were pre-incubated at 37°C for 20 min prior to signal monitoring. An enzyme-free control was tested as the background substrate fluorescent signal. The

progress of the reaction was monitored over 20 min sampling fluorescence at 1 min intervals. The FRET data were plotted with the software SoftMax Pro V5.2 (Molecular Devices, Sunnyvale, CA) to determine the IC_{50} values. The difference in Gibbs free energy ($\Delta\Delta G$) was calculated from the formula $\Delta\Delta G = -RT \ln[IC_{50}(MDR)/IC_{50}(WT)]$, where the $IC_{50}(MDR)$ is the IC_{50} value inhibiting the MDR HIV-1 protease and the $IC_{50}(WT)$ is the IC_{50} value inhibiting the wild-type (WT) HIV-1 protease.

2.4 HIV-1 Protease substrate interference assays

In the enzyme reaction buffer, a regular HIV-1 protease substrate, a heptapeptide, was mixed with the FRET HIV-1 protease substrate (AnaSpec Inc., Fremont, CA) to reach a final concentration of 100 μM regular substrate and 2.5 μM FRET substrate. The nine substrate hepta-peptides were purchase from SynBioSci Corporation, Livermore, CA (Table 6.2). All the hepta-peptides were purified by HPLC to a purity higher than 98%. The procedure of substrate interference assay was similar to the protease inhibition assay except that the inhibitor was replaced by a regular peptide substrate. The HIV-1 protease reaction buffer was 0.1 M sodium acetate, 1.0 M sodium chloride, 1.0 mM ethylenediaminetetraacetic acid (EDTA), 1.0 mM DTT, 10% dimethylsulfoxide (DMSO), and 1mg/ml bovine serum albumin (BSA) at pH 4.7. Upon the addition of the HIV-1 protease, the fluorescent signal emitted through the cleavage of the FRET substrate was recorded using a SpectraMax M5 instrument (Molecular Devices, Sunnyvale, CA, USA) at an excitation wavelength of 340 nm and an emission wavelength of 490 nm. The protease concentration was adjusted to a substrate cleavage velocity of 0.08 RFU/sec in the absence of regular (non-fluorescent) substrate. In the enzyme reaction buffer, a regular

HIV-1 protease substrate was mixed with the FRET HIV-1 protease substrate to reach a final concentration of 100 μM and 2.5 μM, respectively. The average velocity of FRET substrate cleavage in RFU/min was calculated based on fluorescent signals at 30 time points over 30 min. In the control experiment, the average velocity of FRET substrate cleavage was determined in the absence of regular substrate. The result was illustrated by a bar chart representing the ratio of the average FRET substrate cleavage velocity in the presence of a regular substrate over the average FRET substrate cleavage velocity in the absence of a regular substrate.

2.5 Crystallization and diffraction data collection

Inhibitors (tipranavir and darunavir) and substrates (p2/NC and CA/p2) were co-crystallized with the MDR769 82T inactive protease using the hanging drop vapor diffusion method. The expression levels of the active MDR protease variants were generally lower than the expression levels of the inactive enzymes, and the protein amount of the active MDR protease variants was insufficient for crystallization experiments. The protease and ligand were pre-mixed at a molar ratio of 1:20 before crystallization set-up. The protease-ligand solution was then mixed at 2:1 v/v ratio with the precipitant solution. The optimal crystallization condition was obtained from a set of precipitant solutions (0.1 M citric acid/MES/HEPES and 0.8~2.8 M ammonium sulfate at pH 4.0~6.0). The reservoir volume was 750 μl. The MDR 769 82T-p2/NC complex was crystallized from 0.1 M citric acid and 2.4 M ammonium sulfate at pH 5.2 while the MDR 769 82T-CA/p2 complex, MDR 769 82T-darunavir complex, and MDR 769 82T-tipranavir complex were crystallized from 0.1 M MES and 2.4 M ammonium sulfate

at pH 6.0. The reservoir volume was 750 µl. Needle shaped crystals grew to a suitable size for diffraction within a week. The crystals were placed in a 30% w/v glucose cryoprotectant solution before they were frozen in liquid nitrogen. The diffraction data were collected at the Life Sciences Collaborative Access Team (LS-CAT) at the Advanced Photon Source (APS) Sector 21, Argonne National Laboratory (Argonne, IL, USA) and the diffraction data were processed with the HKL2000 program suite (*77*).

2.6 Structure refinement and analysis

Molecular replacement was performed with the CCP4 program Molrep-autoMR (*78*). A previously solved HIV-1 protease structure was used as the search model for molecular replacement. The models of darunavir and tipranavir were built with the program COOT (*79*) using a ligand finding algorithm where the molecular formulas of the inhibitors were provided in SMILES representation. The protease-ligand model was further refined using the program Refmac5. The tipranavir molecules were refined in two orientations with half-occupancy while the darunavir molecule was refined in one orientation with full-occupancy. Restrained anisotropic refinement was used for the tipranavir-MDR 769 82T co-crystal structure. The structure was validated using Procheck V3.4.4 (*80*). The models of tipranavir or darunavir complexed with MDR 769 82T were analyzed for molecular interaction using the program LIGPLOT V4.5.3 (*81*).

The co-crystal structures were deposited in the Protein Data Bank (http://www.pdb.org/). The accession codes for the darunavir-MDR 769 82T complex and the tipranavir-MDR 769 82T complex are 3SO9 and 3SPK, respectively.

2.7 Virtual screening and protein-compound docking studies

The large scale screening was conducted using a fast pre-screening tool, eHiTS LASSO (Ligand Activity in Surface Similarity Order) 2009 (*82*). Compound libraries were downloaded from the ZINC database of commercially-available compounds (http://zinc.docking.org/).

The chemical structure of compound GR346 was drawn in the program ChemBioDraw Ultra 12.0. The three-dimensional structure of GR346 was obtained by energy minimization using the GlycoBioChem PRODRG2 Server (http://davapc1.bioch.dundee.ac.uk/prodrg/). The electrostatic descriptor gasteiger charges were added to atoms of protein and ligand.

To examine the binding of the lead compound from virtual screening, the compound was docked into the binding pocket of HIV-1 protease using the program AutoDock4 (*83*).

2.8 Differential scanning fluorimetry

Differential scanning fluorimetry (DSF) was used to detect the protein melting temperature change after ligand binding (*84*). The bigger the melting temperature change, the tighter the ligand binding. A modified procedure is described as follows. The HIV-1 protease was diluted in the buffer of 100 mM Sodium acetate-AcOH and 150 mM NaCl at pH 5.0 to reach a final concentration of 0.4 mg/ml. The protease solution was pre-incubated with 500 μM protease inhibitors. 5×SYPRO orange dye (stock concentration 5,000×, Sigma-Aldrich, Inc., Saint Louis, Missouri, USA) was added to the protease-inhibitor mixture. The protein-dye solution (100 μl) was heated from 25°C to

82°C. During the heating process, the fluorescence signal was read using a microplate reader (SpectraMax M5, Molecular Devices, Sunnyvale, CA) with the excitation and emission wavelength at 492 nM and 610 nM, respectively.

2.9 Initial model preparation for molecular simulation

The X-ray structures of the protease-substrate complex (MDR 769 in complex with the substrate p2/NC) is solved and used as an initial structure for homology modeling. The complexes of MDR 807-p2/NC, MDR 1385-p2/NC, MDR 3761-p2/NC, and NL43-p2/NC were constructed using SWISS-MODEL (*85*). Similarly, a series of HIV-1 protease-CA/p2 complexes were built based on the MDR 769-CA/p2 co-crystal structure. Based on the catalytic mechanism of HIV-1 protease, Asp 25 was assigned as a protonated state while Asp 25' was assigned as a deprotonated state. All histidine residues were assigned a neutral charge. Protonation states of other amino acid residues were assumed based on the buffer pH in the HIV-1 protease enzymatic assays (pH 4.7). To avoid simulating a catalytic interaction, positional restraints were applied to the scissile peptide bond of the substrate and the β-carboxyl groups of catalytic residue Asp 25 and Asp 25'. The HIV protease-substrate complex was placed into an orthogonal TIP3P water box. The TIP3P is a three-site (three interaction sites) water model, which is one of the most widely used water models. The protease complex was at least 12 Å from the edge of the water box. Na^+ and Cl^- ions were added to neutralize the simulation system.

In the preparation of the NS3/4A HCV protease, the NS3 protease (from residue 1 to residue 181) and NS4A peptide (from residue 21 to residue 32) of HCV genotype 1b, 4a, 5a, and 6a were modeled based on a structure of HCV NS3/4A protease with a ketoamide

inhibitor (PDB: 3KN2) using SWISS-MODEL (*85*). The NS4A peptide binds non-covalently to the NS3 protease domain. Telaprevir was built with ChemBio3D and manually superimposed on the ketoamide inhibitor in the crystal structure. As a common mechanism of ketoamide inhibitors against the HCV protease, telaprevir formed a covalent bond with the hydroxyl oxygen of the catalytic residue Ser 139. The topology and force field parameters for telaprevir were initially generated using ParamChem (https://www.paramchem.org/), the CHARMM General Force Field (CGenFF) program for automatic atom typing and assignment of parameters and charges by analogy. The topology file and parameter file of telaprevir were then manually edited to covalently attach the drug to the Ser 139. Finally, the HCV protease complexes including NS3 protease domain, NS4A peptide, and telaprevir were solvated in a water box where the edge of the box is no less than 12 angstroms away from the protease complex. The system was neutralized with Zn^{2+} and Cl^{-} ions.

2.10 Molecular dynamics simulation

The MD simulations were performed using the parallel computing program Scaling NAno Molecular Dynamics (NAMD) V. 2.7b (*86*). The protease complex models were solvated in a water box using TIP3P models for water molecules. To prevent translational and rotational displacement and to prevent the simulation of the catalytic reaction, positional restraints were applied for the carbonyl group of the substrate scissile peptide bond and the catalytic residue Asp 25 and Asp 25’ of the HIV-1 protease. The cutoff for non-bonded interactions was 10 Å. Particle Mesh Ewald was implemented; the Particle Mesh Ewald method calculates direct-space interactions within a finite distance using a

modification of Coulomb's law, and in reciprocal space using a Fourier transform to build a "mesh" of charges, interpolated onto a grid (*87*). The systems were energy minimized using a conjugate gradient method and gradually heated from 70K to 310K in 200 ps. Simulations were conducted at 310K and 1.0 atm (NPT ensemble) for 10 ns using the CHARMM force field 27 and a timestep of 2 fs.

An example of NAMD configuration commands with annotations is listed as follows.

Adjustable Parameters

structure *input.psf* ; # Specify a protein structure file that contains molecular characteristics required by a particular force field.

coordinates *input.pdb* ; # Specify an atomic coordinate file

set temperature *310* ; # The simulation was conducted at 310 K.

set outputname *output* ; # Name the output files

firsttimestep *0* ; # Start the simulation from step 0.

Simulation Parameters

paraTypeCharmm *on* ; # Turn on CHARMM force field.

parameters *Par_all27_prot_na.prm* ; # Specify the CHARMM27 parameter file which define forces and energies for protein, lipid, DNA, and RNA.

parameters *par_all36_cgenff.prm* ; # Specify the CHARMM General Force Field parameter file for compounds.

temperature *$temperature* ; # Read temperature from "temperature" variable.

reassignFreq *2000* ; # Reassign temperature every 2000 steps.

reassignTemp *70* ; # Increase temperature from 70 K (the liquid nitrogen temperature).

reassignIncr *5* ; # The temperature increment is 5 K.

reassignHold *310* ; # Hold the temperature at 310 K.

 # Force-Field Parameters

exclude *scaled1-4* ; # Four adjacent covalently bonded atoms are excluded from non-bonded calculations.

1-4scaling *1.0* ; # The constant factor of electrostatic interactions for 1-4 atom pairs.

cutoff *12.* ; # The cutoff distance for non-bonded interaction is 12 Å.

switching *on* ; # Smoothing functions are applied to both the electrostatics and van der Waals forces.

switchdist *10.* ; # The smoothing function takes effect when atom distance is between 10 and 12 Å.

pairlistdist *13.5* ; # When distance between atoms is less than 13.5 Å in each cycle, the atom pair is included in pair list for electrostatics and van der Waals interaction calculation.

margin *2.5* ; # Add 2.5 Å extra length in patch dimension.

 # Integrator Parameters

timestep *2.0* ; # The length of each step is 2fs.

rigidBonds *all* ; # The bond between hydrogen and its bonded atoms is constrained.

nonbondedFreq *1* ; # Non-bonded interactions are calculated in every time step.

fullElectFrequency *2* ; # Full electrostatic interactions are calculated in every 2 time step.

stepspercycle 10 ; # The interacting particle lists are updated every 10 time step.

 # Constant Temperature Control

langevin *on* ; # Perform Langevin dynamics to introduce additional damping and random forces to the system.

langevinDamping *5* ; # Langevin damping coefficient is 5/ps

langevinTemp *$temperature* ; # Langevin dynamics adjust the effect on atoms based on the temperature variable.

langevinHydrogen *off* ; # Turn off Langevin dynamics for hydrogen atoms.

 # Constant Pressure Control

useGroupPressure *yes* ; # Pressure is calculated using hydrogen-group based pseudo-molecular virial and kinetic energy which results in smaller fluctuations.

useFlexibleCell *no* ; # Do not allow three orthogonal dimensions of the periodic cell to fluctuate independently.

useConstantRatio *no* ; # Do not maintain a constant ratio of the unit cell in the x-y plane.

langevinPiston *on* ; # Turn on Langevin piston pressure control.

langevinPistonTarget *1.01325* ; # The target pressure for Langevin piston method is 1.01325 bar.

langevinPistonPeriod *100* ; # The barostat oscillation time scale for Langevin piston

method is 100 fs.

langevinPistonDecay *50* ; # The barostat damping time scale for Langevin piston method is 50 fs.

langevinPistonTemp *$temperature* ; # The barostat noise temperature for Langevin piston method is the same as the temperature specified in temperature control.

 # Output

outputName *$outputname* ; Specify the output name for output files.

restartfreq *1000* ; # Prepare the restart files every 1000 steps (2 ps).

dcdfreq *1000* ; # The trajectory file is updated every 1000 steps.

outputEnergies *1000* ; # Output energy to log file every 1000 steps.

outputPressure *1000* ; # Output pressure to log file every 1000 steps.

#Harmonic constraints

constraints *on* ; # Turn on positional restraints.

consexp *2* ; # Use 2 as an exponent for constraint energy function

conskcol *B* ; # Specify the B factor column of the PDB file to use for the harmonic constraint force constant.

consref *constraint.pdb* ; # Input the PDB file indicating constraint atoms.

conskfile *constraint.pdb* ; # Input the PDB file to use for force constants for harmonic constraints.

Extra Parameters

Periodic Boundary Conditions

cellBasisVector1 *80.16 0.0 0.0* ; # Specify the x-dimension of the periodic cell size in Å.

cellBasisVector2 *0.0 63.51 0.0* ; # Specify the y-dimension of the periodic cell size in Å.

cellBasisVector3 *0.0 0.0 64.93* ; # Specify the z-dimension of the periodic cell size in Å.

cellOrigin *5.26 1.74 16.96* ; # Specify the position of the periodic cell center.

wrapWater *on* ; # When water molecules cross a periodic boundary, they are not translated to the other side of the periodic cell.

wrapAll *on* ; # When a molecule cross a periodic boundary, it is not translated to the other side of the periodic cell.

 # PME Particle Mesh Ewald

PME *yes* ; # Turn on Particle Mesh Ewald method, an efficient full electrostatics method in periodic boundary conditions.

PMEGridSizeX *90* ; # Number of grid points in X dimension.

PMEGridSizeY *72* ; # Number of grid points in Y dimension.

PMEGridSizeZ *72* ; # Number of grid points in Z dimension.

Execution Script

 # Minimization

minimize *10000* ; # Run 10,000 step conjugate gradient energy minimization.

reinitvels *$temperature* ; # Reset the temperature according to the temperature variable after minimization

 # Molecular Dynamics

run *7500000* ; # Run molecular dynamics for 7500000 steps, which is equal to 15000 ps (15 ns)

2.10.1 The CHARMM force field

The CHARMM (Chemistry at HARvard Macromolecular Mechanics) force fields represent a set of widely used force fields in molecular dynamics. The CHARMM19 and CHARMM22 force fields are parameterized for proteins, and the CHARMM27 force field is parameterized for lipid, DNA, and RNA. A general force field for drug-like molecules (CGenFF) covers a wide range of chemical groups.

The following equation is the potential energy function from CHARMM22 (*88*): $U_{\text{total}} = U_{\text{bond}} + U_{\text{angle}} + U_{\text{dihedral}} + U_{\text{vdW}} + U_{\text{Coulomb}}$. The components of the potential energy function in this equation contain the energy functions for bond stretching, bond angle bending, torsion (dihedral angle), and nonbonded interaction. The nonbonded interactions include van der Waals (vdW) energy and electrostatic (Coulomb) energy.

$$U_{\text{bond}} = \sum_{\text{bonds } i} k_i^{\text{bond}}(r_i - r_{0i})^2$$

$$U_{\text{angle}} = \sum_{\text{angles } i} k_i^{\text{angle}}(\theta_i - \theta_{0i})^2 \qquad U_{\text{vdW}} = \sum_i \sum_{j>i} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$$

$$U_{\text{dihedral}} = \sum_{\text{dihedral } i} \begin{cases} k_i^{\text{dihe}}[1 + \cos(n_i\phi_i - \gamma_i)], n_i \neq 0 \\ k_i^{\text{dihe}}(0_i - \gamma_i)^2, n = 0 \end{cases} \qquad U_{\text{Coulomb}} = \sum_i \sum_{j>i} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}$$

2.11 Data analysis of molecular dynamics results

Trajectories of MD simulation were visualized and analyzed using the Visual Molecular Dynamics (VMD) program V. 1.91. The superposition of molecular structures was carried out using Pymol. The root-mean-square deviation (RMSD) values were calculated using the VMD RMSD trajectory plug-in. The molecular mechanics energy was calculated using the NAMD energy plug-in in VMD. The final molecular mechanics

energy is the average of molecular mechanics energy of the last 100 frames (the last 200 ps simulation). Previous studies demonstrated that HIV-1 protease ligands show single-maxima probability density function of energy (*13*). Therefore, the last 200 ps simulation represents a relatively stable protease-ligand complex conformation. Hydrophobic solvation energy and electrostatic solvation energy were calculated and averaged based on the 10 snapshot coordinates of the last 200 ps simulation.

The change in desolvation energy during the binding of protease to its substrate was obtained with the following equation, $\Delta G_{solv} = \Delta G_{desolv}^{electrostatic} + \Delta G_{desolv}^{nonpolar}$. The electrostatic desolvation energy was calculated using the Adaptive Poisson-Boltzmann Solver (APBS) program (*89*), a software package for modeling biomolecular solvation through solution of the Poisson-Boltzmann equation (PBE). The electrostatic desolvation energy upon protease-substrate complex formation was calculated using the following equation, $\Delta G_{desolv}^{elec} = \Delta\Delta G_{solv}^{elec} = \Delta G_{solv}^{elec,complex} - \Delta G_{solv}^{elec,protease} - \Delta G_{solv}^{elec,substrate}$. The symbol ΔG_{solv} refers to the difference of solvation energy when a molecule is transferred from water solution (with a dielectric constant of 78.54) to vacuum (with a dielectric constant of 1.00). The symbol $\Delta\Delta G_{solv}^{elec}$ represents the solvation energy change (desolvation energy) upon protease-substrate complex formation. The input for the APBS program was prepared by converting the PDB format coordinate to PQR format through the PDB2PQR server (http://propka.ki.ku.dk), an automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations (*90*). The PDB2PQR server adds charge and radius parameters to existing PDB data. The non-polar desolvation energy was calculated using the equation $\Delta G_{desolv}^{nonpolar} = -\Delta G_{solv}^{nonpolar} = -$

($\gamma\Delta SASA + \beta$), where SASA stands for solvent accessible surface area, γ is a standard value of 0.00542 kcal*mol^{-1}*$Å^{-2}$, and β is 0.92 kcal mol^{-1} (*91*). The change of solvent accessible surface area upon protease-substrate complex formation is calculated using the following equation, $\Delta SASA = SASA^{complex} - SASA^{protease} - SASA^{substrate}$.

2.12 Determination of representative sequences for genotype 4, 5 and 6 HCV NS3/4 proteases

All available NS3/4A protease sequences from the HCV database (http://hcv.lanl.gov) were aligned using the ClustalW2 program (http://www.ebi.ac.uk/Tools/clustalw2) to find a patient isolate sequence that is closest to the consensus sequence of each HCV genotype. That patient isolate sequence was selected as the representative for a particular HCV genotype.

2.13 Expression of the NS3/4A proteases of the HCV genotype 4, 5, and 6

Based on the representative sequences of the HCV genotype 4 to 6, the fusion genes were constructed with the residues 3-181 of NS3 protein and the residues 21-32 of NS4A protein fused to the N-terminus of NS3 protease domain. Therefore, the HCV NS3/4A protease maintains an active form without the need for the addition of a separate NS3/4A peptide. Between NS4A and NS3, there is a four-residue flexible GSGS linker. Since the essential fragment (residue 21-32) of the NS4A peptide has been identified (*92*), and previous studies showed that the fusion protein containing NS4A (residue 21-32), NS3 (residue 3-181) and a tetrapeptide linker (GSGS) exhibited the full proteolytic activity (*93*), the fusion protein is used in this study to facilitate enzyme assays and crystallization

studies later on. For the final constructs, the NS3/4A fusion genes were inserted into the plasmid pET28b(+) which produced a 21-residue thrombin cleavable His-tag (MGSSHHHHHHSSGLVPRGSHM) at the N-terminus. The 216 amino acid fusion proteins (~23 kDa) with NS3 protease domain, a fragment of the NS4A protein, the GSGS linker, and the thrombin cleavable His-tag were expressed and purified by AnaSpec, Inc. (San Jose, CA). The thrombin cleavable His-tag was kept in the final protein since studies conducted at AnaSpec, Inc. showed an increased enzyme activity by keeping the tag.

2.14 HCV protease inhibition essays

The enzyme inhibitory activity of telaprevir, IC_{50}, for HCV NS3/4A protease from genotype 4 to 6 was evaluated. HCV protease FRET substrate S-1 was purchased from AnaSpec, Inc. The fluorescence emitted by the peptide cleavage was monitored by microplate reader (SpectraMax M5, Molecular Devices) using the excitation wavelength 355 nm and the emission wavelength of 495 nm. There is no intermolecular quenching of the FRET peptide observed in the concentration range of 0.5 to 7.0 μM. The optimized concentration of protease was obtained by adjusting the FRET substrate cleavage speed to 0.08 RFUs/sec. The experiments were conducted in a 96-well plate, with each well containing a constant amount of NS3/4A protease in the assay buffer and varying concentrations of the serially diluted telaprevir. After a 1 hour incubation at 37°C to form the tight enzyme-inhibitor complex-telaprevir being a slow, tight-binding inhibitor – the reaction was initiated by the addition of the FRET substrate S1 (final concentrations of 5 μM S1, 2% (vol/vol) DMSO, total assay volume was 100 μl). The reason for

slow-binding inhibition by telaprevir is not known. In contrast, incubation of the HIV-1 protease with any of the licensed drugs prior to substrate addition is not needed. The progress of the HCV NS3 reactions was monitored over 20 min using a sampling rate of 1 min. The FRET substrate cleavage rates (initial velocities) of the NS3/4 protease with serially diluted telaprevir were plotted as data points with the software of SoftMax Pro V. 5.2.

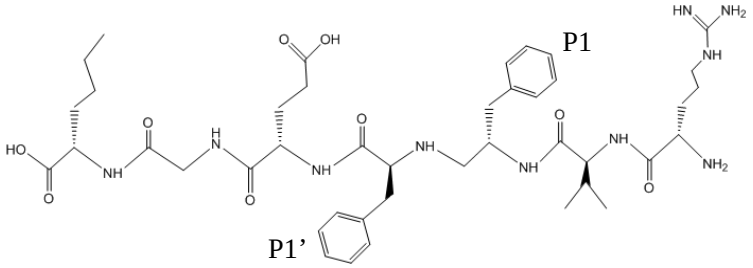
CHAPTER 3 DESIGN AND SYNTHESIS OF A MULTI-DRUG RESISTANT HIV-1 PROTEASE INHIBITOR BASED ON LOPINAVIR

3.1 Introduction

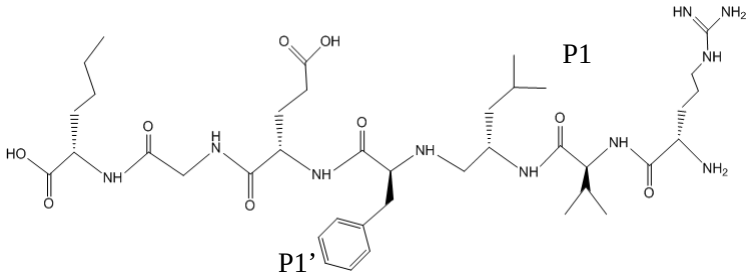
Developing novel HIV-1 protease inhibitors that overcome drug-resistance is still a challenging task. In drug development, it is difficult to keep pace with the emergence of drug resistance. Compared to *de novo* drug design, drug modification and optimization could lower the risk of failure in clinical trials. The modification of available protease inhibitors according to the MDR HIV-1 protease isolate may provide a solution to overcome drug resistance. Lopinavir is a second-generation protease inhibitor, which is derived from ritonavir. The smaller P2 and P2' groups of lopinavir decrease the contact with highly variable residues at the 82 site HIV-1 protease and improve the inhibitory efficacy against drug-resistant mutants of HIV-1 protease (*19*). The MDR HIV-1 protease isolate 769 exhibits cross resistance to HIV-1 protease inhibitors. Though lopinavir inhibits MDR 769 with a low relative resistance, further modification of lopinavir may improve its efficacy against drug resistance mutations in the HIV protease. A mutagenesis experiment has been recently completed in our laboratory. The P1 and P1' groups of the HIV-1 protease substrate CA/p2 possessing a reduced scissile peptide bond have been mutated and tested for HIV-1 protease inhibition. Compared to a phenylalanine-phenylalanine combination at P1 and P1' sites, the combination of leucine

at the P1 site and phenylalanine at the P1’ site exhibits relatively high inhibitory efficacy and reduced relative resistance against MDR769 (unpublished data). The current first line protease inhibitor, darunavir, also contains the asymmetric combination of leucine side chain and phenylalanine side chain at the P1’ and P1 sites (*28*). However, lopinavir carries the symmetric P1 and P1’ groups mimicking the phenylalanine side chain. Therefore, by changing the P1 site of lopinavir from a phenylalanine side chain to a leucine side chain (isobutyl group), the inhibitory potency of the modified lopinavir may increase and the relative resistance against MDR protease may decrease.

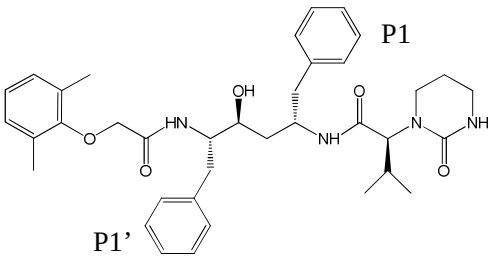
In this chapter, the synthesis, inhibitory potency, and binding mechanism of a lopinavir analog with isobutyl P1 group is discussed. The lopinavir analog lowers the relative resistance against an MDR HIV-1 protease variant. Simulation results indicate that the lopinavir analog induces an asymmetric movement of the 80s loop of the MDR HIV-1 protease and this analog exhibits a stronger interaction energy.



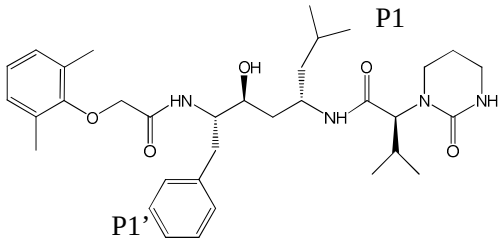
CA/p2 (P1F, P1'F) reduced peptide



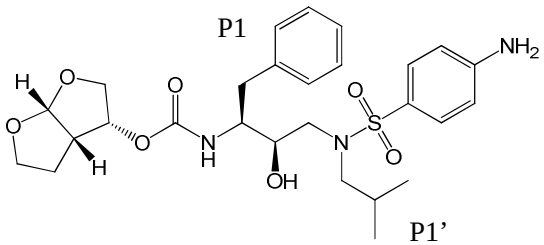
CA/p2 (P1'F) reduced peptide



lopinavir



lopinavir analog



darunavir

Figure 3.1 Chemical structures of compounds discussed in Chapter 3. The reduced peptide bonds, P1, and P1' groups are labeled in the figure.

3.2 Results

3.2.1 Isobutyl P1 lopinavir analog reduced the relative resistance against the MDR HIV-1 protease 769

A lopinavir analog with isobutyl P1 group was chemically synthesized according to the synthesis procedure of lopinavir (Figure 2.1). The modified step was the Grignard reaction, where 1-chloro-2-methyl propane was used to replace benzyl chloride in order to introduce a leucine side chain rather than a phenylalanine side chain to the lopinavir analog. The original lopinavir was synthesized in parallel as a control. The final lopinavir analog was verified using electrospray mass spectrometry (Figure 3.2). The major impurity is an intermediate product.

The IC_{50} of lopinavir analog with an isobutyl P1 group was determined and listed in Table 3.1. The IC_{50} of the synthesized original lopinavir control and the CA/p2 reduced peptide (P1'F) were also measured. The lopinavir analog achieved 50% inhibition of the WT HIV-1 protease and MDR 769 at the concentration of 2.1±0.1 nM and 6.8±0.4 nM, respectively. The inhibition efficacy of the lopinavir analog was higher than the lopinavir control synthesized in parallel which exhibited the IC_{50} values of 2.5±0.5 nM and 16±0.6 nM for the WT HIV-1 protease and MDR 769, respectively. Furthermore, the lopinavir analog decreased the relative drug resistance from 6.4 to 3.2. The CA/p2 reduced peptide still showed the lowest IC_{50} value (4.4±0.2 nM) inhibiting the MDR HIV-1 protease as well as having the lowest relative resistance (1.7 fold).

Table 3.1 IC_{50} results with the lopinavir analog

Compounds (nM)	HIV-1 protease		Relative resistance
	NL4-3	MDR 769	
Synthesized lopinavir analog	2.1±0.1	6.8±0.4	3.2
Synthesized lopinavir control*	2.5±0.5	16±0.6	6.4
Reduced CA/p2 (P1'F) peptide	2.6±0.4	4.4±0.2	1.7
Reduced CA/p2 (P1F, P1'F) peptide	30.7	142	4.6

*The lopinavir control was synthesized in parallel using the same procedures and equipments. The only different reagent is benzyl chloride instead of 1-chloro-2-methyl propane.

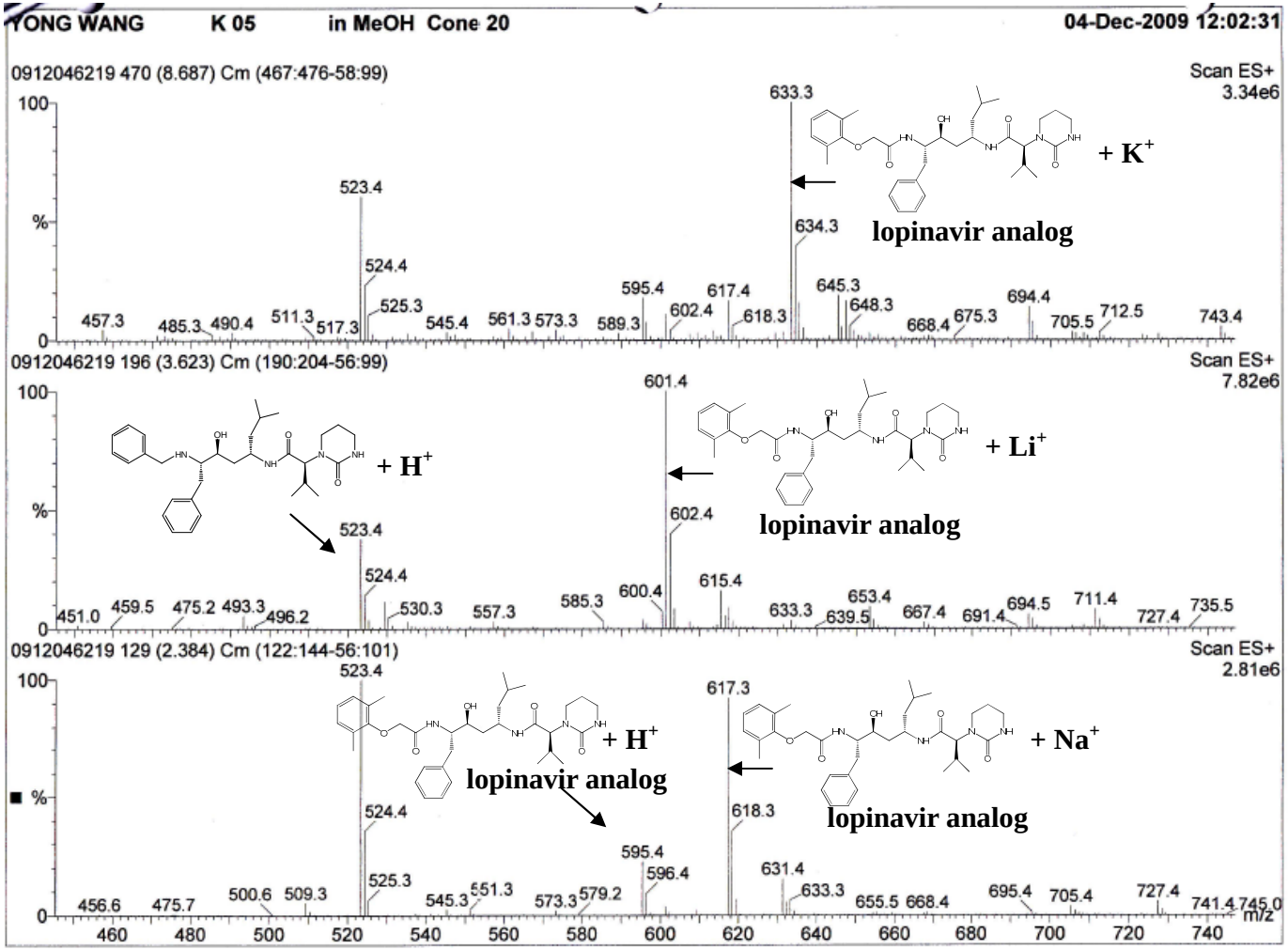


Figure 3.2 Electrospray mass spectrometry results of the lopinavir analog. Samples were analyzed with K^+, Li^+, and Na^+ ions.

3.2.2 Asymmetric P1 and P1’ groups induce asymmetric 80s loop conformations of the MDR HIV-1 protease 769

The combination of isobutyl and benzyl group as P1 and P1' sites could induce an asymmetric movement of the mobile 80s loops of the HIV-1 protease. By examining the co-crystal structures of CA/p2 or darunavir with MDR 769 A82T, the P1 leucine residue of CA/p2 peptide and the P1’ isobutyl group of darunavir brought the Pro 81 residue on the 80s loop of HIV-1 protease closer to the ligand. In Figure 3.3, the MDR 769 82T-CA/p2 complex crystal structure and MDR 769 82T-darunavir complex crystal structure were superimposed to a WT HIV-1 protease-lopinavir complex crystal structure. Compared to the protease 80s loops interacting with the symmetric P1 and P1’ benzyl groups of lopinavir, the P1 leucine of CA/p2 peptide brought the protease 80s loop closer to the P1 group (Figure 3.3 A) while the phenylalanine P1’ group of the CA/p2 peptide induced a similar loop conformation as that of the lopinavir complex (Figure 3.3 B). The asymmetric conformations of protease 80s loops have also been observed in the MDR 769 82T-darunavir complex due to the asymmetric P1 (benzyl) and P1’ (isobutyl) groups of darunavir

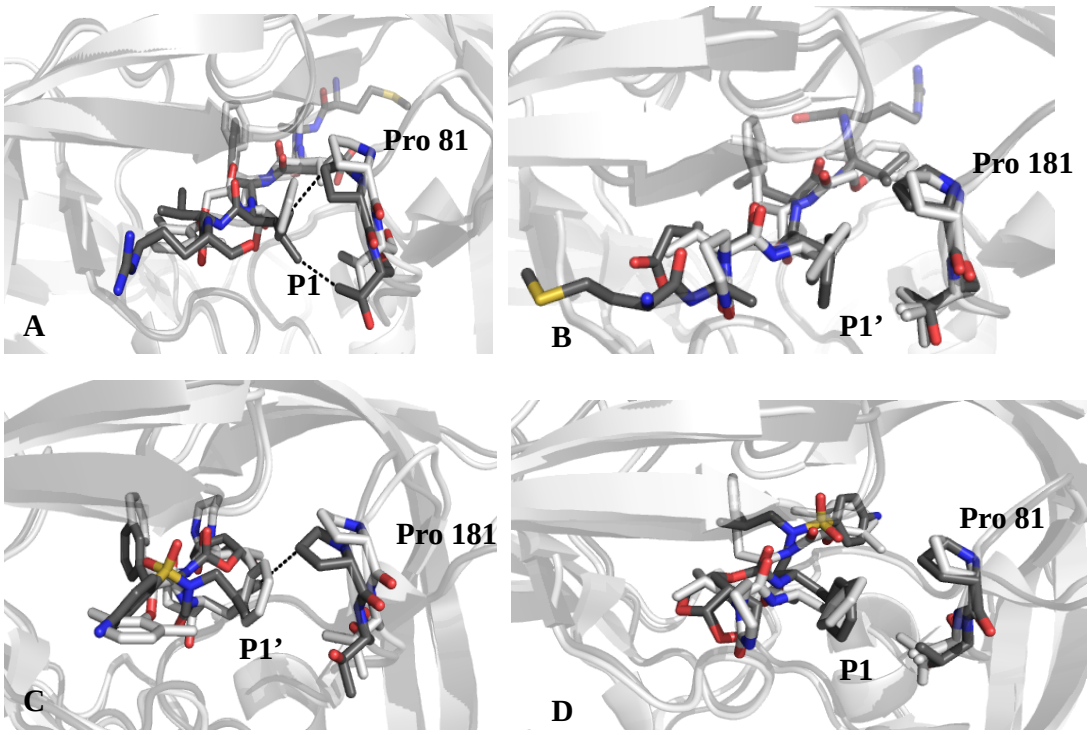


Figure 3.3 The benzyl and isobutyl groups at P1 and P1' sites trigger the asymmetric 80s loop conformation of the MDR HIV-1 protease. (A) The co-crystal structure of CA/p2 with MDR 769 82T (grey) was superimposed on the co-crystal structure of lopinavir with the WT HIV-1 protease (sliver, PDB 2Q5K). The P1 leucine side chain of CA/p2 interacted with the protease and altered the 80s loop conformation of one MDR protease monomer. (B) In the CA/p2-MDR 769 structure (grey), the P1' phenylalanine side chain of CA/p2 did not alter the 80s loop conformation of the MDR HIV-1 protease, compared to the lopinavir-WT HIV-1 protease co-crystal structure (white). (C) The co-crystal structure of darunavir with MDR 769 82T (grey, PDB 3SO9) was superimposed to the co-crystal structure of the lopinavir-MDR 769 82T complex (white). The P1' isobutyl group of darunavir triggered a similar conformation of the 80s loop as that triggered by the P1 leucine side chain of the CA/p2. (D) The 80s loop conformation that interacted with the P1 benzyl group was conserved in the lopinavir-MDR 769 82T

(white) and the darunavir-MDR 769 82T (grey) complex structures.

3.2.3 Simulation analysis of the lopinavir analog binding to the MDR HIV-1 protease

The lopinavir and isobutyl P1 lopinavir analog in complex with MDR 769 were modeled and simulated for 10 ns. The average conformations of the last 200 ps simulation were analyzed for both the lopinavir and lopinavir analog complexes. In the average conformation of the MDR 769-lopinavir complex, the asymmetric conformation of protease 80s loops was observed as well. One 80s loop of the MDR HIV-1 protease 769 moved closer to the isobutyl P1 group of the lopinavir analog (Figure 3.4). The interaction energy calculated between the inhibitor and protease was listed in Table 3.2. The calculated interaction energy suggested that the lopinavir analog interacted with MDR 769 stronger (-99.47 kcal/mol) than lopinavir did (-92.39 kcal/mol). The interaction energy was increased in both electrostatic interactions and van der Waals interactions.

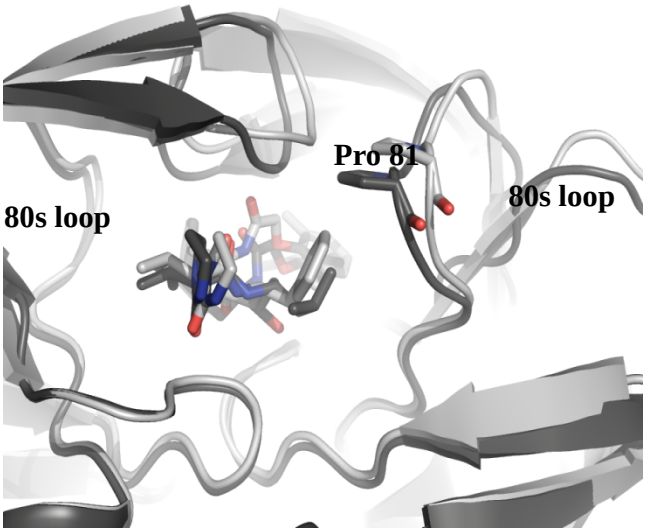


Figure 3.4 Isobutyl P1 lopinavir analog induces asymmetric 80s loop conformation of MDR HIV-1 protease 769. The MDR 769-lopinaivr model is represented in while, and the MDR 769-lopinaivr analog model is represented in black. The HIV-1 proteases

are shown in ribbon representation while the inhibitors and the proline 81 are shown in stick representation.

Table 3.2 Interaction energy of lopinavir and lopinavir analog with the multi-drug resistant HIV-1 protease 769

Compounds	Interaction energy (kcal/mol)		
	Electrostatic	Van der Waals	Nonbonded (total)
lopinavir analog	-37.39±6.55	-62.08±3.57	-99.47±7.45
lopinavir	-32.27±8.57	-60.12±3.76	-92.39±9.34

3.3 Discussion

The drug-resistant mutations may increase the flexibility of HIV-1 protease and therefore alter the active site conformation to be unfavorable for drug binding. The 80s loop is a mutation hot spot of HIV-1 protease, which directly interact with the P1 and P1' group of substrates or drugs. The V82A/C/F/I/L/M/S/T mutations on the loop have been identified (*15, 94*). Both structural analyses and enzymatic results indicate that a smaller P1 or P1' group, such as leucine, brings the 80s loop closer to the ligand and increases binding affinity.

The enzyme inhibitory assays demonstrate the promising role of an isobutyl group at the P1 or P1' position in overcoming drug-resistance. However, due to the limitation of purification techniques, neither the lopinavir analog or lopinavir control reach the potency of the standard lopinavir that was requested from the NIH AIDS Research and Reference Reagent Program. The impurities and stereoisomers mixed in the final products increase the IC_{50} values.

Furthermore, the lopinavir analog does not reach the inhibitory potency of CA/p2 P1’F reduced peptide against the MDR 769 protease. The potency of CA/p2 P1’F reduced peptide may be influenced by the P2’ glutamic acid. Studies have shown that the binding efficacy of reduced peptides with a glutamic acid in the P2’ position are pH-sensitive (*95*). Therefore, the electrostatic interactions of the CA/p2 P1’F reduced peptide is superior to the hydrophobic lopinavir analog, which explains the relatively lower potency of lopinavir analog compared to the CA/p2 P1’F reduced peptide.

The lopinavir analog with an isobutyl group at the P1 position may serve as an HIV-1 protease inhibitor against MDR proteases, especially MDR HIV-1 proteases containing mutations on the 80s loops. Currently, genotypic assays are recommended to assess resistance in the adult and adolescent guidelines of antiretroviral treatment (*28*). Benefiting from protease gene sequences of patience isolates, the choice of optimal regimen could be decided based on the types of drug-resistance mutations. Therefore, there is a demand of developing modified protease inhibitors against various groups of HIV-1 protease variants.

3.4 Author’s contribution

The author synthesized the lopinavir analog and lopinavir control, carried out the HIV-1 protease inhibition assays and molecular dynamics simulations. The analysis of the results including structure superposition was also carried out by the author.

CHAPTER 4 THE HIGHER BARRIER OF DARUNAVIR AND TIPRANAVIR RESISTANCE FOR HIV-1 PROTEASE

4.1 Introduction

The emergence of drug resistance during antiretroviral therapy is a key concern for HIV patients. In patients undergoing combination antiretroviral therapy, protease inhibitors are important contributors to the decrease in the number of circulating virus particles. In total, nine HIV-1 protease inhibitors have been approved by the US Food and Drug Administration (FDA). In the 99 amino acid residue HIV-1 protease, 36 mutation loci have been identified which lead to various levels of resistance to protease inhibitor treatment (*15, 96*). The current HIV-1 protease inhibitors are designed with a hydroxyl group to mimic the transition state of the substrate's scissile peptide bond. Due to the structural similarity of inhibitors, the mutations in HIV-1 protease are commonly associated with cross-resistance to the other inhibitors (*94*).

The clinical multi-drug resistant (MDR) HIV-1 strain 769 was isolated by Palmer *et al.* from patients failing protease inhibitor-containing antiretroviral regimens, and the protease of strain 769, MDR 769, is resistant to all protease inhibitors tested (*97*). As observed in the crystal structure previously solved by our group, the flaps of MDR 769 are farther apart compared to the distance of wild-type (WT) HIV-1 protease flaps (*98*). The threonine mutation at residue 82 of the MDR HIV-1 protease (MDR 769 82T) alters the hydrophobicity of the P1 and P1' binding pockets and could further enhance

cross-drug resistance. The uncomplexed MDR 769 82T crystal structure adopts the wide-open flap conformation as reported earlier for MDR 769 (*98*, *99*).

According to virologic response studies, darunavir and tipranavir show a higher genetic barrier to resistance (*100*). Both inhibitors have been used to treat patients infected with protease inhibitor-resistant viral strains and have effectively inhibited a range of MDR protease isolates (*101-103*). Based on the scoring function of the HIV Drug Resistance Database (http://hivdb.stanford.edu), MDR 769 82T has low resistance to darunavir and high resistance to tipranavir as well as the other seven protease inhibitors (*15*, *96*).

The structural study of the inhibitor-bound MDR HIV-1 protease facilitates the understanding of drug resistance mechanisms. The aim of this study is to test the *in vitro* inhibitory potency of darunavir and tipranavir against MDR 769 82T and to determine the mechanism of overcoming resistance by examining the binding conformation, key contacts, and the stability of inhibitor-protease complexes. The protease inhibition assays demonstrate the decreased susceptibility of MDR 769 82T to all the tested inhibitors and confirm that 82T severely enhances drug resistance. Compared to other protease inhibitors, the higher resistance barrier of darunavir is due to maintaining main chain hydrogen bonds by inhibitor flexibility while the higher resistance barrier of tipranavir is due to tight flap binding.

Table 4.1 Sequences of the HIV-1 protease variants

HIV-1 protease	Sequences[a]
NL4-3	PQITLWKRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI GGFIKVRQYD QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
MDR 769	PQITLWKRPI VTIKIGGQLK EALLDTGADD TVLEEVNLPG RWKPK**L**IGGI GGF**V**KVRQYD QVPIEICGHK VIGTVLVGPT P**A**N**V**IGRNL**M** TQIGCTLNF
MDR 769 82T	PQITLWKRPI VTIKIGGQLK EALLDTGADD TVLEEVNLPG RWKPK**L**IGGI GGF**V**KVRQYD QVPIEICGHK VIGTVLVGPT P**T**N**V**IGRNL**M** TQIGCTLNF

[a]The drug resistant mutations are in bold and polymorphic changes are underlined.

Table 4.2 IC_{50} measurements and relative resistance

Inhibitor	IC_{50} (nM)			Relative resistance		Predicted resistance score[a]	
	WT (NL4-3)	MDR 769	MDR 769 82T	MDR 769	MDR 769 82T	MDR 769	MDR 769 82T
darunavir	0.26 ±0.024	0.74 ±0.20	2.76 ±0.16	2.8	11	10	10
tipranavir	0.24 ±0.18	0.63 ±0.17	3.02 ±0.10	2.6	13	24	64
lopinavir	0.28 ±0.27	0.50 ±0.060	3.18 ±0.23	1.8	11	76	76
amprenavir	0.43 ±0.044	4.82 ±0.90	7.87 ±0.16	11	18	86	86
atazanavir	0.19 ±0.071	2.90 ± 0.16	10.8 ±0.40	15	57	98	103
nelfinavir	1.58 ±0.85	109 ±5.0	448 ±65	69	2.8×10^2	184	184
ritonavir	0.34 ±0.15	60.7 ±3.5	237 ±5.6	180	7.0×10^2	-	-
saquinavir	0.50 ±0.017	294 ±9.4	1.30×10^3 ±87	590	2.6×10^3	114	114
indinavir	0.47±0.19	119 ± 13	307 ±29	250	6.5×10^2	128	128

[a]The resistance scores for darunavir, tipranavir, lopinavir, amprenavir, atazanavir, saquinavir, and indinavir were predicted based on the protease inhibitor plus ritonavir co-administration. The larger scores indicate higher predicted resistance.

4.2 Results

4.2.1 Darunavir and tipranavir presented high barrier of resistance compared to the other protease inhibitors

The nine FDA-approved HIV-1 protease inhibitors were effective against the WT HIV-1 protease, NL4-3, at low nanomolar or sub-nanomolar concentration while the MDR 769 82T was resistant to all nine inhibitors (Table 4.2). The resistance ranged from 11-fold to 2600-fold as compared with the inhibition level against NL4-3. Although 11-13 relative resistance was observed, darunavir and tipranavir were the two most potent inhibitors against MDR 769 82T. Based on the IC_{50} values, the difference in Gibbs free energy ($\Delta\Delta G$) values calculated with the IC_{50} of MDR 769 82T and WT HIV-1 protease were -1.48 kcal/mol and -1.58 kcal/mol for darunavir and tipranavir, respectively.

The 82T mutation of the MDR protease greatly enhanced drug resistance. The relative resistance of the MDR 769 (1.8-590 fold) was approximately 2.6-6.1 times lower compared to that of MDR 769 82T. The 82T mutation caused relatively larger differences in the relative resistance of darunavir, tipranavir, and lopinavir, and the relative resistance was elevated 3.9, 5.0, and 6.1 times, respectively. The IC_{50} values are not fully correlated with predicted values based on the scoring of inhibitor resistance mutations from the HIV Drug Resistance Database since the 82T mutation has not been reported previously as darunavir resistance mutation (*104*).

4.2.2 Darunavir or tipranavir induced flap closure of the MDR HIV-1 protease

The darunavir-MDR 769 82T and tipranavir-MDR 769 82T co-crystal structures

were determined at a resolution of 1.24 Å and 2.87 Å, respectively. The binding of tipranavir or darunavir effectively closed the wide-open flap of the MDR protease and exhibited a conserved binding mode in the MDR protease relative to that in the wild-type protease. The tipranavir-MDR 769 82T complex crystallized in the hexagonal space group $P6_1$. The darunavir-MDR 769 82T complex crystallized in the orthorhombic space group $P2_12_12_1$. The diffraction and refinement statistics are shown in Table 4.3. The structures were deposited in the Protein Data Bank. The PDB IDs for darunavir-MDR 769 82T and tipranavir-MDR 769 82T structures are 3SO9 and 3SPK, respectively. Compared with the binding of darunavir or tipranavir to the WT HIV-1 protease (PDB ID: 3BVB and 2O4P), the butyl group of tipranavir and the P1', P2, and P2' groups of darunavir adopted different conformations relative to their binding to WT HIV-1 protease (Figure 4.1 A and B).

Asymmetric movement of the 80s loop regions was observed in the darunavir-MDR complex. The RMSD of Thr 82 of superimposed protease monomers is 1.318 Å. The asymmetric P1 and P1' group of darunavir induced the asymmetric protease dimer. The smaller isobutyl P1' group of darunavir allowed the 80s loop of the MDR protease to move closer to the inhibitor. According to the superposition of MDR 769 82T and WT HIV-1 protease structures, Pro 81 in one monomer of the MDR protease was 0.7 Å closer to the isobutyl P1' group of darunavir. Thr 82 was 1.6 Å further from the isobutyl P1' group of darunavir compared to Val 82 in WT HIV-1 protease (Figure 4.1 B). The V82T mutation changed the hydrophobicity of the binding pocket in the MDR protease. Therefore, the hydrophilic Thr 82 pointed toward the solvent rather than the hydrophobic

binding pocket. The 80s loop on the other monomer of the MDR protease showed an asymmetric pattern. Both Pro 81 and Thr 82 moved away from the benzyl P1 group of darunavir.

Table 4.3 Crystallographic statistics

Dataset	**Tipranavir-MDR 769 82T co-crystal**	**Darunavir-MDR 769 82T co-crystal**
Data collection		
Space group	$P6_1$	$P2_12_12_1$
Wavelength (Å)	0.979	0.979
Cell constants (Å)	a=63.12 b=63.12 c=83.54	a=28.90 b=66.38 c=90.84
Resolution range (Å)	29.52-1.24 (1.25-1.24)	30.00-2.87 (2.95-2.87)
Number of unique reflections	51613 (1138)	10515 (195)
Completeness (%)	96.7 (70.9)	98.4 (100.0)
Redundancy	9.1 (2.7)	4.4 (4.5)
Mean I/σ (*I*)	13.4 (1.2)	11.4(3.6)
R_{merge} [a]	0.084 (0.679)	0.142 (0.486)
Refinement		
R_{work} (%)[b]	17.40	21.65
R_{free} (%)[b]	22.70	25.62
Number of atoms		
Ligand	42	38
Protease	1514	1514
Solvent	351	142
Average B factor (Å^2)		
Ligand	10.91	35.67
Protease	13.38	18.58
Solvent	37.16	27.95
RMSD bond length (Å)	0.010	0.009
RMSD bond angle (°)	1.37	1.03
Ramachandran plot		
Allowed/generous/disallowed (%)	94.9/5.1/0	92.9/7.1/0

[a] $R_{merge} = \Sigma_{hkl} \Sigma_i |I_i(hkl) - \langle I(hkl) \rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation and $I(hkl)$ is the mean value for its unique reflection.

[b] $R_{work} = \Sigma_{hkl} ||F_o| - |F_c|| / \Sigma_{hkl} |F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes. R_{free} is calculated exactly as R_{work} using a random 5% of the reflections

omitted from refinement.

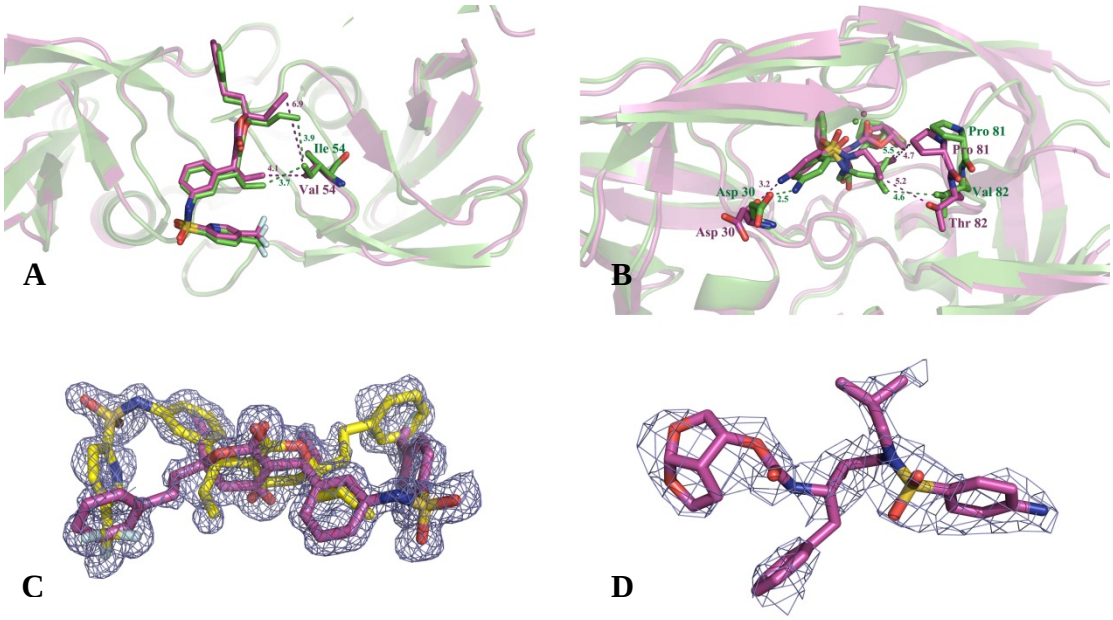


Figure 4.1 Co-crystal structures of darunavir and tipranavir with the MDR protease. (A) The tipranavir-MDR 769 82T structure (magenta) was superimposed on the tipranavir-WT HIV-1 protease structure (green). The I54V mutation is displayed in stick model. The magenta and green dashed lines indicate the distance (Å) between the inhibitor and the MDR HIV-1 protease or the WT HIV-1 protease, respectively. (B) The conformational changes of darunavir binding to MDR 769 82T (magenta) in comparison with the darunavir-WT HIV-1 protease complex (green). (C) The electron density of tipranavir bound to MDR 769 82T protease is shown as an F_o –F_c OMIT map contoured at 2σ**.** Two tipranavir molecules (yellow and magenta) binding in different orientation are presented. (D) The electron density of darunavir bound to MDR 769 82T protease is shown as an F_o –F_c OMIT map contoured at 2σ.

4.2.3 Specific contacts were lost between darunavir or tipranavir and the MDR HIV-1 protease

Hydrophobic interactions or hydrogen bonds were lost between the inhibitors and the MDR protease. According to the LIGPLOT analysis, the number of residues in MDR 769 82T involved in hydrophobic interactions with darunavir or tipranavir decreased by 1~2 residues compared to that of WT HIV-1 protease. Tipranavir lost hydrophobic interactions with Gly 49, Val 32, and Val 84 of the MDR protease but gained additional hydrophobic interactions with Thr 82. Similarly, darunavir failed to maintain hydrophobic interactions with Pro 81 and Ile 84 of the MDR protease but gained hydrophobic contacts with Leu 23. The hydrogen bonding network of tipranavir was generally preserved (Figure 4.2). Tipranavir formed hydrogen bonds with Asn 25, Asp 29, Asp 30, Gly 48, and Ile 50. Except for the hydrogen bonds formed between the hydroxyl group of tipranavir and the catalytic residues of the protease, all other hydrogen bonds were formed on the main chain of the protease. Darunavir lost one hydrogen bond and altered the bonding of another hydrogen bond in the darunavir-MDR 769 82T complex structure. The bicyclic tetrahydrofuran P2 group of darunavir formed a pair of hydrogen bonds with Asp 29 of WT HIV-1 protease. In the MDR protease complex, the dislocation of the P2 group of darunavir causes it to lose one hydrogen bond with Asp 29 (Figure 4.2). The dislocation of the sulfonamide P2’ group lost the hydrogen bond with the carboxyl group of the Asp 29 side chain, but the P2’ group formed a new hydrogen bond with the carbonyl group of the Asp 29 main chain (Figure 4.2).

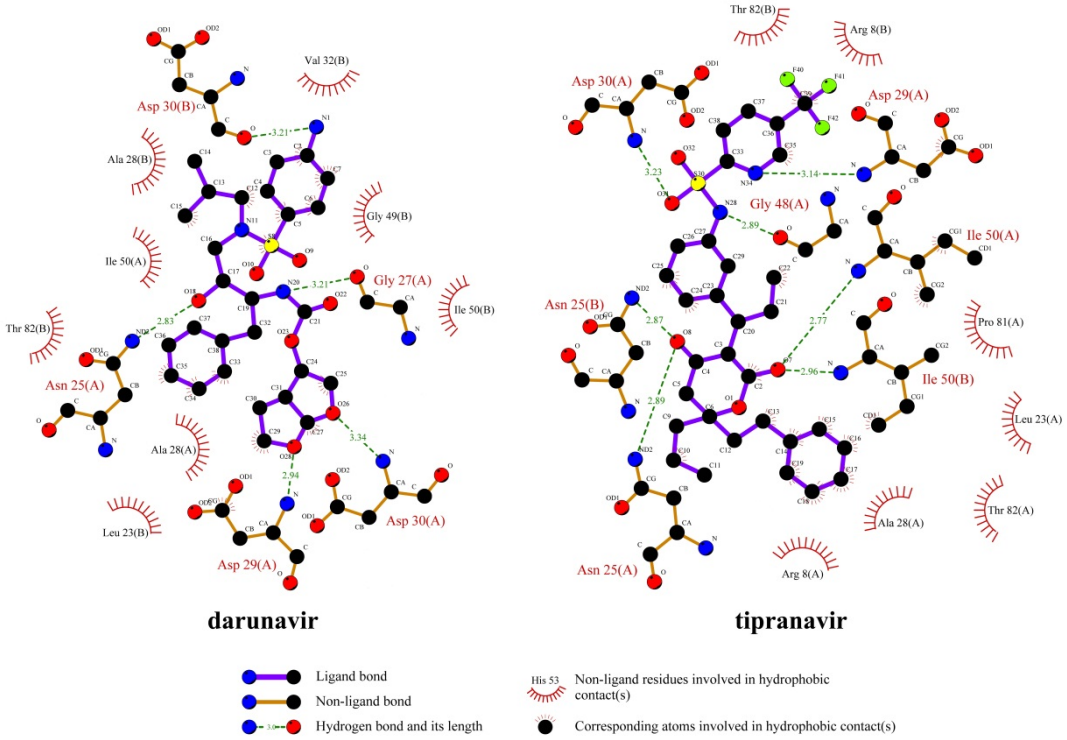


Figure 4.2 Hydrogen bonds and hydrophobic interactions of tipranavir and darunavir with MDR 769 82T protease. Hydrogen bonds are shown as green dashed lines. The residues that form hydrophobic interactions with tipranavir are illustrated by spokes. The interaction map was produced using LIGPLOT V4.5.3.

4.2.4 Tipranavir elevated the melting temperature of the protease-inhibitor complex suggesting enhanced stability

The melting temperatures of the apo-MDR 769 82T and the MDR 769 82T in complex with inhibitors were determined using differential scanning fluorimetry. While eight protease-inhibitor complexes displayed a similar melting temperature to that of the apo-MDR 769 82T (50.8 ± 0.6 °C), the tipranavir-MDR 769 82T complex shifted the melting temperature to a higher value (62.0 ± 2.6 °C). The melting temperature of the darunavir-protease complex was 52.6 ± 0.8 °C. The difference between the melting

temperatures of the two complexes was 9.4 °C suggesting that tipranavir stabilized the protease-inhibitor complex (Figure 4.3).

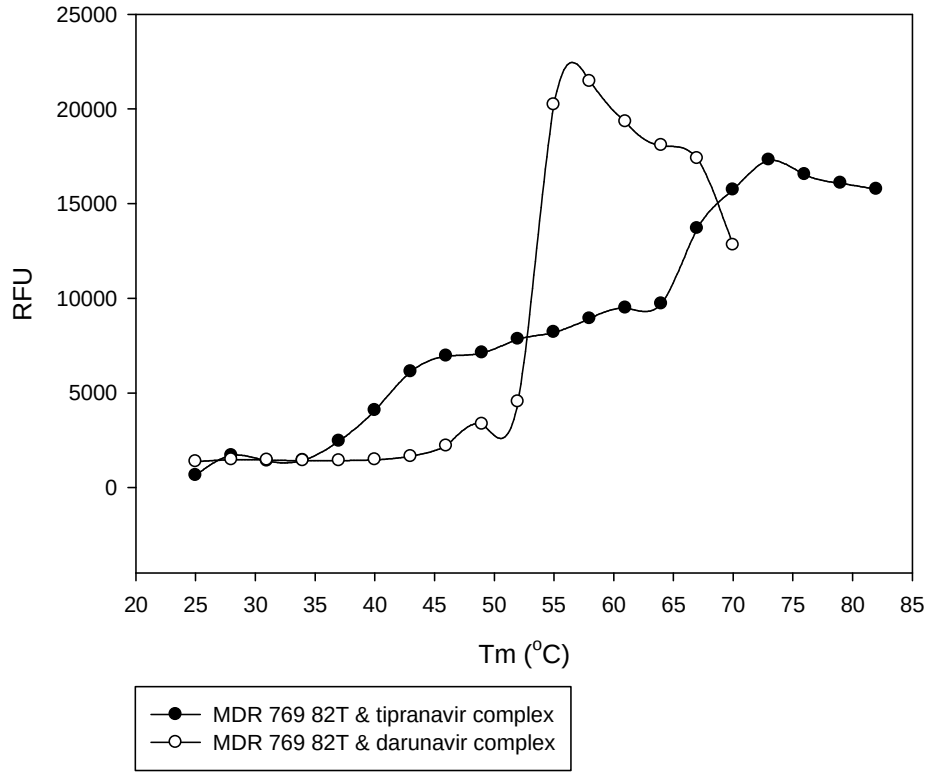


Figure 4.3 The tipranavir-MDR 768 82T complex showed higher denaturation temperature compared to the darunavir-MDR 769 82T complex. The X-axis is temperature in Celsius degree, and the Y-axis is fluorescence change in Relative Fluorescence Units (RFU).

4.3 Discussion

The objective of this study is to investigate how darunavir and tipranavir maintain their inhibitory efficacy against MDR HIV-1 protease. Darunavir preserves its residual inhibitory efficacy by forming altered hydrogen bonds while tipranavir overcomes drug resistance by stabilizing the protease-inhibitor complex through direct interactions with

the protease flaps.

The valine to threonine mutation at residue 82 is an isosteric switch that alters the hydrophilic environment of the S1 and S1’ binding pocket in the protease and greatly enhances the relative resistance to inhibitors. The valine to alanine mutation in MDR 769 is a change from a longer side chain to a shorter side chain, but the hydrophobicity of this mutation is preserved. In both darunavir and tipranavir MDR 769 82T complex structures, the Thr 82 residue moves away from the binding pocket towards the hydrophilic solvent environment exhibiting greater relative resistance than the Ala 82 mutation. Darunavir, tipranavir, and lopinavir inhibit the MDR 769 at sub-nanomolar concentration; however, they only inhibit MDR 769 82T at approximately 3 nM. Therefore, the hydrophobicity change in the binding pocket is more severe than contact loss in the binding pocket.

Darunavir or tipranavir binding to MDR 769 82T effectively triggers the protease flap closure. MDR 769 82T tends to maintain the wide-open form even in the presence of other inhibitors (*105*). The wide-open form of MDR 769 82T is possibly due to the fact that M46L and I54V disrupt the rigid β-sheet of the protease flaps and therefore increase the flexibility of the flaps.

Tipranavir, the only FDA-approved non-peptidomimetic HIV-1 protease inhibitor, does not form the same hydrogen bonds with the flaps through a bridging water molecule as peptidomimetic inhibitors do. Instead, tipranavir forms a hydrogen bond directly between the carbonyl group of tipranavir and the Ile 50 residue on each MDR protease flap. The DSF results indicate that tipranavir greatly stabilizes the inhibitor-protease complex. Therefore, tipranavir may partially overcome drug resistance of the MDR

protease by tightly closing the protease flaps.

Two possible drug design strategies could be considered to improve the potency of inhibitors against MDR proteases based on the advantageous characteristics of darunavir and tipranavir that elevate the resistance barrier to MDR protease. The first strategy is to design chemical groups that make direct contact with the flap residues of MDR proteases rather than mimicking substrates to interact with flaps via a bridging water molecule. Stronger interactions between the inhibitor and protease flaps decrease the probability of the protease-inhibitor complex dissociation. The hydrogen bonds formed directly between the flap residue Ile50 and tipranavir reduce the protease flap flexibility caused by the mutations, demonstrating the stability of the tipranavir-MDR protease complex.

The second strategy to improve the potency of inhibitors against MDR proteases is to design inhibitor interactions with the main chain of the HIV-1 protease by flexible pharmacophores. When the carboxyl group of Asp 29 is far away from the sulfonamide P2’ group of darunavir, the P2’ group preserves a hydrogen bond with the main chain carbonyl group of Asp 29 (Fig. 2). The main chain hydrogen bonds contribute to the ability of darunavir to overcome drug resistance conferred by protease active site mutations (*106*). Other than the P2 group, the darunavir structure is identical to amprenavir. The flexible bicyclic tetrahydrofuran P2 group of darunavir enhances the hydrogen bond forming propensity relative to amprenavir (*27*). According to the darunavir-MDR 769 82T complex, the P2 group of darunavir adopts a different conformation relative to the darunavir-WT HIV-1 protease complex but preserves hydrogen bonds with the main chain of the MDR protease, suggesting that flexible drugs

favorably form contacts in the HIV-1 protease active site cavity and maintain potency.

In summary, the inhibitor-MDR HIV-1 protease complex structures could serve as models for drug optimization. Rather than the *de novo* design of new inhibitors, optimization of existing inhibitors may have a higher chance of success against MDR variants. Minor modifications on the protease inhibitors may restore the contacts without affecting the bioavailability and toxicity of the molecule.

4.4 Author's contribution

The author conducted the protein expression of the inactive MDR 768 82T, active MDR 769 82T, and active MDR 769. The crystallization screening, crystallization setup, structure solution, enzyme assays, DSF experiments, and analysis of results were performed by the author. The crystal diffraction data collection and initial data processing were conducted by Dr. Joseph S. Brunzelle from the Department of Molecular Pharmacology and Biological Chemistry, Feinberg School of Medicine, Northwestern University, Chicago, IL.

CHAPTER 5 IDENTIFICATION OF A LEAD COMPOUND TARGETING A MULTI-DRUG RESISTANT HIV-1 PROTEASE VARIANT USING VIRTUAL SCREENING

5.1 Introduction

The drug-resistance mutations accumulated in the HIV-1 protease during antiretroviral therapy change the protease structure and decrease the efficacy of protease inhibitors. Since most HIV-1 protease inhibitors were designed to mimic the substrate cleavage intermediate, structurally diverse inhibitors are needed to maintain efficacy against the highly flexible protease (*107*). The discovery of new inhibitor scaffolds is the first step to enrich the diversity of HIV-1 protease inhibitors. The HIV-1 protease structures greatly facilitate the discovery of new inhibitors.

Few ligand-free three-dimensional structures of HIV-1 protease have been deposited to the Protein Data Bank. We have previously reported one such structure of an uncomplexed MDR HIV-1 protease, MDR 769 (*108*). As shown in Figure 5.1, the available ligand-free HIV-1 protease can be classified into two classes, wide-open form and curled form (*109*). Considering the flexibility and dynamic movement of the protease, the structures of the uncomplexed HIV-1 protease are as important as the ligand-bound structures needed for successful drug design.

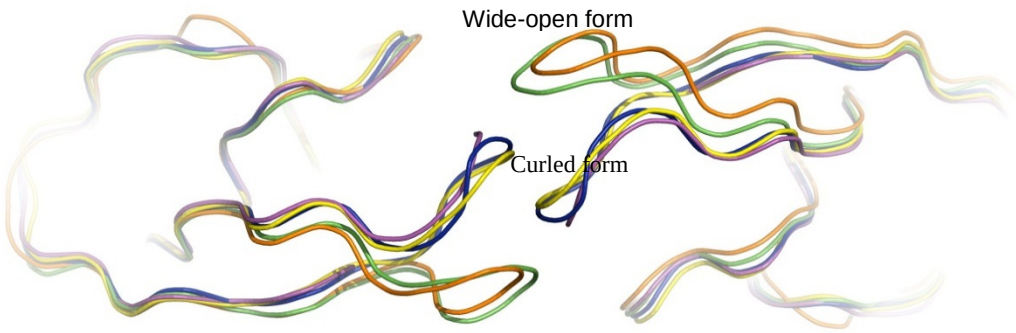


Figure 5.1 The flap conformations of apo HIV-1 protease structures. The structures include the WT HIV-1 protease NL4-3 (PDB ID: 2HB4, blue), the MDR HIV-1 protease with mutations L24I, M46I, F53L, L63P, V77I, and V82A (PDB ID: 2HB2, magenta), the WT HIV-1 protease isolate BRU (PDB ID: 1HHP, yellow), another WT HIV-1 protease (PDB ID: 2PC0, green), and the MDR HIV-1 protease 769 with mutations L10I, M36V, M46L, I54V, I62V, L63P, A71V, V82A, I84V, and L90M (PDB ID: 1TW7, orange)

Understanding the structural diversity of the target protease is essential for inhibitor development. The plasticity of the HIV-1 protease flap region is critical in binding substrates and inhibitors (*109*). The increased flap flexibility may enhance drug cross-resistance by altering the binding pocket of protease inhibitors. Flap movement distorts the binding cavity and therefore reduces the binding affinity to inhibitors. Direct measurement using pulsed double electron-electron resonance (DEER) has confirmed a larger distance between the flaps of MDR HIV-1 protease as compared to the wild-type structure (*110*). Therefore, the wide-open form of HIV-1 protease can serve as a nontraditional model to develop inhibitors targeting the open form protease. A crystal structure has shown that the HIV-1 protease with open-flap conformation binds two

symmetric pyrrolidine diester inhibitors (*111*). The range of opening and conformational flexibility of the protease flaps demonstrate the importance of protein flexibility in structure-based drug design. Using ensembles of protein conformations as receptors for docking is a powerful method to increase accuracy (*112, 113*). Docking inhibitors based on both the open and closed form of HIV-1 MDR protease partially represents the HIV-1 protease dynamics and improves the adaptation of the drug to the flexible movement of the protease.

5.2 Results

The commercially-available compound libraries of TimTec Inc., Adesis Inc., and Enamine Ltd. from the ZINC database (http://zinc.docking.org/) were initially screened using LASSO (Ligand Activity in Surface Similarity Order) 2009 (*82*). Thousands of compounds were selected to generate a smaller library for further validation using eHiTS docking (*82*). The open form MDR HIV-1 protease 769 structure was used as a receptor to identify potential inhibitor scaffolds. One compound, GR346 (Figure 5.2), showed the best docking scores against the open form MDR 769 protease.

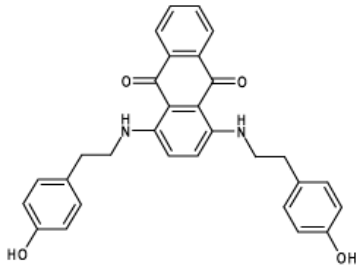
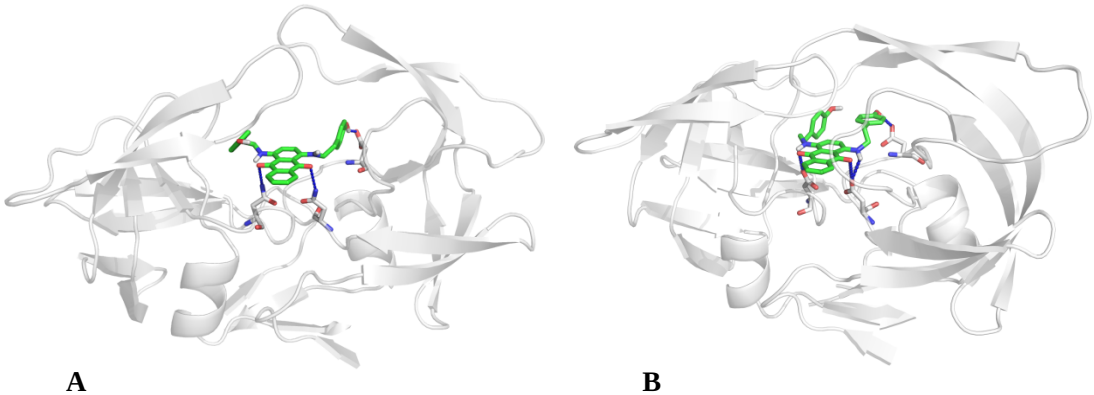


Figure 5.2 Chemical structure of compound GR346.



A B

Figure 5.3 Docking conformation of GR346 in the active site of MDR 769 HIV-1 protease. (A) GR346 was docked to the wide-open form MDR 769. (B) GR346 was docked to the closed form MDR 769 HIV-1 protease. The blue lines represent possible hydrogen bonds.

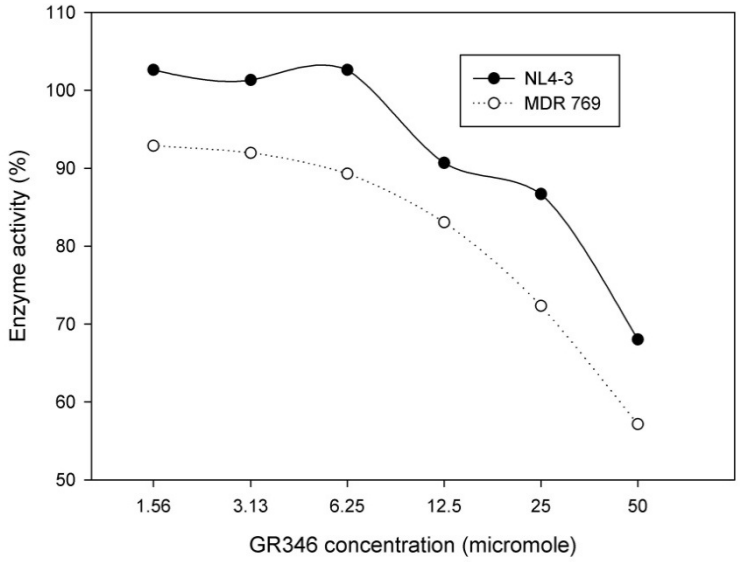


Figure 5.4 HIV-1 protease inhibition by the compound GR346. The MDR 769 protease is inhibited more than the wild type (NL-43) by GR346.

The binding of GR346 was examined by docking it to the closed and wide-open forms of the MDR 769 protease using AutoDock4 (*83*). The docking conformation of GR346 is shown in Figure 5.3. GR346 was originally synthesized as an NAD-dependent

deacetylase sirtuin-2 (SIRT2) inhibitor (IC_{50}=56.7 μM) (*114*). The HIV-1 protease inhibitory efficacy was tested using an *in vitro* enzyme assay (Figure 5.4). GR346 inhibited a larger percentage of MDR 769 enzyme activity at various concentrations compared to its inhibition of the WT HIV-1 protease. At the lower compound concentration (6.25 μM), GR346 did not inhibit the WT HIV-1 protease but it did inhibit the MDR 769 activity by 10%. The wide-open conformation of the MDR 769 flaps allowed the bulky compound, GR346, to easily enter the protease active site cavity. This result suggests the possibility that larger inhibitors could fit better in the expanded active site of MDR HIV-1 protease.

5.3 Discussion

The virtual screening experiments in this chapter have the goal of identifying a novel scaffold as an HIV-1 protease inhibitor. GR346 with micromolar IC_{50} is a good lead compound from screening. The novel protease inhibitor scaffold could be further modified and developed to improve potency. A promising finding is the selective inhibition of GR346 against MDR HIV-1 protease at lower concentration. The protease inhibition assay of GR346 suggests the different ligand preference between MDR HIV-1 protease and WT HIV-1 protease. Thus, it is possible to develop MDR HIV-1 protease specific inhibitors for AIDS patients who fail regular HIV-1 protease inhibitor treatment. The selective binding of GR346 to the MDR HIV-1 protease may be due to the flexibility of the MDR protease flaps. The wider inter-flap distance increases the accessibility of bulky compounds.

Protein flexibility is an important feature in drug design. Designing drugs that make

contacts with the protease flaps to reduce protease flexibility, and targeting the open form protease could be options to overcome drug resistance caused by protease flap conformation and flexibility.

5.4 Author’s contribution

The author performed the virtual screening, docking experiments, enzyme assays, and the analysis of the data.

CHAPTER 6 A HIGHER DESOLVATION ENERGY REDUCES LIGAND RECOGNITION IN MULTI-DRUG RESISTANT HIV-1 PROTEASE

6.1 Introduction

Developing novel HIV-1 protease inhibitors is a current requirement to keep pace with the emergence of drug resistance mutations in the HIV-1 protease. Most of the current HIV-1 protease inhibitors are designed based on substrate mimicking. In addition to mimicking the transition state of the substrate cleavage, inhibitors designed based on the consensus volume of substrates have been developed (*115*). The consensus volume of various substrates is defined as the substrate envelope. The HIV-1 protease substrate envelope has been determined based on the protease-substrate co-crystal structures (*116*). A more realistic dynamic model of the substrate envelope has been refined using molecular dynamics of HIV-1 protease-substrate complex structures (*117*). The study of the substrate binding to various HIV-1 protease variants, especially drug-resistant variants, facilitates the development of drugs inhibiting a broad panel of HIV-1 protease variants. The accumulated drug-resistance mutations in the protease may alter the subsites of the active site cavity, and therefore the inhibitor binding affinity alters accordingly. Studies have shown that the substrate analog inhibitors contact the same set of residues of HIV-1 protease as the natural substrates do (*118*). Therefore, MDR HIV-1 protease variants may also change the binding affinity to the substrates.

Four clinical MDR HIV-1 protease variants, MDR 769, MDR 807, MDR 1385, and MDR 3761, were isolated by Palmer *et al* (*97*), among which MDR 769 exhibited resistance to all the tested inhibitors (*119*) and was successfully crystallized. The accumulation of drug-resistance mutations also alters the structure of the MDR protease variants. The flaps of the apo MDR 769 protease adopt a wide-open conformation (*98*). Based on the available structures and known crystallization conditions, MDR 769 could serve as a model for studying the binding patterns of MDR protease variants.

Peptidomimetic design based on the HIV-1 protease substrates is one of the popular drug design strategies. In our study, four clinical isolates of MDR HIV-1 protease variants were examined. The four MDR HIV-1 protease variants exhibited resistance to all US FDA-approved HIV-1 protease inhibitors, and the substrate recognition ability is also reduced in these proteases. In the experiment of the HIV-1 protease substrate-based peptides competing with a fluorescent substrate, the substrate recognition of various HIV-1 protease variants was illustrated. The results showed different recognition of the substrates p2/NC, CA/p2 P1’F, RH/IN, and RT/RH among the five HIV-1 protease variants. With the co-crystal structures of p2/NC and CA/p2 P1’F with MDR 769 protease, five HIV-1 protease variants in complex with p2/NC and CA/p2 P1’F were modeled and simulated to study the binding of substrates in the protease active site cavity. The simulation results indicated that the MDR proteases need to overcome the higher desolvation energy barrier to form substrate-protease complexes. The binding studies of the heptapeptides p2/NC and CA/p2 P1’F can facilitate the identification of subsite preference of HIV-1 protease variants and the design of drugs inhibiting a broader range

of MDR HIV-1 protease variants.

Table 6.1 Sequences of the HIV-1 protease variants selected

HIV-1 protease	Sequences*
NL4-3	PQITLWKRPL VTIKIGGQLK EALLDTGADD TVLEEMNLPG RWKPKMIGGI
	GGFIKVRQYD QILIEICGHK AIGTVLVGPT PVNIIGRNLL TQIGCTLNF
769	PQITLWKRPI VTIKIGGQLK EALLDTGADD TVLEEVNLPG RWKPK**L**IGGI
	GGF**V**KVRQYD QVPIEICGHK VIGTVLVGPT P**A**N**V**IGRNL**M** TQIGCTLNF
807	PQITLWKRPI VTIKIGGQLK EALLDTGADD TVLEEMNLPG KWKPK**I**IVGI
	GGF**T**KVRQYD NVQIEICGHK VIG**A**VLIGPT P**A**NIIGRNLL TQLGCTLNF
1385	PQITLWKRP**F** VTIKIGGQLK EALLDTGADD TVLEEIDLPG RWKPK**I**IGGI
	GGFIKVKQYD QIPIEICGHK VIGTVLVGPT P**T**NIIGRNM**M** TQLGCTLNF
3761	PQITLWKRPI VAIKVGGQII EALLDTGADD TVLEEMNLPG RWKPK**I**IGGI
	GGFIKVRQYD QIP**V**EICGHK **I**I**T**TVLVGST PVN**V**IGRNL**M** TQLGCTLNF

*The polymorphic changes are underlined. The drug-resistance mutations are in bold.

Table 6.2 Sequences of the nine HIV-1 protease cleavage sites within the HIV-1 Gag-Pol polyprotein

substrate	P3	P2	P1	P1’	P2’	P3’	P4’
MA/CA	Gln	Asn	Tyr	Pro	Ile	Val	Gln
CA/p2*	Arg	Val	Leu	Phe	Glu	Ala	Met
p2/NC	Thr	Ile	Met	Met	Gln	Arg	Gly
NC/p1	Gln	Ala	Asn	Phe	Leu	Gly	Lys
p1/p6	Gly	Asn	Phe	Leu	Gln	Ser	Arg
TF/PR	Phe	Asn	Phe	Pro	Gln	Ile	Thr
PR/RT	Leu	Asn	Phe	Pro	Ile	Ser	Pro
RT/RH	Glu	Thr	Phe	Tyr	Val	Asp	Gly
RH/IN	Lys	Ile	Leu	Phe	Leu	Asp	Gly

The cleavage site is between the P1 and P1’ residue. The CA/p2 contains an alanine to phenylalanine mutation at P1’ position.

6.2 Results

6.2.1 The four clinical MDR HIV-1 protease isolates are resistant to all FDA-approved HIV-1 protease inhibitors

The drug-resistance profiles of the four clinical MDR HIV-1 protease isolates were identified using enzyme inhibition assays. The four MDR protease variants are resistant to all of the FDA-approved HIV-1 protease inhibitors at various levels (Figure 6.1, Table 6.3). In the table, the HIV-1 protease variant NL4-3 represents a WT HIV-1 protease. Regarding to the inhibitory efficacy, the second generation of HIV-1 protease inhibitors (tipranavir, darunavir, lopinavir, and atazanavir) encountered lower relative resistance. Among the four MDR HIV-1 protease variants, MDR 807 and MDR 1385 were more resistant to the second generation HIV-1 protease inhibitors. For the first-line HIV-1 protease inhibitors (*28*), higher resistance was observed to atazanavir compared to darunavir. These results confirmed the *in vitro* resistance of the HIV-1 protease clinical isolates.

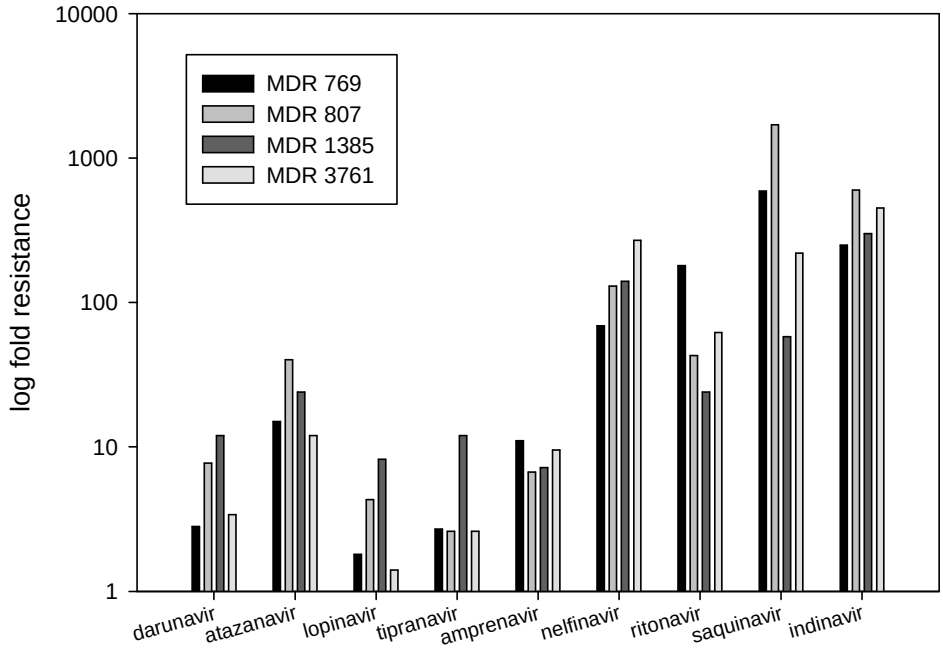


Figure 6.1 Relative resistance of the MDR HIV-1 protease variants. The y-axis is relative resistance on a logarithmic scale while on the x-axis are the names of the FDA-approved HIV-1 protease inhibitors.

Table 6.3 IC_{50} and relative resistance of FDA-approved HIV-1 protease inhibitor against HIV-1 protease variants

HIV-1 Proteases	IC_{50} of HIV-1 protease inhibitors* in nM (relative resistance)								
	darunavir	atazanavir	lopinavir	tipranavir	nelfinavir	amprenavir	saquinavir	indinavir	ritonavir
NL4-3	0.26 (1)	0.19 (1)	0.28 (1)	0.24 (1)	1.6 (1)	0.43 (1)	0.50 (1)	0.47 (1)	0.34 (1)
MDR 769	0.74 (2.8)	2.9 (15)	0.50 (1.8)	0.65 (2.7)	110 (69)	4.8 (11)	290 (590)	120 (250)	61 (180)
MDR 807	2.0 (7.7)	7.6 (40)	1.2 (4.3)	0.63 (2.6)	210 (130)	2.9 (6.7)	850 (1700)	280 (600)	14.6 (43)
MDR 1385	3.0 (12)	4.6 (24)	2.3 (8.2)	2.8 (12)	230 (140)	3.1 (7.2)	29 (58)	140 (300)	8.0 (24)
MDR 3761	0.89 (3.4)	2.2 (12)	0.39 (1.4)	0.63 (2.6)	430 (270)	4.1 (9.5)	110 (220)	210 (450)	21 (62)

*The protease inhibitors were requested from the NIH AIDS Research and Reference Reagent Program (http://www.aidsreagent.org/).

6.2.2 The MDR HIV-1 protease isolates exhibited different substrate binding preference relative to the WT protease

Compared to inhibitors, the higher flexibility of substrates could match the dynamic changes in the protease (*107*). The study of substrate binding facilitates the development of peptidomimetic inhibitors. In the FRET substrate cleavage interference experiments, the five HIV-1 protease variants show varied preferences to the nine heptapeptides. The processing ratio in Figure 6.2 represents the ratio of the average FRET substrate processing velocity in the presence of the heptapeptide over the average FRET substrate processing velocity in the absence of the heptapeptide. The higher ratio indicated less interference by the peptide on FRET substrate processing. Both the peptide and the FRET substrate were in excess molar ratio to the HIV-1 protease, and the peptide concentration was 40-fold higher than the FRET substrate concentration. The rate-limiting step of the HIV-1 protease substrate, catalysis, is a chemical process rather than a physical process (*120*). Therefore, interference was caused by the peptide cleavage process rather than the peptide binding process. Once a protease-substrate complex is formed, the cleavage of FRET substrate is slowed down. The high velocity ratios of the FRET substrate cleavage with or without the regular peptide were due to the lower chance of forming protease-substrate complex with the peptide. The complex formation theory was supported by the enzyme assays using the uncleavable CA/p2 P1’F peptide. The inhibition of the uncleavable peptide is correlated with the peptide competition results. The IC_{50} values of CA/p2 P1’F pseudopeptide for NL4-3, MDR 769, MDR 807, MDR 1385, and MDR 3761 were 2.6 nM, 4.4 nM, 16.0 nM, 32.5 nM, and 55.4 nM, respectively.

The corresponding relative resistance values were 1.7, 6.2, 13, and 21 fold, respectively. The correlation coefficient between the FRET velocity ratios in presence of CA/p2 P1'F and the IC_{50} values of CA/p2 P1'F pseudopeptide was 0.86.

The ratios of heptapeptide interference exhibited the likelihood of protease-substrate complex formation (Figure 6.2). Compared to the WT protease, the MDR protease variants were interfered at lower levels by the presence of regular peptides indicating the impaired ability of the MDR HIV-1 protease to form a complex with the substrates. All five HIV-1 protease variants had higher probability of forming complexes with the CA/p2 P1'F peptide. The presence of CA/p2 P1'F significantly slowed down the processing of the FRET substrate. Especially, the NL4-3 and MDR 769 were completely occupied by CA/p2 P1'F, and the fluorescent signal maintained as the base line of background signal. The MDR 807, MDR 1385, and MDR 3761 had higher residual velocities of FRET substrate processing with the interference of CA/p2 P1'F, suggesting the decreased probability of complex formation. The MDR 807, MDR 1385, and MDR 3761 exhibited a lower probability of forming complex with the substrate p2/NC compared to the WT protease and MDR 769. All four MDR protease variants were unfavorable to form complex with the RT/RH and RH/IN.

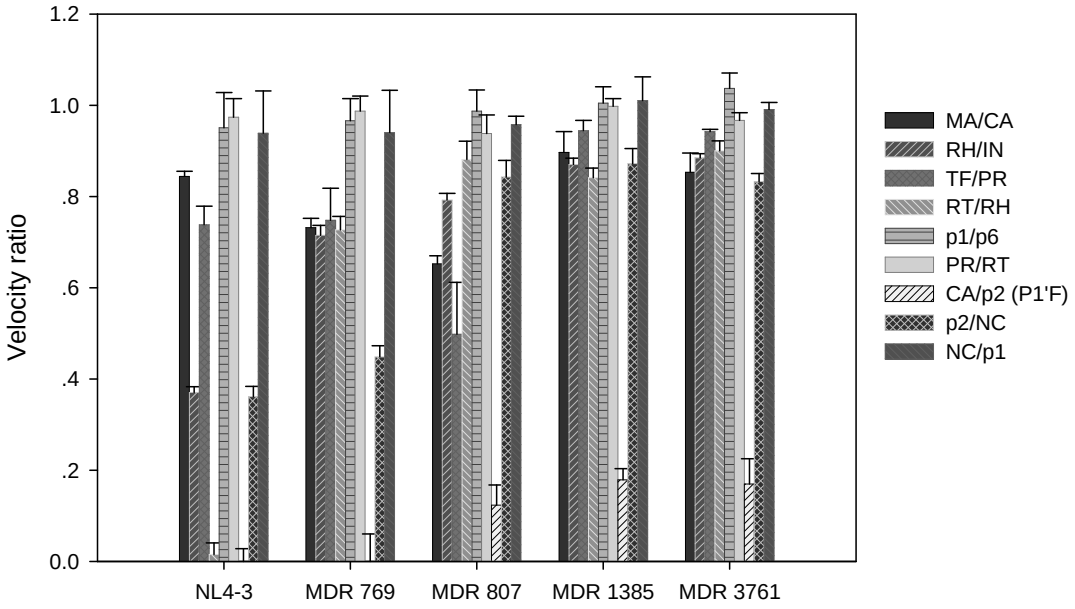


Figure 6.2 Förster resonance energy transfer substrate processing ratio. The bar chart represents the ratio of the average FRET substrate cleavage velocity in the presence of the regular peptide substrate over the average FRET substrate cleavage velocity in the absence of the regular peptide substrate. When the velocity ratio is one, the regular peptide does not affect the FRET substrate cleavage by the protease. When the velocity ratio is zero, the FRET substrate is completely competed out by the regular peptide.

6.2.3 The MDR protease-substrate co-crystal structures were insufficient to explain the different substrate binding behaviors between the MDR and WT HIV-1 proteases

Based on substrate competition experiments, the substrates of interest were selected to co-crystallize with the MDR HIV-1 protease variant. The peptides CA/p2 P1'F and p2/NC were successfully crystallized with MDR 769 82T in $P2_12_12_1$ space group. The CA/p2 P1'F-MDR 769 82T and p2/NC-MDR 769 82T co-crystals diffracted to 2.10 Å

and 2.30 Å, respectively (Table 6.4). By superposing the MDR-substrate complex structures on corresponding WT protease-substrate complexes, the substrate backbones were overlapped in both the MDR and WT complexes, and only minor conformational differences were noticeable in the long flexible side chains (Figure 6.3 A, B). The P1' group of p2/NC in complex with MDR 769 was the major conformational deviation as compared to that of the WT complex. The results suggested that the static structures of binding conformation are insufficient to explain the difference in substrate recognition of HIV-1 protease variants. Dynamic simulations are required to identify the sophisticated differences in substrate binding among HIV-1 protease variants.

Table 6.4 Crystallographic statistics

Dataset	**The MDR 769 82T in complex of substrate CA/p2**	**The MDR 769 82T in complex of substrate p2/NC**
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Wavelength (Å)	0.979	0.979
Cell constants (Å)	a=28.76 b=65.38 c=92.80	a=28.62 b=63.85 c=91.11
Resolution range (Å)	30.00-2.10(2.14-2.10)	30.00-2.30 (2.38-2.30)
Number of unique reflections	10882 (507)	7945 (787)
Completeness (%)	99.9 (99.0)	98.8 (99.5)
Redundancy	7.6 (6.0)	4.0 (4.0)
Mean I/σ (*I*)	13.2 (3.4)	10.0 (2.4)
R_{merge} [a]	0.162 (0.520)	0.114 (0.451)
Refinement		
R_{work} (%)[b]	17.29	20.00
R_{free} (%)[b]	23.99	27.86
Number of atoms		
Ligand	60	56
Protease	1529	1529
Solvent	258	137
Average isotropic B factor (Å^2)		
Ligand	20.28	35.32
Protease	16.83	30.46
Solvent	32.62	45.32
RMSD bond length (Å)	0.008	0.009
RMSD bond angle (°)	1.06	1.26
Ramachandran plot		
Allowed/generous/disallowed (%)	100/0/0	99.0/1.0/0

[a] $R_{merge} = \Sigma_{hkl} \Sigma_i |I_i(hkl) - \langle I(hkl) \rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation and $I(hkl)$ is the mean value for its unique reflection.

[b] $R_{work} = \Sigma_{hkl} ||F_o| - |F_c|| / \Sigma_{hkl} |F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes. R_{free} is calculated exactly as R_{work} using a random 5% of the reflections omitted from refinement.

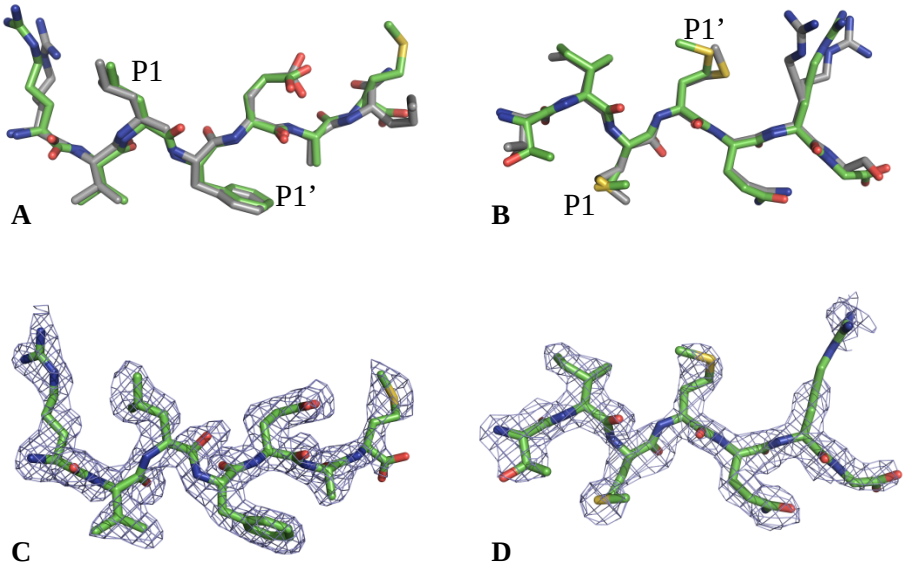


Figure 6.3 Substrate conformation illustrating binding to the MDR protease. (A) The CA/p2 P1'F binding to the MDR 769 82T (green) comparing to its binding to a WT protease (grey, PDB ID: 1A8K). (B) The p2/NC binding to the MDR 769 82T (green) comparing to its binding to a WT protease (grey, PDB ID: 1KJ7). (C) The electron density of CA/p2 P1'F. The mesh is an *Fo-Fc* OMIT map at 2.0 σ. (D) The electron density of p2/NC. The mesh is an *Fo-Fc* OMIT map at 2.0 σ.

6.2.4 The desolvation energy required by the MDR HIV-1 protease variants to form protease-substrate complexes correlated with the substrate binding assay

Homology complex models of MDR 769, MDR 807, MDR 1385, MDR 3761, and NL4-3 with p2/NC or CA/p2 were built to analyze the dynamic interactions between substrate and protease. After 10 ns molecular dynamics simulation, the HIV-1 protease complex models became relatively stable (Figure 6.4, 6.5), and the movement of substrates was analyzed for the last 40 ps of simulation. The RMSD of each residue of p2/NC or CA/p2 was plotted in Figure 6.6. The lower RMSD standard deviation suggests

a more stable substrate-protease complex. The RMSD standard deviation values of CA/p2 binding to MDR proteases were comparable to that of CA/p2 binding to the WT HIV-1 protease. Only the Met at P4' site was unstable in the MDR-CA/p2 complexes. This result indicates that CA/p2 stabilizes MDR protease complexes and could serve as a template to be further developed as a resistance overcoming peptidomimetic inhibitor. The RMSD standard deviation values of p2/NC showed that P1 to P3 residues were stable in both the WT HIV-1 protease and MDR proteases while the P1' to P3' residues were more unstable in the MDR proteases compared to that in the WT HIV-1 protease. The P4 Gly, the tail of p2/NC, underwent large movements.

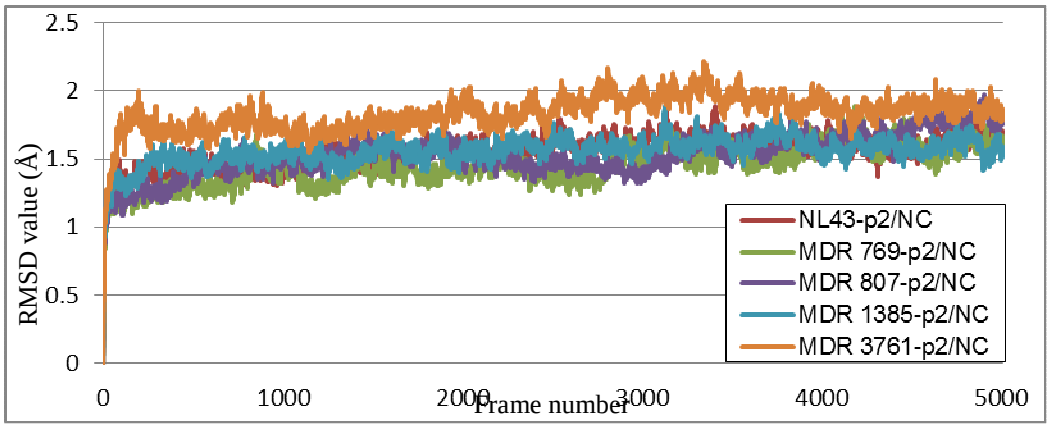


Figure 6.4 RMSD values of the HIV-1 protease-p2/NC complexes.

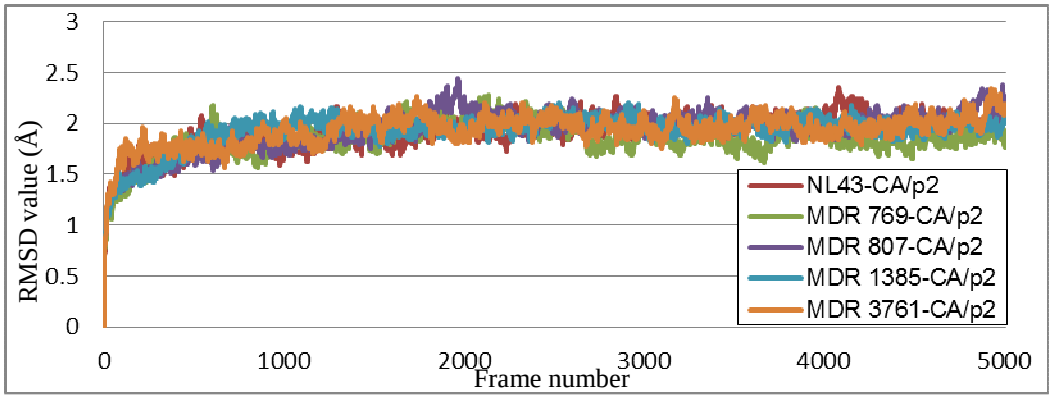


Figure 6.5 RMSD values of the HIV-1 protease-CA/p2 complexes.

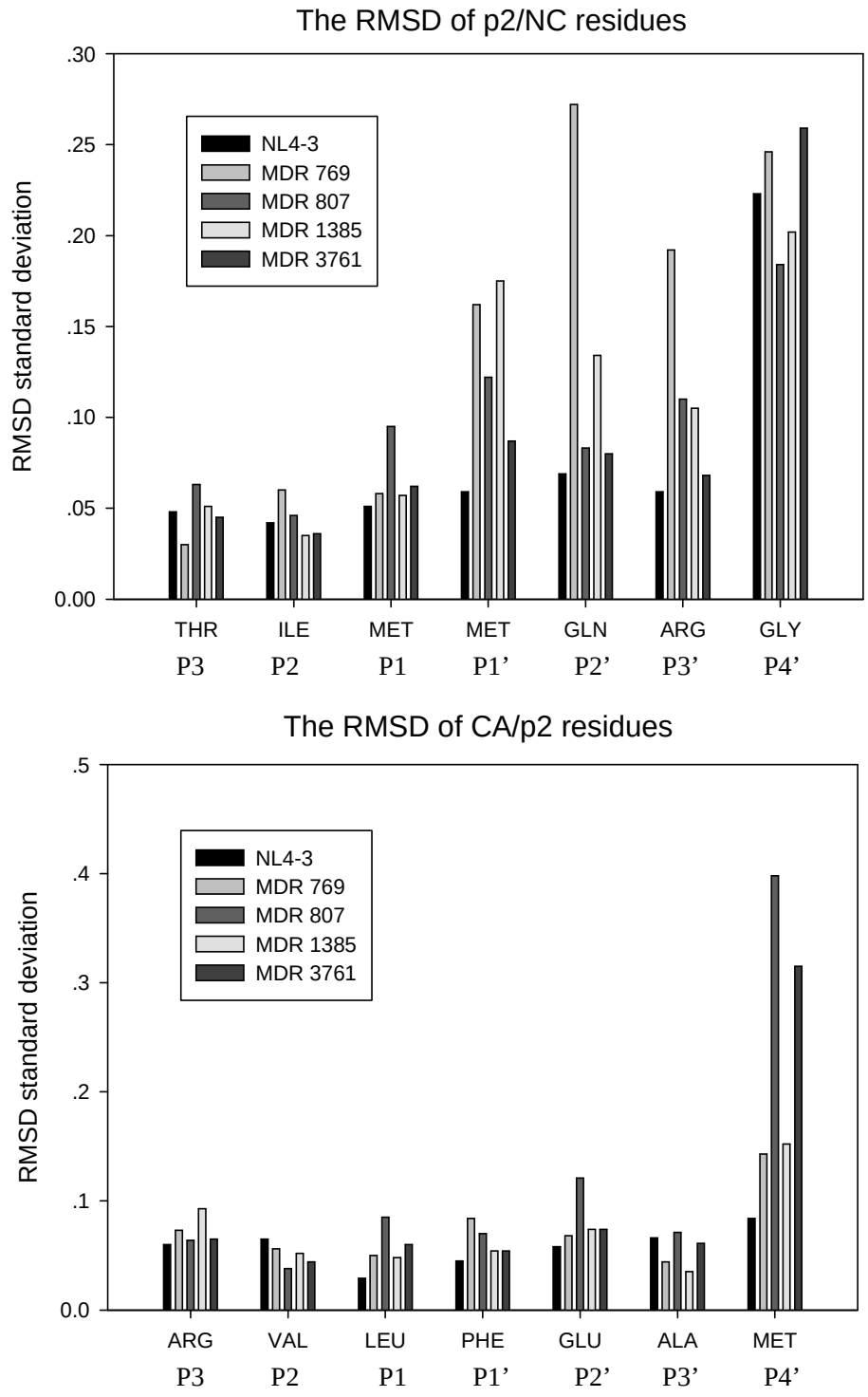


Figure 6.6 RMSD standard deviation of substrate residues in the active site of HIV-1 protease variants.

The interaction energy and desolvation energy between the substrates (p2/NC and CA/p2) and HIV-1 protease variants were calculated based on the molecular mechanics/Poisson–Boltzmann solvent-accessible surface area (MM-PBSA) method (*121*). The interaction energies of the MDR 807-p2/NC and MDR 1385-p2/NC complexes were stronger than those of other protease variant complexes. The stronger interactions were contributed by the stronger electrostatic interaction energies. The pattern of CA/p2 binding energies were different form that of p2/NC. The CA/p2 peptide bound to MDR proteases were stronger resulting in stronger electrostatic interaction energies.

The electrostatic desolvation energy is an unfavorable contribution to the binding free energy of HIV-1 protease-substrate complexes. Compared to the electrostatic desolvation energy, the non-polar desolvation energy was not a large component for the binding of HIV-1 protease to p2/NC or CA/p2. The desolvation energy changes for MDR proteases in complex with CA/p2 were higher than that for the WT complex. The p2/NC complexes with MDR 807, MDR 1385, and MDR 3761 were calculated to have higher desolvation energy than MDR 769 and NL4-3 complexes. Generally, MDR proteases required to overcome higher desolvation barriers to bind substrates.

Table 6.5 Energy analysis of the HIV-1 protease-p2/NC complex

Terms of binding free energy	HIV-1 protease				
	NL4-3	MDR 769	MDR 807	MDR 1385	MDR 3761
ΔG_{int}^{elec}	-103.7±11.9	-109.0±13.8	-136.5±22.1	-134.5±18.6	-82.5±11.9
ΔG_{int}^{vdW}	-78.1±3.4	-73.4±3.8	-77.3±4.0	-69.4±3.5	-77.6±3.4
ΔG_{int}	-181.7±11.6	-182.4±13.3	-213.9±21.5	-203.9±17.6	-160.1±12.0
ΔG_{desolv}^{elec}	176.0±8.9	184.2±9.6	203.9±17.8	223.2±17.7	261.6±9.1
$\Delta G_{desolv}^{nonpolar}$	-7.6±0.1	-7.6±0.2	-8.0±0.2	-7.4±0.2	-7.6±0.1
ΔG_{desolv}	168.4±8.9	176.6±9.6	195.9±18	215.8±18	254.0±9.1

Table 6.6 Energy analysis of the HIV-1 protease-CA/p2 complex

Terms of binding free energy	HIV-1 protease				
	NL4-3	MDR 769	MDR 807	MDR 1385	MDR 3761
ΔG_{int}^{elec}	-158.7±27.6	-191.7±31.9	-193.8±25.1	-206.4±20.6	-169.3±18.5
ΔG_{int}^{vdW}	-67.8±4.4	-66.6±4.4	-69.0±3.7	-59.3±4.4	-66.1±3.9
ΔG_{int}	-226.5±25.9	-258.3±31.5	-262.8±24.6	-265.7±19.5	-235.5±17.5
ΔG_{desolv}^{elec}	281.0±24.3	282.1±24.5	313.8±14.1	298.4±18.8	332.3±16.6
$\Delta G_{desolv}^{nonpolar}$	-7.6±0.1	-8.0±0.2	-7.5±0.2	-6.8±0.1	-7.2±0.1
ΔG_{desolv}	273.4±24	274.1±25	306.3±14	291.6±19	325.1±17

The interaction energies (ΔG_{int}) between the substrate and HIV-1 proteases are the combination of electrostatic interaction energy change (ΔG_{int}^{elec}) and van der Waals

interaction energy change (ΔG_{int}^{vdW}). The interaction energies of the p2/NC or CA/p2 with the MDR protease variants were not weaker compared to the interaction energy with the WT HIV-1 protease (Table 6.5, 6.6). Except for the interaction energy between p2/NC and MDR 3761, the interaction energies with all the other MDR protease variants were stronger than the WT protease interaction energy, which does not explain the drug resistance and the results of substrate competition experiments. However, the energy barrier of desolvation correlated well with velocity ratios in substrate competition experiments. The desolvation energy (ΔG_{desolv}) includes the non-polar desolvation energy ($\Delta G_{desolv}^{nonpolar}$) and the electrostatic desolvation energy (ΔG_{desolv}^{elec}). The MDR proteases require a higher desolvation energy to remove the water shell in the active site to form complexes with the substrate (Table 6.5, 6.6). Pearson's correlation coefficient between the velocity ratios for the p2/NC peptide (Figure 6.2) and the desolvation energies of the p2/NC-MDR protease complexes is 0.78. Pearson's correlation coefficient for the CA/p2 dataset is 0.81.

6.3 Discussion

The study of MDR HIV-1 protease substrate preference in binding and catalysis improves the understanding of MDR HIV-1 protease evolution as well as the effective inhibition of the MDR protease variants. Novel drugs could be designed or optimized according to the favorable binding of substrates to the MDR HIV-1 protease.

The identification of the substrate sub-site preference of MDR HIV-1 proteases helps in the design of drugs that specifically inhibit MDR HIV-1 proteases. The MDR HIV-1 protease inhibitors may benefit patients who fail the regular protease inhibitor-containing

antiretroviral therapy. According to the movement of p2/NC residues as measured by RMSD in the protease active site (Figure 6.6), the P1’ Met is not favorable in MDR 769 and MDR 807, and P2’ Gln is not favorable in MDR 769 either. The active site blocking behavior of CA/p2 and its interaction energy with MDR proteases suggests that the CA/p2 peptide easily forms a strong complex with HIV-1 proteases. The tighter CA/p2 binding to the MDR protease variants as well as the reduced movement of CA/p2 in the MDR protease active site also supports the idea that the MDR HIV-1 protease-specific inhibitors against a panel of MDR HIV-1 protease variants could be developed using the CA/p2 as an initial template for drug-optimization.

During the viral maturation, the processing of Gag and Gag-pol polyprotein is an ordered process (*122*). The WT HIV-1 protease should maintain varying binding affinities to the different substrates. However, in the substrate interference assays, the FRET substrate cleavage interference pattern of the WT HIV-1 protease is not preserved in the assays of MDR proteases. MDR proteases acquire drug-resistance by decreasing the probability of binding to inhibitors. It is also possible for the mutations and polymorphic substitutions to decrease the ability of MDR HIV-1 protease to bind substrates. The change in regulated substrate site processing could be a mechanism to compensate for the loss of viral fitness and could be a mechanism to enlarge the MDR HIV-1 variant quasispecies populations in an infected individual. However, the short substrate peptides used in this study may not reflect the processing order of the polyproteins cleavage sites because the structural context and protease accessibility were not evaluated.

The desolvation energy is an important component for substrate and inhibitor recognition. The higher unfavorable desolvation energies of CA/p2-MDR complexes explain the resistance to the uncleavable CA/p2 reduced peptide. In the dynamic process of ligand binding, desolvation energy plays an important role in the ligand entry and protease-ligand complex formation. The MDR proteases required a higher energy to desolvate the substrate and the protease active site. The higher desolvation barrier for MDR proteases increases the difficulty of protease-ligand complex formation. The input structures for simulation were modeled based on the closed form MDR 769. Both the crystal structure and direct measurement using pulsed double electron-electron resonance (DEER) have confirmed a larger distance between the flaps of MDR 769 as compared to the wild-type protease structure (*110*, *123*). The wide-open conformation of the MDR protease in the solution exposes more active site regions to solvent which may further elevate the desolvation barrier. Since the calculation of interaction energy and desolvation energy were performed using different methods, it is not appropriate to compare directly between the two datasets. However, the trends of energy difference among HIV-1 protease complexes are good indicators for interpretation.

In conclusion, the difficulty in protease-ligand complex formation is increased for MDR protease due to the high desolvation energy barrier. Drug design purely based on the enzyme of inhibited state may not be able to reproduce the dynamic interactions during the formation of enzyme-inhibitor complex. The results in this chapter suggest that the desolvation of the HIV-1 protease active site is an important step in protease-ligand complex formation as well as drug resistance. Since it is computationally costly to

simulate the process of inhibitor-enzyme recognition, desolvation energy could serve as a simplified parameter to be considered in drug design. The evaluation of desolvation energy could be utilized as one standard to screen drug candidates.

6.4 Author's contribution

The author conducted the protein expression of the inactive MDR 768 82T, active MDR 769, active MDR 807, active MDR 1385, and active MDR 3761 HIV-1 proteases. The crystallization screening, crystallization setups, structure solution, enzyme inhibition assays, substrate competition assays, molecular dynamics simulations, and the analysis of the results were performed by the author. The crystal diffraction data collection and initial data processing were done by Dr. Joseph S. Brunzelle from the Department of Molecular Pharmacology and Biological Chemistry, Feinberg School of Medicine, Northwestern University, Chicago, IL.

CHAPTER 7 STUDY OF TELAPREVIR POTENCY AGAINST NON-GENOTYPE 1 HEPATITIS C VIRUS PROTEASES

7.1 Introduction

Hepatitis C Virus (HCV) infection is more prevalent than HIV infection; HCV has infected approximately 170 million people worldwide (*124*). About 60%~80% of the persons infected with HCV have a chronic infection (*125*). To eliminate the HCV caused public health burden, effective treatments are needed. The classic hepatitis C viral therapy involves a combination drug cocktail; pegylated interferon-α-2a or pegylated interferon-α-2b is paired with ribavirin in a 24-to-48 week regimen. The length of the treatment is contingent upon the nature of viral genotype. The prospects of treatment for hepatitis C have improved recently, particularly with the advent of such drugs as Merck & Co., Inc.'s boceprevir (FDA approved on May 13, 2011) and Vertex Pharmaceuticals' telaprevir (FDA approved on May 23, 2011). Both telaprevir and boceprevir are potent HCV protease inhibitors. The HCV serine protease NS3-4A is a heterodimer comprised of two HCV-encoded proteins, the catalytic subunit NS3 and the activation subunit NS4A. Furthermore, the HCV NS3/4A protease employs a serine protease hydrolytic mechanism. Zinc ions are instrumental in maintaining the structural integrity of the NS3, although they do not play a catalytic role as they are situated at a distance from the active site. Serine protease function is quite varied, but all enzymes employ a catalytic, active site

serine residue that plays a central role in cleaving the substrate of choice. HCV protease cleaves at four major junctions on the viral polyprotein precursor, which include the NS4A/NS4B, NS4B/NS5A, NS5A/NS5B linkages, and a self-cleavage at NS3/NS4A. The HCV serine protease NS3/4A is broadly considered to be among the most attractive targets for anti-viral, oral drugs. Telaprevir inhibits HCV serine protease via covalent bonding to the serine hydroxyl oxygen. Since telaprevir was developed based on genotype 1 HCV protease, its efficacy with the other genotypes of HCV proteases is not fully understood.

The focus of this study is to illustrate a comparison of drug resistance, among different HCV genotypes, via molecular modeling and dynamics. Data gleaned from the simulation can expand the present knowledge base surrounding the virus treatment by providing new insight into yet untested interactions between these new inhibitors and various HCV genotypes. Telaprevir was initially modeled with the HCV serine protease NS3/4A of several different genotypes and subsequently subjected to molecular dynamics that simulated the interaction between protease and inhibitor. The molecular dynamics simulation allows a glimpse into the movement and interactions of the HCV protease-telaprevir complex over time.

7.2 Results

7.2.1 HCV NS3/4A protease expression and crystallization screening

The constructs for HCV genotype 4a, 5a, and 6a proteases were designed for NS3/4A fusion proteases. The sequences of the NS3 protease domain and the NS4A peptide for each HCV genotypes were selected based on the closest similarity to the consensus

sequence of a particular HCV genotype. The representative sequences for genotype 4a, 5a, and 6a HCV NS3/4 proteases were determined and listed in Table 7.1.

The HCV NS3/4A protease from genotype 4a and 6a were made by AnaSpec, Inc. (Fremont, CA). The expression of HCV genotype 5a protease was unsuccessful due to low solubility. The HCV genotype 4a protease, whose concentration is higher compared to that of HCV genotype 6a protease, was sent to Hauptman Woodward Institute (Buffalo, NY) for a high-throughput crystal-growth screening. However, no crystallization hits were obtained from the screening results. Moreover, no crystal grew in the grid screening of previously published crystallization conditions of HCV genotype 1b NS3/4A protease.

Table 7.1 Representative NS3/4A fusion protease sequences for HCV genotypes 4a, 5a, and 6a

Genotype	Strain	Accession	NS3/4 fusion protease Seq.*
4a	01-09	DQ418782	**MGSSHHHHHHSSGLVPRGSHM**GSVVIVGRVVLSGSGSITAYAQQTRGLFSTIITSLTGRDTNENCG EVQVLSTATQSFLGTAVNGVMWTVYHGAGSKTISGPKGPVNQMYTNVDQDLVGWPAPPGVKSLTP CTCGASDLYLVTRHADVVPVRRRGDTRGALLSPRPISTLKGSSGGPLLCPMGHAAGLFRAAVCTRG VAKAVDFVPVESLETTMRS
5a	EUH1480	Y13184	**MGSSHHHHHHSSGLVPRGSHM**GSVAIVGRIILSGSGSITAYAQQTRGVLGAIVLSLTGRDKNEAEG EVQFLSTATQTFLGICINGVMWTLFHGAGSKTLAGPKGPVVQMYTNVDKDLVGWPSPPGKGSLTRC TCGSADLYLVTRHADVIPARRRGDTRASLLSPRPISYLKGSSGGPIMCPSGHVVGVFRAAVCTRGVA KALEFVPVENLETTMRS
6a	6a67	DQ480520	**MGSSHHHHHHSSGLVPRGSHM**GCVVICGRITLTGSGSITAYAQQTRGLVGTIVTSLTGRDKNEVEG EVQVVSTATQSFLATSINGVMWTVYHGAGSKTLAGPKGPVCQMYTNVDKDLVGWPSPPGARSLTP CTCGSSDLYLVTREADVIPARRRGDNRAALLSPRPISTLKGSSGGPIMCPSGHVVGLFRAAVCTRGVA KSLDFIPVENMETTMRS

*The N terminal sequence in bold is the His tag and thrombin cleavage site from the pET28 vector.

7.2.2 The telaprevir derivative retained sufficient potency to inhibit genotype 4 and 6 HCV NS3/4A proteases

Enzyme IC_{50} of HCV proteases from genotypes 2 and 3 were measured by Dr. Ladislau C. Kovari, during his sabbatical with Vertex Pharmaceuticals. Enzyme assays showed the potency of telaprevir against HCV NS3/4A protease of genotype 2 was similar to that of telaprevir against HCV genotype 1 proteases, but telaprevir exhibited up to 10-fold less activity against genotype 3 proteases (Figure 7.1).

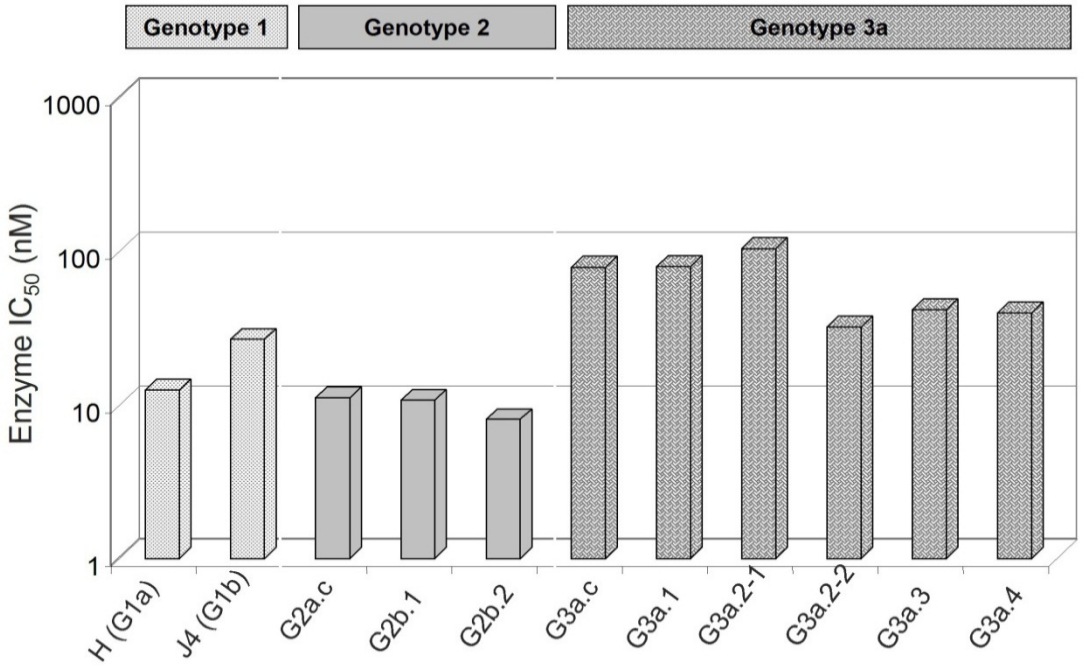


Figure 7.1 Genotype 1, 2, or 3 HCV protease sensitivities to telaprevir.

To further investigate the potency of telaprevir in the treatment of HCV other than genotype 1, FRET substrate cleavage assays were performed to measure the IC_{50} against genotype 4 and 6 HCV proteases. The treatment efficacy of inhibitor against HCV genotypes from 4 and 6 can be inferred from the *in vitro* enzyme assays. Vertex Pharmaceuticals kindly provided us a second generation inhibitor based on telaprevir

named PI-1 (Figure 7.2). Inhibitor PI-1 is a specific, competitive, reversible, peptidomimetic inhibitor of HCV NS3/4A protease, which is structurally similar to telaprevir. PI-1 also shares the inhibition mechanism. The α-ketoamide of PI-1 forms a reversible covalent bond with the catalytic serine in the active site. The inhibitory efficacy of PI-1 against NS3/4A proteases of HCV genotype 1a, 4a, and 6a was tested. The IC_{50} of PI-1 against genotype 1 HCV NS3/4A protease was determined as 5.54 nM. Decreased inhibitory efficacy of PI-1 against genotype 4 and 6 HCV proteases was observed. The IC_{50} values of PI-1 against genotype 4 and 6 HCV protease were 7.19 nM and 13.9 nM, respectively (Figure 7.3). The inhibitory efficacy of PI-1 against genotype 4 HCV protease is comparable to that against genotype 1 HCV protease. The decreased potency of PI-1 against genotype 6 HCV protease is only approximately 2-fold.

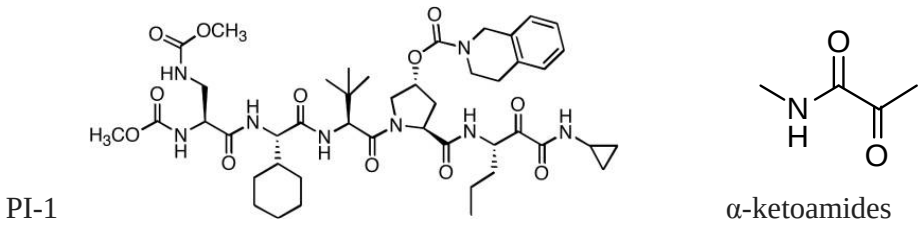


Figure 7.2 Chemical structure of the telaprevir analog, PI-1.

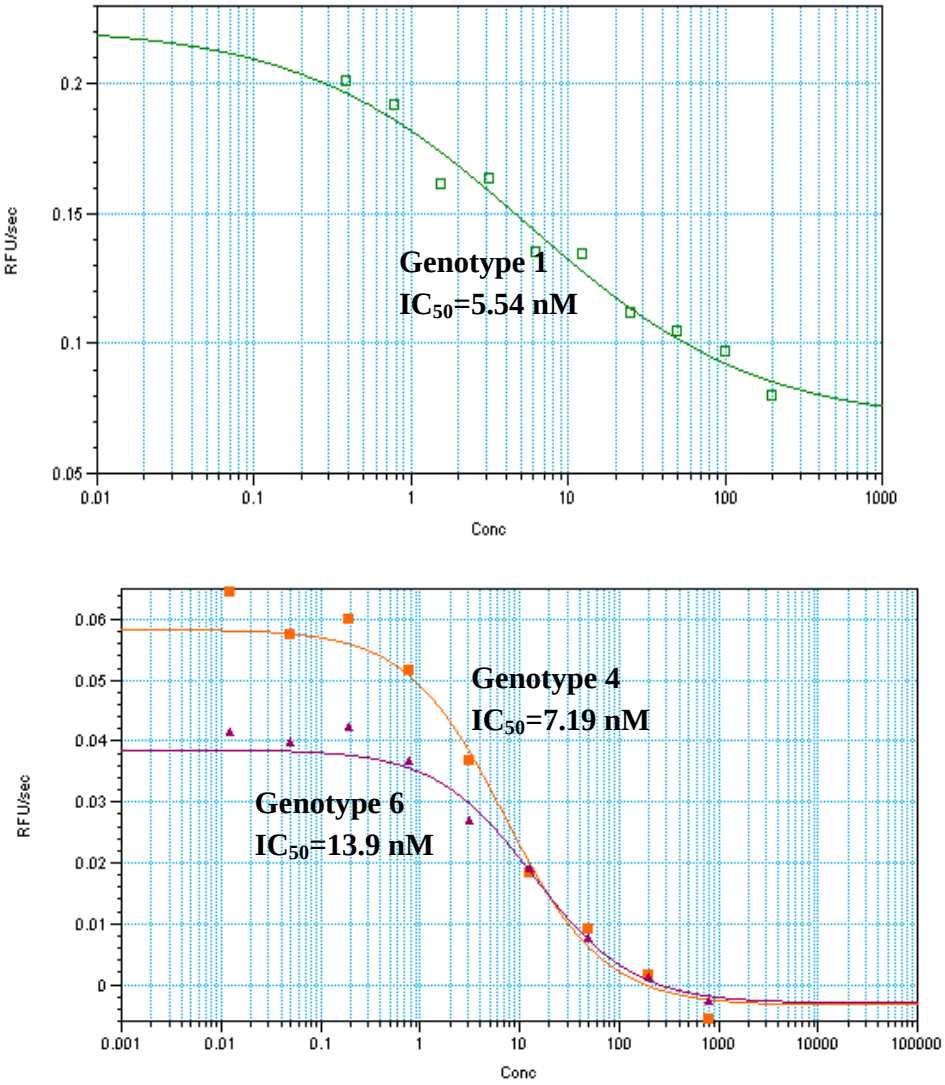


Figure 7.3 Inhibitor PI-1 IC_{50} against Genotype 1, 4, and 6 HCV proteases. The IC_{50} values were calculated using the SoftMax Pro Software V.5 (Molecular Devices Inc., Sunnyvale, CA). The IC_{50} values are indicated adjacent to the curves. The top panel is the compound inhibition experiment of the WT HIV-1 protease while the bottom panel is the compound inhibition experiment of the MDR HIV-1 protease. The curve data points presented in this figure are average numbers of triplicate readings. The Y-axis represents the Relative Fluorescence Units (RFU) per second. The X-axis is the compound concentrations in nanomolar values. The empty square symbol represents the

genotype 1 HCV protease activity curve while the solid square symbol and solid triangle symbol represent the genotype 4 and 6 HCV protease activity curve, respectively. The IC50 calculations were performed by fitting a four-parameter sigmoidal curve to the data points. To calculate the IC50 value, we used the equation y = (A-D)/[1+(x/C)^B]+D, where A and D are the Y-values corresponding to the asymptote at the lowest and highest concentration, respectively; B is a coefficient describing the slope of the curve (Hill coefficient); and C is the IC_{50} value.

7.2.3 Telaprevir exhibits less movement when it binds to HCV genotype 1 and genotype 4a proteases

The HCV NS3/4A protease of genotypes 4a, 5a, and 6a were modeled and simulated for 10 ns. Using the same methodology, the HCV genotype 1b NS3/4A protease was simulated as a control. After a 10-ns simulation, the interaction of telaprevir with the catalytic triad of HCV genotype 1 NS3/4A protease was illustrated in Figure 7.4. The telaprevir is covalently attached to the Ser 139 hydroxyl group and forms a hydrogen bond with His 57.

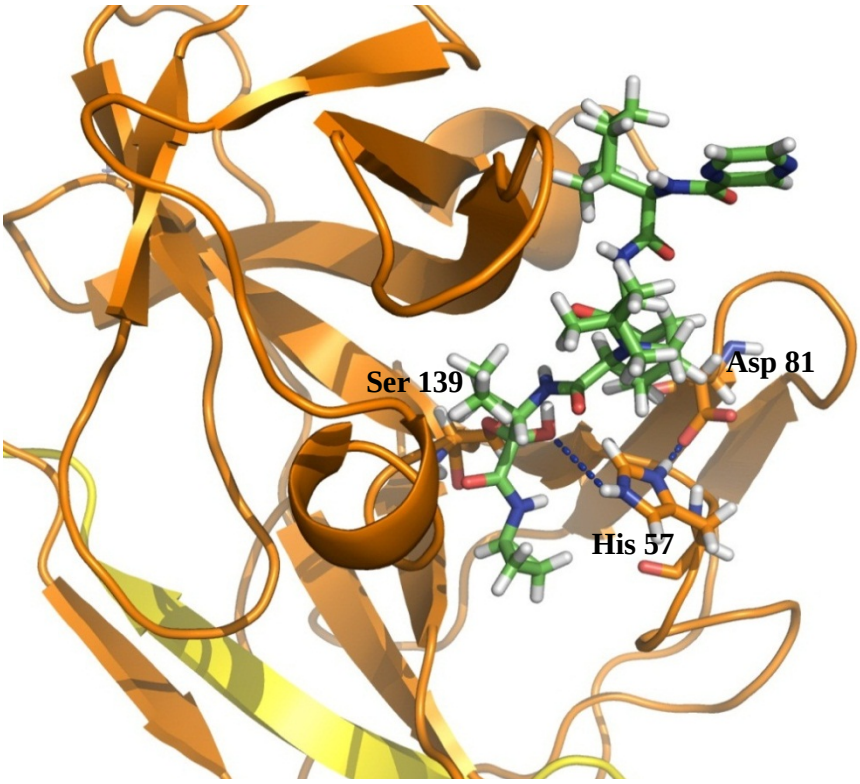


Figure 7.4 Interaction of telaprevir with the catalytic triad of HCV genotype 1 NS3/4A protease. The catalytic triad, His 57, Asp 81, and Ser 139 are shown in stick model. Telaprevir is represented in a green stick model. The orange cartoon model is the NS3 protease domain and the yellow cartoon model is the cofactor NS4A.

Based on the RMSD values throughout the simulation process (Figure 7.5), the HCV proteases of genotypes 1b and 6a had the least amount of movement from the initial models while overall RMSD values of the HCV genotype 4a and 5a proteases were 0.2 Å larger. The RMSD values of protease-inhibitor complexes suggested the structural similarity between genotypes 1b and 6a proteases is greater than that between genotype 1b protease and genotype 4a or 5a protease. Regarding the movement of telaprevir (Figure 7.6), the RMSD values of telaprevir binding to HCV genotype 4a proteases were larger than those of telaprevir binding to genotype 1, 5, or 6 HCV

proteases.

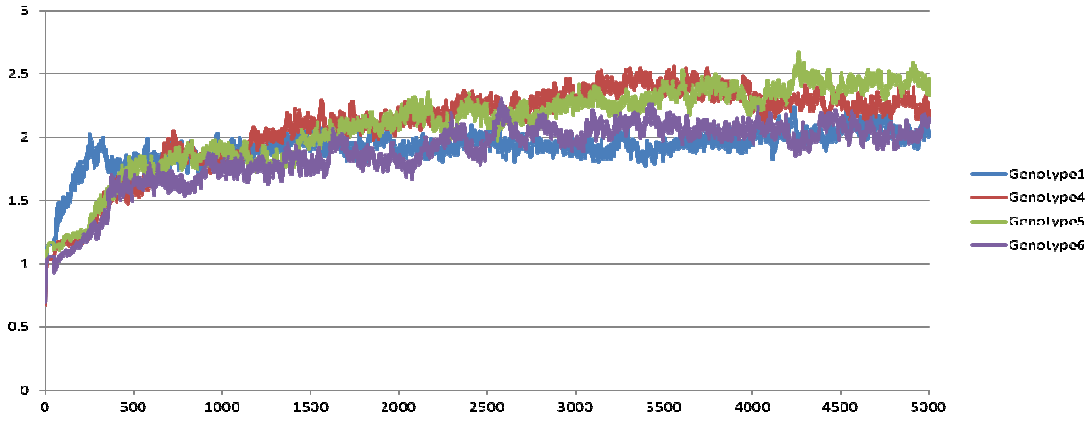


Figure 7.5 RMSD values of HCV NS3/4A protease-telaprevir complexes during the 10-ns simulation.

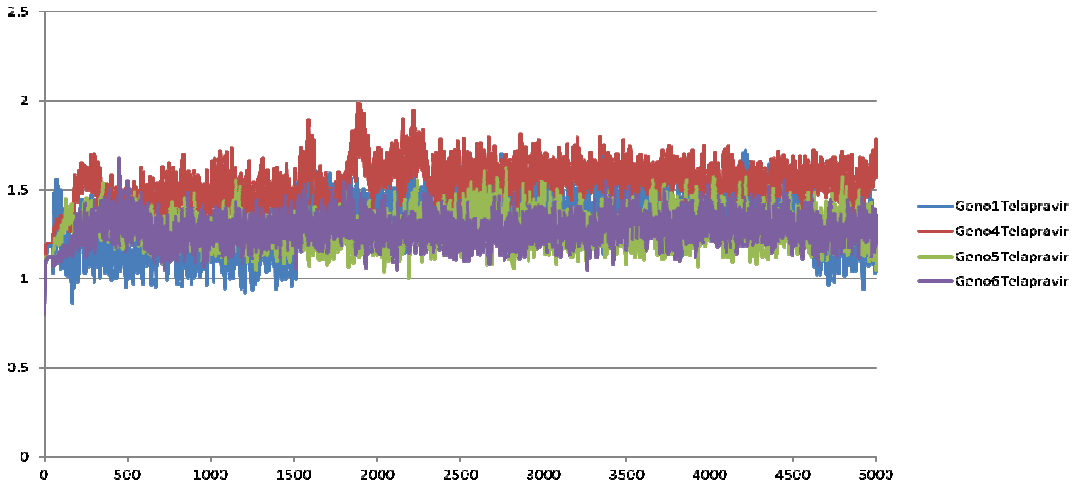


Figure 7.6 RMSD values of telaprevir binding to various HCV NS3/4A proteases during the 10-ns simulation.

Telapravir displayed the most relative stability when complexed with NS3/4A protease genotypes 1 and 4 than with genotypes 5 and 6. This conclusion is supported by considering the standard deviation of average RMSD taken from the final 200 frames of the protease-ligand complex trajectory files (final 0.4 ns of the 10-ns simulation, Figure 7.7). A relatively lower standard deviation between plotted points (frames) of the molecular dynamic corresponds to a relatively greater measure of protease-ligand

complex stability.

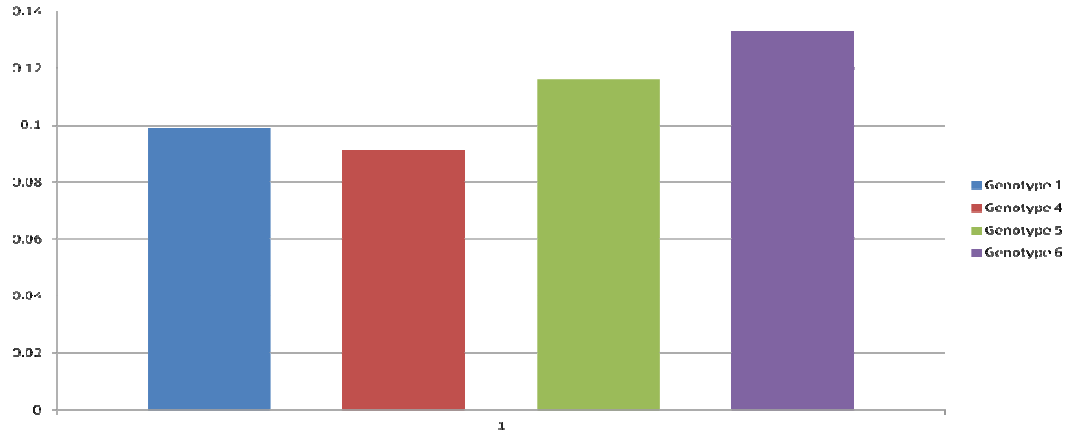


Figure 7.7 Standard deviation of the telaprevir average RMSD during the last 400-ps of the simulation.

7.2.4 The binding between telaprevir and HCV NS3/4A genotype 1 protease residues is superior to telaprevir complexes with the proteases of HCV genotypes 4, 5, and 6

Analysis of protease-ligand interactions is a direct way to evaluate ligand binding. The Ligplot software (*81*) was used to formulate and visually represent the interactions between the residues of a given NS3/4A protease genotype and the associated ligand, telaprevir. The interaction maps of telaprevir with four different genotypes of HCV proteases are represented in Figure 7.8. The genotype 1 protease-ligand complex displayed the most favorable bonding profile, showing eight hydrogen bonds compared with seven in both the genotype 5 and genotype 6 complexes and six in the genotype 4 complex. These data suggest that active site binding between telaprevir and the implicated NS3/4A genotype 1 protease residues is superior to telaprevir complexes with the remaining studied genotypes (4, 5 and 6).

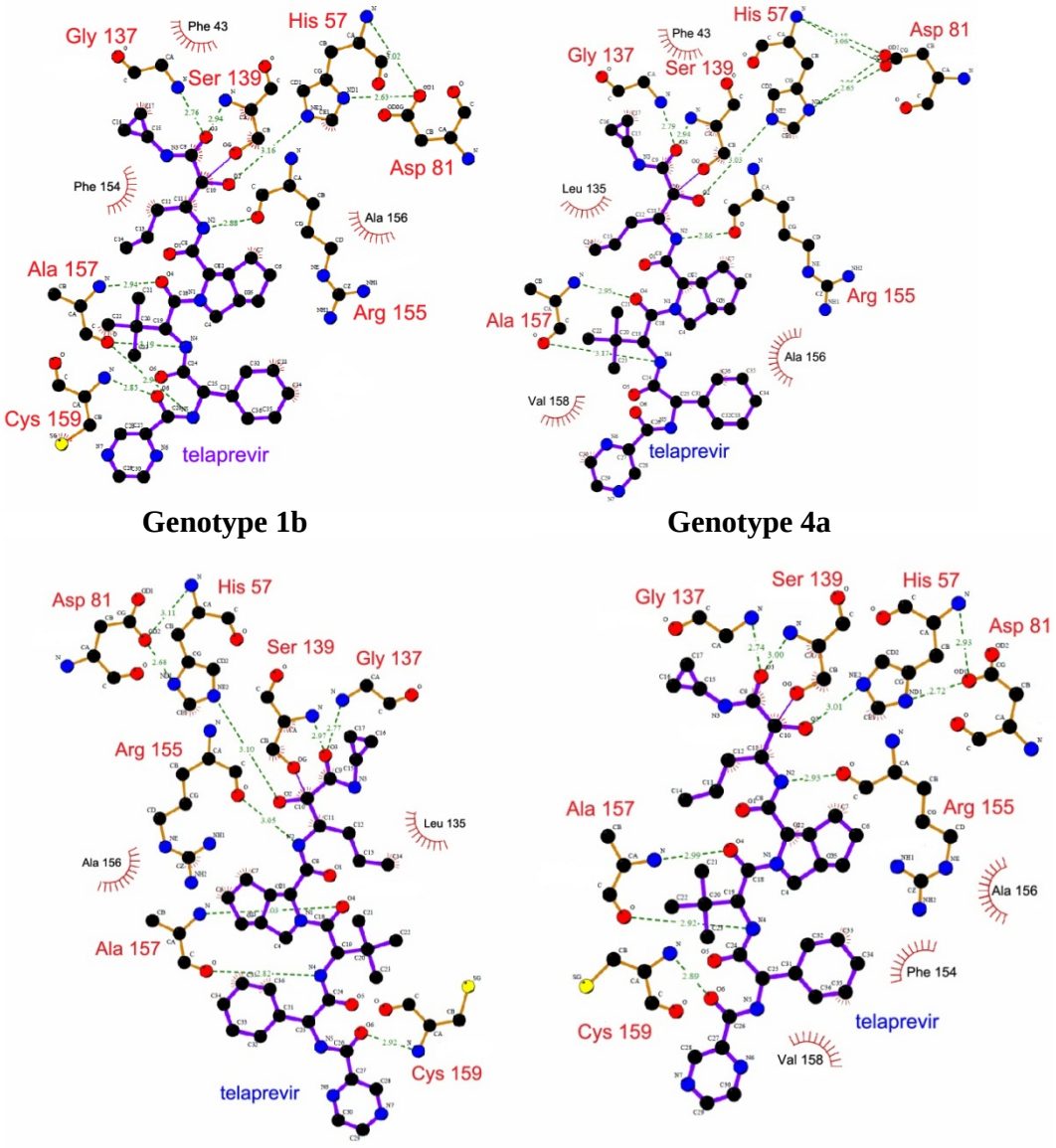


Figure 7.8 Ligplot analysis of telaprevir with the proteases from various HCV genotypes. Hydrogen bonds appear in dashed, green line and hydrophobic interactions are depicted with red, pronged semicircles.

Table 7.2 Analysis results using the Protein Interfaces, Surfaces and Assemblies server

Telaprevir complexes with HCV NS3/4A proteases	Buried area ($Å^2$)	ΔG^{int} (kcal/mol)	Number of hydrogen bonds with inhibitor
Genotype 1	475.9	-2.18	8
Genotype 4	492.6	-1.62	6
Genotype 5	480.2	-0.87	7
Genotype 6	475.5	-0.76	7

When genotype 1 NS3/4A protease-telaprevir complex is compared with telaprevir binding to genotype 4, 5, and 6 proteases, the genotype 4 NS3/4A protease had fewer hydrogen bonds with the ligand telaprevir, but more van der Waals interactions relative to the other genotypes. The genotype 4 protease complex displayed one fewer hydrogen bonding residue than do genotypes 1, 5 and 6 HCV proteases (Table 7.2). The genotype 4 protease complex lacked a hydrogen bond between Cys 159 and the inhibitor, whereas genotypes 1, 5 and 6 all displayed a hydrogen bond between the Cys 159 residue and telaprevir. The HCV protease-telaprevir interactions were further analyzed using PDBe PISA server (*126*). In the PISA analysis, the covalent bond between the HCV protease and telaprevir was manually removed. Telaprevir binding to the genotype 4 HCV protease had the largest buried ligand area. In the Table 7.2, ΔG^{int} indicates the solvation free energy gain upon complex formation, which suggested that the genotype 1 HCV protease-telaprevir is the tightest complex among the genotype 1, 4, 5, and 6 complexes. According to buried ligand area and ligand RMSD deviation results, the genotype 4

HCV protease-telaprevir RMSD data suggest that telaprevir shows comparable binding of between the 1 and 4 genotypes.

Table 7.3 Interaction energy between telaprevir and HCV proteases

Telaprevir complexes with HCV NS3/4A proteases	Electrostatic interaction energy (kcal/mol)	Van der Waals interaction energy (kcal/mol)	Nonbonded interaction energy (kcal/mol)
Genotype 1	-87.89	-55.12	-143.01
Genotype 4	-77.62	-56.88	-134.50
Genotype 5	-78.21	-55.78	-134.00
Genotype 6	-81.18	-55.69	-136.87

The interaction energy calculation of the HCV protease-telaprevir complexes supported the results of PISA analysis. Genotype 1 HCV protease-telaprevir complex showed the strongest total interaction energy (-143.01 kcal/mol, Table 7.3). The Genotype 4 HCV protease-telaprevir complex was stronger in van der Waals interactions but weaker in electrostatic interactions, which is correlated with the larger ligand buried area and loss of one hydrogen bond in Table 7.2. In summary, based on the energy calculations, telaprevir inhibited genotype 1 HCV protease with the best potency but retained sufficient potency to effectively inhibit non-genotype 1 HCV proteases.

7.3 Discussion

The development of telaprevir greatly benefits the hepatitis C patients. Since telaprevir has only been approved for the treatment of patients infected with genotype 1 HCV, more basic researches need to be conducted prior to the initiation of clinical

studies to investigate the potency of telaprevir against non-genotype 1 HCV strains. The first section of this study is the testing of *in vitro* inhibitory efficacy of the telaprevir analog against genotype 4 and 6 HCV proteases. The second section of the study is the determination of the structures of genotypes 4, 5, and 6 HCV proteases in complex with telaprevir using homology modeling and molecular dynamics.

According to the enzyme assays with the telaprevir analog, PI-1, and simulation results of the HCV protease-telaprevir complex, the binding patterns of telaprevir to genotype 4, 5, and 6 HCV proteases are conserved. A preliminary conclusion is that telaprevir possesses sufficient potency to treat genotype 4, 5, and 6 HCV strains. Therefore, it is possible to expand the patient target population of telaprevir or its derivatives to those infected with non-genotype 1 HCV variants. A weakness for the enzyme assays is that the second generation inhibitor, a telaprevir derivative, was used rather than original telaprevir. The potency of the telaprevir derivative in clinical tests is unclear.

The structural study of telaprevir binding to NS3/4A proteases of different genotypes may lead to the developing treatments for patients infected with non-genotype 1 HCV. To improve telaprevir efficacy with the other HCV genotypes, one strategy might be to structurally modify telaprevir to restore the protease binding network, such as the loss of hydrogen bond with the protease residue Cys 159. Another strategy is to extend the flexible end (the P4 and P5 groups) of telaprevir to make stable contacts with the HCV protease, which reduces the binding instability of the inhibitor.

The telaprevir binding to HCV NS3/4A protease is a two-step process. Telaprevir

first forms a weaker non-covalent collision complex with the HCV protease, and then a covalent bond is formed between telaprevir and the catalytic serine of the HCV protease. In this simulation study, the HCV protease-telaprevir complexes are modeled as the covalently attached complex. Whether telaprevir is able to successfully form the non-covalent collision complex with non-genotype 1 HCV proteases has not been examined. In future experiments, the ligand binding affinity of non-covalent HCV protease-telaprevir complexes will be studied to further explore the potential of telaprevir in the treatment of patients with non-genotype 1 HCV.

7.4 Author's contribution

The author performed the enzyme assays and set up the molecular dynamics simulations. The HCV protease expression experiments were carried out by the scientists at AnaSpec, Inc. (Fremont, CA, USA). The author and his lab member, Samuel J. Reiter, analyzed the data. Samuel J. Reiter produced the figures of the RMSD plot.

CHAPTER 8 FUTURE DIRECTIONS

Drug discovery is still a trial-and-error testing process. Rational design and *in silico* evaluation may accelerate this process. We will propose new modifications of lopinavir and evaluate them by simulating the inhibitor-protease interactions. The promising candidate molecules will be synthesized through collaboration with medicinal chemists. The goal is to develop inhibitors that target with equal potency both the WT HIV-1 protease and the MDR HIV-1 protease.

The drug resistance is caused by multiple steps of dynamic interactions. Crystal structures plus molecular dynamics are powerful tools to understand the complicated mechanisms. In the Protein Data Bank, there are still a limited number of apo HIV-1 protease structures. Increasing the diversity of the apo HIV-1 protease structures and the HIV-1 protease receptor structures for a given protease inhibitor will improve our knowledge of protease flexibility upon drug binding. Furthermore, we will simulate the movement of the protease flaps that is triggered by inhibitors, possibly linking protein flexibility with drug resistance.

Regarding the HCV project, we will study the drug resistant of HCV protease mutations to boceprevir, another newly approved HCV protease inhibitor. Boceprevir inhibits the HCV protease catalytic serine through the α-ketoamide function group as described earlier for telaprevir. During the clinical trial, several mutations that cause boceprevir resistance have been identified. The goal is to crystallize the boceprevir

resistant HCV NS3/4A protease mutants with or without boceprevir as starting models for molecular dynamic simulations.

REFERENCES

1. UNAIDS. (2010) UNAIDS Report on the global AIDS epidemic 2010, In *Global Report*, Geneva.
2. Pomerantz, R. J., and Horn, D. L. (2003) Twenty years of therapy for HIV-1 infection, *Nat Med 9*, 867-873.
3. Shehu-Xhilaga, M., Crowe, S. M., and Mak, J. (2001) Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity, *J Virol 75*, 1834-1841.
4. Marin, M., Rose, K. M., Kozak, S. L., and Kabat, D. (2003) HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation, *Nat Med 9*, 1398-1403.
5. Popov, S., Rexach, M., Zybarth, G., Reiling, N., Lee, M. A., Ratner, L., Lane, C. M., Moore, M. S., Blobel, G., and Bukrinsky, M. (1998) Viral protein R regulates nuclear import of the HIV-1 pre-integration complex, *Embo J 17*, 909-917.
6. He, J., Choe, S., Walker, R., Di Marzio, P., Morgan, D. O., and Landau, N. R. (1995) Human immunodeficiency virus type 1 viral protein R (Vpr) arrests cells in the G2 phase of the cell cycle by inhibiting p34cdc2 activity, *J Virol 69*, 6705-6711.
7. Jacotot, E., Ravagnan, L., Loeffler, M., Ferri, K. F., Vieira, H. L., Zamzami, N., Costantini, P., Druillennec, S., Hoebeke, J., Briand, J. P., Irinopoulou, T., Daugas, E., Susin, S. A., Cointe, D., Xie, Z. H., Reed, J. C., Roques, B. P., and Kroemer,

G. (2000) The HIV-1 viral protein R induces apoptosis via a direct effect on the mitochondrial permeability transition pore, *J Exp Med 191*, 33-46.

8. Schubert, U., Bour, S., Ferrer-Montiel, A. V., Montal, M., Maldarell, F., and Strebel, K. (1996) The two biological activities of human immunodeficiency virus type 1 Vpu protein involve two separable structural domains, *J Virol 70*, 809-819.

9. Kestler, H. W., Ringler, D. J., Mori, K., Panicali, D. L., Sehgal, P. K., Daniel, M. D., and Desrosiers, R. C. (1991) Importance of the nef gene for maintenance of high virus loads and for development of AIDS, *Cell 65*, 651-662.

10. Popovic, M., Sarin, P. S., Robert-Gurroff, M., Kalyanaraman, V. S., Mann, D., Minowada, J., and Gallo, R. C. (1983) Isolation and transmission of human retrovirus (human t-cell leukemia virus), *Science 219*, 856-859.

11. Matthews, T., Salgo, M., Greenberg, M., Chung, J., DeMasi, R., and Bolognesi, D. (2004) Enfuvirtide: The first therapy to inhibit the entry of HIV-1 into host CD4 lymphocytes, *Nat Rev Drug Discov 3*, 215-225.

12. Hitchcock, C. A. (2005) The discovery and exploratory development of Maraviroc (UK-427,857): A novel CCR5 antagonist for the treatment of HIV, *Retrovirology 2*.

13. Hazuda, D. J., Felock, P., Witmer, M., Wolfe, A., Stillmock, K., Grobler, J. A., Espeseth, A., Gabryelski, L., Schleif, W., Blau, C., and Miller, M. D. (2000) Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells, *Science 287*, 646-650.

14. Sato, M., Motomura, T., Aramaki, H., Matsuda, T., Yamashita, M., Ito, Y., Kawakami, H., Matsuzaki, Y., Watanabe, W., Yamataka, K., Ikeda, S., Kodama, E., Matsuoka, M., and Shinkai, H. (2006) Novel HIV-1 integrase inhibitors derived from quinolone antibiotics, *J Med Chem 49*, 1506-1508.

15. Rhee, S. Y., Gonzales, M. J., Kantor, R., Betts, B. J., Ravela, J., and Shafer, R. W. (2003) Human immunodeficiency virus reverse transcriptase and protease sequence database, *Nucleic Acids Res 31*, 298-303.

16. van de Vijver, D. A., Wensing, A. M., Asjo, B., Bruckova, M., Bruun Jorgensen, L., Camacho, R., Horban, A., Linka, M., Lazanas, M., Loveday, C., Macrae, E., Nielsen, C., Paraskevis, D., Poljak, M., Puchhammer-Stockl, E., Ruiz, L., Schmit, J. C., Stanczak, G., Stanojevic, M., Vandamme, A. M., Vercauteren, J., Zazzi, M., Bacheler, L., Lecocq, P., Villacian, J., and Boucher, C. A. (2010) HIV-1 drug-resistance patterns among patients on failing treatment in a large number of European countries, *Acta Dermatovenerol Alp Panonica Adriat 19*, 3-9.

17. Krohn, A., Redshaw, S., Ritchie, J. C., Graves, B. J., and Hatada, M. H. (1991) Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (R)-hydroxyethylamine isostere, *J Med Chem 34*, 3340-3342.

18. Kempf, D. J., Marsh, K. C., Kumar, G., Rodrigues, A. D., Denissen, J. F., McDonald, E., Kukulka, M. J., Hsu, A., Granneman, G. R., Baroldi, P. A., Sun, E., Pizzuti, D., Plattner, J. J., Norbeck, D. W., and Leonard, J. M. (1997) Pharmacokinetic enhancement of inhibitors of the human immunodeficiency

virus protease by coadministration with ritonavir, *Antimicrob Agents Chemother 41*, 654-660.

19. Sham, H. L., Kempf, D. J., Molla, A., Marsh, K. C., Kumar, G. N., Chen, C. M., Kati, W., Stewart, K., Lal, R., Hsu, A., Betebenner, D., Korneyeva, M., Vasavanonda, S., McDonald, E., Saldivar, A., Wideburg, N., Chen, X., Niu, P., Park, C., Jayanti, V., Grabowski, B., Granneman, G. R., Sun, E., Japour, A. J., Leonard, J. M., Plattner, J. J., and Norbeck, D. W. (1998) ABT-378, a highly potent inhibitor of the human immunodeficiency virus protease, *Antimicrob Agents Chemother 42*, 3218-3224.

20. Gonzalez de Requena, D., Gallego, O., de Mendoza, C., Corral, A., Jimenez-Nacher, I., and Soriano, V. (2003) Indinavir plasma concentrations and resistance mutations in patients experiencing early virological failure, *AIDS Res Hum Retroviruses 19*, 457-459.

21. St Clair, M. H., Millard, J., Rooney, J., Tisdale, M., Parry, N., Sadler, B. M., Blum, M. R., and Painter, G. (1996) In vitro antiviral activity of 141W94 (VX-478) in combination with other antiretroviral agents, *Antiviral Res 29*, 53-56.

22. Rusconi, S., La Seta Catamancio, S., Citterio, P., Kurtagic, S., Violin, M., Balotta, C., Moroni, M., Galli, M., and d'Arminio-Monforte, A. (2000) Susceptibility to PNU-140690 (Tipranavir) of human immunodeficiency virus type 1 isolates derived from patients with multidrug resistance to other protease inhibitors, *Antimicrob Agents Chemother 44*, 1328-1332.

23. Larder, B. A., Hertogs, K., Bloor, S., van den Eynde, C. H., DeCian, W., Wang, Y., Freimuth, W. W., and Tarpley, G. (2000) Tipranavir inhibits broadly protease inhibitor-resistant HIV-1 clinical samples, *AIDS 14*, 1943-1948.

24. Food and Drug Administration. (2006) Important safety information about Aptivus (tipranavir), Food and Drug Administration.

25. Robinson, B. S., Riccardi, K. A., Gong, Y. F., Guo, Q., Stock, D. A., Blair, W. S., Terry, B. J., Deminie, C. A., Djang, F., Colonno, R. J., and Lin, P. F. (2000) BMS-232632, a highly potent human immunodeficiency virus protease inhibitor that can be used in combination with other available antiretroviral agents, *Antimicrob Agents Chemother 44*, 2093-2099.

26. Lefebvre, E., and Schiffer, C. A. (2008) Resilience to resistance of HIV-1 protease inhibitors: profile of darunavir, *AIDS Rev 10*, 131-142.

27. Tie, Y., Boross, P. I., Wang, Y. F., Gaddis, L., Hussain, A. K., Leshchenko, S., Ghosh, A. K., Louis, J. M., Harrison, R. W., and Weber, I. T. (2004) High resolution crystal structures of HIV-1 protease with a potent non-peptide inhibitor (UIC-94017) active against multi-drug-resistant clinical strains, *J Mol Biol 338*, 341-352.

28. Panel on Antiretroviral Guidelines for Adults and Adolescents. (2011) Guidelines for the use of antiretroviral agents in HIV-1-infected adults and adolescents, (Department of Health and Human Services, Ed.), p 41.

29. Sham, H. L., Zhao, C., Li, L., Betebenner, D. A., Saldivar, A., Vasavanonda, S., Kempf, D. J., Plattner, J. J., and Norbeck, D. W. (2002) Novel lopinavir

analogues incorporating non-Aromatic P-1 side chains--synthesis and structure--activity relationships, *Bioorg Med Chem Lett 12*, 3101-3103.

30. Prabu-Jeyabalan, M., King, N. M., Nalivaika, E. A., Heilek-Snyder, G., Cammack, N., and Schiffer, C. A. (2006) Substrate envelope and drug resistance: crystal structure of RO1 in complex with wild-type human immunodeficiency virus type 1 protease, *Antimicrob Agents Chemother 50*, 1518-1521.

31. Ala, P. J., DeLoskey, R. J., Huston, E. E., Jadhav, P. K., Lam, P. Y., Eyermann, C. J., Hodge, C. N., Schadt, M. C., Lewandowski, F. A., Weber, P. C., McCabe, D. D., Duke, J. L., and Chang, C. H. (1998) Molecular recognition of cyclic urea HIV-1 protease inhibitors, *J Biol Chem 273*, 12325-12331.

32. Hoffmann, C. (2010) ART 2010/2011: The Horizon and Beyond, In *HIV 2010 - A Textbook* (Christian Hoffmann, J. K. R., Ed.), Medizin Fokus Verlag, Hamburg, Germany.

33. Koshland, D. E. (1995) The Key-Lock Theory and the Induced Fit Theory, *Angew Chem Int Edit 33*, 2375-2378.

34. Barlow, D. J., and Thornton, J. M. (1983) Ion-pairs in proteins, *J Mol Biol 168*, 867-885.

35. Kassel, D. B., Green, M. D., Wehbie, R. S., Swanstrom, R., and Berman, J. (1995) HIV-1 protease specificity derived from a complex mixture of synthetic substrates, *Anal Biochem 228*, 259-266.

36. Prabu-Jeyabalan, M., Nalivaika, E., and Schiffer, C. A. (2002) Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal

structures of six substrate complexes, *Structure 10*, 369-381.

37. Altman, M. D., Ali, A., Reddy, G. S., Nalam, M. N., 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, *J Am Chem Soc 130*, 6099-6113.

38. Suguna, K., Padlan, E. A., Smith, C. W., Carlson, W. D., and Davies, D. R. (1987) Binding of a reduced peptide inhibitor to the aspartic proteinase from Rhizopus chinensis: implications for a mechanism of action, *Proc Natl Acad Sci U S A 84*, 7009-7013.

39. Hyland, L. J., Tomaszek, T. A., Jr., and Meek, T. D. (1991) Human immunodeficiency virus-1 protease. 2. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism, *Biochemistry 30*, 8454-8463.

40. Chatfield, D. C., and Brooks, B. R. (1995) Hiv-1 protease cleavage mechanism elucidated with molecular-dynamics simulation, *J Am Chem Soc 117*, 5561-5572.

41. Veerapandian, B., Cooper, J. B., Sali, A., Blundell, T. L., Rosati, R. L., Dominy, B. W., Damon, D. B., and Hoover, D. J. (1992) Direct observation by X-ray analysis of the tetrahedral "intermediate" of aspartic proteinases, *Protein Sci 1*, 322-328.

42. Ishima, R., and Louis, J. M. (2008) A diverse view of protein dynamics from

NMR studies of HIV-1 protease flaps, *Proteins 70*, 1408-1415.

43. Purcell, R. H. (1994) Hepatitis C virus: historical perspective and current concepts, *FEMS Microbiol Rev 14*, 181-191.

44. Alter, M. J., and Mast, E. E. (1994) The epidemiology of viral hepatitis in the United States, *Gastroenterol Clin North Am 23*, 437-455.

45. Renault, P. F., and Hoofnagle, J. H. (1989) Side effects of alpha interferon, *Semin Liver Dis 9*, 273-277.

46. Pileri, P., Uematsu, Y., Campagnoli, S., Galli, G., Falugi, F., Petracca, R., Weiner, A. J., Houghton, M., Rosa, D., Grandi, G., and Abrignani, S. (1998) Binding of hepatitis C virus to CD81, *Science 282*, 938-941.

47. Scarselli, E., Ansuini, H., Cerino, R., Roccasecca, R. M., Acali, S., Filocamo, G., Traboni, C., Nicosia, A., Cortese, R., and Vitelli, A. (2002) The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus, *Embo J 21*, 5017-5025.

48. Lozach, P. Y., Lortat-Jacob, H., de Lavalette, A. D., Staropoli, I., Foung, S., Amara, A., Houles, C., Fieschi, F., Schwartz, O., Virelizier, J. L., Arenzana-Seisdedos, F., and Altmeyer, R. (2003) DC-SIGN and L-SIGN are high affinity binding receptors for hepatitis C virus glycoprotein E2, *J Biol Chem 278*, 20358-20366.

49. Gardner, J. P., Durso, R. J., Arrigale, R. R., Donovan, G. P., Maddon, P. J., Dragic, T., and Olson, W. C. (2003) L-SIGN (CD 209L) is a liver-specific capture receptor for hepatitis C virus, *Proc Natl Acad Sci U S A 100*, 4498-4503.

50. Agnello, V., Abel, G., Elfahal, M., Knight, G. B., and Zhang, Q. X. (1999) Hepatitis C virus and other flaviviridae viruses enter cells via low density lipoprotein receptor, *Proc Natl Acad Sci U S A 96*, 12766-12771.

51. Saunier, B., Triyatni, M., Ulianich, L., Maruvada, P., Yen, P., and Kohn, L. D. (2003) Role of the asialoglycoprotein receptor in binding and entry of hepatitis C virus structural proteins in cultured human hepatocytes, *J Virol 77*, 546-559.

52. Dorner, M., Horwitz, J. A., Robbins, J. B., Barry, W. T., Feng, Q., Mu, K., Jones, C. T., Schoggins, J. W., Catanese, M. T., Burton, D. R., Law, M., Rice, C. M., and Ploss, A. (2011) A genetically humanized mouse model for hepatitis C virus infection, *Nature 474*, 208-211.

53. Moradpour, D., Penin, F., and Rice, C. M. (2007) Replication of hepatitis C virus, *Nature reviews. Microbiology 5*, 453-463.

54. Ji, H., Fraser, C. S., Yu, Y., Leary, J., and Doudna, J. A. (2004) Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA, *Proc Natl Acad Sci U S A 101,* 16990-16995.

55. Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M., and Rice, C. M. (1993) A second hepatitis C virus-encoded proteinase, *Proc Natl Acad Sci U S A 90,* 10583-10587.

56. Lindenbach, B. D., and Rice, C. M. (2005) Unravelling hepatitis C virus replication from genome to function, *Nature 436*, 933-938.

57. Penin, F., Dubuisson, J., Rey, F. A., Moradpour, D., and Pawlotsky, J. M. (2004) Structural biology of hepatitis C virus, *Hepatology 39*, 5-19.

58. Tan, S. L., Pause, A., Shi, Y., and Sonenberg, N. (2002) Hepatitis C therapeutics: current status and emerging strategies, *Nat Rev Drug Discov 1*, 867-881.

59. Bartenschlager, R., Ahlborn-Laake, L., Mous, J., and Jacobsen, H. (1993) Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions, *J Virol 67*, 3835-3844.

60. Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M., and Rice, C. M. (1993) Characterization of the hepatitis C virus-encoded serine proteinase: determination of proteinase-dependent polyprotein cleavage sites, *J Virol 67*, 2832-2843.

61. Hijikata, M., Mizushima, H., Tanji, Y., Komoda, Y., Hirowatari, Y., Akagi, T., Kato, N., Kimura, K., and Shimotohno, K. (1993) Proteolytic processing and membrane association of putative nonstructural proteins of hepatitis C virus, *Proc Natl Acad Sci U S A 90*, 10773-10777.

62. Steinkuhler, C., Urbani, A., Tomei, L., Biasiol, G., Sardana, M., Bianchi, E., Pessi, A., and De Francesco, R. (1996) Activity of purified hepatitis C virus protease NS3 on peptide substrates, *J Virol 70*, 6694-6700.

63. Tai, C. L., Chi, W. K., Chen, D. S., and Hwang, L. H. (1996) The helicase activity associated with hepatitis C virus nonstructural protein 3 (NS3), *J Virol 70*, 8477-8484.

64. Bartenschlager, R., Ahlborn-Laake, L., Mous, J., and Jacobsen, H. (1994) Kinetic and structural analyses of hepatitis C virus polyprotein processing, *J Virol 68*, 5045-5055.

65. Failla, C., Tomei, L., and De Francesco, R. (1994) Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus nonstructural proteins, *J Virol 68*, 3753-3760.

66. Kim, J. L., Morgenstern, K. A., Lin, C., Fox, T., Dwyer, M. D., Landro, J. A., Chambers, S. P., Markland, W., Lepre, C. A., O'Malley, E. T., Harbeson, S. L., Rice, C. M., Murcko, M. A., Caron, P. R., and Thomson, J. A. (1996) Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide, *Cell 87*, 343-355.

67. Lin, C., and Rice, C. M. (1995) The hepatitis C virus NS3 serine proteinase and NS4A cofactor: establishment of a cell-free trans-processing assay, *Proc Natl Acad Sci U S A 92*, 7622-7626.

68. Yan, Y., Li, Y., Munshi, S., Sardana, V., Cole, J. L., Sardana, M., Steinkuehler, C., Tomei, L., De Francesco, R., Kuo, L. C., and Chen, Z. (1998) Complex of NS3 protease and NS4A peptide of BK strain hepatitis C virus: a 2.2 A resolution structure in a hexagonal crystal form, *Protein Sci 7*, 837-847.

69. Lin, C., Kwong, A. D., and Perni, R. B. (2006) Discovery and development of VX-950, a novel, covalent, and reversible inhibitor of hepatitis C virus NS3.4A serine protease, *Infect Disord Drug Targets 6*, 3-16.

70. Reesink, H. W., Zeuzem, S., Weegink, C. J., Forestier, N., van Vliet, A., van de Wetering de Rooij, J., McNair, L., Purdy, S., Kauffman, R., Alam, J., and Jansen, P. L. (2006) Rapid decline of viral RNA in hepatitis C patients treated with VX-950: a phase Ib, placebo-controlled, randomized study, *Gastroenterology*

131, 997-1002.

71. Stauber, R. E., and Kessler, H. H. (2008) Drugs in development for hepatitis C, *Drugs 68*, 1347-1359.

72. Lau, J. Y., Davis, G. L., Kniffen, J., Qian, K. P., Urdea, M. S., Chan, C. S., Mizokami, M., Neuwald, P. D., and Wilber, J. C. (1993) Significance of serum hepatitis C virus RNA levels in chronic hepatitis C, *Lancet 341*, 1501-1504.

73. Taylor, W., Luong, Y. P., Rao, B. G., Brennan, D. L., Fulghum, J. R., Lippke, J., Perni, R. B., Kwong, A. D., and Lin, C. (2004) In *The 11th International symposium on hepatitis C virus and related viruses: Molecular Virology, Pathogenesis and Antiviral Therapy*, Heidelberg, Germany.

74. Welsch, C., Domingues, F. S., Susser, S., Antes, I., Hartmann, C., Mayr, G., Schlicker, A., Sarrazin, C., Albrecht, M., Zeuzem, S., and Lengauer, T. (2008) Molecular basis of telaprevir resistance due to V36 and T54 mutations in the NS3-4A protease of the hepatitis C virus, *Genome Biol 9*, R16.

75. Sham, H. L., Betebenner, D. A., Herrin, T., Kumar, G., Saldivar, A., Vasavanonda, S., Molla, A., Kempf, D. J., Plattner, J. J., and Norbeck, D. W. (2001) Synthesis and antiviral activities of the major metabolites of the HIV protease inhibitor ABT-378 (Lopinavir), *Bioorg Med Chem Lett 11*, 1351-1353.

76. Welch, M., Govindarajan, S., Ness, J. E., Villalobos, A., Gurney, A., Minshull, J., and Gustafsson, C. (2009) Design parameters to control synthetic gene expression in Escherichia coli, *PLoS One 4*, e7002.

77. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data

collected in oscillation mode, In *Methods in Enzymology* (Carter, C. W., and Sweet, R. M., Eds.), pp 307-326, Academic Press, Inc., New York, NY.

78. COLLABORATIVE COMPUTATIONAL PROJECT NUMBER 4. (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallogr D Biol Crystallogr 50*, 760-763.

79. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr D Biol Crystallogr 60*, 2126-2132.

80. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J Appl Cryst 26*, 283-291.

81. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions, *Protein Eng 8*, 127-134.

82. Reid, D., Sadjad, B. S., Zsoldos, Z., and Simon, A. (2008) LASSO-ligand activity by surface similarity order: a new tool for ligand based virtual screening, *J Comput Aid Mol Des 22*, 479-487.

83. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, *J Comput Chem 30*, 2785-2791.

84. Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability, *Nat Protoc 2*, 2212-2221.

85. Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling, *Bioinformatics 22*, 195-201.

86. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD, *J Comput Chem 26*, 1781-1802.

87. Darden, T., York, D., and Pedersen, L. (1993) Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems, *J Chem Phys 98*, 10089-10092.

88. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J Phys Chem B 102*, 3586-3616.

89. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome, *Proc Natl Acad Sci U S A 98*, 10037-10041.

90. Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., and Baker, N. A. (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations, *Nucleic Acids Res 35*, W522-525.

91. Sitkoff, D., Sharp, K. A., and Honig, B. (1994) Accurate Calculation of Hydration Free-Energies Using Macroscopic Solvent Models, *J Phys Chem-Us 98*, 1978-1988.

92. Butkiewicz, N. J., Wendel, M., Zhang, R., Jubin, R., Pichardo, J., Smith, E. B., Hart, A. M., Ingram, R., Durkin, J., Mui, P. W., Murray, M. G., Ramanathan, L., and Dasmahapatra, B. (1996) Enhancement of hepatitis C virus NS3 proteinase activity by association with NS4A-specific synthetic peptides: identification of sequence and critical residues of NS4A for the cofactor activity, *Virology 225*, 328-338.

93. Taremi, S. S., Beyer, B., Maher, M., Yao, N., Prosise, W., Weber, P. C., and Malcolm, B. A. (1998) Construction, expression, and characterization of a novel fully activated recombinant single-chain hepatitis C virus protease, *Protein Sci 7*, 2143-2149.

94. Rhee, S. Y., Taylor, J., Fessel, W. J., Kaufman, D., Towner, W., Troia, P., Ruane, P., Hellinger, J., Shirvani, V., Zolopa, A., and Shafer, R. W. (2010) HIV-1 Protease Mutations and Protease Inhibitor Cross Resistance, *Antimicrob Agents Chemother 54*, 4253-4261.

95. Beck, Z. Q., Hervio, L., Dawson, P. E., Elder, J. H., and Madison, E. L. (2000) Identification of efficiently cleaved substrates for HIV-1 protease using a phage display library and use in inhibitor development, *Virology 274*, 391-401.

96. Shafer, R. W. (2006) Rationale and uses of a public HIV drug-resistance database, *J Infect Dis 194 Suppl 1*, S51-58.

97. Palmer, S., Shafer, R. W., and Merigan, T. C. (1999) Highly drug-resistant HIV-1 clinical isolates are cross-resistant to many antiretroviral compounds in current clinical development, *AIDS 13*, 661-667.

98. Logsdon, B. C., Vickrey, J. F., Martin, P., Proteasa, G., Koepke, J. I., Terlecky, S. R., Wawrzak, Z., Winters, M. A., Merigan, T. C., and Kovari, L. C. (2004) Crystal structures of a multidrug-resistant human immunodeficiency virus type 1 protease reveal an expanded active-site cavity, *J Virol 78*, 3123-3132.

99. Yedidi, R. S., Proteasa, G., Martinez, J. L., Vickrey, J. F., Martin, P. D., Wawrzak, Z., Liu, Z., Kovari, I. A., and Kovari, L. C. (2011) Contribution of the 80s loop of HIV-1 protease to the multidrug-resistance mechanism: crystallographic study of MDR769 HIV-1 protease variants, *Acta Crystallogr D Biol Crystallogr 67*, 524-532.

100. Tenore, S. B., and Ferreira, P. R. (2009) The Place of protease inhibitors in antiretroviral treatment, *Braz J Infect Dis 13*, 371-374.

101. Temesgen, Z., and Feinberg, J. (2007) Tipranavir: a new option for the treatment of drug-resistant HIV infection, *Clin Infect Dis 45*, 761-769.

102. Ghosh, A. K., Dawson, Z. L., and Mitsuya, H. (2007) Darunavir, a conceptually new HIV-1 protease inhibitor for the treatment of drug-resistant HIV, *Bioorg Med Chem 15*, 7576-7580.

103. Streeck, H., and Rockstroh, J. K. (2007) Review of tipranavir in the treatment of drug-resistant HIV, *Ther Clin Risk Manag 3*, 641-651.

104. Liu, T. F., and Shafer, R. W. (2006) Web resources for HIV type 1

genotypic-resistance test interpretation, *Clin Infect Dis 42*, 1608-1618.

105. Liu, Z., Wang, Y., Brunzelle, J., Kovari, I. A., and Kovari, L. C. (2011) Nine crystal structures determine the substrate envelope of the MDR HIV-1 protease, *Protein J 30*, 173-183.

106. Ghosh, A. K., Chapsal, B. D., Weber, I. T., and Mitsuya, H. (2008) Design of HIV protease inhibitors targeting protein backbone: an effective strategy for combating drug resistance, *Accounts of chemical research 41*, 78-86.

107. Ozer, N., Schiffer, C. A., and Haliloglu, T. (2010) Rationale for more diverse inhibitors in competition with substrates in HIV-1 protease, *Biophys J 99*, 1650-1659.

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

109. Heaslet, H., Rosenfeld, R., Giffin, M., Lin, Y. C., Tam, K., Torbett, B. E., Elder, J. H., McRee, D. E., and Stout, C. D. (2007) Conformational flexibility in the flap domains of ligand-free HIV protease, *Acta Crystallogr D Biol Crystallogr 63*, 866-875.

110. Galiano, L., Ding, F., Veloro, A. M., Blackburn, M. E., Simmerling, C., and Fanucci, G. E. (2009) Drug pressure selected mutations in HIV-1 protease alter flap conformations, *J Am Chem Soc 131*, 430-431.

111. Bottcher, J., Blum, A., Dorr, S., Heine, A., Diederich, W. E., and Klebe, G. (2008) Targeting the open-flap conformation of HIV-1 protease with pyrrolidine-based

inhibitors, *ChemMedChem 3*, 1337-1344.

112. Bolstad, E. S., and Anderson, A. C. (2009) In pursuit of virtual lead optimization: pruning ensembles of receptor structures for increased efficiency and accuracy during docking, *Proteins 75*, 62-74.

113. Huang, S. Y., and Zou, X. (2007) Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking, *Proteins 66*, 399-421.

114. Kiviranta, P. H., Leppanen, J., Kyrylenko, S., Salo, H. S., Lahtela-Kakkonen, M., Tervo, A. J., Wittekindt, C., Suuronen, T., Kuusisto, E., Jarvinen, T., Salminen, A., Poso, A., and Wallen, E. A. (2006) N,N'-Bisbenzylidenebenzene-1,4-diamines and N,N'-Bisbenzylidenenaphthalene -1,4-diamines as Sirtuin Type 2 (SIRT2) Inhibitors, *J Med Chem 49*, 7907-7911.

115. Nalam, M. N. L., Ali, A., Altman, M. D., Reddy, G. S. K. K., Chellappan, S., Kairys, V., Ozen, A., Cao, H., Gilson, M. K., Tidor, B., Rana, T. M., and Schiffer, C. A. (2010) Evaluating the Substrate-Envelope Hypothesis: Structural Analysis of Novel HIV-1 Protease Inhibitors Designed To Be Robust against Drug Resistance, *J Virol 84*, 5368-5378.

116. Prabu-Jeyabalan, M., Schiffer, C. A., and Nalivaika, E. (2002) Substrate shape determines specificity of recognition for HIV-1 protease: Analysis of crystal structures of six substrate complexes, *Structure 10*, 369-381.

117. Ozen, A., Haliloglu, T., and Schiffer, C. A. (2011) Dynamics of Preferential Substrate Recognition in HIV-1 Protease: Redefining the Substrate Envelope, *J*

Mol Biol 410, 726-744.

118. Luque, I., Todd, M. J., Gomez, J., Semo, N., and Freire, E. (1998) Molecular basis of resistance to HIV-1 protease inhibition: A plausible hypothesis, *Biochemistry 37*, 5791-5797.

119. Wang, Y., Liu, Z., Brunzelle, J. S., Kovari, I. A., Dewdney, T. G., Reiter, S. J., and Kovari, L. C. (2011) The higher barrier of darunavir and tipranavir resistance for HIV-1 protease, *Biochem Biophys Res Commun 412*, 737-742.

120. Szeltner, Z., and Polgar, L. (1996) Rate-determining steps in HIV-1 protease catalysis - The hydrolysis of the most specific substrate, *J Biol Chem 271*, 32180-32184.

121. Srinivasan, J., Cheatham, T. E., Cieplak, P., Kollman, P. A., and Case, D. A. (1998) Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate - DNA helices, *J Am Chem Soc 120*, 9401-9409.

122. Pettit, S. C., Moody, M. D., Wehbie, R. S., Kaplan, A. H., Nantermet, P. V., Klein, C. A., and Swanstrom, R. (1994) The P2 Domain of Human-Immunodeficiency-Virus Type-1 Gag Regulates Sequential Proteolytic Processing and Is Required to Produce Fully Infectious Virions, *J Virol 68*, 8017-8027.

123. Vickrey, J. F., Logsdon, B. C., Proteasa, G., Palmer, S., Winters, M. A., Merigan, T. C., and Kovari, L. C. (2003) HIV-1 protease variants from 100-fold drug resistant clinical isolates: expression, purification, and crystallization, *Protein Expr Purif 28*, 165-172.

124. Modi, A. A., and Liang, T. J. (2008) Hepatitis C: a clinical review, *Oral Dis 14*, 10-14.

125. Di Bisceglie, A. M. (2000) Natural history of hepatitis C: its impact on clinical management, *Hepatology 31*, 1014-1018.

126. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state, *J Mol Biol 372*, 774-797.

ABSTRACT

DRUG RESISTANCE MECHANISMS AND DRUG DESIGN STRATEGIES FOR HUMAN IMMUNODEFICIENCY VIRUS AND HEPATITIS C VIRUS PROTEASES

by

YONG WANG

May 2012

Advisor: Dr. Ladislau C. Kovari

Major: Biochemistry and Molecular Biology

Degree: Doctor of Philosophy

The antiviral drug development has improved steadily to treat the infections of human immunodeficiency virus (HIV) and hepatitis C virus (HCV) which represent heavy public health burdens. The viral protease plays an indispensable role in viral maturation and therefore becomes one of the most important targets for drug design. Nine HIV-1 protease inhibitors and two HCV protease inhibitors have been developed and approved by the U.S. Food and Drug Administration. However, mutations in the protease decrease reduce the efficacy the drugs. In this study, the enzyme assays indicate that darunavir and tipranavir exhibit the most potent inhibition against the multi-drug resistant (MDR) HIV-1 protease, which is supported by the co-crystal structures of the MDR protease-darunavir complex and the MDR protease-tipranavir complex. The MDR HIV-1 protease not only decreases the susceptibility to drugs but also impedes the formation of protease-substrate complex. Molecular dynamics simulation results show that the MDR HIV-1 protease needs to conquer a higher desolvation energy barrier to

bind the substrate. Besides the study of drug resistance mechanisms, two drug discovery methods have been carried out in the study. One method is the modification of a current drug, lopinavir. The potency of the lopinavir analog against the MDR HIV-1 protease increases. The other method is the identification of a novel HIV-1 protease inhibitor scaffold to increase the structural diversity of inhibitors. In the HCV section, the study focuses on an HCV NS3/4A protease inhibitor, telaprevir, which is approved for the treatment of patients infected with the HCV genotype 1. The enzyme assays and molecular dynamic studies suggest that telaprevir may retain sufficient potency to treat the non-genotype 1 HCV strains.

AUTOBIOGRAPHICAL STATEMENT

Education

07/2007-12/2011	Ph.D. candidate in Biochemistry and Molecular Biology Department of Biochemistry and Molecular Biology, Wayne State University School of Medicine
09/2004-07/2007	M.S. in Pathogen Biology National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention
09/2000-07/2004	B.S. in Biotechnology School of Life Sciences, Sichuan University

Publications

1. **Wang Y**, Liu Z, Brunzelle JS, Kovari IA, Dewdney, TG, and Kovari LC. The higher barrier of darunavir and tipranavir resistance for HIV-1 protease. *Biochem Biophys Res Commun.* 2011;412(4):737-42.

2. Liu Z, **Wang Y**, Yedidi RS, Brunzelle JS, Kovari IA, Sohi J, Kamholz J, and Kovari LC. Crystal structure of the extracellular domain of human myelin protein zero. *Proteins*. 2011. doi: 10.1002/prot.23164

3. Liu Z, **Wang Y**, Brunzelle J, Kovari IA, and Kovari LC. Nine crystal structures determine the substrate envelope of the MDR HIV-1 protease. *Protein J*. 2011;30(3):173-83.

4. **Wang Y**, Pan X. Proteus spp., *Encyclopedia of Food Safety*. Kidlington: Elsevier, 2012

5. Wang J, Ha Y, Gao C, **Wang Y**, Yang Y, and Chen H. The prevalence of canine *Leishmania infantum* infection in western China detected by PCR and serological tests. *Parasites & Vectors* 2011; 4(1):69.

6. **Wang Y**, Yang Y, Wang J, Bao Y, Guan L, Gao C, and Shi F. Molecular characterization of *Leishamania* isolates from China by inter-simple sequence repeat polymerase chain reaction. *Parasitol Res*. 2010;106(6):1385-94.

7. **Wang Y**, Wang J. Application of molecular methods in identification and phylogeny of *Leishmania*. *Int J of Med Parasit Dis*. 2006; 33(5): 239-43.