12-18-2014

Crystallographic Analysis and Molecular Modeling Studies of HIV-1 Protease and Drug Resistant Mutants

Chen-Hsiang Shen

CRYSTALLOGRAPHIC ANALYSIS AND MOLECULAR MODELING STUDIES

OF HIV-1 PROTEASE AND DRUG RESISTANT MUTANTS

by

CHEN-HSIANG SHEN

Under the Direction of Irene T. Weber, Ph.D.

ABSTRACT

HIV-1 protease (PR) is an effective target protein for drugs in anti-retroviral therapy (ART). Using PR inhibitors (PIs) in clinical therapy successfully reduces mortality of HIV infected patients. However, drug resistant variants are selected in AIDS patients because of the fast evolution of the viral genome. Structural, kinetic and MD simulations of PR variants with or without substrate or PIs were used to better understand the molecular basis of drug resistance. Information obtained from these extensive studies will benefit the design of more effective inhibitor in ART.

Amprenavir (APV) inhibition of PR_{WT}, and single mutants of PR_{V32I}, PR_{I50V}, PR_{I54M}, PR_{I54V}, PR_{I84V} and PR_{L90M} were studied and X-ray crystal structures of PR variants complexes with APV were solved at resolutions of 1.02-1.85 Å to identify structural alterations. Crystal structures of PR_{WT}, PR_{V32I} and PR_{I47V} were solved at resolutions of 1.20-1.40 Å. Reaction intermediates were captured in the substrate binding cavity, which represent three consecutive steps in the catalytic reaction of HIV PR. HIV-1 PR20 variant is a

multi-drug resistant variant from a clinical isolate and it is of utility to investigate the mechanisms of resistance. The crystal structures of PR20 with inactivating mutation D25N have been determined at 1.45-1.75 Å resolution, and three distinct flap conformations, open, twisted and tucked, were observed. These studies help understand molecular basis of drug resistance and provide clues for design of inhibitors to combat multi-drug resistant PR.

The evaluation of electrostatic force in MD simulations is the computationally intensive work, which is of order $\theta(N^2)$ with integration of all atom pairs. AMMP invokes Amortized FMM in summation of electrostatic force, which reduced work load to $\theta(N)$. A hybrid, CPU and GPU, parallel implementation of Amortized FMM was developed and improves the elapsed time of MD simulation 20 fold faster than CPU based parallelization.

INDEX WORDS: HIV-1 protease, drug resistance, amprenavir, crystal structure, kinetics, parallel computing, CUDA, OpenMP.

CRYSTALLOGRAPHIC ANALYSIS AND MOLECULAR MODELING STUDIES OF HIV-1 PROTEASE AND DRUG RESISTANT MUTANTS

by

CHEN-HSIANG SHEN

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2014

CRYSTALLOGRAPHIC ANALYSIS AND MOLECULAR MODELING STUDIES

OF HIV-1 PROTEASE AND DRUG RESISTANT MUTANTS

by

CHEN-HSIANG SHEN

Committee Chair: Irene T. Weber

Committee: Robert W. Harrison

John E. Houghton

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

December 2014

ACKNOWLEDGEMENTS

I would like to express my most sincere thanks to my advisor Dr. Irene Weber. I am deeply indebted for her guidance and advice in the past several years. She taught me how to think scientifically and gave me this opportunity to develop my career in science. Her positive attitude and constant encouragements definitely made major impacts in my life.

I have special thanks to my committee members: Dr. Robert Harrison, who shared interesting projects, offered great helps in understanding research in computer science, and gave me priceless advices; Dr. John Edgar Houghton, for his supportive guidance, encouragements, and insightful discussions about the research.

I give special thanks to Dr. Johnson Agniswamy and Yuanfang Wang who helped my research from initial training to sharing projects with insightful discussions and valuable suggestions. My special thanks to Dr. Andrey Kovalevsky for his influence in my life to choose science as my career. Thanks also go to my lab members Dr. Eric Chang, Dr. Ying Zhang, Dr. Hongmei Zhang, Dr. Guoxing Fu, Daniel Kneller and Andres Wong. I am very grateful to all faculty, staff and friends in the Department of Biology at the Georgia State University for their help, friendship, and support. I am also very appreciative to the staff in the research computing at Georgia State University for the assistance of using high performance computer. Special thanks to the staff at the SER-CAT beamline at Argonne National Laboratory for the assistance during X-ray data collection. My studies were supported in part by Georgia Research Alliance, Georgia Cancer Coalition, National Institute of Health and Molecular Basis of Disease Program at Georgia State University.

At last, I am indebted to my devoted family. My loving wife, Han Ting, has been a true and great supporter and has unconditionally loved me during my good and bad times. She has been non-judgmental of me and instrumental in instilling confidence. She has faith in me and my intellect even when I didn't

have faith in myself. I truly thank Han Ting for sticking by my side, taking care of my lovely daughter Felicitas, even when I was irritable and depressed. I am also grateful for my parents for their great understanding, endless patience and support. Without them, I would not have made it this far.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS v

TABLE OF CONTENTS vii

LIST OF TABLES xii

LIST OF FIGURES xiii

LIST OF ABBREVIATIONS xv

1 INTRODUCTION 17

1.1 HIV-1 Life Cycle 17

1.2 Structural Features of HIV-1 PR 21

1.3 Catalytic Mechanism of HIV-1 PR 25

1.4 Design of HIV-1 Protease Inhibitors 28

1.5 Viral Mutations and Drug Resistance 30

1.6 Theory of molecular dynamics simulations 34

1.7 Molecular Dynamics Simulation and HIV PR 36

1.8 Parallel Computing 37

1.9 Rational for the studies 41

2 AMPRENAVIR COMPLEXES WITH HIV-1 PROTEASE AND ITS DRUG RESISTANT MUTANTS ALTERING HYDROPHOBIC CLUSTERS 44

2.1 Abstract 44

2.2 Introduction 45

2.3 Experimental Procedures....48

2.3.1 Protein Expression and Purification.....48

2.3.2 Kinetic Assays....48

2.3.3 Crystallographic Analysis....48

2.4 Results....49

2.4.1 APV inhibition of PR and mutants....49

2.4.2 Crystal Structures of APV Complexes....51

2.4.3 PR interactions with APV and the influence of alternative conformations....54

2.4.4 Effects of Mutations on PR Structure and Interactions with APV....59

2.4.5 Comparison of the mutant complexes with APV and SQV....65

2.5 Discussion....70

2.6 Acknowledgment....73

3 CAPTURING THE REACTION PATHWAY IN NEAR-ATOMIC-RESOLUTION CRYSTAL STRUCTURES OF HIV-1 PROTEASE....74

3.1 ABSTRACT....74

3.2 Introduction....74

3.3 Experimental Procedures....77

3.3.1 Protein Preparation and Crystallization of HIV-1 PR_{WT}, PR_{V32I}, and PR_{I47V}....77

3.3.2 X-ray Data Collection and Refinement....78

3.4 Results....79

3.4.1 Crystallographic Analysis....79

3.4.2 Structural Changes around the Mutated Residues 81

3.4.3 PR_{WT}–TI Interaction 83

3.4.4 Interactions of PR_{I47V} with Two Products 83

3.4.5 Interactions of PR_{V32I} with the P5–P1 Product 86

3.4.6 Interactions with Catalytic Residues 86

3.5 Discussion 89

3.5.1 Implications for the Reaction Pathway 89

3.6 Acknowledgment 92

4 Dynamic Variation in the Flaps of an Extreme Drug Resistant HIV Protease Variant 93

4.1 Introduction 93

4.2 Materials and Methods 95

4.2.1 Preparation of PR20 with D25N mutation 95

4.2.2 Protein crystallization, X-ray data collection and structure determination 95

4.2.3 Molecular dynamics simulations of wild-type PR and PR20 96

4.3 Results 97

4.3.1 Crystal structures of Ligand-free $PR20_{D25N}$ 97

4.3.2 Flaps exhibit diverse conformations 100

4.3.3 Unusual tucked flap conformation 105

4.3.4 MD simulations 105

4.3.5 Cluster Analysis of simulations 106

4.3.6 Correlated motions extracted from MD simulations..... 108

4.4 Discussion..... 110

4.5 Acknowledgment..... 112

5 A HYBRID IMPLEMENTATION OF PARALLEL AMORTIZED FAST MULTIPOLE ALGORITHM FOR THE MOLECULAR MODELING PROGRAM: AMMP..... 113

5.1 Abstract..... 113

5.2 Introduction..... 113

5.3 Specification of Algorithms..... 114

5.3.1 Algorithm for update:..... 115

5.3.2 Algorithm for non-update:..... 116

5.3.3 Amortized Scalar FMM Algorithm:..... 117

5.4 Parallelization of AMMP..... 119

5.4.1 The OpenMP AMMP..... 119

5.4.2 The openMP-CUDA AMMP..... 120

5.5 Results..... 120

5.5.1 Sample preparation for simulations..... 120

5.5.2 Run-time Analyses..... 121

5.5.3 Stability of simulation by parallelized AMMP..... 123

5.6 Discussion..... 123

5.7 Acknowledgment..... 125

6 OVERALL SUMMARY AND DISCUSSION..... 126

7 REFERENCES 130

8 APPENDICES 141

Appendix A 141

Appendix B 142

Appendix C 152

Appendix D 162

Appendix E 173

LIST OF TABLES

Table 2-1 Kinetic parameters for substrate hydrolysis and inhibition of amprenavir. 50

Table 2-2 Crystallographic Data Collection and Refinement Statistics 52

Table 2-3 Hydrogen bond interactions between HIV-1 protease and APV. 58

Table 2-4 C-H···O interactions between HIV-1 protease and APV. 60

Table 3-1 Crystallographic Data Collection and Refinement Statistics 80

Table 4-1 Crystallographic Data Collection and Refinement Statistics 98

Table 5-1 Run time analysis of ATPase and HIV protease of two different implementations. 124

LIST OF FIGURES

Figure 1-1 The life cycle of HIV. 18

Figure 1-2 The known structures of HIV-1 proteins and protein fragments [17]. 20

Figure 1-3 The structure of HIV-1 PR dimer complexed with DRV. 22

Figure 1-4 Schematic diagram of a substrate (P4-P3') bound to HIV-1 PR (S4-S3') subsites. 22

Figure 1-5 Top view of different conformations of flaps are shown: 24

Figure 1-6 The proposed general acid-base catalysis mechanism for aspartic PR [45]. 27

Figure 1-7 Currently approved FDA drugs for HIV PR. 29

Figure 1-8 List of mutations that have been associated with resistance to PIs [97]. 33

Figure 1-9 Distributed Memory Model. 38

Figure 1-10 Shared Memory Model. 38

Figure 1-11 Parallel programming with Compute Unified Device Architecture. 40

Figure 2-1 46

Figure 2-2 2Fo–Fc electron density maps for the mutated residue in: 55

Figure 2-3 Inhibitor binding site in PR_{WT}–APV. 56

Figure 2-4 The interactions of mutated residues in 62

Figure 2-5 Structural differences between APV and SQV complexes. 67

Figure 2-6 Interactions of Ile50, Ile54' and Thr80'. 69

Figure 3-1 Structure of the HIV-1 PR dimer in a green backbone representation. 76

Figure 3-2 Mutations alter internal hydrophobic contacts. 82

Figure 3-3 Electron density maps for the peptide intermediates. 84

Figure 3-4 Hydrogen bond interactions: 85

Figure 3-5 Minor conformations of TI or products superposed on major conformations showing that main chain atoms retain similar positions. 87

Figure 3-6 Hydrogen bond interactions with the catalytic residues Asp25 and -25': 87

Figure 3-7 Scheme of the reaction pathway. 90

Figure 4-1 99

Figure 4-2 101

Figure 4-3 Interactions of Ile50 with PR20 residues. 102

Figure 4-4 Dimer conformations of $PR20_{D25N}$, PR_{WT}open (2PC0) and PR20open(3UF3). 104

Figure 4-5 Trajectories of the MD simulations. 107

Figure 4-6 Superposition of averaged structures of each cluster calculated for simulations: . 107

Figure 4-7 Dynamic Cross Correlation Maps show correlated motions of: 109

Figure 5-1 The elapsed time of HIV protease and ATPase simulated on DELL workstation and IBM x3850. 122

Figure 5-2 The parallel efficiencies are calculated from the simulations with various numbers of cores included in computation. 122

Figure 5-3 Time course plot of the RMSD draws by the superposition main chain atoms of the whole protein to the initial crystallographic structure. 124

LIST OF ABBREVIATIONS

Å	Angstrom
AIDS	Acquired Immuno Deficiency Syndrome
APV	Amprenavir
Cα	alpha carbon
CUDA	Compute Unified Device Architecture
DRV	Darunavir
FMM	Fast Multipole Method
GPU	Graphic Processing Unit
HAART	Highly Active AntiRetroviral Therapy
HIV	Human Immunodeficiency Virus
OpenMP	Open Multi-Processing
PDB	Protein Data Bank
PR	Protease
PI	Protease Inhibitor
PR20	PR with 20 mutations
PR_{WT}	wild type HIV-1 protease
PR_{V32I}	PR with V32I mutation
PR_{I50V}	PR with I50V mutation
PR_{I54M}	PR with I54M mutation
PR_{I54V}	PR with I54V mutation
PR_{I84V}	PR with I84V mutation
RMS	root mean square
RT	Reverse Transcriptase

SQV	Saquinavir
Bis-THF	*bis*-tetrahydrofuran
TI	tetrahedral intermediate
µl	microliter

1 INTRODUCTION

1.1 HIV-1 Life Cycle

Human immunodeficiency virus type 1 (HIV-1) is the pathogen, first reported in 1984, of acquired immune deficiency syndrome (AIDS) [1]. Estimations from UNAIDS in 2013 show that globally, 35.3 (32.2–38.8) million people live with HIV and each year there are 2.3 million new HIV infections [2]. HIV is a member of the retroviruses family and there are two major subtypes, HIV-1 and HIV-2. The originally identified virus is the type 1 HIV and is more virulent than type 2 [3]. The major type causing AIDS around world is HIV-1.

HIV attacks and destroys specific CD4 T cells of the human's immune system, causing a significant deterioration of the body's defenses and reduced ability to fight various diseases. The general steps of life cycle of HIV are shown in Figure 1-1. Viral particle consists of proteins and genome encoded by a ~9-kb positive–sense RNA. The life cycle begins from the viral envelope glycoprotein gp120 of HIV recognizing and binding to receptor CD4 on host cell surface [4]. This induces conformational changes of the viral envelope glycoprotein gp120 and gp41 and fuses viral particle with the host cell membrane [5]. The genomic RNA, two copies of single-strand RNA, of HIV is released into the cytosol of the host cell and is reversed transcribed into viral cDNA by viral RT [6]. Pre-integration complex formed by viral cDNA, MA, IN and other viral proteins is further imported into nucleus by the help of viral accessory protein, Vpr [7, 8]. Viral cDNA is inserted into host genome with the help of viral IN. Subsequently the replication of the genome of infected host cell also replicates viral genome. The viral genome codes three polyproteins: Gag, Gag-Pol and Env, which are afterward proteolyzed into functional and structural proteins by viral and human proteases. Gag and Gag-Pol polyproteins share the same starting codon, however, ribosomal frame shift between NC and p6 during translation of Gag polyprotein results in Gag-Pol polyprotein [9]. The uncleaved viral polyproteins and viral RNA are assembled at the interior of the cell membrane and

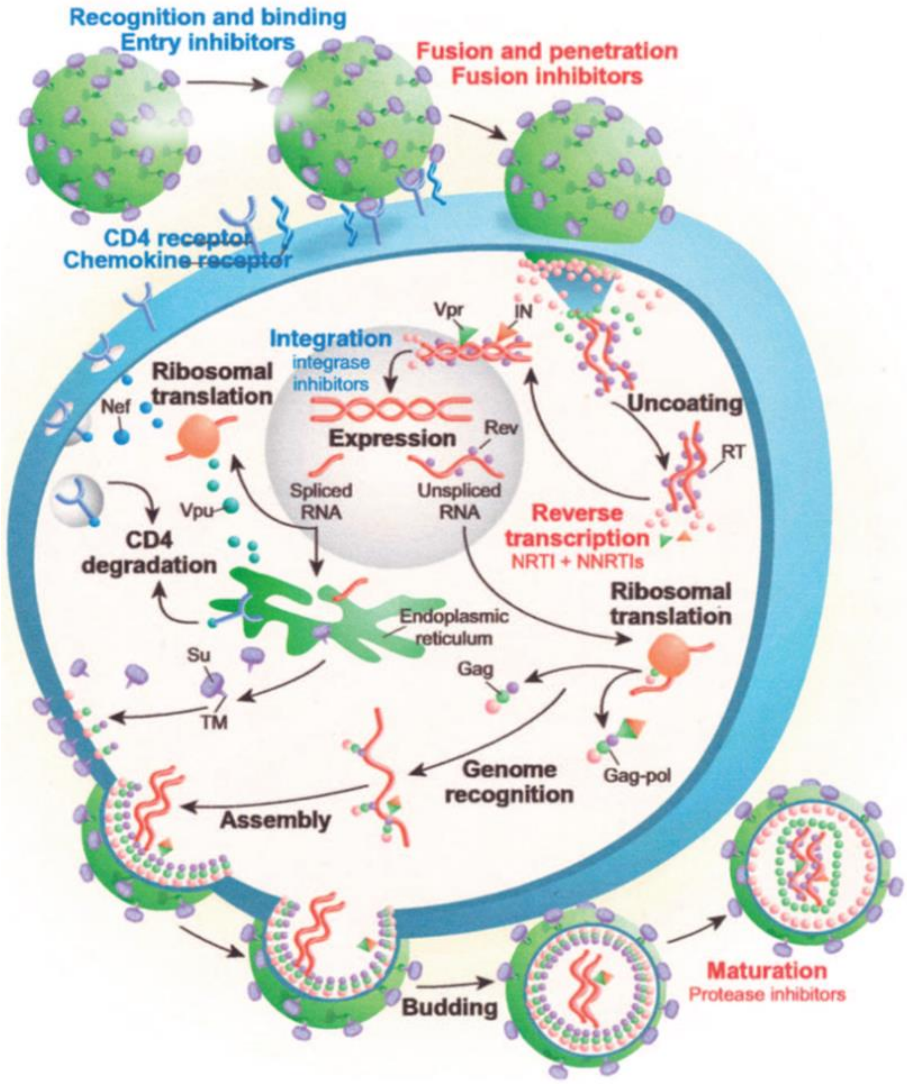


Figure 1-1 The life cycle of HIV.
General essential steps in HIV-1 life cycle [10].

bud from the surface of membrane as immature virions. At this stage, the virus cannot infect other host cells. The PR precursor in the Gag-Pol polyprotein hydrolyzes itself by folding and dimerizing in the immature virion to release mature PR [11, 12]. Active HIV PR cleaves Gag and Gag-Pol polyproteins into individual proteins. The rates of cleavage differ up to hundred fold between different cleavage sites, and the correct order of cleavage of Gag and Gag-Pol polyproteins into individual proteins is necessary for maturation of virions [13]. Thus, PR is a valuable drug target since inhibition of PR catalytic activity results in immature noninfectious virions [14, 15].

The general structural features of the mature HIV-1 virion are shown in Figure 1-2. The core and the outer membrane envelope are composed of four Gag proteins and two Env proteins which correspond to matrix (MA), capsid (CA), nucleocapsid (NC), p6, surface unit glycoprotein (SU), and transmembrane (TM). Essential enzymatic functions are carried out by three Gag-Pol proteins which are protease (PR), reverse transcriptase (RT), and integrase (IN). These enzymes are encapsulated within the virus particle. In addition to these proteins, there are many accessory proteins found, such as Vif, Vpr, Nef, Tat, Rev [16, 17].

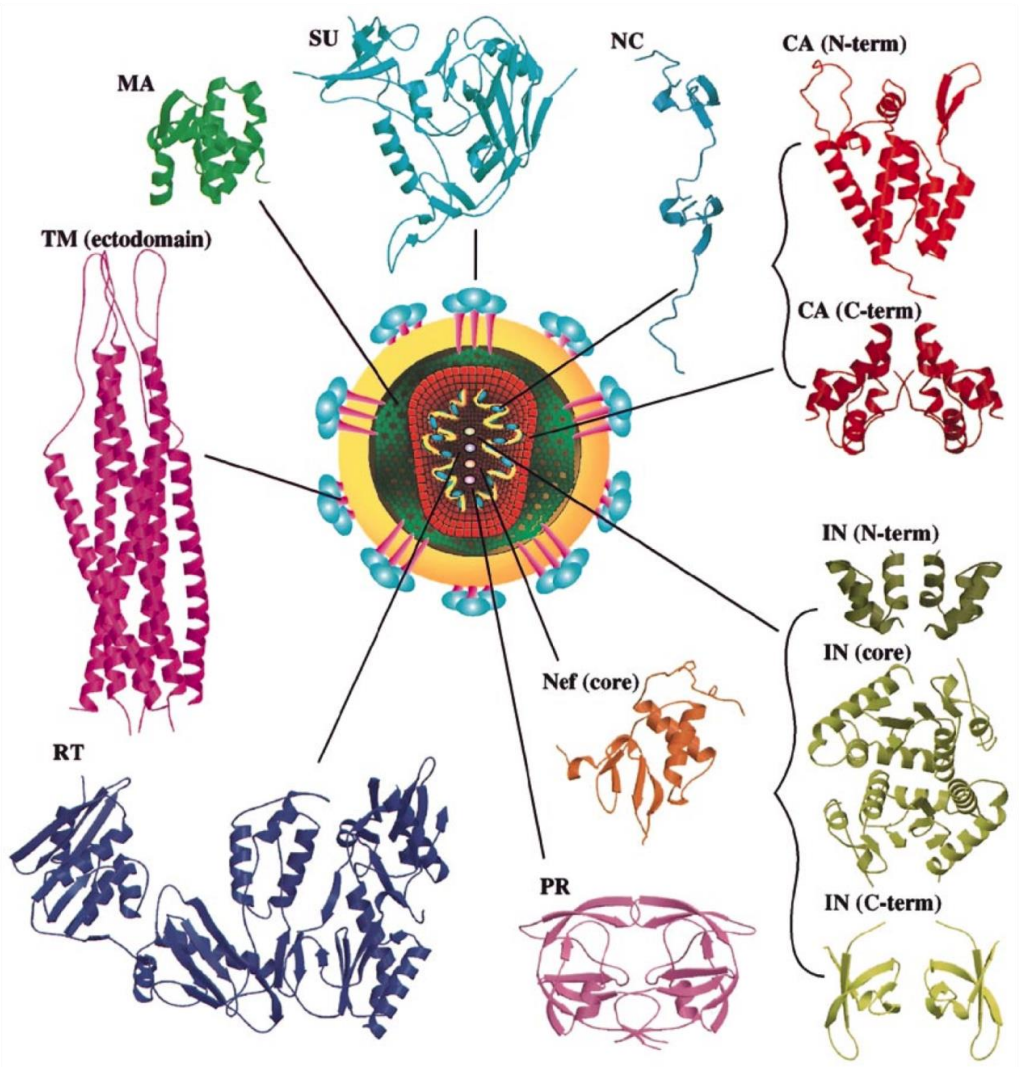


Figure 1-2 The known structures of HIV-1 proteins and protein fragments [17].

1.2 Structural Features of HIV-1 PR

HIV-1 PR is the enzyme responsible for the cleavage of viral Gag and Gag-Pol polyproteins into mature, functional proteins. Currently, more than 400 structures of HIV-1 PR complexed with or without various inhibitors have been deposited to the Protein Data Bank server (http://www.pdb.org). The structural features of PR are shown in Figure 1-3. The dimeric aspartic protease is composed of residues 1-99 and 1'-99'[18]. The substrate binding cavity formed by residues 8, 23-30, 32, 45-50, 53, 56, 76, 80-82 and 84 accommodates and recognizes 7 amino acids of peptide substrate. The conserved catalytic triplets, Asp25-Thr26-Gly27, from both subunits provide the key elements for formation of the enzyme active site[18]. In the substrate binding cavity, Asp25, Gly27, Asp29 and Gly48 provide hydrogen bonds with main chain atoms of substrate[19]; residues with hydrophobic interactions with substrate are Leu23, Gly27, Ala28, Val32, Lys45, Ile47, Met46, Gly48, Gly49, Ile50, Phe53, Leu76, Thr80, Pro81, Val82 and Ile84[20]; polar side chains or distal main chain groups in longer peptides can interact with Arg8, Asp30 and Lys45[21]. The amino acid side chains of substrate peptide, P4 to P3', and corresponding substrate binding site, subsites S4 to S3' are shown in Figure 1-4 [22]. Important residues in each substrate binding subsite are also shown in the figure. Based on composition of residues around each subsite, S1 and S1' subsites are hydrophobic and prefer hydrophobic amino acids in corresponding P1 and P1' subsites[22]. S2 and S2' subsites are hydrophobic and can bind polar residues at P2 and P2'. The rest of subsites at both ends are solvent accessible, thus the preferred amino acids in P3-P4 and P3'-P4' are more varied.

Binding of substrate or inhibitor goes together with large structural changes on two glycine-rich flaps, residues 45 to 55 from each subunit. Mutagenesis study on Gly-rich region, Gly48, Gly49, Gly51 and Gly52, of the flaps show this region is crucial for PR activities [23]. Disrupting flexibility of flaps could lead to the reduced catalytic activities of the PR[19]; For the apo form of PR, significant structural changes on flaps are observed in various techniques, such as X-ray crystallography, pulsed electron paramagnetic resonance (EPR) spectroscopy and molecular dynamic studies[24, 25]. Three main categories, closed,

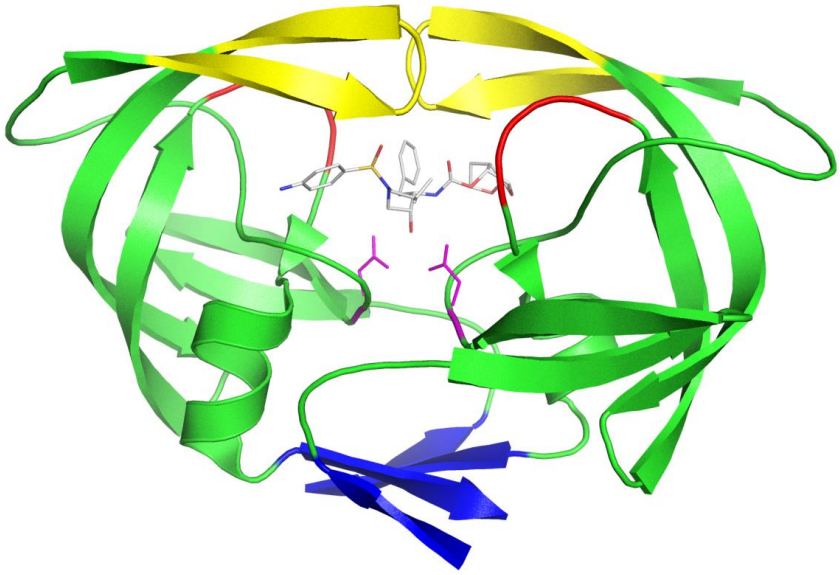


Figure 1-3 The structure of HIV-1 PR dimer complexed with DRV.
Conserved N- and C- terminal β-sheets, flaps 80's loop are represented in blue, yellow and red, respectively. Catalytic dyad of Asp25, Asp25' is indicated as magenta sticks. DRV is located in the center of substrate binding cavity and is represented in grey.

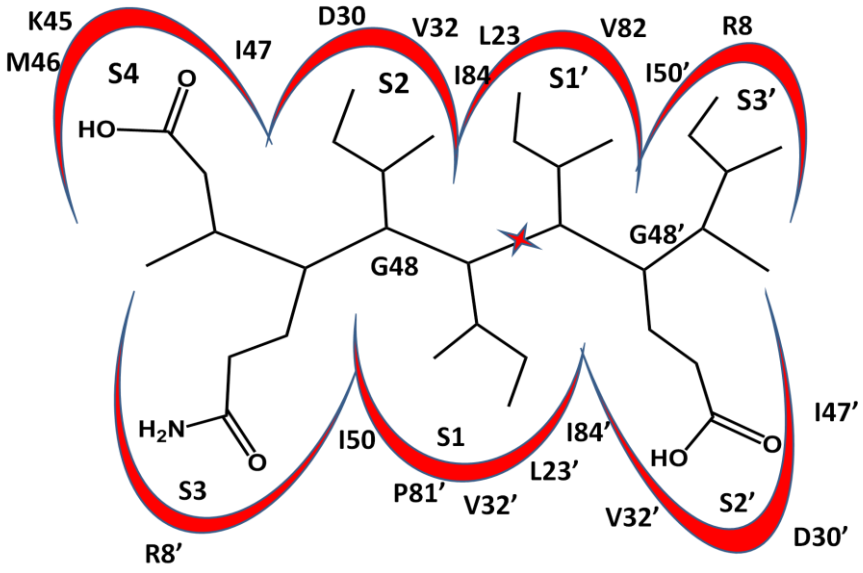


Figure 1-4 Schematic diagram of a substrate (P4-P3') bound to HIV-1 PR (S4-S3') subsites.
The scissile bond is indicated by a red star. Substrate binding subsites are colored as red arcs. Residues contributing to the substrate binding cavity are labeled

semi-open and open conformation, of flap are defined based on observation in different crystal structures of the apo-forms of PR (Figure 1-5) [25]. Closed flap is typically observed in PR complexed with substrate / inhibitor; thus it prevents entrance of substrate as well as leaving of captured substrate. The tips of the closed flaps are interlocked over the bound substrate or inhibitor, and the main chain Cα of Ile50 and Gly51 lie parallel to each other in the tips of the two flaps. In the semi-open conformation the tips of the flaps switch their interactions compared to the closed conformation. In the open conformation, tips of the flaps are shifted up and separated from each other. The tips of the opened flaps lose hydrogen bond and Van der Waals interactions to the opposite flap and 80' loop, residues 78 to 85, thus the active site cavity is exposed for the access of substrate / inhibitor.

Unlike cellular aspartic proteases where the active form is monomeric, the dimeric form of HIV-1 PR is crucial for its proteolytic activities [26]. The homodimeric form of PR is stabilized via noncovalent interactions at the dimer interface between residues from two subunits [27]. Four-stranded terminal β-sheet of each monomer (residues 1-4, 96-99, 1`-4` and 96`-99`), tips of the flaps, catalytic triplets and salt bridges between D29 and R87 of one monomer and R8' of the opposite monomer are important elements in HIV-1 PR dimer interface [28]. Todd *et al.* using thermodynamic analysis to evaluate dimer interface of HIV-1 PR showed that the beta sheet formed by the four termini contributes 75% of the total Gibbs energy, the active site residues (T26, G27 and D29) are also important contributors for dimer stability, and the lesser contributors are flap residues (G49, I50 and G51) [29]. Site direct mutagenesis of F99A disrupts the dimeric interface and produces monomeric PR[30]. Crystallographic analysis of HIV-1 PR also shows 50% of the intermonomeric ionic and hydrogen bond interactions are contributed by the C-termini of each monomer [28]. Because PR activity depends on dimerization, residues are highly conserved at the dimer interface with low mutation rates even after inhibitor treatments [31, 32]. Thus, instead of targeting on the substrate binding cavity, the N- and C-terminal residues are suggested as second target region for rational drug development against retrovirus [28]. Schramm et al. designed short lipopeptides that mimic

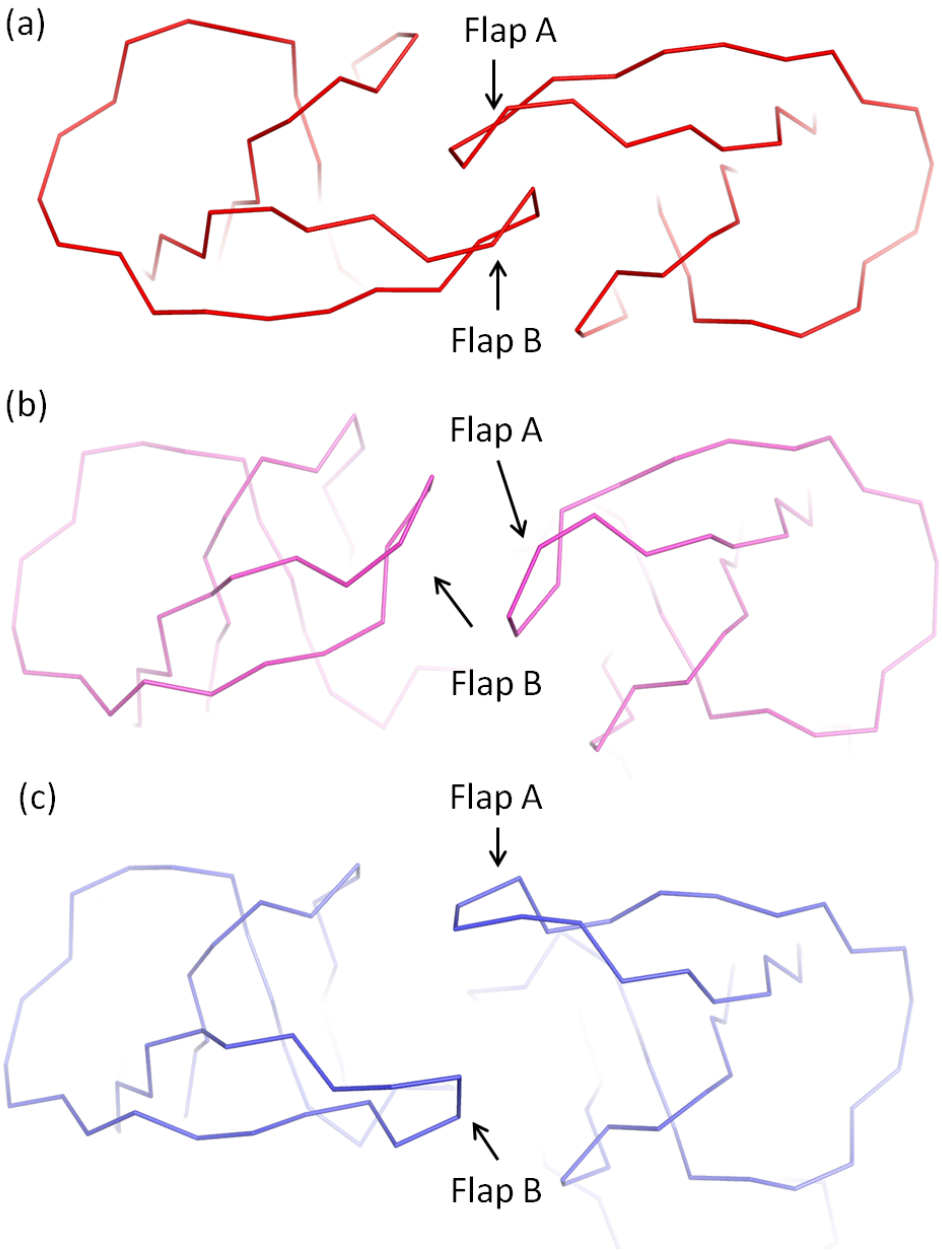


Figure 1-5 Top view of different conformations of flaps are shown:
(a) wild type PR complexed with DRV shown as closed flap(PDBID: 2IEN, colored red), (b) unliganded PR with F53L mutation represented as semi-opened flap (PDB ID: 2G69, colored magenta), (c) unliganded wild type PR shown as opened flap (PDB ID: 1HHP, colored blue).

interface between terminal peptide fragment, and these peptide display ki values in the nanomolar range [33]. Bannwarth et al. synthes tripeptides targeting dimer interface, the dimerization inhibitor successfully inhibits dimerization of PR[34]. Nevertheless, no dimerization inhibitor has been approved by FDA [35].

1.3 Catalytic Mechanism of HIV-1 PR

The HIV-1 PR activity is blocked by pepstatin, a natural product that selectively inhibits aspartic protease family members, indicating that HIV-1 PR is a member of aspartic protease family [36, 37] . The substitution of Asp25 to Asn25 leads to the production of immature viral particles, which shows that the Asp25 of each subunit is the residue critical for catalytic activity [38].

The peptide bond of substrate is hydrolyzed by catalytic Asp dyad, and the catalytic process is affected by the atoms surrounding the Asp25 / Asp25' within the protein microenvironment. Several studies have been done experimentally or theoretically to study reaction mechanism of peptide cleavage by HIV-1 PR. ^{18}O-exchange mass spectrometry experiment with HIV-1 PR suggests a nucleophile water attacks carbonyl carbon at scissile bond and forms reversible and metastable reaction intermediate, gem-diol structure [39]. Experiments done by kinetics method and NMR show that two aspartate 25 groups in the dimer have distinct protonization states [40, 41]. The combined neutron crystallography and X-ray crystallography study on HIV PR complexed with transition-state inhibitor KNI-272 and with perdeuterated PR with APV indicated different protonation of the catalytic aspartic dyad [42, 43]. However, the type of inhibitor captured in active site cavity might affect the location of hydrogen atom and the hydrogen bond at catalytic aspartic dyad. Still, the protonation state of catalytic aspartic dyad might be different when hydrolysable substrate is captured in active site cavity, and the protonation state may change during the course of the reaction.

Different proteolytic models have been proposed to illustrate the catalytic mechanism of aspartic protease based on experimental results. However, the detailed catalytic mechanism is still controversial. The general acid and general base mechanism, proposed by Suguna in 1987, is widely accepted

mechanism for aspartic PR family [44]. The proposed mechanism is based on observation of a 1.8 Å crystal structure of *Rhizopys chinensis* PR complexed with a substrate analog. In the first step of general acid-base mechanism, one of the aspartate groups in the protonated state acts as an acid to polarize carbonyl oxygen at scissile bond, while the other aspartate group is deprotonated and acts as a base to attack the lytic water (Figure 1-6). The lytic water subsequently attacks carbonyl carbon at scissile bond and forms a metastable tetrahedral intermediate (TI). One proton originated from TI further transfers to the amide nitrogen at scissile bond. The protonated nitrogen reduces stability of TI and causes peptide bond breakage producing two products. Hyland proposed a similar model, based on general acid / general base model, for HIV-1 PR[41].

In addition to the proposed lytic water molecule bridging the catalytic aspartic dyad, a second water molecule connects two tips of the flaps. This water molecule makes hydrogen bond interactions between two carbonyl oxygens of the substrate at both sides of the scissile bond and amide nitrogen of Ile50 of each monomer. This water molecule has been observed in most of the dimer structures in the closed conformation and has been proposed to be involved in the catalytic reaction [19].

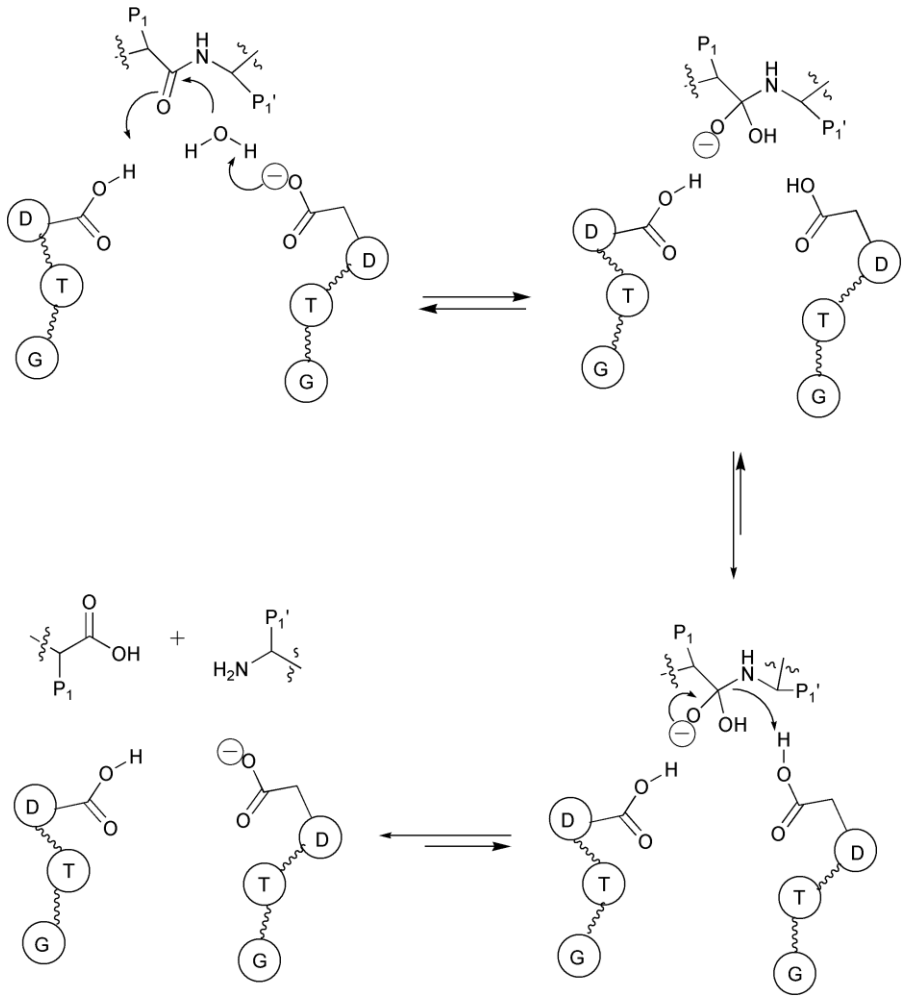


Figure 1-6 The proposed general acid-base catalysis mechanism for aspartic PR [45].

1.4 Design of HIV-1 Protease Inhibitors

HIV-1 PR is responsible for maturation of viral particle. In viral maturation, the viral precursor polyproteins Gag and Gag-Pol are cleaved into the structural and enzymatic proteins by HIV-1 PR [18, 46]. Thus, PR is one of the target enzymes to inhibit HIV life cycle and the use~~s~~ of PIs in antiretroviral treatment has been successful in extending life span of HIV infected patients[47]. Presently, 9 drugs, saquinavir, ritonavir, indinavir, nelfinavir, lopinavir, atazanavir, fosamprenavir, tipranavir, and darunavir, are approved by FDA (Figure 1-7) [48]. Except tipranavir, the rest of PIs are peptidomimetic agents. The first designs of PIs benefited from the designs of inhibitors of eukaryotic aspartic proteases, which mimic the transition state diol structure formed during hydrolysis of the peptide bond [49]. Thus, except fos-amprenavir is a phosphate ester pro-drug, 8 inhibitors have a hydroxyl moiety in order to block catalytic activities. X-ray crystallographic structures of PR complexed with substrate analog solved in 1989, and later, guided the design of PIs to improve the binding affinity and specificity [50-52]. Structure-guided drug design is one of the powerful approaches in drug development while three dimensional structures of target protein are available or predicted accurately[53]. Structural analyses of PR complexed with substrate analogs show that residues 25-29 and 48-50 have hydrogen bonds and hydrophobic interactions to substrates or inhibitors [54, 55]. Saquinavir, approved in 1995, was the first PI and keeps many features of substrates [56, 57]. Since that time, drug resistant mutations emerged quickly after SQV applied in clinic. Structural analysis of mutant PRs that are resistant to SQV reveals altered the interactions between PR and inhibitors [56]. Similar alterations on PR structures are also observed for drug resistant PR complexed with indinavir, ritonavir or nelfinavir, respectively [58-60]. Nevertheless, other disadvantages appear in first generation inhibitors in clinic, such as large molecular weight, poor bioavailability, faster plasma clearance rates, low tolerance, and toxicity [48]. The X-ray structural analysis of HIV-1 PR and its drug resistant variants reveals that the structures of main chain atoms of residues are stable between mutants and wild type PR[61]. Additionally, some amino acids are conserved in PR [32], because

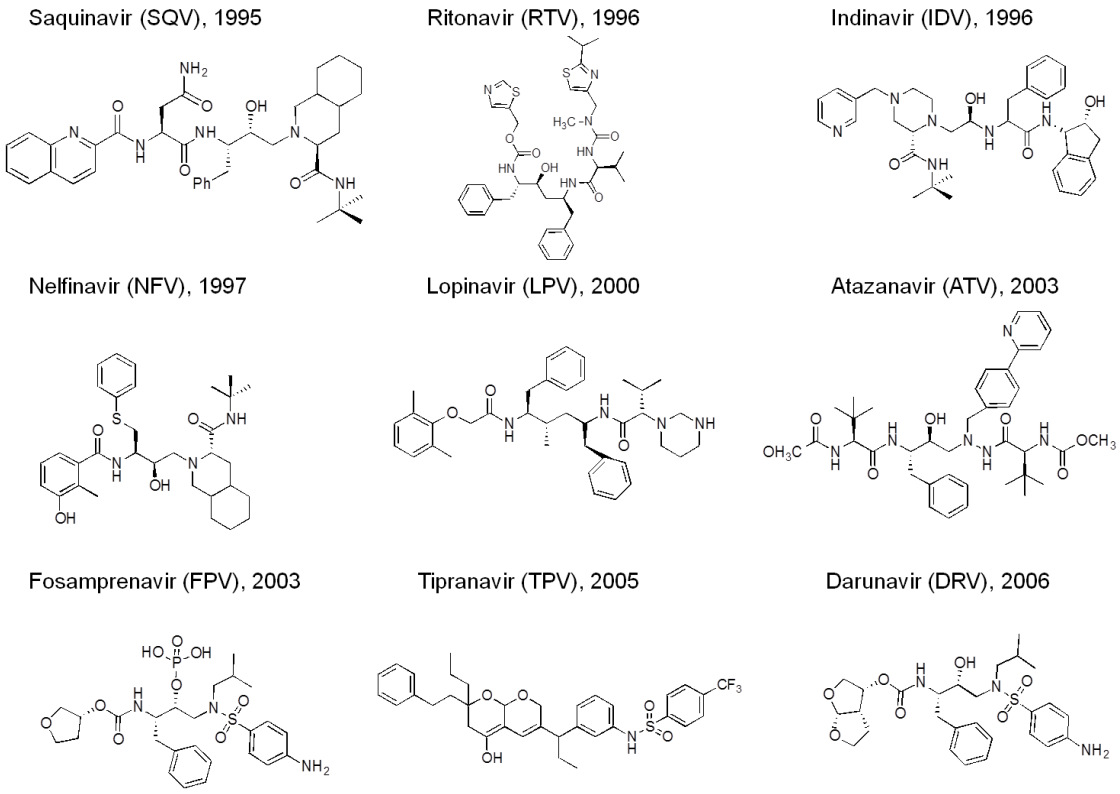


Figure 1-7 Currently approved FDA drugs for HIV PR.

Based on the structural similarities to the peptide substrates, these PIs can be grouped into first generation PIs that keep peptide-like properties from the substrate (SQV, RTV, IDV and NFV) and second generation PIs that have fewer structural similarities to substrate. (LPV, ATV, FPV, TPV, DRV) [48]

substitutions on conserved residues may inactivate PR [30, 62]. Therefore, the next generation inhibitor DRV was designed based on Dr. Ghosh's backbone-binding concept by increasing the interactions with main chain atoms, conserved amino acids that are critical for PR activities and lower peptidomimetic similarities [61, 63]. Based on this strategy, backbone atoms and conserved residues in S2-S2' subsites will be the target to maximize the hydrogen bond interactions with inhibitor [64]. Sulfonamide isostere and *bis*-tetrahydrofuran (bis-THF) moieties are introduced in the P2 /P2' group of Darunavir (DRV). In wild type PR, the binding affinity of DRV is hundred fold higher than for APV, ATV, LPV [65]. FDA approved DRV for clinical use in 2006 and it was successfully applied in AIDS therapy [66, 67]. Analyses of X-ray structures of wild type PR complexed with DRV show that DRV has more hydrogen bond interactions with PR than seen in wild type PR complexed with SQV and these polar interactions are maintained in drug resistant mutants [58, 68-70]. In summary, the effectiveness of DRV in combating HIV/AIDS demonstrates the success of using the backbone-binding strategy for design of inhibitors.

1.5 Viral Mutations and Drug Resistance

The evolution of HIV happens naturally as sequence polymorphism of HIV is observed prior the drug treatment due to the absence of a 3'-5'exonuclease proofreading activity in HIV RT [71-73]. Approximately half of the PR residues show sequence polymorphisms that explain why the drug resistant mutations exist before PI treatment in some infections [74]. Thus, in reality, there is no single wild type PR because it is the mixture of various HIV sequences. Furthermore, concurrent infection of HIV with different sequences might boost the sequence diversity through genetic recombination [75, 76]. At the end, the dynamic equilibrium of these minor variants in HIV infected patients is changed to select drug resistant mutations under drug pressure. Drug resistant mutations in PR not only alter the binding to the drug but also can ultimately change the enzymatic properties such as catalytic activity or stability, furthermore, mutations on the processing sites of Gag and Gag-Pol polyproteins compensate for drug resistant mutations that lower catalytic activity [77-81]. In addition, drug resistant mutations on PR

sequence are not limited to the substitution of residues and include amino acid insertion on substrate and PR sequence [82-85], however, fewer studies have focused on the impact of drug resistance after insertion or deletion of PR sequences.

The development of vaccines to prevent HIV infections still is extremely challenging [86, 87]. Currently antiretroviral therapy depends on Highly Active Anti-retroviral Therapy (HAART), introduced in 1996. HAART successfully reduces death rates with HIV infection, as the consequence, there are more patients living with HIV [88]. HAART uses the combinations from more than 20 different drugs, including inhibitors of the HIV-1 enzymes, RT, PR and IN, and inhibitors of cell entry and fusion. Nevertheless, the effectiveness of HAART cannot fully eradicate HIV from infected persons, because infected host cell DNA contains viral genome and many drugs cannot cross into the brain which has a reservoir of HIV infected cells.[89]. In addition, the key challenge of antiretroviral therapy is the rapid evolution of the drug resistant mutants because these mutants escape from the drug treatment.[90]. Thus, the gathering of sequences of drug resistant mutants will be critical for researchers and clinicians to establish correlations between mutant and resistance. Several HIV drug resistance databases have been built in order to provide prevalence and trend of resistant mutations of clinical inhibitors [91, 92].

Drug resistance is selected by combining several resistance mutations, many drug resistant mutations identified in clinical isolates significantly increase resistance to a PI. For licensed PIs, more than 60 mutations have been observed in 36 residues and reported as drug resistant mutations by the International AIDS Society-US panel (Figure 1-8) [31]. Extensive studies have been done to seek the molecular basis of drug resistance. Based on the association of the resistance mutations to PIs, these selected mutations can be classified as 15 major mutations, causing drug resistance to one or more PIs themselves, and 21 accessory mutations, appearing with other mutations [93]. Also the location of the mutation on PR can be identified as active site, flap region, dimer interface and distal mutations that alter residues not in the active site cavity and flap. Currently, crystallographic analyses by comparing wild type

PR and its single mutants have identified the molecular mechanisms for drug resistance to PI which are: 1) mutations that reduced interactions with inhibitor: the sites of these mutations are L23, D30, V32, M46, I47, G48, I50, V82, and I84. These mutations surround the active site cavity and lead to the decreased binding with various inhibitors. 2) mutation that shifts main chain atoms: various mutations on V82 in the active site cavity are found in resistant mutants [94]. Structural studies of V82A mutant display shifting on main chain atoms to accommodate the inhibitors [60, 68, 95]. Nevertheless, the structural changes on main chain atoms do not fully compensate for the loss of hydrophobic interactions with inhibitor due to substitution of a smaller side chain [21, 68, 95]. 3) mutations altering the dimer interface: Dimeric PR is essential for PR activity, the reduced stability of PR as well as reduced interactions between two subunits were observed in L24I, I50V and F53L [18, 96]. Thus, the mutations on dimer interface are considered as one of mechanisms to produce drug resistance. 4) active site cavity alterations from the changes in distal mutations [20]. Unlike the mutations listed above, distal mutations normally cause subtle and variable effects [58, 60, 97]. For instance, structural changes caused by mutations G73S, N88S and L90M alter the interactions or network of interactions leading to the active site, thus having an indirect interference on the interaction with inhibitors [18, 96, 98, 99].

Different PIs select for different mutations, although there are overlaps. Generally, a single drug resistant mutation does not show significant resistance to specific PI in phenotype, accumulation of more resistant mutations is required to produce high level resistance and alter the phenotype. Multi drug resistance in clinical isolate may result from the combined usage of different PIs in HAART, which causes extreme multidrug resistance. Previous studies of single and double mutations indicate that the structural alternations caused by single mutants are usually observed in the double mutants, too. However, the other properties of PR may not be conserved [58, 100]. Clinical isolate PR variant MDR769 possesses 10 resistant mutations at L10I, M36V, M46L, I54V, I62V, L63P, A71V, V82A, I84V, and L90M, and reveals 2000-

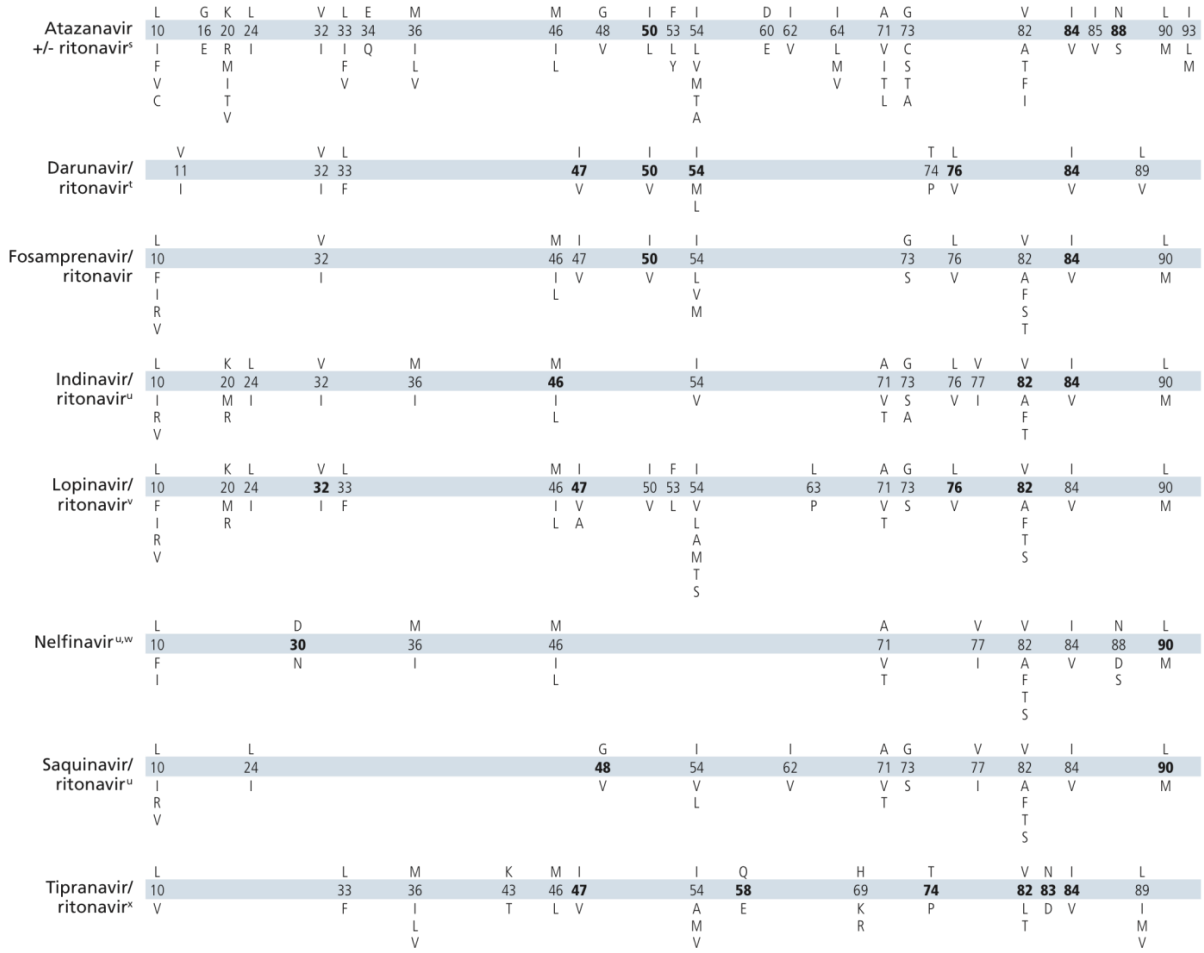


Figure 1-8 List of mutations that have been associated with resistance to PIs [101].
The bold number indicates major mutations for the corresponding PI.

and 700- fold reduced susceptibility to DRV and RTV compared to wild type PR [59, 102]. Dierynck et al. did the kinetic study based on surface plasmon resonance sensor for five clinical isolated multi drug resistant variants that possess 10 - 14 resistant substitutions, these mutants display poorer binding affinity due to the faster dissociation rate [103]. Compared to wild type PR, PR20 harboring 20 mutations [Q7K, L10F, I13V, I15V, D30N, V32I, L33F, E35D, M36I, S37N, I47V, I54L, Q58E, I62V, L63P, A71V, I84V, N88D, L89T and L90M] increases dissociation constants(K_L) of DRV and SQV by 8,000 and 2,000 fold, respectively [104]. The drug resistance mechanisms are not easy to identify in variants with multiple mutations, the structural changes might be contributed by the combination of major and minor mutations. X-ray crystal structures of MDR769 show altered PR conformation and reveal unusual wide open flaps and expanded S1/S1' and S3/S3' subsites [102, 105]. Hydrogen bonds and van der Waals interactions are lost between MDR769 and RTV [59]. The alteration of hydrophobic pocket in substrate binding cavity is reported in structures of MDR769 V82T complexed with DRV and TPV [106]. PRP51 with 14 amino acid substitutions that selected for resistance of DRV, which display wide open flap conformation, reveals unusual binding orientation of DRV in substrate binding cavity [107]. Structural analyses of multi drug resistant mutations indicate that combination of multiple mutations simultaneously show compensating structural changes, altered interactions with inhibitor and significant widening of the substrate binding cavity.

1.6 Theory of molecular dynamics simulations

It is difficult to analyze properties of the individual components in the complicated system, for instance macromolecule complex, because biological system is a macro system including numerous protein and solvent molecules. Molecular Dynamics (MD) simulation which also referred to as "in silico experiment" or "virtual experiment" uses computational method that models biological macromolecules as a micro system in order to study the molecular motion, vibration and thermodynamics. Thus, MD simulation represents a different way to observe phenomena and predict properties for proteins that cannot easily be observed experimentally. The first molecular dynamic studies applied in biological system

were the simulation of bovine pancreatic trypsin inhibitor, which appeared in 1977[108]. Now, more and more researchers use MD simulation to address problems, such as molecular motion, vibration and thermodynamics of proteins, and to guide wet laboratory experiment and/or to explain the experimental results.

Based on Born-Oppenheimer approximation, an atom can be treated separately as the electrons and nucleus [109]. The lighter electrons are approximated and represented as potential energy, and the simulation of heavier nucleus representing the mass of the atom follows classical Newtonian dynamics. Force fields are used to define potential energy of each atom in the system, and each atom type has specific parameters for its force fields that are obtained by empirical fitting from experimental data. Force fields can be simply expressed as bonding and nonbonding potential energies in biological and chemical systems:

$$V_{total} = \Sigma(V_{bonding} + V_{nonbonding})$$

$$V\ (conformations) = E_{bonds} + E_{angles} + E_{torsion} + E_{vdw} + E_{electrostatic}$$

The van der Waals interaction is calculated by the Lennard-Jones potential and paired ionic interaction is evaluated from Coulomb's law. The chemical and physical properties of proteins make the number of bond, angle and torsion terms a fixed number and the steps for calculating these forces are fixed and relative to the size of system. The nonbonded interactions calculated from all pairs of atoms are built dynamically since the atom moves to a new position after each time step, and the step for direct summation of long range interaction is the square of the number of atoms. Thus, calculations on nonbonded interaction are the most time-consuming work for MD simulation. The “cut-off radius” is introduced to restrict calculation and save CPU time. This method calculates nonbonded interactions for the atoms located within a specific distance. However, truncating the energy function leads to an unrealistic variation in energy [110, 111]. Two numerical algorithms, Ewald method using Fourier

transform and fast multiple method (FMM) invoking multipole expansion, are applied to compute long range nonbonded terms to minimize energy drift and CPU usage [112, 113].

The atomic positions of each atom in the system are computed by the integration of the equations of motion. Algorithms like the Leap-frog algorithm, computes the position of an atom at time

$$r(t + \Delta t) = r(t) + v(t + ½ \Delta t) \Delta t$$

The velocities are calculated from

$$v(t + ½ \Delta t) = v(t – ½ \Delta t) + a(t) \Delta t$$

Where, *a* is the acceleration calculated from the energy expression based on Newton's law. Velocities are explicitly calculated in Leap-frog algorithm; however, they are not calculated at the same time as the positions. Thus, large fluctuations in energy and violate conservation of energy are made by small variations in predict atomic locations. Predict and correct algorithm is proposed that calculations include midpoint of forces to improve the accuracy of new atomic positions of each atom in the system, which improves stability of molecular dynamics simulation [111].

$$r(t + \Delta t) = r(t) + (v(t-½ \Delta t)+v(t + ½ \Delta t)) \Delta t/2$$

1.7 Molecular Dynamics Simulation and HIV PR

Molecular dynamics simulation is suitable for studying various molecular basis questions that cannot be directly observed by experimental instruments. Toth et al. used molecular dynamics to study the mechanism of flap opening and propose that polar interactions between the flaps are responsible for stabilizing the semi-open flap [114]. Previous study by Hamelberg et al. showed that the formation of loop conformation on flap is contributed by its fast trans−cis isomerization of the Gly-Gly ω-bonds and this loop conformation is critical to flexibility of flap [115]. Combined studies of NMR and MD simulation show that the drug resistant mutations located on flap region alter enzyme kinetic property in turn-over rate and drug binding [116]. Scott et al. discover the tips of the flap are curling in 10 ns MD simulation, which refers as drug resistant mechanism [117]. Harrison et al. report all atom simulation of PR / substrate complex to

explore protonation state of two catalytic Asp [118]. Maturation of HIV PR is studied by using explicit solvent MD simulation, which indicates that dimerization is the rate-determining step [119]. Additionally, atoms in the system are assigned with kinetic energy under certain temperature, therefore changes on binding energy can be used to help the design of novel inhibitor as well as the studies on drug resistant mutations [120-124].

1.8 Parallel Computing

Parallel computing is a type of computation that uses multiple resources concurrently to solve computational problem. Currently, three programming models, shared memory, distributed memory and GPU-accelerated computing, have been broadly applied for parallelizing MD algorithms [125-129]. Distributed memory system refers to a computing system that is composed of several individual nodes (Figure 1-9). Node is the building block of large system. Each node has its own processor and memory space, each processor is limited to address local memory. Several nodes are connected through variety of networks to build larger system and data as well as memory space is passed from node to node. Message Passing Interface (MPI) is popular library for programming in distributed system [130]. The advantage of distributed system is that it is easy to build a scalable many processors system, for example the system with more than 1,000 processors. The disadvantages of using MPI to parallelize serial code are that it is complex to split data into different nodes and to transfer data between nodes; thus it is difficult design high performance parallel algorithm in distributed system.

Shared memory system is the system in which software uses multiple cores or threads working with the same memory space simultaneously (Figure 1-10 A). Parallel computing by shared memory model has been broadly applied for parallelizing MD algorithms. Open Multi-Processing (OpenMP) is the industry standard library and is easily implemented for most computationally intensive loops by adding directives

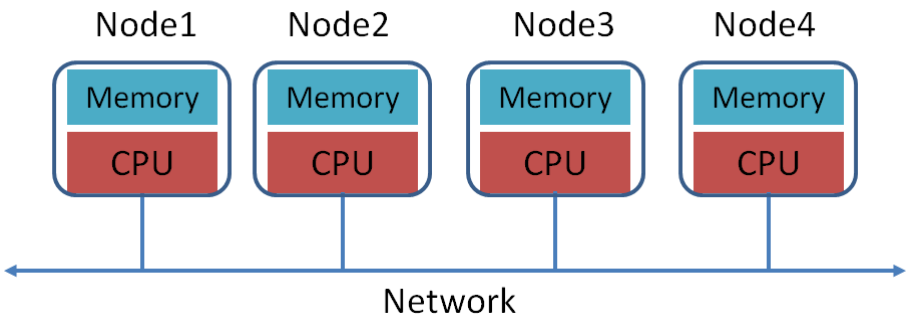


Figure 1-9 Distributed Memory Model.
Each node has its own CPU and memory, nodes are connected by high speed network.

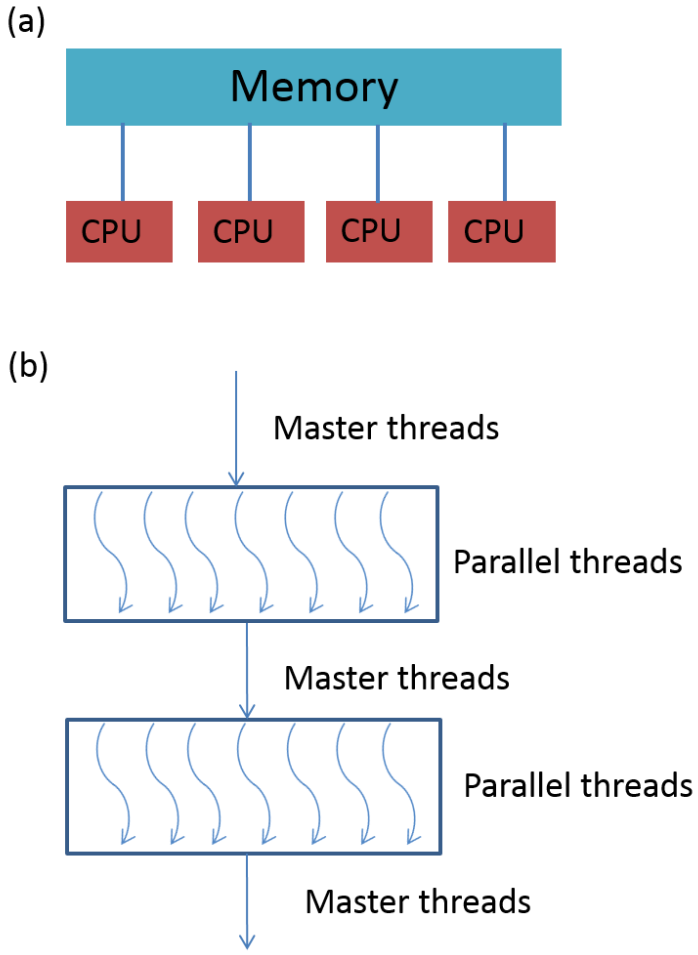


Figure 1-10 Shared Memory Model.
(a) The basic unit in shared memory model has multiple processing units, all the processing units maintain the same memory space. (b) fork-join parallelism starts from single thread as serial code, and threads are created as needed to process data in parallel.

to the sequential code [131]. The OpenMP program is executed as “fork-join” model (Figure 1-10 B). The starting thread is the master thread that executes sequentially until the start of the loop region. The loop region is divided to several small loops and a number of threads are generated to execute each small loop in parallel, which is the scenario of fork method. Each small loop is assigned by one thread and has unique start and stop point. The join scenario happens while the team of threads is destroyed after finishing massive computation and merging data into main memory space. Thus, the master thread takes the control of program again and executes sequential code until the next loop region. The advantage of using OpenMP is that it uses directives to guide compiler to parallelize job. Thus, the sequential code is easier to transform to parallel version since the details of the parallelism is controlled by the compiler. However, it is expensive to build a system with more than 100 cores, which restricts the scalability of performance.

In additional to CPU based parallel computing, recently, performance of parallel computing are greatly improved by using general purpose graphics processing unit (GP-GPU). Traditionally, GPU is designed to provide real time 3D effects for computational graphics, for example, in gaming and building 3D models. Now, GPU is wildly applied in scientific calculation because of the extraordinary computational capability provided by GPU. GPU-accelerated computing applies modern graphics card which equipped with CUDA cores (Compute Unified Device Architecture) or Stream processors to accelerate calculation. Parallel programming with GPU is a heterogeneous model, namely data is processed by CPU and GPU. The programming model of parallel programming with GPU is heterogeneous programming (Figure 1-11 A). One thread on CPU controls the flow of the program. Hundreds of thousands of threads on GPU are created with identical kernel function, these threads are grouped into blocks and a collection of thread blocks is a grid that runs on GPU for data processing (Figure 1-11 B). CPU and GPU maintain their own memory space separately; thus, developers write code to move data between host and device, the sequential codes and some light loops are performed on CPU, computationally intensive loops are executed by thousands of threads on GPU concurrently. The disadvantage for GPU-accelerated computing

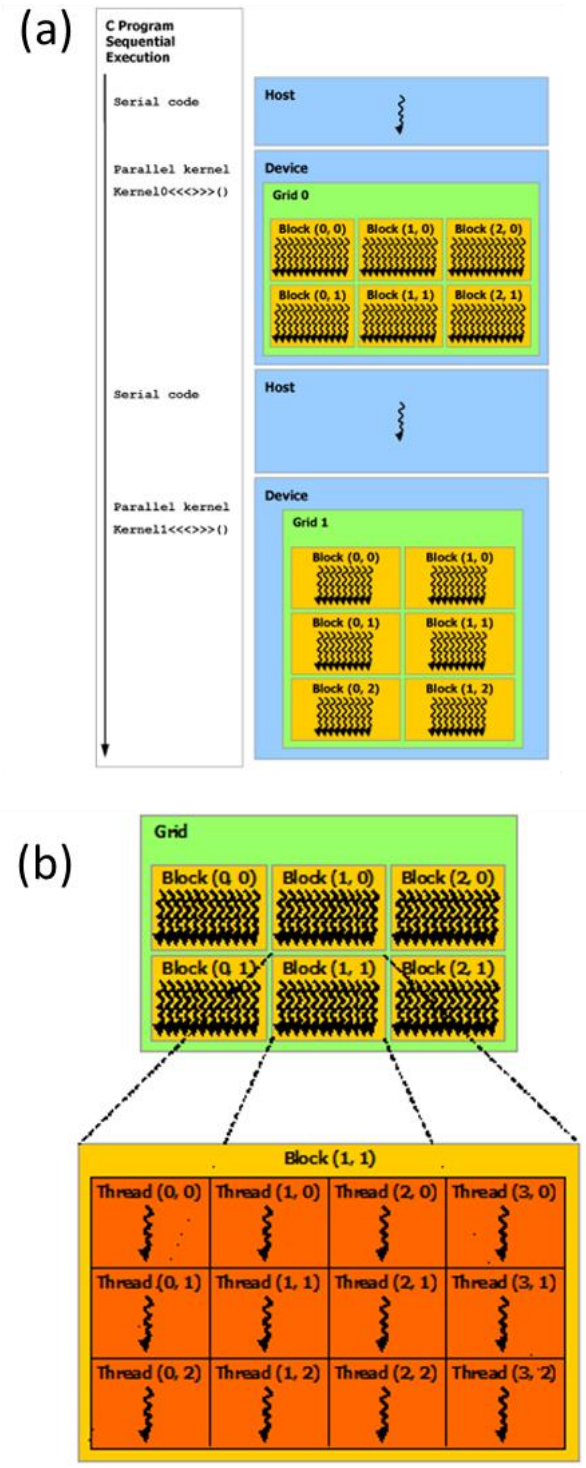


Figure 1-11 Parallel programming with Compute Unified Device Architecture.
(a) Heterogeneous Programming using CPU and CUDA cores. (b) Diagram of thread blocks in GPU. From the website: http://docs.nvidia.com/cuda/cuda-c-programming-guide/index.html#programming-model

is low portability. The code written with CUDA library is only workable for NVIDIA's GPU.

In addition, three programming models do not need to be considered as three different options for the implementation of parallel algorithm. A cluster of nodes that equipped with multi cores and GPUs can be constructed, and scalable MD code can be executed on cluster by mixing parallel computing models to maximize performance.

1.9 Rational for the studies

Inhibition of PR activity produces immature and noninfectious virions, which suggests HIV-1 PR is a good target for anti-retrovirus therapy [14, 15]. To date, nine PR inhibitors (PIs) have been approved for AIDS therapy. However, drug resistance arises after HIV exposure to drugs. In addition to the high mutation rate of HIV-1 PR, reported drug resistance related mutations occupy ~40% of its sequence [132]. Therefore, it is required to understand how PR acquires drug resistance through these mutations to develop new PIs with high effectiveness against drug resistant mutants. Here, **the overall goal is focused on understanding the structural changes in HIV-1 PR and its mutant complexes with inhibitors or substrates.** Our overall research hypothesis is: **Structural changes in mutants are responsible for altered properties of PR and resistance to drugs.**

The designs of amprenavir (APV) and saquinavir (SQV) are based on different concepts, maximizing hydrophilic interactions and maximizing hydrophobic interactions, respectively. This radical difference in design is also reflected by the different patterns of drug resistant mutations for APV and SQV [132]. Here we propose comparative analysis of structural changes in drug resistant mutants complexed with different inhibitors to further understand how mutants respond to different PIs. HIV-1 PR and mutants will be crystallized with APV in order to compare to SQV structures obtained from previous studies of our laboratory.

Aspartyl proteases play important roles in a number of biological processes, but the proteolytic mechanism is still not fully understood [45]. The presence of a metastable reaction intermediate is proposed based on isotope exchange experiment [39]. X-ray crystallographic analysis is a powerful tool to study catalytic mechanisms [133]. However, previous research done by X-ray crystallography provides limited information due to low resolution [134, 135]. Reaction intermediates observed in crystal structures of HIV protease and its mutants will be analyzed in order to better understand the proteolytic mechanism.

More than 400 HIV-1 PR structures have been deposited on Protein Data Bank (PDB) server. Interestingly, there are only few structures of clinically isolated mutants showing high levels of resistance to various PIs. Clinical isolate mutant, PR_{20}, harboring 20 substitutions significantly reduces binding affinity to clinical inhibitors [103]. The crystallographic study of structural changes of PR_{20} complexed with APV and SQV will aid in understanding the molecular basis of high drug resistance.

The significant increase in dissociation rate to various PIs indicates that these 20 mutations alter the molecular properties of PR_{20} [103]. However, the crystallographic analysis of inhibitor bound mutants might only show small structural changes, which might not fully reflect the structural properties caused by the mutations. MD simulations imitating the atomic motion in solution allow insight into structural motions. MD simulation will be applied to study changes in molecular motion of PR_{20} relative to wild-type enzyme.

Proteins are macromolecules, thus all atom MD simulations are computationally expensive. In practice, the size of the system and the number of iterations limit most MD simulations to the nanosecond scale. However, this time scale is shorter than many important protein processes, such as folding and flap motion, that are microsecond to millisecond events [136, 137]. The performance of molecular mechanics and dynamics program, AMMP, was improved by invoking amortized Fast Multipole Method, a multi-time step algorithm[111]. However, it is still a challenge to use AMMP in a large system or longtime simulation

because AMMP only uses a single core in the calculation. In addition to applying numerical algorithms to speedup simulation, modern parallel techniques, multi-core processors and graphics cards are different approaches to accelerate the performance. Therefore, the design of an efficient parallel algorithm is an effective way to improve the calculation capability for large scale and long timescale MD simulations.

2 AMPRENAVIR COMPLEXES WITH HIV-1 PROTEASE AND ITS DRUG RESISTANT MUTANTS ALTERING HYDROPHOBIC CLUSTERS

(Published: Shen CH, Wang YF, Kovalevsky AY, Harrison RW, Weber IT. 2010. Amprenavir complexes with HIV-1 protease and its drug-resistant mutants altering hydrophobic clusters. FEBS J. 2010 Sep;277(18):3699-714)

2.1 Abstract

The structural and kinetic effects of amprenavir (APV), a clinical HIV protease (PR) inhibitor, were analyzed with wild type enzyme and mutants with single substitutions of V32I, I50V, I54V, I54M, I84V and L90M that are common in drug resistance. Crystal structures of the APV complexes at resolutions of 1.02 to 1.85 Å reveal the structural changes due to the mutations. Substitution of the larger side chains in PR_{V32I}, PR_{I54M} and PR_{L90M} resulted in formation of new hydrophobic contacts with flap residues, residues 79 and 80, and Asp25, respectively. Mutation to smaller side chains eliminated hydrophobic interactions in the PR_{I50V} and PR_{I54V} structures. The PR_{I84V}-APV complex had lost hydrophobic contacts with APV, the PR_{V32I}-APV complex showed increased hydrophobic contacts within the hydrophobic cluster, and the PR_{I50V} complex had weaker polar and hydrophobic interactions with APV. The observed structural changes in PR_{I84V}-APV, PR_{V32I}-APV and PR_{I50V}-APV were related to their reduced inhibition by APV of 6-, 10- and 30-fold, respectively, relative to wild type PR. The APV complexes were compared with saquinavir (SQV) complexes. The PR dimers had distinct rearrangements of the flaps and 80's loops that adapt to the different P1' groups of the inhibitors while maintaining contacts within the hydrophobic cluster. These small changes in the loops and weak internal interactions produce the different patterns of resistant mutations for the two drugs.

2.2 Introduction

Currently, about 33 million people worldwide are estimated to be infected with human immunodeficiency virus (HIV) in the AIDS pandemic [138]. Despite the effectiveness of highly active anti-retroviral therapy (HAART) the virus cannot be fully eradicated [89]. Furthermore, development of vaccines has been extremely challenging [86]. HAART uses more than 20 different drugs, including inhibitors of the HIV-1 enzymes, reverse transcriptase (RT), protease (PR) and integrase, as well as inhibitors of cell entry and fusion. The major challenge limiting current therapy is the rapid evolution of drug resistance due to the high mutation rate caused by the absence of a proof-reading function in HIV RT [71].

HIV-1 PR is the enzyme responsible for the cleavage of the viral Gag and Gag-Pol polyproteins into mature, functional proteins. PR is a valuable drug target since inhibition of PR activity results in immature noninfectious virions [14, 15]. PR is a dimeric aspartic protease composed of residues 1-99 and 1'-99'. The conserved catalytic triplets, Asp25-Thr26-Gly27, from both subunits provide the key elements for formation of the enzyme active site. Inhibitors and substrates bind in the active site cavity between the catalytic residues and the flexible flaps comprising residues 45-55 and 45'-55' [52].

Amprenavir (APV) was the first HIV-1 PR inhibitor (PI) to include a sulfonamide group (Figure 2-1 A). Similar to other PIs, APV contains a hydroxyethylamine core that mimics the transition state of the enzyme. Unlike the first generation PIs, such as saquinavir (SQV), APV was designed to maximize hydrophilic interactions with PR [139]. The sulfonamide group increases the water solubility of APV (60 μg/mL) compared to SQV (36 μg/mL) [140]. The crystal structures of PR complexes with APV [139, 141] and SQV[95, 142] demonstrated the critical PR-PI interactions.

HIV-1 resistance to PIs arises mainly from accumulation of PR mutations. Conservative mutations of hydrophobic residues are common in PI resistance, including V32I, I50V, I54V/M, I84V and L90M that are the focus of this study [143]. The location of these mutations in the PR dimer structure is shown in Figure 2-1 B.

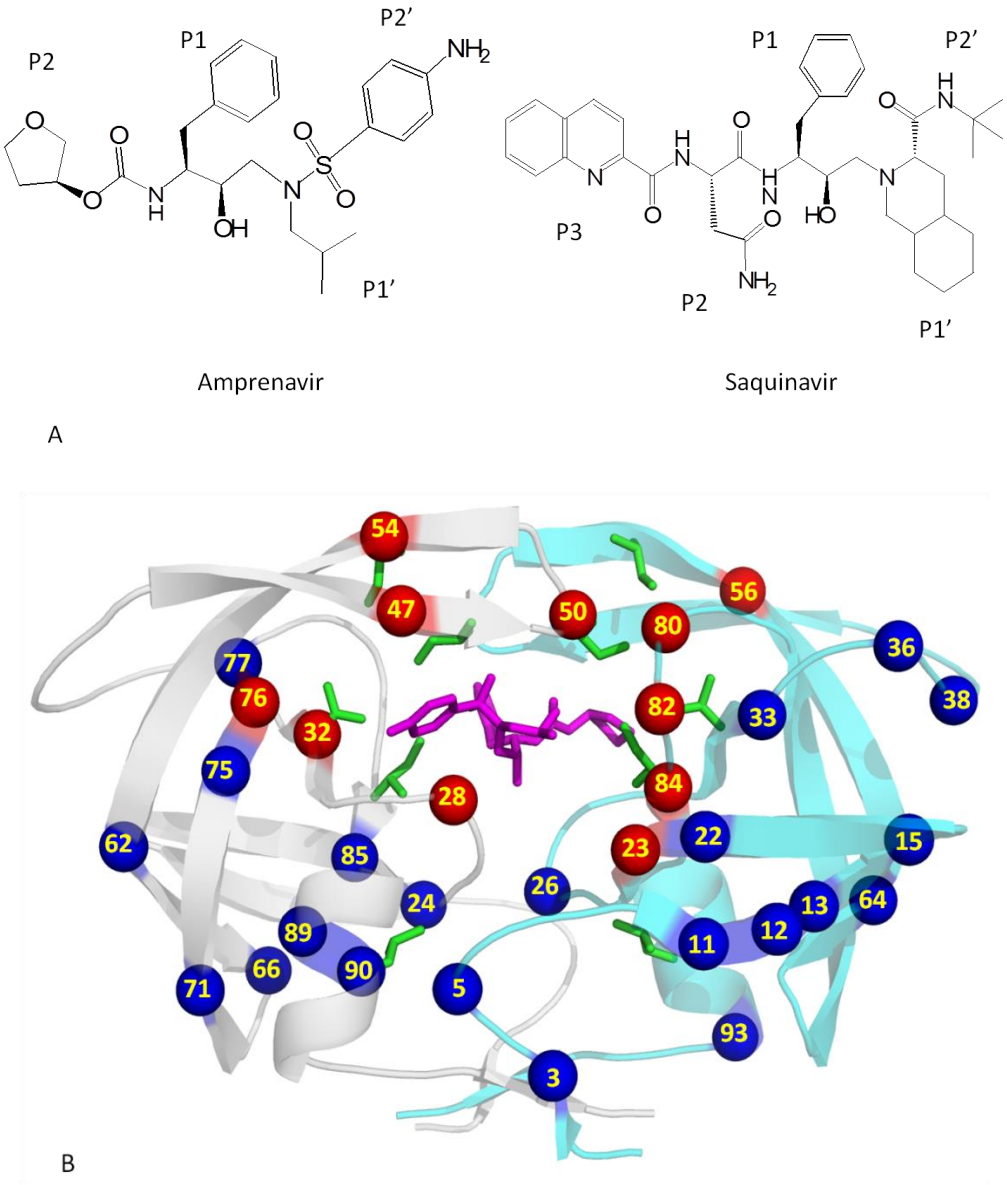


Figure 2-1
(A) The chemical structures of APV and SQV. (B) The structure of PR dimer with the sites of mutation Val32, Ile50, Ile54, Ile84 and Leu90 indicated by green sticks for side chain atoms in both subunits. Amino acids are labeled in one subunit only. APV is shown in magenta sticks. The amino acids in the inner hydrophobic cluster are indicated by numbered red spheres, and the amino acids in the outer hydrophobic cluster are shown as blue spheres.

Multi-drug-resistant mutation V32I, which alters a residue in the active site cavity, appears in about 20% of patients treated with APV[144] and is associated with high levels of drug resistance to lopinavir/ritonavir [143]. Ile50 and Ile54 are located in the flap region, which is important for catalysis and binding of substrates or inhibitors [96, 139]. Mutations of flap residues can alter the protein stability or binding of inhibitors [70, 96, 145, 146]. PR with mutation I50V shows 9-fold worse inhibition by DRV relative to wild type enzyme [147], and 50- and 20- fold decreased inhibition by IDV and SQV [70, 146]. Unlike Ile50, Ile54 does not directly interact with APV, but mutations of Ile54 are frequent in APV resistance and the I54M mutation causes 6-fold increased IC_{50} [148]. Mutation I54V appears in resistance to IDV, LPV, nelfinavir (NFV) and SQV [143]. I54V in combination with other mutations, especially V82A [149, 150], decreases the susceptibility to PI therapy [70]. I84V, which is located in the active site cavity, significantly reduces drug susceptibility to APV [151]. L90M is commonly found during PI treatment [144] and is resistant to all currently used PIs, with major effects on NFV and SQV [143].

Mutations of hydrophobic residues are found in more than half of drug resistant mutants [143, 152] and several of these mutations show altered PR stability [60, 146]. Hydrophobic interactions play an important role in protein stability. Aliphatic groups reportedly contribute about 70% of the hydrophobic interactions in proteins [153]. Removing a methyl group in the protein hydrophobic core affects protein folding and decreases the protein stability in mutant proteins [154]. In HIV PR, two clusters of methyl groups have been identified; one inner cluster surrounding the active site cavity and the second cluster in an outer hydrophobic core, as shown in Figure 2-1 B [152]. Drug resistant mutations V32I, I50V, I54V/M, and I84V belong to the inner cluster around the active site, while L90M is in the second outer cluster.

In order to establish a better understanding of the mechanism of resistance to APV, atomic and high resolution crystal structures have been determined of APV complexes with wild type PR and its mutants containing single substitutions of Val32, Ile50, Ile54, Ile84 and Leu90. HIV-1 PR mutations can have distinct effects on the binding of different inhibitors. Therefore, the structural effects of APV and SQV were

compared for wild type PR and mutants PR_{I50V}, PR_{I54M}, and PR_{I54V} complexes, using previously reported SQV complexes [70, 95]. Exploring the changes in PR due to binding of two different inhibitors will give insight into the mechanisms of resistance and help in the design of new inhibitors.

2.3 Experimental Procedures

2.3.1 Protein Expression and Purification.

The HIV-1 PR and mutants were constructed with five mutations Q7K, L33I, L63I, C67A, and C95A to prevent cysteine-thiol oxidation and diminish autoproteolysis [155]. The expression, purification and refolding methods are described in [97, 155].

2.3.2 Kinetic Assays

The fluorogenic substrate Abz-Thr-Ile-Nle-p-nitro-Phe-Gln-Arg-NH_2, where Abz is anthranilic acid and Nle is norleucine, (Bachem) with sequence derived from the p2/NC cleavage site of the Gag polyprotein was used in kinetic assays. Proteases were diluted in reaction buffer (100 mM MES, pH 5.6, 400mM sodium chloride, 1 mM EDTA, and 5% glycerol). 10 µL diluted enzyme was mixed with 98 L reaction buffer and 2µL Dimethyl Sulfoxide (DMSO) or APV (dissolved in DMSO) and incubated at 37 ºC for 5 min. Then, the reaction was initialized with addition of 90 µL substrate. The reaction was monitored over 5 min in the POLARstar OPTIMA microplate reader at wavelengths of 340 nm and 420 nm for excitation and emission. Data analysis used the program SigmaPlot 9.0 (SPSS Inc., Chicago, IL). Km and kcat values were obtained by standard data-fitting with the Michaelis-Menten equation. The Ki value was obtained from the IC50 values estimated from an inhibitor dose-response curve using the equation Ki = (IC50-[E]/2)/(1+[S]/Km), where [E] and [S] are the PR and substrate concentrations.

2.3.3 Crystallographic Analysis

Inhibitor APV (from AIDS reagent program) was dissolved in DMSO by vortex-mixing. The mixture was incubated on ice prior to centrifugation to remove any insoluble material. The inhibitor was mixed

with protein in molar ratios from 2:1 to 10:1. The crystallization trials employed the hanging drop method using equal volumes of enzyme-inhibitor and reservoir solution. PR_{WT}-APV was crystallized from 0.1M MES, pH 5.6, and 0.6-0.8 M sodium chloride. Crystals of PR_{V32I}-APV, PR_{I50V}-APV, and PR_{I54M}-APV were grown from 0.1M sodium acetate, pH 4.6-5.4, and 0.6-1.2 M sodium chloride. The other mutant complexes were crystallized using 0.1M sodium acetate, pH 3.8-4.2, and 0.1-0.2 M sodium iodide. Single crystals were mounted on fiber loops with 20 to 30 % (v/v) glycerol as cryoprotectant in the reservoir solution. X-ray diffraction data were collected at the SER-CAT beamline of the Advanced Photon Source, Argonne National Laboratories. Diffraction data were integrated, scaled, and merged using the HKL2000 package [156]. PR_{WT}-APV, PR_{V32I}-APV and PR_{I50V}-APV were solved by molecular replacement program Phaser [157] with structure 2QCI[158] as the starting model. The other complexes were solved by MOLREP [159], using 2F8G as the starting model [147]. The crystal structures were refined using SHELX-97 [160], except that the lower resolution structure of PR_{I84V}-APV was refined with REFMAC 5.2 [161]. The diffraction-data precision indicator (DPI) was used for determining the accuracy in the atomic positions [54]. The molecular graphics program COOT was used for map display and model building [162]. Structural figures were made by PyMol [163]. The structures were compared by superimposing their C_α atoms and using HIVAGENT [164] to calculate the distance between two atoms. The cut-off distances for different interactions were as described in [69].

2.4 Results

2.4.1 APV inhibition of PR and mutants

The kinetic parameters and inhibition constants of APV for PR_{WT} and the drug-resistant mutants PR_{V32I}, PR_{I50V}, PR_{I54M}, PR_{I54V}, PR_{I84V} and PR_{L90M} are shown in Table 2-1. The lowest catalytic efficiency (k_{cat}/K_m) values were seen for PR_{V32I} and PR_{I50V}, with 30% and 10% of the PR_{WT} value, respectively. PR_{L90M} showed a surprisingly high 11-fold increase in catalytic efficiency, whereas the other mutants were similar to PR_{WT}.

Table 2-1 Kinetic parameters for substrate hydrolysis and inhibition of amprenavir.

	K_m (μM)	k_{cat} (/min)	k_{cat}/K_m (μM/min)	Relative k_{cat}/K_m	K_i (nM)	Relative K_i
WT*	30±5	190±20	6.5±1.3	1.0	0.15±0.04	1
V32I	65±6	120±10	1.8±0.2	0.3	1.5±0.2	10
I50V*	109±8	68±5	0.6±0.03	0.1	4.5±0.6	30
I54M*	41±5	300±40	7.3±0.8	1.1	0.50±0.06	3
I54V*	43±6	130±20	3.1±0.9	0.5	0.41±0.05	3
I84V	73±6	320±30	4.4±0.5	0.7	0.9±0.2	6
L90M	13±2	950±120	73±13	11.2	0.16±0.01	1

* K_m and k_{cat} values from [70]

Error in k_{cat}/K_m is calculated as (A/B) +/- $(1/B^2)$(square root($B^2a^2 + A^2b^2$)), where A is k_{cat}, a is k_{cat} error, B is K_m, and b is error in K_m.

The k_{cat}/K_m values for PR_{L90M} appear to depend on the substrate, however, as only a modest three-fold increase relative to PR_{WT} was observed using a different substrate with the sequence derived from the MA/CA rather than the p2/NC cleavage site [147]. The six mutants and PR_{WT} were assayed for inhibition by APV (Table 2-1). APV showed subnanomolar inhibition with a K_i of 0.16 nMfor PR_{WT} and PR_{L90M}. PR_{I54M} and PR_{I54V} showed modestly increased (three-fold) relative K_i values. The largest increases in K_i of six-, 10- and 30-fold were observed for PR_{I84V}, PR_{V32I} and PR_{I50V}, respectively, relative to PR_{WT}. The substantially decreased inhibition of PR_{V32I} and PR_{I50V} suggested the loss of interactions with APV.

2.4.2 Crystal Structures of APV Complexes

The crystal structures of PR and drug-resistant mutants PR_{V32I}, $PR_{I50V,}$ PR_{I54M}, PR_{I54V}, PR_{I84V} and PR_{L90M} were determined in their complexes with APV at resolutions of 1.02–1.85 Å to investigate the structural changes. The crystallographic data are summarized in Table 2-2. All structures were determined in space group $P2_12_12$. The asymmetric unit contains one PR dimer of residues 1–99 and 1'–99' as well as APV. The lowest resolution structure of PR_{I84V} was refined to an R-factor of 0.20 with isotropic B-factors and solvent molecules. The other structures were refined at 1.50–1.02 Å resolution to R-factors of 0.12–0.16, including anisotropic B-factors, hydrogen atoms and solvent molecules. PR_{WT} had the highest resolution and lowest R-factor, concomitant with the lowest average B-factors for the protein and inhibitor atoms.

Because of the high resolution of the diffraction data, all structures except for PR_{I84V}–APV, were modeled with more than 150 water molecules, ions and other small molecules from the crystallization solutions, including many with partial occupancy (Table 2-2). The solvent molecules were identified by the shape and intensity of the electron density and the potential for interactions with other molecules. The nonwater-solvent molecules were: a single sodium ion, three chloride ions, two partial glycerol molecules in PR_{WT}–APV; one sodium ion, three chloride ions in PR_{V32I}–APV; three sodium ions, seven chloride ions, two partial acetate ions in PR_{I50V}–APV; one sodium ions, three chloride, two partial acetate ions in PR_{I54M}–

Table 2-2 Crystallographic Data Collection and Refinement Statistics

	PR	**PR_{V32I}**	**PR_{I50V}**	**PR_{I54M}**	**PR_{I54V}**	**PR_{I84V}***	**PR_{L90M}**
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit cell dimensions: (Å)							
a	58.11	57.77	57.95	58.12	57.50	59.51	57.94
b	85.97	86.13	86.01	85.91	86.00	86.88	85.91
c	46.42	46.28	46.21	46.10	45.95	45.44	46.10
Resolution range (Å)	50-1.02	50-1.20	50-1.29	50-1.16	50-1.50	50-1.85	50-1.35
Unique reflections	113,227	66,626	55,569	73,638	37,010	18,138	50,443
[b]R_{merge} (%) overall[a]	5.7 (38.2)	8.1 (44.2)	7.0 (40.2)	7.2 (35.7)	6.0 (46.2)	9.7 (34.5)	5.5 (46.2)
I/σ(I) overall[a]	15.3 (2.6)	11.3 (2.5)	15.2 (2.3)	20.1 (2.1)	16.8 (2.4)	15.8 (5.8)	17.9 (2.5)
Completeness (%) overall[a]	95.8 (65.0)	91.6 (62.7)	93.9 (70.4)	91.8 (58.9)	99.7 (99.2)	93.2 (76.6)	97.8 (97.3)
Data range for refinement (Å)	10-1.02	10-1.20	10-1.29	10-1.16	10-1.50	10-1.85	10-1.35
[c]R (%)	12.4	16.4	15.5	15.4	14.9	19.9	14.3
[d]R_{free} (%)	14.2	20.1	19.3	18.8	19.7	23.6	19.9
No. of solvent atoms (total occupancies)	292 (207.3)	151 (129.8)	177 (143.6)	242 (221.5)	152 (128.5)	84 (84)	211 (202.5)
RMS deviation from ideality							
Bonds (Å)	0.017	0.013	0.012	0.016	0.010	0.014	0.012
Angle distance (Å)	0.036	0.031	0.030	0.033	0.029	1.546[e]	0.030
Average B-factors (Å^2)							
Main-chain atoms	10.8	16.0	14.4	14.2	23.2	25.6	20.3
Side-chain atoms	14.8	21.4	20.7	20.8	28.8	28.3	23.6
Inhibitor	10.5	16.9	17.1	17.8	28.5	23.7	16.1
Solvent	20.8	25.6	24.3	36.1	47.0	49.5	39.9
Relative occupancy of APV	0.7/0.3	-	0.6/0.4	-	-	-	-
RMS deviation with PR	-	0.154	0.187	0.328	0.261	0.381	0.192
RMS deviation from SQV complex (Å)	0.87	-	0.285	0.356	0.317	0.356	-

*Refined using REFMAC 5.0, all others were refined with SHELX-97

[a]Values in parentheses are given for the highest resolution shell.

[b]$R_{merge} = \Sigma_{hkl}|I_{hkl} - \langle I_{hkl}\rangle|/\Sigma_{hkl}I_{hkl}$.

[c]$R = \Sigma|F_{obs}-F_{cal}|/\Sigma F_{obs}$.

[d]$R_{free} = \Sigma_{test}(|F_{obs}|-|F_{cal}|)^2/\Sigma_{test}|F_{obs}|^2$.

[e]degree

APV; 19 iodide ions in PR_{I54V}–APV; 33 iodide ions in PR_{I84V}–APV; and 19 iodide ions in PR_{L90M}–APV. However, many iodide ions had partial occupancy. They were identified by the high peaks in electron density maps, abnormal B-factors and contact distances of 3.4–3.8 Å to nitrogen atoms.

Alternative conformations were modeled for residues in all crystal structures. Alternative conformations were modeled for a total of 48, 13, 28, 11, 1, 8 residues in PR_{WT}–APV, PR_{V32I}–APV, PR_{I50V}–APV, PR_{I54M}–APV, PR_{I54V}–APV, PR_{I84V}–APV and PR_{L90M}–APV structures, respectively. APV was observed in two alternative orientations related by a rotation of 180° in the complexes with PR_{WT} and PR_{I50V} with relative occupancies of 0.7/0.3 and 0.6/0.4, respectively. The highest resolution structure, PR_{WT}–APV, showed the most alternative conformations for main chain and side chain residues. Several residues in the active site cavity showed two alternative conformations and were refined with the same relative occupancies as for APV. Surface residues with longer flexible side chains, such as Trp6, Arg8, Glu21, Glu34, Ser37, Lys45, Met46, Lys55, Arg57, Gln61 and Glu65, were refined with alternative conformations. Also, some internal hydrophobic residues, such as Ile64, Leu97, showed a second conformation for the side chain. At the other extreme, the lowest resolution structure of PR_{I84V}–APV showed only one residue, Leu97, with an alternative side chain conformation. In all the structures, the two catalytic Asp25 residues showed negative difference density around the carboxylate oxygens. This phenomenon might be caused by radiation damage in the carboxylate side chains, especially due to their location at the active site, as described in [165].

The accuracy in the atomic positions was evaluated by the diffraction data precision indicator (DPI), which is calculated in sfcheck from the resolution, R-factor, completeness and observed data [166]. The highest resolution structure of PR_{WT}–APV had the lowest DPI value of 0.02 Å, whereas the lowest resolution structure of PR_{I84V}–APV had the highest DPI value of 0.13 Å (Table 2-2). We estimate that significant differences in interatomic distances should be at least three-fold larger than the DPI value [52]. Hence, structural changes > 0.06 Å are significant for PR_{WT}–APV and > 0.4 Å for PR_{I84V}–APV at the two

extremes of resolution. The quality of the crystal structures is illustrated by the 2Fo–Fc electron density maps for the mutated residues (Figure 2-2). The mutated residues had single conformations, except for the side chains of Met54, Val54 and Met90 in one subunit that were refined with relative occupancies of 0.6/0.4, 0.7/0.3 and 0.5/0.5, respectively. Overall, the mutants and wild-type enzyme had very similar structures, probably because they shared the same crystallographic unit cell. The PR_{I54M}, PR_{I54V} and PR_{I84V} complexes had RMS deviations for the C_α atoms ranging from 0.26 to 0.38 Å compared with the wild-type structure. The structures of PR_{V32I}, PR_{I50V} and PR_{L90M} were more similar to PR_{WT} with RMS deviations of 0.15–0.19 Å for the main chain atoms.

2.4.3 PR interactions with APV and the influence of alternative conformations

The atomic resolution crystal structure of PR_{WT}–APV was refined with two differently populated conformations for the inhibitor and several residues forming the binding site with relative occupancies of 0.7/0.3 (Figure 2-3 A). Residues Arg8, Asp30, Val32, Lys45, Gly48, Ile50 and Pro81 showed alternative conformations in both subunits, and Asp25' had two alternative conformations for the side chain. Alternative conformations were also refined for the main chain of residues 24', 29', 30, 30', 31, 31', 48, 48', 79' and 80' around the inhibitor binding site. Moreover, the conserved water molecule between the flaps and the inhibitor showed two alternative positions. Similar, although less extensive, disorder in the inhibitor binding site has been observed in other atomic resolution crystal structures of this enzyme [69, 95]. In fact, the highest resolution structure reported to date (0.84 Å) of PR_{V32I} with DRV comprised two distinct populations for the entire dimer with inhibitor and one conformer contained an unusual second binding site for DRV [69]. Moreover, a similar asymmetric arrangement of Asp25/25' with a single conformation for Asp25 and two conformations for Asp25' was observed in the crystal structure of PR_{WT}–GRL0255A [167]. Only single conformations were apparent for APV in the mutant protease structures, with the exception of PR_{I50V}. However, the mutant structures were refined with lower resolution data where alternative conformations may be less clearly resolved than for the PR_{WT}–APV structure.

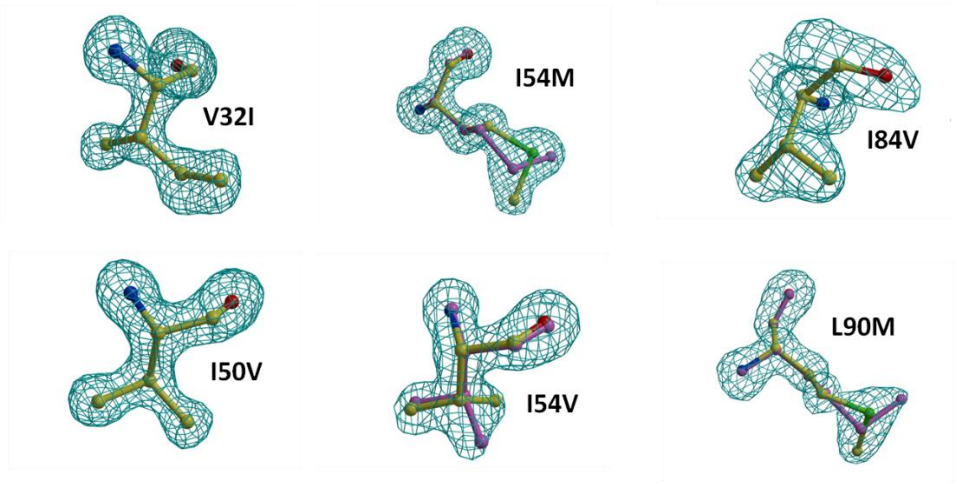


Figure 2-2 2Fo–Fc electron density maps for the mutated residue in:
PR_{V32I}, PR_{I50V}, PR_{I54M}, PR_{I54V}, PR_{I84V} and PR_{L90M} complexes with APV.

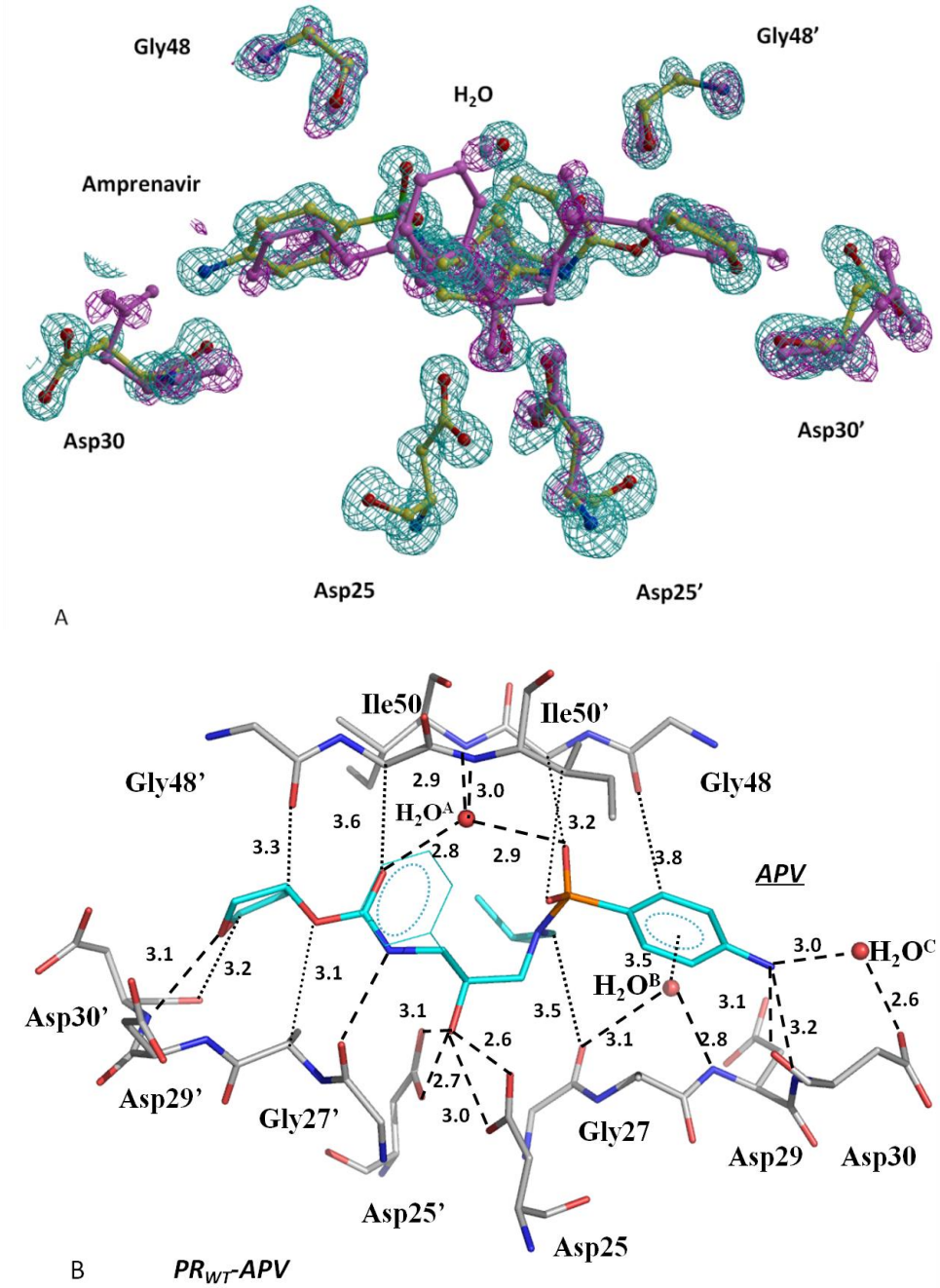


Figure 2-3 Inhibitor binding site in PR_{WT}–APV.

(A) APV and PR residues in the binding site with alternative conformations. Omit maps for major (green) and minor (magenta) conformations of APV, interacting PR residues Asp25, Gly48 and Asp30 from both subunits, and the conserved flap water are contoured at a level of 3.5 σ. (B) Hydrogen bond, C-H···O and H_2O···π interactions between PR (gray) and APV (cyan). Hydrogen bond interactions are indicated by dashed lines. C-H···O and H_2O···π interactions are indicated by dotted lines.

APV interactions with PR_{WT} were analyzed in terms of the hydrogen bond, C-H···O and H_2O···π interactions, as described for the PR_{V32I} complex with DRV [69]. The polar interactions of the major conformation of APV with PR_{WT} are illustrated in Figure 2-3 B. The central hydroxyl group of APV formed strong hydrogen bond interactions with the carboxylate oxygens of the catalytic residues Asp25 and Asp25'. APV formed four direct hydrogen bonds with the main chain amide of Asp30', the carbonyl oxygen of Gly27', and the amide and carbonyl oxygen of Asp30. Water molecules make important contributions to the binding site. The flap water molecule (H_2O^A in Figure 2-3 B), which is conserved in almost all PR–inhibitor complexes, formed a tetrahedral arrangement of hydrogen bonds connecting the amide nitrogen atoms of Ile50/50' in the flap region with the sulfonamide oxygen and the carbamate carbonyl oxygen of APV. The second conserved water (H_2O^B) bridged APV and the PR main chain by hydrogen bonds to the carbonyl of Gly27 and the amide of Asp29 and a H_2O···π interaction with the aniline group of APV. The interactions of H_2O^B are conserved in PR complexes with DRV and antiviral inhibitors based on the same chemical scaffold [158, 167]. The third water, H_2O^C, which is conserved in these APV complexes and in DRV complexes, mediated hydrogen bond interactions between the carboxylate of Asp30 and aniline NH_2 of APV. Also, several C-H···O interactions linked the PR main chain to APV: the carbonyl oxygens of Gly48' and Asp30' interacted with the tetrahydrafuran (THF) moiety, Gly27 carbonyl oxygen with the isopropyl group, Gly48 carbonyl oxygen with the aniline ring, the sulfonamide oxygens of APV to Gly49 and Ile50', APV carbonyl oxygen with Gly49' and APV oxygen to Gly27' (Figure 2-3). The C-H···O interactions formed by the PR amides and carbonyl oxygens mimic the conserved hydrogen bond interactions observed in PR complexes with peptide analogs [21, 168].

The minor APV conformation refined with 0.3 relative occupancy lay in the opposite orientation to the major conformation and interacted with the opposite subunits of the PR dimer. The minor conformation of APV retained almost identical hydrogen bond, C-H···O and H_2O···π interactions to the major conformation, with the following exceptions (Table 2-3). The hydrogen bond between the aniline

Table 2-3 Hydrogen bond interactions between HIV-1 protease and APV.

	PR [a](.7/.3)	PR_{V32I}	PR_{I50V} [a] (.6/.4)	PR_{I54M}	PR_{I54V}	PR_{I84V}	PR_{L90M}
APVN1-H_2O^C/	3.0/ -	2.9	2.7/ -	2.9	3.0	3.3	3.1
H^2O_C- Asp30OD2/ APVN1-Asp30'OD2	2.6/3.4	2.5	3.0/3.1	2.7	2.7	3.0	2.6
APVN1-Asp30N/ APVN1-Asp30'N	3.2/2.9	3.3	3.2/3.1	3.2	3.3	3.3	3.3
APVN1-Asp30O/ APVN1-Asp30'O	3.1/(3.6)	3.1	3.4/3.5	3.3	3.3	3.3	3.2
APVO26-Asp30'N/ APVO26-Asp30N	3.1/(3.7)	3.2	3.2/(3.7)	3.2	3.3	3.3	3.1
APVO26-Asp29'N/ APVO26-Asp29N	3.4/(3.6)	3.4	3.5/(3.7)	3.4	3.4	3.4	3.4
APVN20-Gly27'O/ APVN20-Gly27O	3.1/3.3	3.2	3.2/3.3	3.1	3.1	3.2	3.2
H_2O^B-Gly27O/ H_2O^B-Gly27'O	3.1/3.0	3.2	3.2/3.1	3.0	2.9	3.1	3.2
H_2O^B-Asp29N/ H_2O^B-Asp29'N	2.8/3.0	2.9	2.9/3.0	3.0	3.0	3.3	2.8
APVO18-Asp25'OD2/ APVO18-Asp25OD2	3.1/2.8	3.2	3.1/3.0	3.2	3.1	3.3	3.2
APVO18-Asp25'OD1/ APVO18-Asp25OD1	2.7/2.8	2.7	2.7/2.9	2.6	2.7	2.7	2.6
APVO18-Asp25OD2/ APVO18-Asp25'OD2	2.6/3.1	2.5	2.5/2.5	2.5	2.3	2.6	2.5
APVO18-Asp25OD1/ APVO18-Asp25'OD2	3.0/2.3	2.9	2.9/2.5	2.9	2.7	2.9	2.9
H_2O^A-Ile50N/ H_2O^A-Ile50'N	2.9/2.7	3.0	2.9/2.9	3.0	3.0	2.9	3.0
H_2O^A-Ile50'N/ H_2O^A-Ile50N	3.0/3.5	3.0	3.0/3.0	2.9	3.0	3.1	3.0
H_2O^A-APVO10	2.8/2.5	2.9	2.9/2.3	2.9	2.9	2.8	2.8
H_2O^A-APVO22	2.9/3.0	2.8	2.6/3.3	2.8	2.7	2.8	2.9
APVO26-Asp30'OD1/ APVO26-Asp30OD1	(3.8)/3.3	(3.7)	(3.6)/2.6	(3.9)	(3.8)	(3.9)	(3.9)

nitrogen of APV and the carbonyl oxygen of Asp30 was lost in the minor APV conformation (distance increased to 3.6 Å). The water-mediated interaction between the APV aniline nitrogen and the carboxylate group of Asp30 was replaced by a weak direct hydrogen bond (distance of 3.4 Å). The hydrogen bond of the THF oxygen with the amide of Asp30′ was lost in the minor APV conformation (distance increased to 3.7 Å). Instead, the THF oxygen of APV formed a new interaction with the carboxylate group in the minor conformation of Asp30. The water interaction with the amide of Ile50′ was weakened (distance of 3.5 Å). The C-H···O interaction between the carbonyl oxygen of APV and the C_α of Gly49 was lost in the minor conformation of APV. Some of these differences probably reflect the lower occupancy and greater positional error in the minor conformation. Variability in the interactions of Asp30/30′ due to flexibility of the side chains has been observed in other PR complexes [147]. Overall, the minor conformation of APV showed one less hydrogen bond, one less C-H···O interaction and weaker interactions than the major conformation had with PR.

2.4.4 Effects of Mutations on PR Structure and Interactions with APV

The structures of the mutants and PR_{WT} complexes with APV were compared in order to identify any significant changes. Overall, the polar interactions between APV and PR were well maintained in the mutant complexes. In these seven complexes the distances between nonhydrogen atoms were observed to be in the range of 2.3–3.3 Å for hydrogen bonds and 3.2–3.8 Å for C-H···O interactions (Table 2-3 and Table 2-4). The estimated error in atomic position is ~ 0.05 Å in structures at 1.0–1.2 Å resolution compared with the higher estimated errors of 0.10–0.15 Å in structures at 1.5–1.8 Å resolution [52], such as the complexes of PR_{I54V}–APV and PR_{I84V}–APV. Structural changes are detailed below for the mutant complexes with respect to the major conformation in PR_{WT}–APV. Generally, the changes in the mutants involved hydrophobic C-H···H-C contacts or C-H···O polar interactions, although shifts of main chain atoms were observed in some cases. The ideal distances between nonhydrogen atoms are considered to be 3.0–

Table 2-4 C-H···O interactions between HIV-1 protease and APV.

	PR [a](.7/.3)	PR_{V32I}	PR_{I50V} [a] (.6/.4)	PR_{I54M}	PR_{I54V}	PR_{I84V}	PR_{L90M}
APVC6-Gly48O/ APVC6-Gly48'O	3.8/3.1	3.7	3.5/3.0	3.5	3.6	3.5	3.6
APVO9-Ile50'C APVO9-Ile50C	3.4/4.0	3.5	3.4/3.9	3.6	3.5	3.7	3.4
APVO10-Gly49CA/ APVO10-Gly49'CA	3.5/2.8	3.5	3.2/2.8	3.3	3.2	3.2	3.4
APVC12-Gly27O/ APVC12-Gly27'O	3.5/3.1	3.5	3.4/3.5	3.5	3.4	3.6	3.5
APVO22-Gly49'Ca/ APVO22-Gly49Ca	3.6/(4.6)	3.8	3.5/(4.1)	3.7	3.7	3.7	3.7
APVO23-Ala28'Ca/ APVO23-Ala28 Ca	3.5/3.7	3.6	3.6/3.6	3.6	3.6	3.7	3.6
APVC24-Gly48'O/ APVC24-Gly48O	3.3/4.0	3.5	3.6/4.0	3.5	3.5	3.7	3.4
APVC27-Asp30'O/ APVC27-Asp30O	3.2/3.5	3.2	3.2/3.4	3.3	3.3	3.2	3.3

Distances between hydrogen donor and acceptor atoms are shown in Å. Atoms separated by more than 4.0 Å (in parentheses) are not considered to form C-H···O interactions.
[a] PR and PR_{I50V} complexes had two alternate conformations of APV.

3.7 Å for C-H···O interactions and 3.8–4.2 Å for van der Waals interactions, as described in [69]. The structural differences are described separately for each mutant.

Val32 is an important part of the S2 pocket in the active site cavity and forms van der Waals interactions with inhibitors. In the PR_{WT}–APV structure, Val32 forms hydrophobic contacts with Ile47, Ile56, Thr80 and Ile84, whereas Val32' interacts with Thr80' and Ile84'. Mutation of Val to Ile, which adds one methyl group, can reduce the volume of the active site cavity and alter the hydrophobic interactions in the cluster. The mutant with Ile32 did not show significant alterations in the main chain conformation or the interactions with APV. However, the $C_{\delta 1}$ methyl of the Ile side chain provided new van der Waals contacts with other hydrophobic side chains. Ile32 formed new hydrophobic contacts with the side chains of Val56, Leu76 and the main chain atoms of residues 77–78, and Ile32' showed new interactions with the side chains of Ile47', Ile50 and Val56' in the flaps (Figure 2-4 A). The flaps can exist in an open conformation in the absence of inhibitor and a closed conformation when inhibitor is bound. The interactions of residue 32 differ in the closed and open conformations; Val32 has no hydrophobic contacts with flap residues in the PR–APV structure, whereas Val32 forms hydrophobic contacts with Ile47 in the open conformation structure [169]. The flexibility of the flaps is probably altered in PR_{V32I} by the new hydrophobic contacts of Ile32/32', which is expected to contribute to the three-fold reduced catalytic activity and the 10-fold decreased APV inhibition of the PR_{V32I} mutant relative to wild-type enzyme (Table 2-1).

Ile50 is located at the tip of the flap on each PR monomer, where its side chain forms hydrophobic interactions with inhibitors. In the wild-type enzyme, Ile50/Ile50' interacts with Pro81'/Pro81 and Thr80'/Thr80 in the 80's loop, as well as Ile47'/47 and Ile54'/54 in the flaps. The $C_{\delta 1}$ methyl of the Ile50 side chain forms C-H···O interactions with the hydroxyl oxygen of Thr80' and carbonyl oxygen of Pro79', and the $C_{\delta 1}$ of Ile50' interacts with the hydroxyl of Thr80. Mutation from Ile50 to Val shortens the side chain by a methyl group, which eliminates the C-H···O interaction with the hydroxyl oxygen of Thr80' and van der Waals contact with Ile54' (Figure 2-4B). In the other subunit, mutation to Val50' eliminates the C-

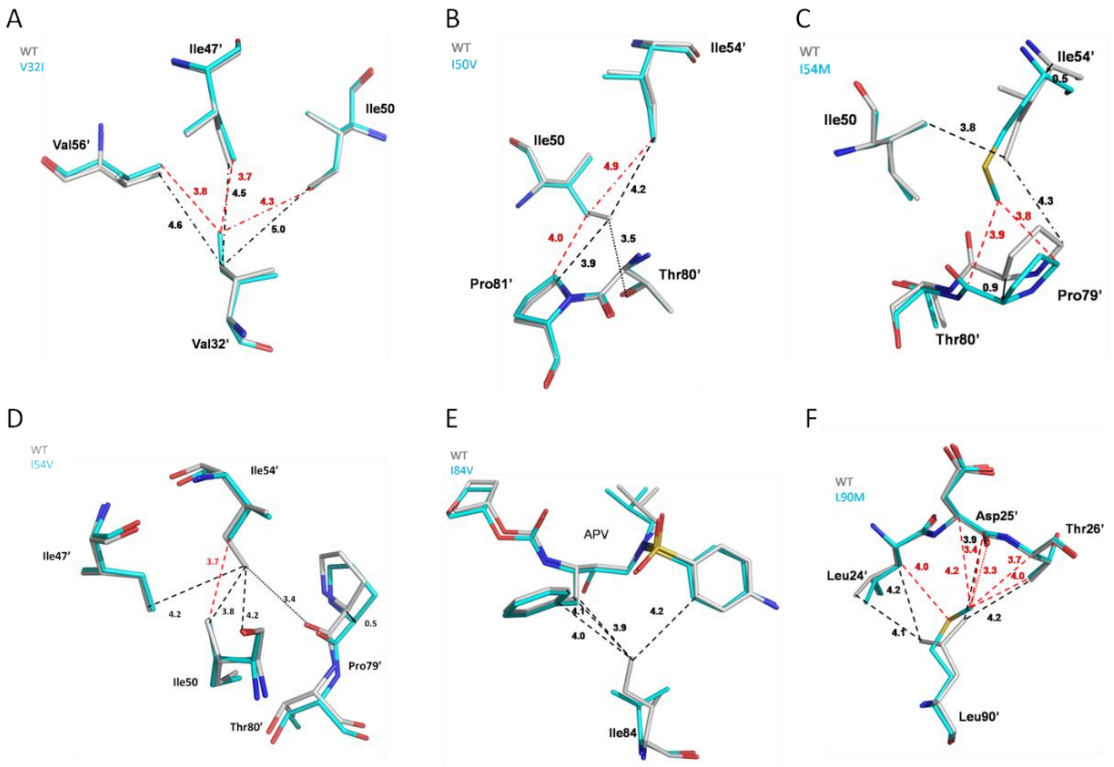


Figure 2-4 The interactions of mutated residues in

(A) PR_{V32I}–APV, (B) PR_{I50V}–APV, (C) PR_{I54M}–APV, (D) PR_{I54V}–APV, (E) PR_{I84V}–APV and (F) PR_{L90M}–APV. The gray color corresponds to PR_{WT}–APV and the cyan color indicates the mutant complex. Dashed lines indicate van der Waals interactions and dotted lines show C-H···O interactions. Interatomic distances are shown in Å with black lines indicating PR_{WT} and red lines indicating the mutant. Interatomic distances of > 4.3 Å are shown in dash-dot lines to indicate the absence of favorable interaction.

H···O interaction with the hydroxyl of Thr80 and a hydrophobic contact with Pro81. The APV in PR_{I50V} complex had two alternative conformations with 0.6/0.4 relative occupancy. The APV showed a more elongated hydrogen bond than seen in the PR_{WT} complex between the aniline group and the carbonyl oxygen of Asp30, with an interatomic distance of 3.4 Å for the major conformation and 3.5 Å for the minor APV conformation. Val50' also lost hydrophobic interactions with the THF group of APV. The minor conformation of APV showed similar changes in interactions with Asp30/30' as described for the minor APV conformation in the PR_{WT} complex. Overall, the observed structural changes in PR_{I50V}–APV were the loss of two C-H···O interactions and van der Waals contacts, the elongated hydrogen bond and reduced hydrophobic contacts with APV. PR_{I50V} showed a large decrease in sensitivity to APV shown by the 30-fold drop in the relative inhibition coupled with 10-fold decreased catalytic efficiency, which suggests the importance of Ile50. Loss of the C-H···O interaction of Val50 with Thr80' has not been previously described. Thr80 is a conserved residue in the PR sequences and its hydroxyl forms a hydrogen bond with the carbonyl oxygen of Val82, which contributes hydrophobic interactions with the inhibitors. Moreover, the hydroxyl group of Thr80 was shown to be important for PR activity using site-directed mutagenesis where only mutation to Ser retaining the hydroxyl group, and not to Val or Asn, maintained enzymatic activity [62]. These lines of evidence, taken together, strengthen the suggestion that loss of the C-H···O interaction of residue 50 with the hydroxyl of Thr80, as well as loss of hydrophobic contacts with inhibitor, are important for the decreased catalytic activity and APV inhibition of PR_{I50V}.

Ile54 is another flap residue that forms hydrophobic interactions with Ile50' and residues 79–80, although it has no direct contact with inhibitor. Mutation I54M introduces a longer side chain, and the nearby main chain atoms have shifted relative to their positions in PR_{WT} (Figure 2-4 C). Compared with PR_{WT}, the C_α of Met54 moved by 0.7 Å, and the longer Met side chain pushed residues 79, 80 and 81 away by 0.7–1.4 Å. In the other subunit, the C_α of Met54' moved by 0.5 Å towards Pro79', and there was a correlated motion of Pro79' of 0.9 Å relative to its position in PR–APV. The longer Met54/54' side chains

formed more hydrophobic contacts with Pro79/79′ and Thr80/80′ in PR_{I54M} relative to those of PR_{WT}. Overall, the Ile54 to Met mutation improved contacts within the hydrophobic cluster, although the interatomic distances to residues 79–80/79′–80′ were increased. Similar structural changes were observed in the PR_{I54M} complexes with DRV and SQV [70]. Despite these correlated changes between the main chain atoms of the flaps and 80′ loops, this mutant was similar to PR_{WT} in catalytic efficiency and had only three-fold reduced inhibition by APV.

In contrast to PR_{I54M}, mutation I54V substitutes the shorter Val in PR_{I54V}. In PR_{WT}, the $C_{\delta1}$ of Ile54 interacted with Ile50′, Val56, Pro79 and Thr80, whereas the $C_{\delta1}$of Ile54′ showed van der Waals interactions with Ile47′, Ile50′, Val56′ and Pro79 and one C-H···O interaction with the carbonyl oxygen of Pro79′. The shorter Val side chain in the mutant resulted in loss of several van der Waals contacts with the adjacent residues, thus decreasing the stability of the hydrophobic cluster formed by flap residues 47, 54 and 50′ (Figure 2-4 D). No C-H···O interaction was possible with Pro79′, which was associated with a shift of ∼ 0.5 Å in Pro79′ increasing the separation of the flap and 80′s loop. The mutation I54V decreased the hydrophobic interactions within the flaps and with Pro79. However, PR_{I54V} showed similar K_i value and only a three-fold reduced activity relative to the wild-type enzyme.

Ile84 forms part of the S1/S1′ subsites of PR, and mutation to Val84 removes a methylene moiety, which can reduce interactions with substrates and inhibitors. In PR_{WT}, van der Waals contacts were found between $C_{\delta1}$ of Ile84 and the benzyl and aniline moieties of APV and from $C_{\delta1}$ of Ile84′ to the isopropyl group of APV. These interactions were lost in the PR_{I84V} mutant structure as the interatomic distances increased to > 4.3 Å (Figure 2-4 E). The loss of hydrophobic contacts with APV is consistent with the modest change of six-fold in K_i value for PR_{I84V}.

Leu90 is located in the short alpha helix outside of the active site cavity, although it extends close to the main chain of the catalytic Asp25. Mutation of Leu90 to Met substituted a longer side chain and introduced new van der Waals contacts with residues Asp25-Thr26. Moreover, the long Met90/90′ side

chains formed close C-H···O interactions with the carbonyl oxygen of the catalytic Asp25 and Asp25' (Figure 2-4 F). The alternative conformations of the Met90 side chain were arranged as described previously [60]. The new interactions of Met90/90' with the catalytic aspartates and adjacent residues are presumed to play an important role in the observed 11-fold increase in catalytic activity, as described previously [60, 147]. The increased catalytic efficiency of the PR_{L90M} mutant is mainly due to an almost five-fold higher k_{cat}. On the other hand, the K_m of the mutant is only about half that of the wild-type enzyme. Therefore, the new interactions of Met90 with the catalytic residues that are absent in the wild-type structure may minimally affect the binding of substrate, but at the same time dramatically lower the activation barrier for substrate hydrolysis, leading to substantial improvement of the PR_{L90M} catalytic activity. No change, however, was detected in the APV inhibition of PR_{L90M}.

2.4.5 Comparison of the mutant complexes with APV and SQV.

The structures of PR complexes with APV or SQV were analyzed in order to understand their distinct drug resistance profiles. PR_{WT}–APV was compared with PR_{WT}–SQV (2NMW) solved at 1.16 Å resolution in a different unit cell and space group $P2_12_12_1$ [95]. The mutant APV complexes reported here were compared with the published SQV complexes of PR_{I50V}–SQV (3CYX), PR_{I54M}–SQV (3D1X), PR_{I54V}–SQV (3D1Y) and PR_{I84V}–SQV (2NNK) refined at resolutions of 1.05–1.25 Å in the isomorphous unit cell and identical space group $P2_12_12$ as for all the APV complexes [70, 95]. No SQV complexes have been reported for mutants PR_{V32I} and PR_{L90M}. A lower resolution (2.6 Å) crystal structure has been reported for the SQV complex with the double mutant $PR_{G48V/L90M}$ [170], in which Met90 showed interactions similar to those seen in the structure of PR_{L90M}–APV. To analyze how the PR conformation alters to fit the different inhibitors, all structures were superimposed on the PR_{WT}–APV structure. The superposition was tested for both possible arrangements of the two subunits in the asymmetric dimer of PR, i.e. superimposing residues 1–99 and 1'–99' with 1–99 and 1'–99', as well as with the opposite subunit arrangement of 1'–99' and 1–99. The arrangement with the lowest RMS deviation was used in further comparison.

Interestingly, the major conformation of SQV has the opposite orientation to that of APV for the superimposed dimers with the lowest RMS values. PR_{WT}–SQV had the highest RMS deviation value of 0.87 Å on C_α atoms due to the different space groups, whereas the RMS deviations for the mutant complexes with SQV were lower, ranging from 0.29 to 0.36 Å, as usual for two structures in the same space group (Table 2-2).

The corresponding pairs of wild-type and mutant complexes with the two inhibitors were compared (Figure 2-5 A). The structures of PR_{WT}–SQV and PR_{WT}–APV showed larger RMS deviations of > 1.0 Å for residues in surface loops, probably arising from altered lattice contacts due to the different space groups, as reported previously [60]. Moreover, the two subunits in the dimer showed asymmetric deviations due to nonidentical lattice contacts as well as the presence of different asymmetric inhibitors. Changes in residues 52'–56', 79'–81' in the active site cavity are assumed to reflect variation in the interactions with the two inhibitors, whereas the lower deviations of catalytic triplet residues 25–27/25'–27' reflect their important function. The pairs of mutant complexes were determined in isomorphous unit cells with less overall variation so that changes are more likely to arise from different interactions with APV and SQV. PR_{I50V} had the fewest RMS differences between the two inhibitor complexes with a peak of 1.3 Å for Phe53'. PR_{I54V} and PR_{I84V} showed the largest change for Pro81' of 1.9 and 1.6 Å, respectively, whereas PR_{I54M} showed the maximum RMS deviation of 1.2 Å at residue 54'. Three regions were analyzed in more detail due to their flexibility and proximity to inhibitors and mutations: the flaps, the 80's loops and the hydrophobic clusters formed by residues Ile47, Ile54, Thr80, Ile84 and Ile50' from the opposite subunit.

The conformation of the flaps segregated into two categories corresponding to the APV complexes and the SQV complexes (Figure 2-5 B). The coordinated changes in the flaps were most obvious for residues 50–51 and 50'–51' at the tips of the flaps. The flap residues 50 and 51 showed differences in C_α position of 0.5–0.9 Å between the complexes with APV or SQV. Differences of 0.6–0.8 Å at Gly51' and 0.1–0.4 Å at

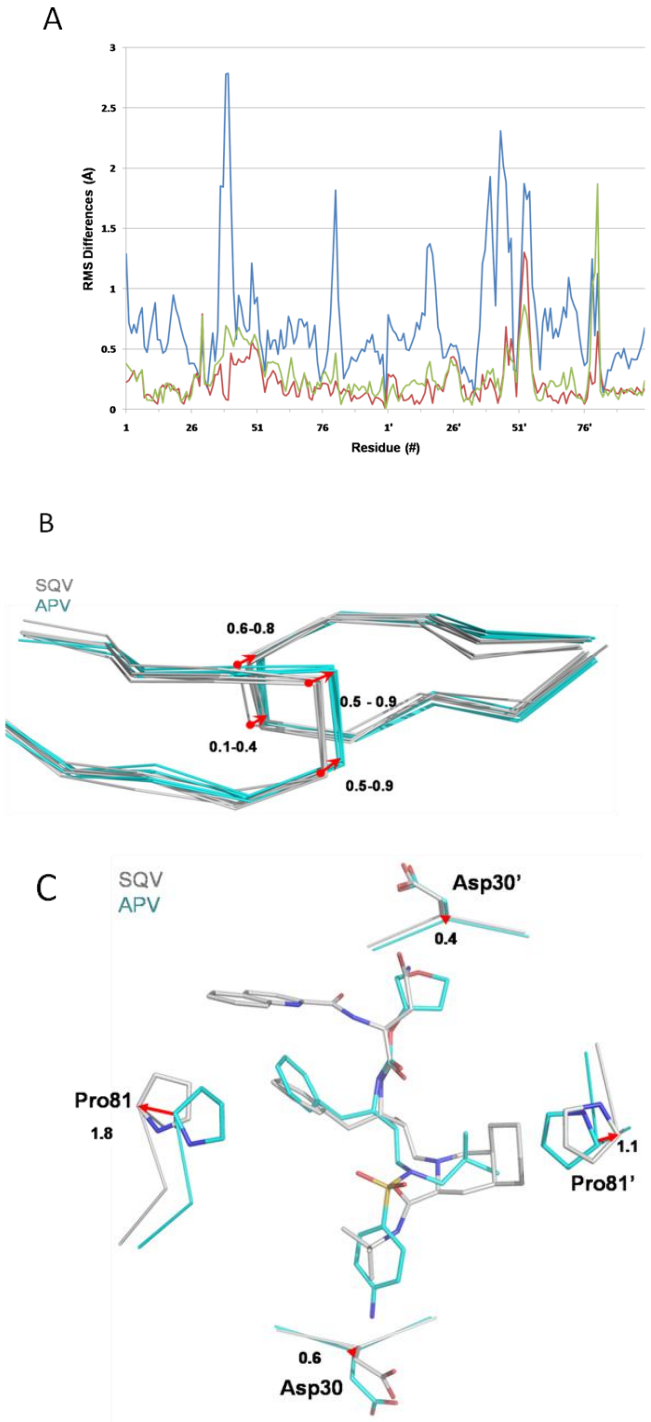


Figure 2-5 Structural differences between APV and SQV complexes.
(A) The RMS difference (Å) per residue is plotted for C_α atoms of SQV complexes compared with the corresponding APV complexes: PR_{WT} (blue line), PR_{I50V}(red line) and PR_{I54V} (green line). (B) Comparison of the flap regions in the structures. The complexes with APV are in cyan, and the complexes with SQV are in gray. The arrow indicates the shifts between C_α atoms at the residues 50 and 51 in the PR complexes with the two inhibitors. (C) The width across the S1–S1' subsites increases in PR_{WT}–SQV relative to PR_{WT}–APV. Similar changes were seen for the mutant complexes, except for PR_{I50V}.

Ile/Val50' were seen in the flap from the other subunit. The different flap conformations are probably related to the larger chemical groups at P2 and P1' in SQV compared with those in APV (Figure 2-1). Changes in the 80's loops, which have been described as intrinsically flexible [60, 95, 171] and function in substrate recognition [172, 173], were assessed using the distance between the C_α atoms of Pro81 and Pro81' to reflect alterations in the S1/S1' subsites. Pro81 and Pro81' were separated by 17.6–19.4 Å in the APV complexes, whereas these residues were 0.7–2.5 Å further apart in the SQV complexes (separations of 18.5–20.5 Å). The comparison of wild-type complexes is shown in Figure 2-5 C. PR_{I50V} complexes had the smallest distance between Pro81 and Pro81', whereas the greatest separation was observed for PR_{I54M} complexes, probably due to close contacts of the longer Met54/54' side chains with the 80's loops in both inhibitor complexes. The distance between Pro81 and Pro81' in the other structures was $\sim$ 2.5 Å longer in the SQV complexes compared with the APV complexes, which corresponds to the increment of the width across the S1/S1' pockets caused by binding of the big decahydroisoquinoline P1' group of SQV instead of the smaller P1' group in APV (Figure 2-5 C)[95].

The more rigid region of Asp30/30' showed smaller shifts of $\sim$ 0.5 Å for the C_α atoms. The different P2 and P2' groups, THF in APV or Asn in SQV at P2 and aniline in APV or t-butyl group in SQV at P2', are accommodated by these shifts, as shown in Figure 2-5 C.

The side chain interactions within the inner hydrophobic cluster were analyzed in PR_{WT}, PR_{I50V}, PR_{I54M} and PR_{I54V} complexes with SQV and APV. Overall, the main chains of the flaps were shifted relative to the 80' loops in APV complexes compared with SQV complexes. The hydrophobic cluster around the active site was formed by Ile47, Ile54, Thr80 and Ile84 from one subunit and Ile50' from the opposite subunit, as well as Val32 in a more rigid region in PR_{WT}. Differences in the side chain interactions are described. In PR_{WT}–APV the $C_{\gamma1}$ of Ile50 made good van der Waals contacts with Ile54', but the side chains were further apart in PR_{WT}–SQV (Figure 2-6 A). One C-H···O interaction between $C_{\delta1}$ of Ile50/Ile50' and the hydroxyl of Thr80'/Thr80 was conserved in PR_{WT}–APV, PR_{WT}–SQV and PR_{I50V}–SQV. The structure of PR_{I50V}–APV,

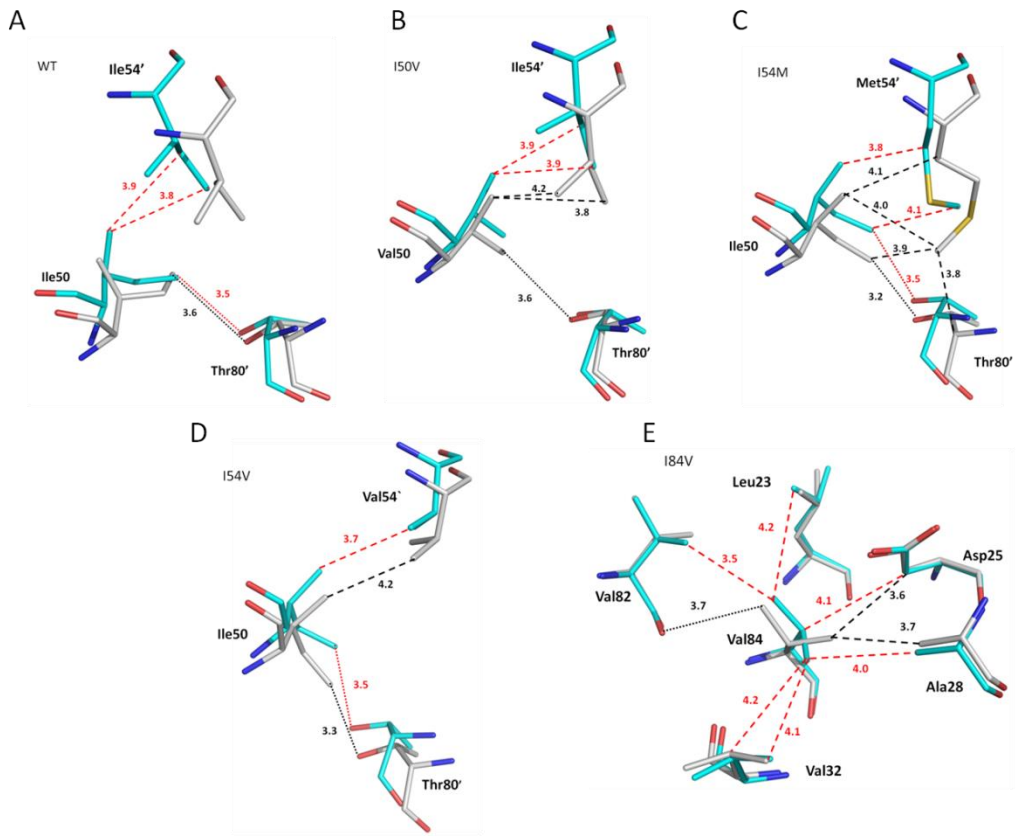


Figure 2-6 Interactions of Ile50, Ile54' and Thr80'.

(A) PR_{WT}–APV compared with PR_{WT}–SQV. (B) PR_{I50V}–APV compared with PR_{I50V}–SQV. (C) PR_{I54M}–APV compared with PR_{I54M}–SQV. (D) PR_{I54V}–APV compared with PR_{I54V}–SQV. (E) PR_{I84V}–APV compared with PR_{I84V}–SQV. Dashed lines indicate van der Waals contacts with interatomic distances in Å. Dotted lines indicate C-H···O interactions. Black lines indicate interactions in SQV complexes. Red lines indicate interactions in APV complexes.

however, lost this C-H···O interaction of residue 50 with the hydroxyl of Thr80', which reduced the interactions between the flap and the 80's loop (Figure 2-6 B). Mutation of Thr80 to Val80, which eliminates the C-H···O interaction with Ile50 $C_{\delta 1}$, significantly reduces the catalytic activity and binding affinity of SQV [62]. Apart from the different interactions between PR and inhibitors, the loss of the C-H···O contact between residue 50 and Thr80 in PR_{I50V}–APV and not in PR_{I50V}–SQV appears to correlate with the observed drug resistance. Mutation I50V is associated strongly with resistance to APV, but not to SQV.

In the PR_{I54M} and PR_{I54V} complexes with both inhibitors, however, similar interactions were observed among the side chains of residues 50, 54' and 80' (Figure 2-6 C,D). Val84 in PR_{I84V}–SQV had fewer hydrophobic contacts than in PR_{I84V}–APV, but it had one C-H···O interaction between Cγ2 of Val84 and the carbonyl oxygen of Val82 (Figure 2-6 E). In contrast to the significant conformational shifts in main chain atoms, the side chains in this hydrophobic cluster generally have rearranged to maintain internal hydrophobic contacts in the complexes with both inhibitors.

2.5 Discussion

We have described an atomic resolution crystal structure of PR_{WT} with APV, and analyzed structural changes in the APV complexes with mutants PR_{V32I}, PR_{I50V}, PR_{I54M}, PR_{154V}, PR_{I84V} and PR_{L90M}. The mutated residues contribute to an inner hydrophobic cluster around the substrate binding cavity, with the exception of Leu90, which is located near the backbone of the catalytic Asp25 in the outer hydrophobic cluster (Figure 2-1 B). Studies of the patterns of resistance mutations and molecular dynamics simulations have suggested the importance of these hydrophobic mutations in drug resistance [144, 174]. Our analysis showed that interactions within the inner hydrophobic cluster containing residues 32, 47, 54 and 50 were frequently altered relative to those in the wild-type enzyme. Mutations to larger side chains in PR_{V32I}, PR_{I54M} and PR_{L90M} resulted in the formation of new hydrophobic contacts with flap residues, residues 79 and 80, and Asp25, respectively. Mutation to smaller side chains caused loss of internal hydrophobic interactions in the PR_{I50V} and PR_{I54V} structures. PR_{I84V}, PR_{V32I} and PR_{I50V} showed reduced APV inhibition by

6-, 10- and 30-fold, respectively, relative to PR_{WT}, which is consistent with the observed structural changes. The PR_{I84V}–APV complex had lost hydrophobic contacts with APV, the PR_{V32I}–APV complex showed increased hydrophobic contacts with the flaps that probably restricted the flexibility needed for catalysis, and the PR_{I50V} complex had weaker interactions with APV. Ile54 had no direct contact with APV, which is consistent with the relatively small changes in inhibition, catalytic properties, protease stability and structure shown by the I54 mutants. No compensating changes were identified elsewhere in the hydrophobic core. In PR_{L90M}, the longer side chain of Met90/90' lies close to the main chain of the catalytic aspartates forming new van der Waals contacts and a C-H···O interaction. These new contacts with Asp25/25' near the dimer interface correlate with the reduced stability and altered catalytic parameters of this mutant, as described in studies with indinavir and DRV [60, 147]. No evidence was found, however, that PR_{L90M}–APV had substantially altered the volume of the S1/S1' substrate binding pockets, unlike the $PR_{G48V/L90M}$–SQV structure, which showed reduced volume for the S1/S1' subsites relative to the wild-type complex [170]. Also, the structure of PR_{G48V}–DRV showed reduced volume of the active site cavity relative to the wild-type complex, consistent with a major effect of the G48V rather than the L90M mutation in reducing the S1/S1' volume in $PR_{G48V/L90M}$–SQV [70]. Reduced interactions with inhibitors and conformational adjustments of the flaps and 80's loops were observed in our previous studies of these mutants with other inhibitors [60, 68-70, 146, 147]. These structural changes are expected to contribute to drug resistance. However, crystal structures show only a static picture of the effect of mutations, whereas changes in protein dynamic and thermodynamic properties also probably contribute to resistance. Future studies using molecular dynamics simulations and calorimetric analysis will help to address the changes in these other properties.

Individual mutations have distinct effects on the protease structure and activity. However, drug-resistant clinical isolates generally accumulate multiple mutations. Previous studies of double and single mutants suggested that the structural changes due to a single mutation were retained in the double

mutant, although other properties did not combine in predictable ways [100]. A variety of structural and biochemical mechanisms have been reported for different combinations of mutations, including: compensating structural changes, altered interactions with inhibitor and substantial opening of the active site cavity [105, 175, 176]. Clearly, further studies are needed to assess the effects of the multiple protease mutations that are observed clinically.

Comparison of the APV complexes with the corresponding SQV complexes for PR_{WT}, PR_{I50V}, PR_{I54V} and PR_{I84V} showed changes in the conformation of the flexible flaps and 80's loops. Despite the conformational changes in main chain atoms, the internal side chains generally rearranged to preserve the internal hydrophobic contacts in the complexes with both inhibitors. The structural changes can be correlated with the type of inhibitor. In particular, the separation of Pro81 and Pro81' was significantly smaller in the complexes with APV compared with the equivalent SQV complexes, which reflects the smaller size of the P1' group in APV relative to that of SQV. Also, the flexible side chains of Asp30 and Asp30' accommodate diverse functional groups at P2 and P2' of SQV and APV at the surface of the PR active site cavity. The functional group can be critical for a tight binding inhibitor. For example, DRV was derived from APV by changing THF to bis-THF, which introduces more hydrogen bonds with PR main chain atoms and dramatically increases the potency on drug-resistant HIV [177].

APV was designed to include several hydrophilic interactions with PR, and SQV optimizes hydrophobic interactions with PR[95]. Both of the inhibitors have been classified as peptidomimetic inhibitors, which mimic PR–substrate interactions and block enzyme activity, although APV has only a single CO-NH peptide bond compared with three in SQV [178]. Resistant mutations within the hydrophobic clusters frequently involve small changes such as the addition or deletion of a methylene group [143]. Mutations within the hydrophobic cluster have the potential to alter the flap dynamics or stability of PR as well as the binding of inhibitors, as shown in our structural analysis. The current drugs target the active site cavity and demonstrate strong binding to the catalytic Asps [179]. The hydrophobic

pockets around the substrate binding site have been proposed as an alternative drug target [180, 181]. However, our structural analysis showed that side chains in the hydrophobic cluster can rearrange readily to maintain the PR structure and activity, suggesting that this region is a poor target for drugs.

The accuracy of high and atomic resolution crystal structures is critical for deciphering how mutations and inhibitors alter the PR structure. Here, we describe how PR recognizes the inhibitors APV and SQV by structural rearrangements of the two beta-hairpin flaps and two 80's loops, and how the same mutation results in different structural changes with the two drugs. Comparison of the structures provides insight into why I50V is a major drug-resistant mutation observed on exposure to APV, but appears less critical in resistance to SQV therapy. Apart from the different interactions between PR and inhibitors, the absence of the C-H···O contact between flap residue 50 and Thr80 in PR_{I50V}–APV, and its presence in PR_{I50V}–SQV and PR_{WT} complexes, contributes to the drug resistance of this mutation. The conclusion is that small rearrangements of the PR loops enclosing the inhibitor combined with changes in weak internal interactions produce the distinct patterns of resistant mutations for the two drugs.

2.6 Acknowledgment

CS was supported in part by the Georgia State University Research Program Enhancement award. The research was supported in part by the National Institutes of Health grant GM062920. We thank the staff at SER-CAT beamline at the Advanced Photon Source, Argonne National Laboratory, Argonne, IL, USA, for assistance during X-ray data collection. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under contract no. W-31-109-Eng-38.

3 CAPTURING THE REACTION PATHWAY IN NEAR-ATOMIC-RESOLUTION CRYSTAL STRUCTURES OF HIV-1 PROTEASE

(Published: Shen CH, Tie Y, Yu X, Wang YF, Kovalevsky AY, Harrison RW, Weber IT. Capturing the reaction pathway in near-atomic-resolution crystal structures of HIV-1 protease. Biochemistry. 2012 51(39): 7726-32.)

3.1 ABSTRACT

Snapshots of three consecutive steps in the proteolytic reaction of HIV-1 protease (PR) were obtained in crystal structures at resolutions of 1.2−1.4 Å. Structures of wild-type protease and two mutants (PRV32I and PRI47V) with V32I and I47V substitutions, which are common in drug resistance, reveal the gem-diol tetrahedral intermediate, the separating N- and C-terminal products, and the C-terminal product of an autoproteolytic peptide. These structures represent three stages in the reaction pathway and shed light on the reaction mechanism. The near-atomic-resolution geometric details include a short hydrogen bond between the intermediate and the outer carboxylate oxygen of one catalytic Asp25 that is conserved in all three structures. The two products in the complex with mutant PR_{I47V} have a 2.2 Å separation of the amide and carboxyl carbon of the adjacent ends, suggesting partial cleavage prior to product release. The complex of mutant PR_{V32I} with a single C-terminal product shows density for water molecules in the other half of the binding site, including a partial occupancy water molecule interacting with the product carboxylate end and the carbonyl oxygen of one conformation of Gly27, which suggests a potential role of Gly27 in recycling from the product complex to the ligand-free enzyme. These structural details at near-atomic resolution enhance our understanding of the reaction pathway and will assist in the design of mechanism-based inhibitors as antiviral agents.

3.2 Introduction

The human immunodeficiency virus type 1 protease (HIV-1 PR) acts as a dimer of two identical 99-amino acid subunits to process the viral Gag and Gag-Pol polyproteins into functional proteins (Figure

3-1 A) [52]. The indispensable function of PR in replication of infectious virus makes it an important target for antiretroviral therapy. However, the efficacy of PR inhibitors decreases over time because of the evolution of drug resistance, primarily by mutations in the PR[20]. Even treatment with highly active antiretroviral therapy (HAART) does not completely eliminate resistant virus. Therefore, there is a continuing need for new PR inhibitors to combat drug resistance. Improved knowledge of the PR reaction intermediates will help in the design of novel mechanism-based inhibitors.

The aspartyl protease family is widely distributed in a variety of organisms, and its members participate in diverse biological functions; however, the detailed proteolytic mechanism is not fully understood[45]. Several experimental or theoretical studies have addressed the reaction mechanism of peptide cleavage by aspartyl proteases. ^{18}O exchange mass spectrometry experiments with HIV-1 PR suggest that the peptide hydrolysis reaction proceeds via the formation of a reversible and metastable gem-diol reaction intermediate[39]. Recent analysis by neutron crystallography has provided critical information about the location of hydrogen atoms in the active site of HIV-1 PR complexed with an inhibitor, which does not contain the gem-diol reaction intermediate [42]. X-ray crystallographic analysis of reaction intermediates trapped in the enzyme structure has proven to be a powerful tool for probing catalytic mechanisms [133]. Veerapandian et al. used the aspartic protease endothiapepsin complexed with a renin inhibitor mimicking both hydroxyls in the putative intermediate to study the proteolytic mechanism [182, 183]. X-ray structures were reported for N-terminal and C-terminal peptide products bound to PRs from HIV-1 and the closely related simian immunodeficiency virus[184]. Other studies soaked tethered HIV-1 PR crystals with a substrate peptide to trap different components of the reaction, including the tetrahedral intermediate [135, 185]. Our group has reported the highest resolution of 1.5 Å for the structures of a tetrahedral reaction intermediate in the wild-type and mutant HIV-1 PR [134]. In these examples, however, the structural information is limited by the resolution and the disorder observed frequently for peptide intermediates.

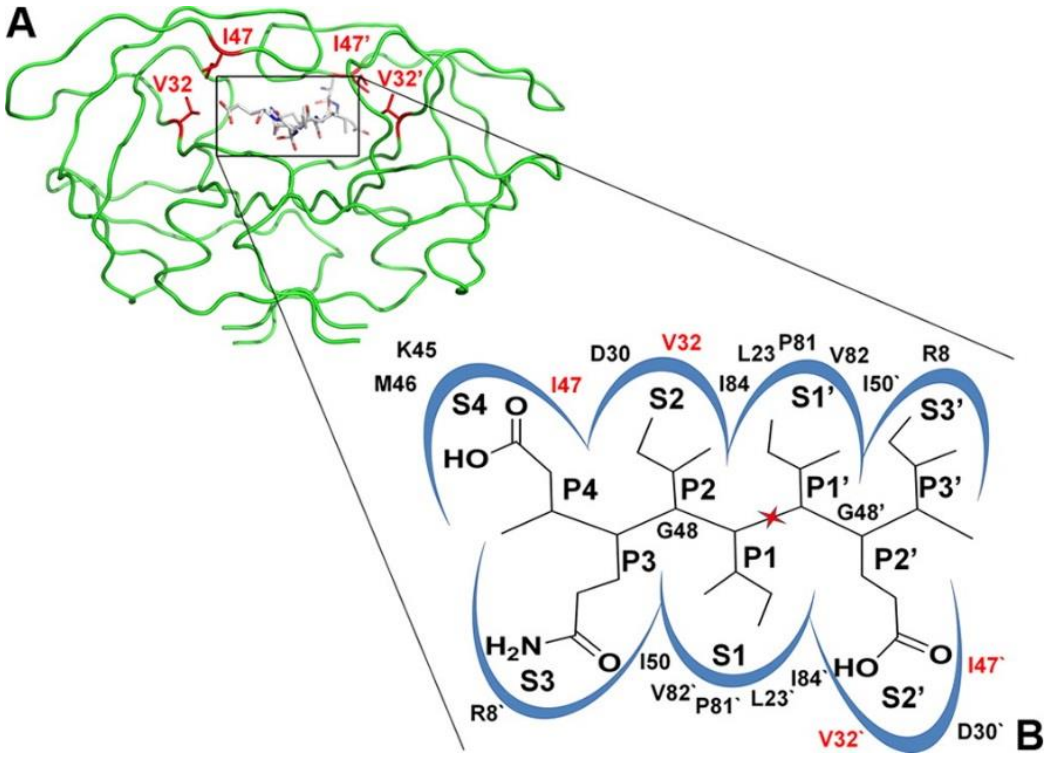


Figure 3-1 Structure of the HIV-1 PR dimer in a green backbone representation.
(A) The sites of mutations Val32 and Ile47 are shown as red sticks for the side chain atoms in both subunits, with the prime indicating the "second" subunit. The tetrahedral intermediate (TI) peptide is shown as sticks colored by atom type. (B) Schematic illustration of the substrate binding site of HIV-1 PR. The peptide DQIIxIEI (P4–P3') is shown in the S4–S3' subsites of the PR dimer. The scissile peptide bond is indicated by the red star. PR residues contributing to the binding site are indicated.

More recently, our studies have focused on understanding the influence on the reaction intermediates of selected mutations found in drug resistance. Moreover, we identified a 1.2 Å resolution structure of the wild-type PR with a tetrahedral intermediate. PR recognizes peptide substrates of at least six residues binding in subsites S3–S3' within the active site cavity (Figure 3-1 B)[52]. Mutations of PR residues Val32 and Ile47 were selected because they contribute hydrophobic interactions with substrates or inhibitors (Figure 3-1 A), and mutations at these sites are common in drug resistance[94]. Multidrug resistant mutation V32I appears in ~20% of patients treated with amprenavir[144] and is associated with high levels of drug resistance to lopinavir and ritonavir[94]. Drug resistant mutation I47V is located in the flexible flap and interacts with the inhibitor. Mutation I47V is associated with resistance to darunavir, lopinavir, tipranavir, and ritonavir in therapy [94, 145].

We describe crystal structures of PR and its mutants with different reaction intermediates and the implications for the proteolytic mechanism. The crystal structures of wild-type HIV-1 PR (PR_{WT}) and its mutants containing the single substitutions I47V (PR_{I47V}) and V32I (PR_{V32I}) were refined at near-atomic resolutions of 1.2–1.4 Å and, by serendipity, illustrate three different steps in the hydrolytic reaction. Peptide products and the reaction intermediate corresponding to an autoproteolytic cleavage site were observed in the structures, as described previously [134]. These near-atomic-resolution crystal structures provide more accurate information for the catalytic mechanism and the design of next-generation antiviral inhibitors.

3.3 Experimental Procedures

3.3.1 Protein Preparation and Crystallization of HIV-1 PR_{WT}, PR_{V32I}, and PR_{I47V}

The mutants were constructed and expressed in bacteria, and the protein was purified as described previously [134]. The crystallization trials employed the hanging drop method using equal volumes of enzyme/inhibitor and reservoir solution. PR and mutant proteins at 2.2 mg/mL were mixed

with the inhibitor or peptide (dissolved in DMSO) at a molar ratio of 1:5 and incubated on ice for 30 min prior to centrifugation to remove any insoluble material. PR_{WT} was crystallized from 0.1 M sodium acetate buffer (pH 4.8), 0.41 M potassium chloride, and an investigational inhibitor. PR_{V32I} was crystallized from 0.06 M sodium acetate buffer (pH 5.6), 0.67 M sodium chloride, and a synthetic peptide. PR_{I47V} crystals were grown from 0.05 M sodium acetate buffer (pH 5.0), 1.2 M sodium formate, and 2.5% PEG8000.

3.3.2 X-ray Data Collection and Refinement

Single crystals were mounted on fiber loops with 25% (v/v) glycerol as a cryoprotectant in the reservoir solution. X-ray diffraction data were collected at the SER-CAT beamline of the Advanced Photon Source, Argonne National Laboratories (Argonne, IL). Diffraction data were integrated, scaled, and merged using the HKL2000 package[156]. Structures of PR_{WT}, PR_{V32I}, and PR_{I47V} were determined by the molecular replacement program Molrep [157] using structures 3B7V, 1FG6 and 2F8G as the respective starting models[97, 134, 147]. Refinement was conducted using SHELX-97[160]. No electron density was seen for the inhibitor used in the crystallization solution for PR_{WT}; however, the gem-diol intermediate of autoproteolysis of PR residues 59–67 (YDQIIxIEIA) fit the observed density. The synthetic peptide sequence did not fit the electron density in PR_{V32I} and was replaced by the single C-terminal product of residues 59–63 (YDQII). Tyr59 at P5 was refined as Ala because of the poor electron density for its side chain atoms in PR_{WT} and PR_{V32I}. PR_{I47V} was refined with both N-terminal and C-terminal products extending in opposing directions (residues 60–63 with residues 64 and 65 in one orientation and residues 61–63 with residues 64–66 in the other). The structures were refined with anisotropic atomic displacement parameters (*B* factors). Hydrogen atoms were added at the final stages of the refinement. The molecular graphics program COOT was used for map display and model building[162]. Structural figures were made using PyMol [163]. The structures were compared by superimposing their Cα atoms and using HIVAGENT[164] to calculate the distance between two atoms. The cutoff distances for different types of interactions were as described in ref [69].

3.4 Results

3.4.1 Crystallographic Analysis

The crystal structures of PR_{WT}, PR_{V32I}, and PR_{I47V} were determined in the same space group$P2_12_12$, and the crystallographic statistics are summarized in Table 3-1. The asymmetric units include one PR homodimer, and the residues in the two subunits are labeled 1–99 and 1'–99'. The diffraction data extend to a resolution of 1.2 Å for PR_{WT}, 1.3 Å for PR_{I47V}, and 1.4 Å for PR_{V32I}, and the structures were refined to *R* factors of 14.4–17.5%. The majority of protein residues and solvent molecules showed clear electron density in all the structures. The tetrahedral intermediate and cleavage products of PR residues 59–67 (YDQII*IEIA, where the asterisk indicates the cleavage site between P1 and P1' residues) matched the electron density in the active site cavities of the PR_{WT}, PR_{I47V}, and PR_{V32I} structures. These peptides are thought to derive from very slow autoproteolysis, because the L63I substitution almost eliminates a site of autoproteolytic cleavage [186]. The PR_{WT} dimer included two partial occupancy gem-diol intermediates: residues 59–65 (designated P5–P2') were fit in one conformation and 61–67 (P3–P4') for the conformation in the opposite orientation with relative occupancies of 0.4 and 0.5, respectively. The two mutants trapped the product peptides. A single C-terminal product containing residues 59–63 (P5–P1) was seen in PR_{V32I}. Ala was refined instead of Tyr59 at P5 because of weak electron density for the side chain in PR_{WT} and PR_{V32I}. PR_{I47V} was refined with both N- and C-terminal products in two alternate conformations comprising residues 60–63 (P4–P1) with residues 64 and 65 (P1'–P2') at 0.5 occupancy and residues 61–63 (P3–P1) with residues 64–66 (P1'–P3') at 0.4 occupancy. The side chains of P4 Asp and P3 Glu were not visible in the electron density. The solvent, consisting of water molecules, sodium ions, chloride ions, and glycerol, was modeled with 204, 181, and 140 molecules in PR_{WT}, PR_{I47V}, and PR_{V32I}, respectively. Alternate conformations were refined for 26 residues in PR_{WT}, 13 residues in PR_{I47V}, and 7 residues in PR_{V32I}. Generally, alternate conformations were seen for the longer side chains on surface residues. A few internal residues showed alternate conformations for the side chain or main chain: Ile84,

Table 3-1 Crystallographic Data Collection and Refinement Statistics

	PR_{WT}	**PR_{I47V}**	**PR_{V32I}**
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit cell dimensions: (Å)			
A	58.41	58.08	58.06
B	86.20	86.30	86.14
C	46.36	46.35	46.30
Resolution range (Å)	10-1.2	10-1.31	50-1.4
Unique reflections	66188	55879	45833
R_{merge} (%) overall (final shell)	9.7 (39)	6.3 (51)	7.2 (43)
I/σ(I) overall (final shell)	15.03 (2.2)	27.8 (2.1)	18.3 (2.1)
Completeness (%) overall (final shell)	94.3 (56.8)	98.8 (90.6)	93.0 (98.7)
Data range for refinement (Å)	10-1.2	10-1.31	10-1.4
R (*%*)	0.14	0.15	0.17
R_{free} (*%*)	0.18	0.18	0.23
No. of solvent atoms (total occupancies)	206 (191)	181 (169.5)	140 (134)
RMS deviation from ideality			
Bonds (Å)	0.015	0.012	0.010
Angle distance (Å)	0.034	0.031	0.029
Average B-factors ($Å^2$)			
Main-chain atoms	19.2	17.1	16.8
Side-chain atoms	24.9	20.4	23.3
Peptide Intermediate	54.1	53.5	32.2
Solvent	36.4	33.3	30.5
Intermediate Peptide	(Y)[a]DQIIxIE	(D)(Q)II+IE	(Y)DQII
Peptide Occupancy	0.5/0.4	0.5/0.4	1.0

[a] Parentheses indicate peptide residues refined as alanine because of poor electron density for the longer side chain.

Leu97, Ile33', and Ile84' of PR_{WT}, Gly27, Ile84, and Ile33' of PR_{V32I}, and Val47, Ile64, Ile84, Ile33', and Ile84' of PR_{I47V}. The backbone structures were almost identical for the three dimers with low pairwise rmsd's of 0.15–0.17 Å for all Cα atoms.

3.4.2 Structural Changes around the Mutated Residues

The two drug resistant mutations are conservative substitutions of hydrophobic residues in the substrate binding cavity. The side chain of Val32 in the PR_{WT} structure forms van der Waals contacts with the internal hydrophobic cluster comprising residues Ile50', Ile47, Ile56, Leu76, Thr80, and Ile84, and similar interactions were seen in the other subunit. The longer Ile32 side chain in the PR_{V32I} mutant has the potential to form new van der Waals contacts within the cluster, as described for the PR_{V32I} complex with amprenavir[187]. In one subunit, the Cδ1 methyl of the Ile32 side chain is directed toward the hydrophobic cluster, providing new van der Waals contacts with several hydrophobic side chains. In the other subunit, however, the Cδ1 methyl of Ile32' rotates to form new hydrophobic contacts only with Ile47' and Ile50 (Figure 3-2 A). The 3.3 Å distance between the Cδ1 methyl groups of Ile32' and Ile47' is unusually short for a C–H···H–C interaction, which may indicate a destabilizing interaction[188]. Ile32 forms a van der Waals contact with Ile at the P2 position in the product peptide similar to the contact seen for Val32 in PR_{WT}, while no product is seen in the other subunit.

The hydrophobic side chain of Ile47 forms internal hydrophobic contacts and interacts with P2 and P2' residues of the reaction intermediate in both subunits of the PR_{WT} complex. These hydrophobic interactions are retained in the PR_{I47V} mutant. In PR_{WT}, Cδ1 of Ile47 interacts with Ile50', Val32, Val56, and Leu76, while Cδ1 of Ile47' shows van der Waals interactions with the corresponding residues in the other subunit and with Ile54'. Mutation I47V substitutes the shorter Val side chain and eliminates van der Waals contacts with adjacent residues, thus probably decreasing the stability of the hydrophobic cluster in PR_{I47V} (Figure 3-2 B).

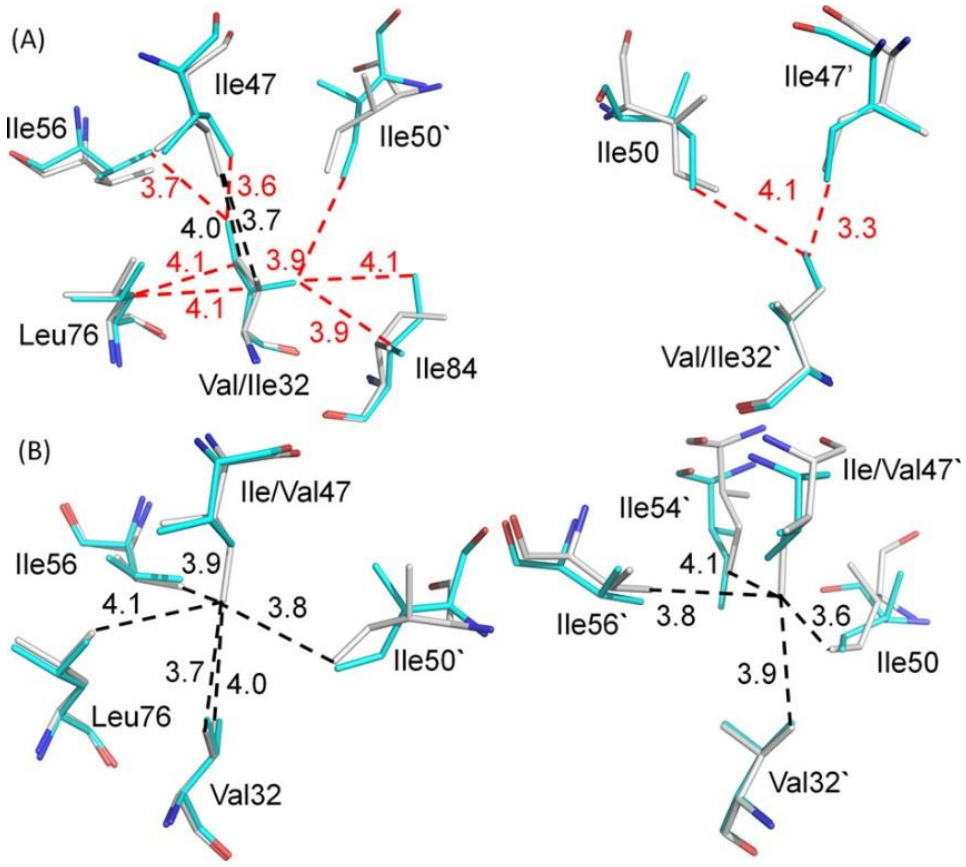


Figure 3-2 Mutations alter internal hydrophobic contacts.
Residues are shown from superimposed structures of PR_{WT} with (A) PR_{V32I} and (B) PR_{I47V}. The two subunits are shown in the left and right panels, respectively. The PR_{WT} residues are colored gray for carbon atoms, while the mutants are colored cyan. The van der Waals interactions are indicated by dashed lines in black for PR_{WT} and red for the mutants with interatomic distances in angstroms. The mutation of I47 to a smaller side chain in PR_{I47V} eliminates hydrophobic contacts seen for Ile47 in PRWT. The opposite effect occurs with substitution of the large side chain in PRV32I.

3.4.3 PR_{WT}–TI Interaction

The tetrahedral intermediate (TI) of residues Y_{59}DQII*IEIA_{67} (where the asterisk indicates the gem-diol group) was observed in the omit $F_o - F_c$ electron density map in PR_{WT} in two alternate conformations extending in opposite directions (Figure 3-3 A), as found frequently in PR complexes with peptide analogues [21, 97]. Residues P3–P4′ were fit in one conformation and residues P5–P2′ in the opposite orientation, with occupancies refined to 0.5 and 0.4, respectively. The omit electron density ($F_o - F_c$) map for the gem-diol structure clearly indicates four hydroxyl oxygens (Figure 3-3 B). The two alternate conformations of the TI peptide formed essentially identical interactions with PR_{WT}. Interactions with the catalytic Asp25 and -25′ are described later. The main chain atoms of the TI peptide formed hydrogen bond interactions with residues Gly27, Asp29, Gly48, Gly27′, Asp29′, Asp30′, and Gly48′ (Figure 3-4 A). The side chain of Glu at P2′ showed hydrogen bond interactions with Asp29 and Asp30. A shorter interaction of 2.4 Å seen between the carboxylate side chains of P2 Glu and Asp30 is consistent with protonation of P2 Glu, as described in other crystal structures of PR with peptide analogues [21, 97]. Additional stabilizing interactions include water-mediated hydrogen bonds and hydrophobic contacts between PR_{WT} and the peptide.

3.4.4 Interactions of PR_{I47V} with Two Products

The PR_{I47V} structure revealed two alternate conformations of both the N- and C-terminal products extending in opposite directions (Figure 3-3 C). The occupancy of residues P4–P1 and P1′–P2′ in one conformation was 0.5, and residues P3–P1 and P1′–P2′ were refined with an occupancy of 0.4 for the opposite orientation. The alternate conformations maintained similar interactions with the two subunits of PR_{I47V}. Interactions with the catalytic Asp25 and -25′ are described below. Four hydrogen bond interactions connected the main chain atoms of product P4–P1 with residues Gly27, Asp29, and Gly48, and two hydrogen bond interactions linked main chain atoms of product P1′–P2′ with residues Gly27′ and

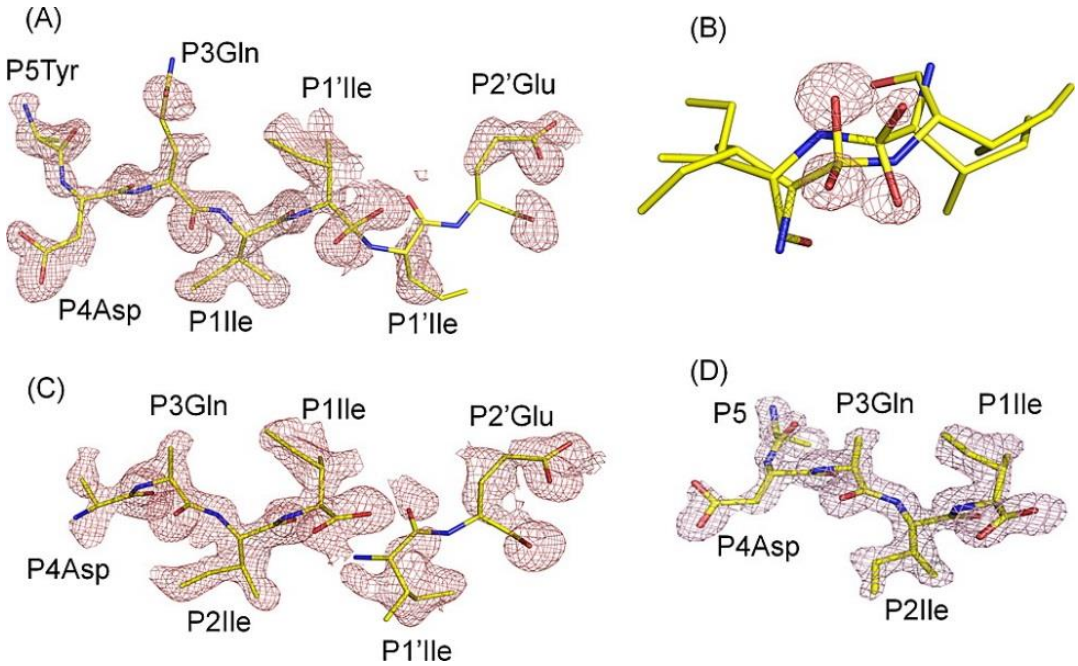


Figure 3-3 Electron density maps for the peptide intermediates.

(A) Omit map (*F*o – *F*c) for the major conformation of TI in the PRWT structure. The contour level is 2.0σ. (B) *F*o – *F*c omit map for the hydroxyl oxygen atoms of P1 Ile in the major and minor conformations contoured at 3.5σ. (C) Omit electron density map (*F*o – *F*c) for the major conformation of product peptides in the PRI47V structure. The contour level is 2.0σ. (D) *F*o – *F*c omit map for the peptide C-terminal product of PRV32I contoured at 2.5σ. The corresponding stereofigures are given in Figure S1A–D of the Supporting Information.

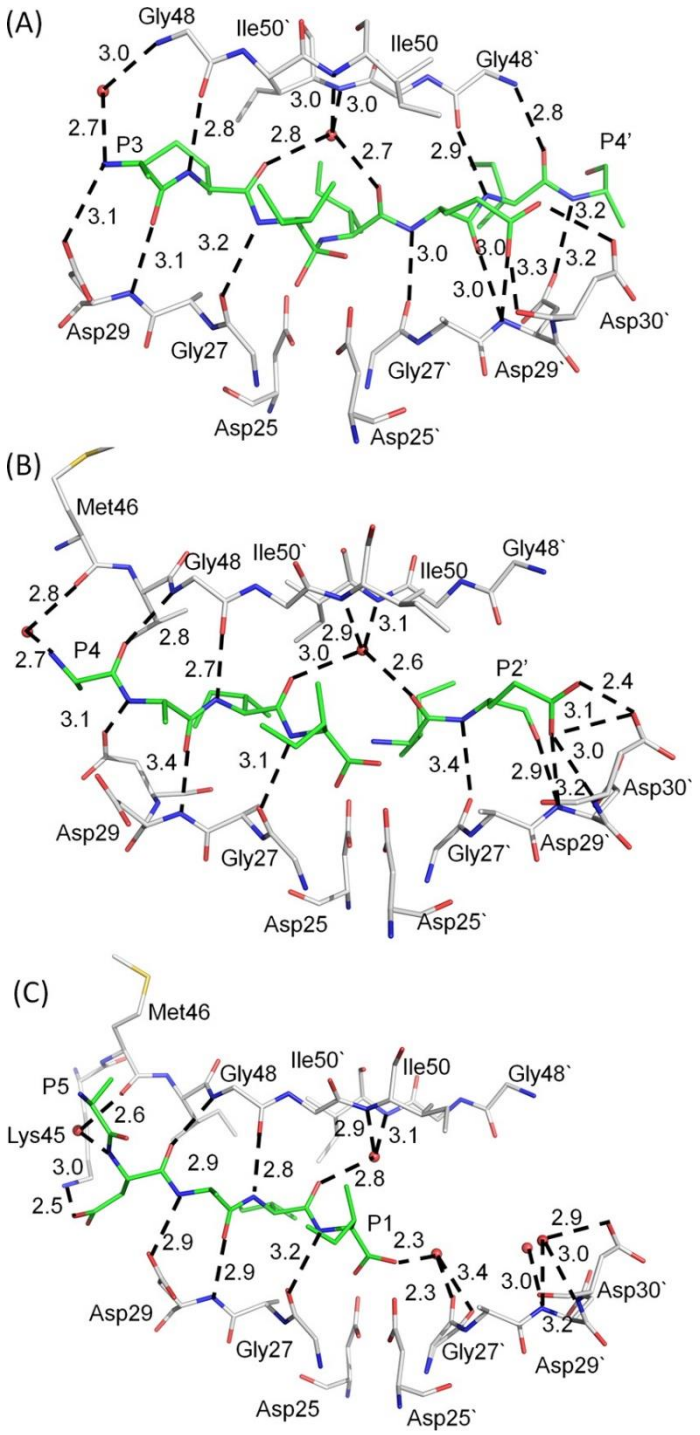


Figure 3-4 Hydrogen bond interactions:
(A) the tetrahedral intermediate with PR_{WT}, (B) the two products with the PR_{I47V} mutant, and (C) the single product with the PR_{V32I} mutant. Only the major conformation is shown for the peptide intermediate or products in panels A and B. Hydrogen bond interactions are indicated by dashed lines with distances in angstroms. Interactions of Asp25 and -25' have been omitted for the sake of clarity. PR is shown with gray carbons, and TI and product peptides are shown with green carbons. The water and carbonyl oxygen of Gly27' in panel C were refined with a partial occupancy of 0.45.

Asp29′ (Figure 3-4 B). The side chain carboxylate oxygen atoms of P2′ Glu interact with Asp29′ and Asp30′, as described for P2′ in the TI peptide.

3.4.5 Interactions of PR_{V32I} with the P5–P1 Product

The PR_{V32I} structure showed clear electron density for the single peptide product of residues P5–P1, except for the side chain of P5 (Figure 3-3 D). Individual water molecules were fit in the other subunit because no peptide product was visible in the electron density. The polar interactions of the single product with PR_{V32I} are shown in Figure 3-4 C. Two waters near the P2′ position formed hydrogen bond interactions to the main chain of Asp29′ and main chain and side chain of Asp30′, resembling the interactions of the P2′ Glu in the other peptides. Another water molecule at an occupancy of 0.5 interacts with both alternate conformations of the carbonyl of Gly27′ and with the product carboxyl terminus. Hydrogen bond interactions connect the main chain of the product to PR_{V32I} residues Gly27, Asp29, and Gly48, as observed in the other complexes. A new salt bridge was seen between the side chains of P4 Asp and Lys45, and the P4 main chain amide had a water-mediated interaction with the carbonyl oxygen of Met46. Hydrophobic interactions were also observed between the side chains of the product and side chains of PR_{V32I}.

3.4.6 Interactions with Catalytic Residues

The interactions with the catalytic residues are essentially identical in the major and minor conformations of the TI and two product peptides as shown by the superposition in Figure 3-5. In the TI complex with PR_{WT}, the gem-diol structure lies between P1 Ile and P1′ Ile (Figure 3-6 A). The O1 hydroxyl of P1 Ile forms a very short 2.3 Å hydrogen bond to the outer carboxylate Oδ2 atom of Asp25′, while the O2 hydroxyl interacted with all four carboxylate oxygens of Asp25 and -25′. The nitrogen of the gem-diol-amine moiety had a hydrogen bond interaction with Oδ2 of Asp25.

The PR_{WT}–TI interactions are comparable to those described in the 1.0 Å resolution crystal structure of endothiapepsin, a fungal aspartic proteinase, complexed with a gem-diol analogue[183]. In

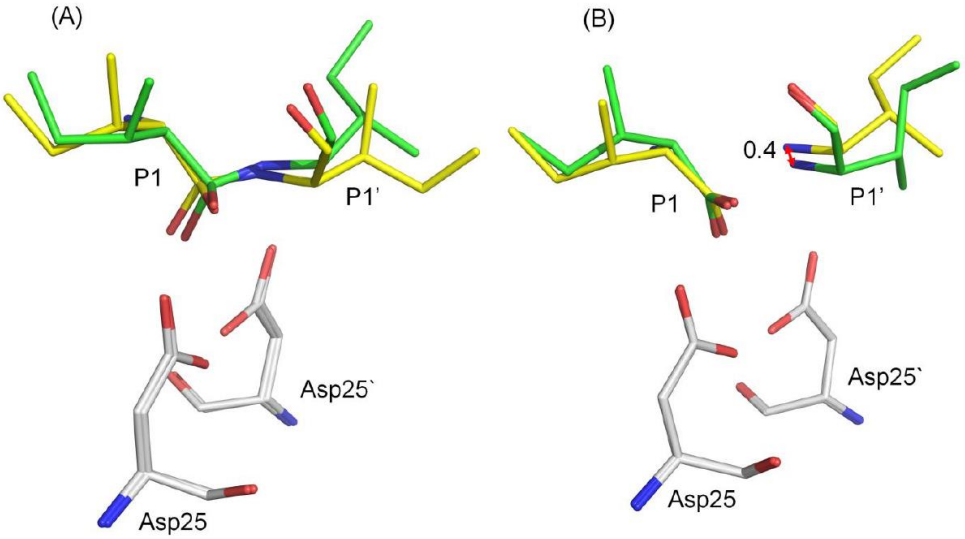


Figure 3-5 Minor conformations of TI or products superposed on major conformations showing that main chain atoms retain similar positions.
Gray color represents PR. Green color represents TI or product refined with 0.5 occupancy. Yellow color indicates TI or products with 0.4 occupancy. A) PR_{WT}-TI. B) PR_{I47V} with two products. The main chain atoms of P1 and P1' have small RMSD values of 0.2 Å and 0.4 Å, respectively, in both A) and B).

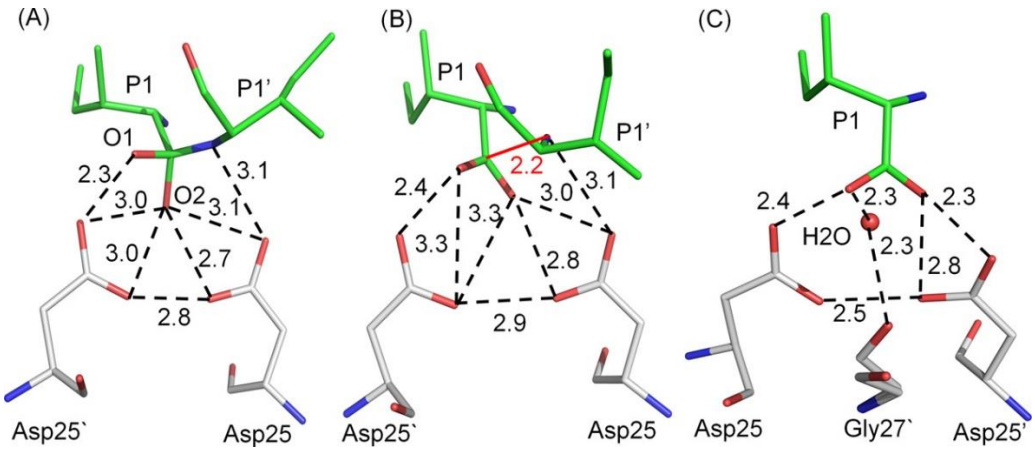


Figure 3-6 Hydrogen bond interactions with the catalytic residues Asp25 and -25':
(A) PR_{WT}–TI complex, (B) PR_{I47V} with both products, and (C) PRV32I with a single product. Hydrogen bond interactions are indicated by dashed lines with distances in angstroms. The PR is colored gray, and TI and products are colored green. The short separation of 2.2 with distances in angstroms between the N- and C-terminal products is indicated by the red line in panel B. Only one conformation (occupancy of 0.5) of Gly27' and interacting water is shown in panel C.

endothiapepsin, one short hydrogen bond interaction was seen between the O1 hydroxyl and Oδ2 of Asp35 with a distance of 2.5 Å; the O2 hydroxyl formed two hydrogen bond interactions with Oδ1 and Oδ2 of Asp219 with distances of 2.6 and 3.0 Å, respectively; and one hydrogen bond interaction with Oδ2 of Asp35. Unlike PR_{WT}, the hydrogen bond interaction between the amide nitrogen of the gem-diol analogue and the catalytic aspartate was absent in the endothiapepsin complex.

The inhibitor KNI-272 has been used as a transition state mimic in neutron diffraction studies to locate important hydrogen atoms at the catalytic site of HIV PR [42]. Although the hydroxymethylcarbonyl isostere of KNI-272 is not identical to the gem-diol of TI, the hydrogen bond interactions of O2 of KNI-272 are similar to those of O2 of P1 Ile with both catalytic aspartates. Also, O4 of KNI-272 resembles O1 of TI in forming a single hydrogen bond interaction to Oδ2 of Asp25, although the hydrogen bond is not particularly short.

The two product peptides bound in PR_{I47V} are partially separated with a 2.2 Å distance between the P1′ amide nitrogen and the P1 carbonyl carbon (Figure 3-6 B), whereas the standard C–N separation of 1.3–1.4 Å occurs in the TI complex. The 2.2 Å distance may indicate that there is still some bonding interaction present between the N atom of the amino product and the C atom of the carboxylate product. The N-terminus of P1′ Ile has one hydrogen bond interaction with Oδ2 of Asp25. One carboxylate oxygen of P1 Ile forms a short 2.4 Å hydrogen bond interaction with the outer Oδ2 atom of Asp25′, as seen for the tetrahedral intermediate. The other carboxylate oxygen of P1 has hydrogen bond interactions with the four aspartate carboxylate oxygens. This analysis suggests that the structure of PR_{I47V} with two products represents the stage immediately after hydrolysis of the peptide bond.

In PR_{V32I}, a partial (0.5) occupancy water (Figure 3-6 C) in the catalytic site forms hydrogen bond interactions linking the carboxylate oxygen of P1 and the alternate conformations of the carbonyl oxygen of Gly27′. One carboxylate oxygen of P1 forms a short hydrogen bond with Oδ2 of Asp25 with a distance of 2.4 Å, and the other carboxylate oxygen of the product forms two hydrogen bond interactions (2.3 and

2.8 Å) with Oδ1 and Oδ2 of Asp25′. These hydrogen bond interactions of the P1 carboxylate group are similar to those reported in other structures with synthetic peptide products [189]. This complex appears to represent the reaction step after release of the N-terminal product from the catalytic site. The alternate positions of the carbonyl oxygen of Gly27′ may reflect conformational changes occurring upon release of the N-terminal product, because this oxygen usually interacts with an amide of the bound substrate peptide, as shown in the PR_{WT}–TI complex and the PR_{I47V} complex with two products.

3.5 Discussion

3.5.1 Implications for the Reaction Pathway

Reaction intermediates have been described in a variety of enzyme crystal structures, as reviewed in ref [190]. Several metastable intermediates can sometimes exist along a reaction path indicating the catalytic process proceeds through a series of energy barriers. In the case of HIV PR, a number of intermediates can be trapped by simply crystallizing the protein in the presence of a peptide substrate, which suggests that the energy barriers of the hydrolysis reaction may be of similar height. Thus, it is entirely possible that the reaction pathway of the peptide hydrolysis by PR does not include a single rate-limiting step. The three new crystal structures represent three consecutive steps in the proteolytic reaction of HIV PR and provide improved geometric details because of the near-atomic-resolution X-ray data. The majority of the interactions with the bound peptides are conserved in the wild-type enzyme and the mutants, in agreement with the report that mutants share similar transition states with wild-type PR [191]. The interactions observed in the new structures have been mapped on the scheme for the reaction (Figure 3-7**Error! Reference source not found.**). The hydrogen atoms around the catalytic sites cannot be identified in X-ray structures at this resolution; however, their locations were deduced from the short interactions. One Asp25 is protonated and the other deprotonated in most states, in agreement with

neutron diffraction studies of an inhibitor complex [42]. In the absence of substrate, the active site cavity of PR generally contains water, or possibly

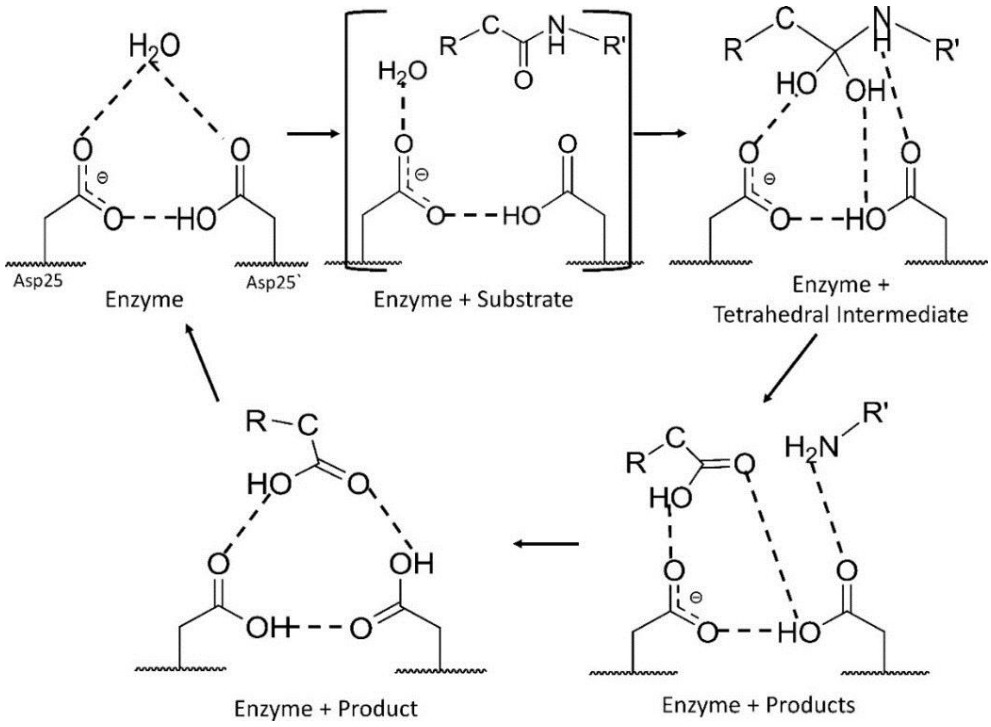


Figure 3-7 Scheme of the reaction pathway.
The interactions of the peptide intermediates with the catalytic residues are illustrated on the basis of the new crystal structures. The minimal number of hydrogen atoms and hydrogen bonds is indicated based on the interaction distances in the crystal structures. Crystal structures have been described for four stages. Only the active enzyme with bound substrate has not been seen, as indicated by the large brackets.

a metal cation, interacting with the two catalytic aspartates [192]. This water reacts with the bound peptide to form the gem-diol intermediate, represented by the structure of the PR_{WT}–TI complex (Figure 3-6 A). The tetrahedral intermediate dissociates into two products in the step shown in the PR_{I47V} complex (Figure 3-6 B). Then, the N-terminal product is released, in association with rotation of the carbonyl group of Gly27', as shown in the PR_{V32I} structure (Figure 3-6 C). This state with a single product is shown with protonation of both Asp25 and -25' to reflect the short interactions observed in the crystal structure, which suggests diprotonation in the simplest interpretation, as found in theoretical studies of PR with some inhibitors[193]. Finally, the C-terminal product is released, and the enzyme recycles to the first step. All three intermediate stages retain the short 2.3–2.4 Å hydrogen bond, which may be a low-barrier hydrogen bond, of the hydroxyl group of the peptide intermediate with one of the catalytic aspartates, as reported in lower-resolution crystal structures [135, 185, 189]. Moreover, the amide of the cleaved bond in the gem-diol intermediate and the freed amino terminus of the product also form hydrogen bond interactions with the carboxylate of Asp25.

The conserved catalytic Asp25-Thr26-Gly27 triplets are important for the activity and dimerization of PR. Previous X-ray and neutron diffraction studies have shown hydrogen bond interactions with the catalytic Asp25 and -25' similar to those described in our new structures [42, 134]. The position of hydrogen atoms, however, cannot be determined unambiguously in these X-ray structures, and further studies by neutron crystallography will be important in determining the protonation state of the catalytic residues in the reaction steps. Our structures suggest a new role for rotation of Gly27 to facilitate release of product as well as its known role in binding the main chain amides of peptides [21]. Analysis of these structures emphasizes the essential roles of residues 25–30 and flap residues 48–50 in binding peptide substrates and products and provides details of the geometry around the catalytic site at near-atomic resolution. The critical interactions can be targeted in the design of new antiviral inhibitors for resistant HIV.

3.6 Acknowledgment

Data were collected at Southeast Regional Collaborative Access Team (SER-CAT) beamline 22ID at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at http://www.ser-cat.org/members.html. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract W-31-109-Eng-38.

4 DYNAMIC VARIATION IN THE FLAPS OF AN EXTREME DRUG RESISTANT HIV PROTEASE VARIANT

(In-preparation: Chen-Hsiang Shen, Yu-Chung Chang, Johnson Agniswamy, Robert W. Harrison and Irene T. Weber)

4.1 Introduction

Drug resistance is a critical barrier to successful treatment of HIV/AIDS. The human immunodeficiency virus type 1 (HIV-1) protease (PR) is a retroviral protease composed of two identical subunits of 99 amino acids. PR proteolytic activity is essential for maturation of the virion through processing viral precursor polyproteins, Gag and Gag-Pol, into functional and structural proteins [20]. HIV-1 PR is an important and effective target for antiviral drug therapy. PR inhibitors are more effective than reverse transcriptase inhibitors in preventing cell-free and cell to cell transmission of the virus [194]. Drug resistance arises quickly in HIV because of the lacking of proof reading function of the viral reverse transcriptase, in addition to the fast replication of the virus[195]. The introduction of highly active antiretroviral therapy (HAART) greatly improved the survival of HIV infected people. However, the viral replication is merely temporally repressed by HAART, and life time administration of drugs is essential. Thus, compliance to HAART is critical to alleviate the emergence of resistance [196, 197].

Multidrug resistant viral strains are common in infected patients, to date, 36 missense mutations with resistance to one or more of the nine FDA approved PR inhibitors have been identified[198]. We have studied a highly drug resistant variant of PR from a clinical isolate, which bears 20 mutations (PR20) that alter different structural regions of the catalytically active dimer [103]. Compared to wild-type enzyme, PR20 shows higher dimer dissociation constant (k_d) and ~13 fold higher K_m value for substrate.[199] Importantly, PR20 shows drastically decreased affinity of greater than three orders of magnitude compared to the wild type PR for current drugs [199]. The stability of the core monomer structure was enhanced for PR20 relative to the wild type enzyme.[200]. Altered intersubunit interactions and unusually

widely separated flaps were observed in an X-ray crystallographic study of PR20 in the presence and absence of inhibitors.[201]

The flaps comprising amino acid residues 45-55 are an important part of the PR dimer structure. The two interacting flaps exist in a closed conformation when inhibitor or substrate is bound, while in the apo-enzyme, the flaps can adjust into an open conformation, distant from the catalytic site. Three main categories, defined as closed, semi-open, and open conformations, have been described for different structures of the apo-enzyme.[25] The observation of widely separated flaps in drug resistance mutants such as PR20 and MDR769[105] suggests that altered flap flexibility may contribute to resistance. Therefore, we have explored the conformational variation of the flap regions using crystallographic studies of the PR20 mutant with the inactivating D25N mutation ($PR20_{D25N}$) in combination with molecular dynamics (MD) simulations of the apo enzymes. Introduction of the D25N mutation alone increased the equilibrium dimer dissociation constant by more than 100-fold and decreased the binding affinity of darunavir (DRV) by about 10^6 fold relative to the wild type PR, nevertheless, no significant alteration was observed in the inhibitor-bound dimer structure or the interactions with DRV.[179]

Our new crystal structures of ligand-free $PR20_{D25N}$ showed various flap conformations. In particular, an unusual “tucked” conformation with one flap penetrating into the active site cavity was observed in one crystal structure, and also occurred in the MD simulations of the dimers. Extreme flap conformations were found for $PR20_{D25N}$ relative to wild type enzyme in both the crystal structures and the MD simulations consistent with the idea that increased flap mobility contributes to drug resistance in this highly resistant variant.

4.2 Materials and Methods

4.2.1 Preparation of PR20 with D25N mutation

Plasmid DNA encoding PR (subtype B of group M) with 20 mutations Q7K, L10F, I13V, I15V, D30N, V32I, L33F, E35D, M36I, S37N, I47V, I54L, Q58E, I62V, L63P, A71V, I84V, N88D, L89T and L90M (termed PR20)[201] cloned between the Nde1 and BamH1 sites of pET11a vector (Novagen, San Diego, CA) was used to introduce the D25N mutation by the Quick-Change mutagenesis kit (Stratagene). This construct ($PR20_{D25N}$) was transformed into *E. coli* BL-21 (DE3; Stratagene) for protein expression, purification, and folding as described [201].

4.2.2 Protein crystallization, X-ray data collection and structure determination

Protein crystals were grown by the hanging drop vapor diffusion method. Crystals of $PR20_{D25N}$ were obtained by mixing 1 µL of protein (2.1 mg/mL) and 1 µL of reservoir solution under two different conditions: 0.2M magnesium chloride and 20% PEG 3350 at pH 5.9; and 0.9M sodium chloride and 0.2 M sodium acetate at pH 4.8. The crystals were cooled in a mixture of the mother liquor and 30% glycerol for X-ray data collection.

Diffraction data were collected at 100 K on beamline 22-ID of the Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source, Argonne National Laboratory. The data were integrated and scaled with HKL2000 [156]. The structure designated $PR20_{D25N}$ open was solved by molecular replacement with the wild-type HIV-1 PR in complex with p1/p6 (2AOI) as the starting model by PHASER[202]. The $PR20_{D25N}$open structure was refined using Refmac5.5 with TLS [161, 203], and the model building was carried out in COOT [162]. The second crystal structure of $PR20_{D25N}$ was solved by PHASER using 3UCB as the starting model [201, 202], and CNS and Refmac5.5 were used for refinement including anisotropic B factors [161, 204]. This crystal structure contains two $PR20_{D25N}$ dimers in the asymmetric unit. The flaps in both dimers were deleted initially and rebuilt during refinement. The solvent molecules were identified from the shape of the electron density, and interatomic distances and angles

consistent with hydrogen bond interactions with other molecules. Secondary-structure matching (SSM) was used for superposition of protein structures [205]. Molecular figures were prepared with PyMOL [163].

4.2.3 Molecular dynamics simulations of wild-type PR and PR20

The starting models for MD simulations were the DRV complexes of PR20 (3UCB)[201] and wild type PR (2IEN)[68] with resolutions of 1.4Å and 1.3Å, respectively. DRV was removed from the starting models for MD runs termed PR_{MD} and $PR20_{MD}$. A second set of simulations termed PR_{D25N}MD and PR_{20D25N}MD was prepared by substituting Asparagine for the catalytic Aspartic acids. Water molecules in the crystal structures were included in the MD simulation. Furthermore, the PR dimer was solvated with 50 sodium ions, 50 chloride ions and about nine thousand water molecules randomly generated to fill the free space within a 10Å shell of the protein. The temperature was set to 300 K and constrained by Nose constraints, as described in [118]. Initially, the system was equilibrated for the randomly generated solvent molecules prior to MD simulation. The program, AMMP, was used for the MD simulations with the TUNA potential set [206].

It is expected that the correct protonation state will produce the highest affinity of PR for inhibitor [207]. Calculations on PR_{WT} (PDB 3NU3) complexed with amprenavir [187], showed the lowest calculated binding energy ($\Delta G_{bind} = \Delta G_{PI+WAT}$) when the O$\delta$2 of Asp25 and O$\delta$1 of Asp25' were protonated. This protonation state agrees with 20ns the MD simulations performed by AMBER and analyzed by the Poisson-Boltzmann surface area (MM-PBSA) method [208]. This protonation state was used for the MD simulations. Simulations were performed with a NVT ensemble (atom number (N), volume (V) and temperature (T) are fixed) and an amortized fast multipole algorithm for the calculation of non-bonded terms. Initially, the system was equilibrated for randomly generated solvent molecules prior to MD simulation, and the substrate binding site was visually checked for full occupation by water molecules. A total of 1000 frames, one for every 10 ps, were saved from the 10 ns simulation. Each frame was

superimposed on the starting model to remove rotational and translational motion from the conformations before calculating the RMSD values and averaged structure.

Trajectories of the MD simulations are represented by a large number of snapshots, and analyzing these snapshots by hand is a daunting task. Therefore two statistical analyses were applied to the simulation. The major conformations of the protein were extracted using k-means clustering, and the individual variation was analyzed with a cross-correlation coefficient. The optimal number of clusters was decided by evaluating within the group sum of squares [209], and trajectories were partitioned by applying k-means algorithm that recursively assigns data points to its nearest centroid until the cluster converges [210]. The cross-correlation coefficient obtained by calculation of displacement vector of any two atoms i and j by $C_{ij} = (\Delta r_i \cdot \Delta r_j) / \sqrt{\sum \Delta r_i^2 \cdot \Delta r_j^2}$, where Δr is the displacement from the mean position of the atom[211]. Cross correlation map was plotted by gnuplot 4.6.

4.3 Results

4.3.1 Crystal structures of Ligand-free PR20$_{D25N}$

Two crystal structures were solved for ligand-free PR20$_{D25N}$ comprising three distinct dimers with varied flap conformations. The two crystal structures were refined to resolutions of 1.6 and 1.45Å and R-factors of 19.3% and 13.8%, as summarized in the crystallographic statistics shown in Table 4-1. One structure termed PR20$_{D25N}$open was refined with one dimer (residues 1-99 and 1'-99') in the asymmetric unit. The second crystal structure contained two dimers per asymmetric unit exhibiting different flap conformations: one dimer (PR20$_{D25N}$twist) showed twisted flap conformations relative to those of PR20open, and the other dimer (PR20$_{D25N}$tuck) had one flap tucked or inserted into the active site cavity and the opposite flap raised out of the cavity (Figure 4-1 A). Overall RMS values calculated for Cα atoms are: 1.5 Å (0.9 Å) for comparison of the PR20$_{D25N}$open and PR20$_{D25N}$twist dimers, 1.1 Å (0.8 Å) for PR20$_{D25N}$open and PR20$_{D25N}$tuck, and 1.4 Å (0.3 Å) for PR20$_{D25N}$twist and PR20$_{D25N}$tuck. [Values inside the

Table 4-1 Crystallographic Data Collection and Refinement Statistics

	$PR20_{D25N}$ Open	$PR20_{D25N}$ (2 dimers)
Space group	$P4_1$	$P\ 12_11$
Unit cell dimensions: (Å)		
a b c	45.52 45.52 104.14	54.15 48.57 69.68
	α = β = γ =90	α = γ = 90 β = 99.15
Resolution range (Å)	41.71-1.75	50-1.45
Unique reflections	21,444	59,088
R_{merge} (%) overall (final shell)	6.5(41.4)	4.9(31.6)
I/σ(I) overall (final shell)	17.9(4.5)	27.3(4.2)
Completeness (%) overall (final shell)	97.3 (98.6)	93.0 (65.2)
Data range for refinement (Å)	41.71-1.75	45.9-1.45
R (%)	19.3	14.0
R_{free} (%)	23.4	19.2
No. of solvent atoms (total occupancies)	110 (110)	417 (415.5)
RMS deviation from ideality		
Bonds (Å)	0.027	0.026
Angle distance (Degree)	2.125	2.039
Average B-factors ($Å^2$)		
Main-chain atoms	11.7	14.6
Side-chain atoms	15.7	18.6
Solvent	17.2	30.9

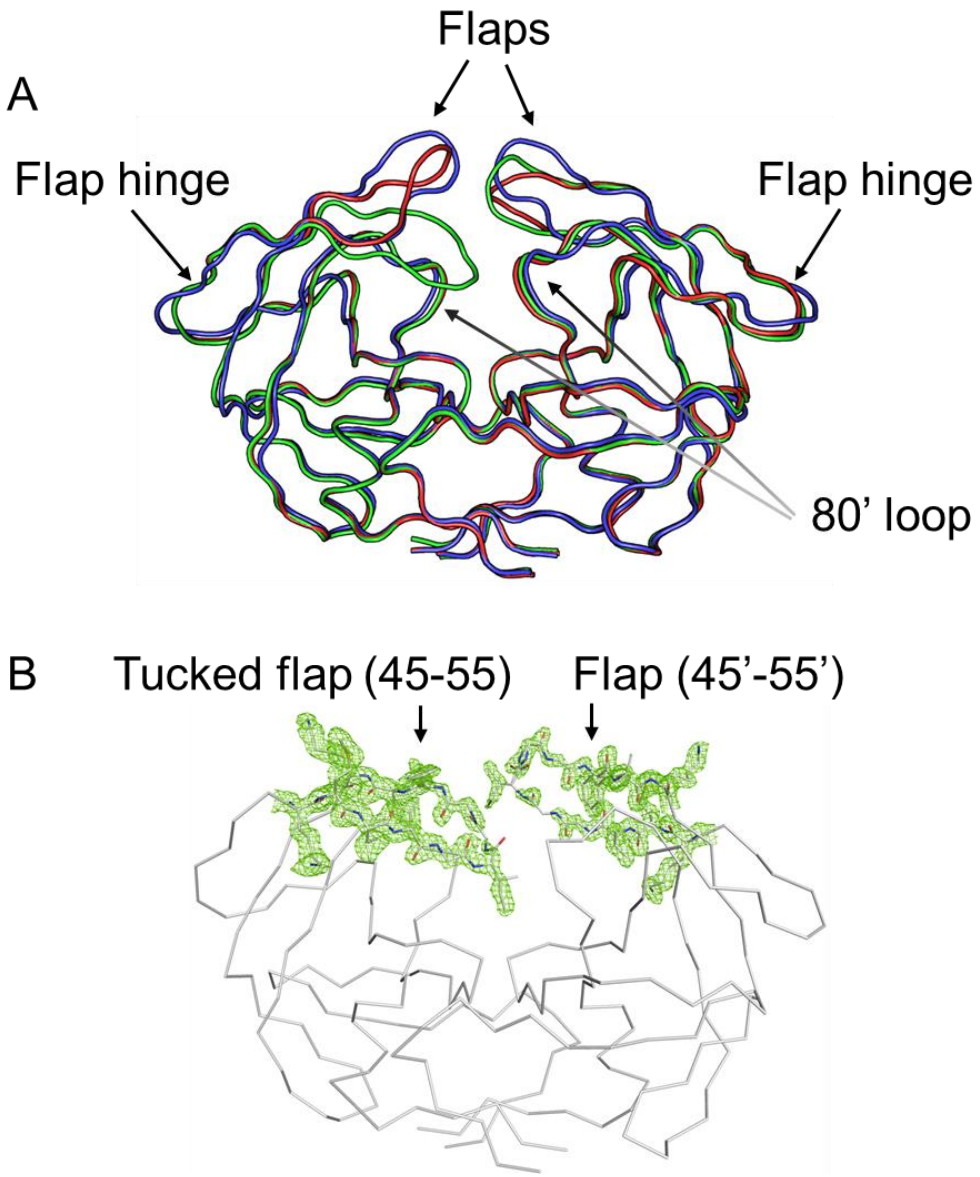


Figure 4-1

(A) Superposition of ligand-free dimers of $PR20_{D25N}$open (blue), $PR20_{D25N}$twist (red) and $PR20_{D25N}$tuck (green) showing various flap conformation. (B) 2 F_o-F_c electron density map contoured at 1 σ.level for the flap region in the crystal structure of $PR20_{D25N}$tuck.

parentheses were calculated without the flap residues.]

The positional disorder in the backbone of each dimer was assessed by plotting the B-factor averaged over the main chain atoms per residue (Figure 4-2). The highest variation among the structures was seen for the flap residues. The largest B-factors of about 40 $Å^2$ occurred in one flap (residues 45'-55') in the $PR20_{D25N}$twist dimer and in both flaps of the $PR20_{D25N}$tuck structure, while the flap residues in the $PR20_{D25N}$open dimer had the maximum B-factor of about 20 $Å^2$. The B factors showed similar peaks in both subunits of the three structures at the four termini in the dimer and the variable surface loops of residues 33–38, 65–70, and 76–83.

4.3.2 Flaps exhibit diverse conformations

Three categories of flap conformations, open, semi-open and closed, have been described for different crystal structures of the apo-forms of PR [25]. PR-inhibitor complexes usually show the closed conformation of the flaps. In the closed conformation, the tips of the two flap interlock over the active site cavity, so that the main chain Cα atoms of Ile50 and Gly51 lie parallel to each other and perpendicular to the catalytic aspartates (Figure 4-3 A). In the semi-open conformation, the tips of the flaps lie parallel to each other but show reversed contacts compared to the closed conformation. In the open conformation, the tips of the flaps are detached from each other and located further from the catalytic aspartates than in the closed conformation. An open conformation of the flap is required for the substrate or inhibitor to enter the active site cavity.

In the $PR20_{D25N}$open structure, the tips of the flaps are separated widely and raised vertically compared to the flaps in the closed conformation (Figure 4-1A). No van der Waals contacts occur between the atoms of the two flaps. The minimum separation of 7.5 Å occurs between the carbonyl oxygen of Ile50 and the Cγ2 atom of Ile50' in the two flaps of $PR20_{D25N}$open. Besides losing contacts between the two flaps, Ile50 also lacks van der Waals interactions with Pro81', which are generally retained in both open and closed conformation dimers[105]. The shortest distances between the side chain atoms of Ile50/Ile50'

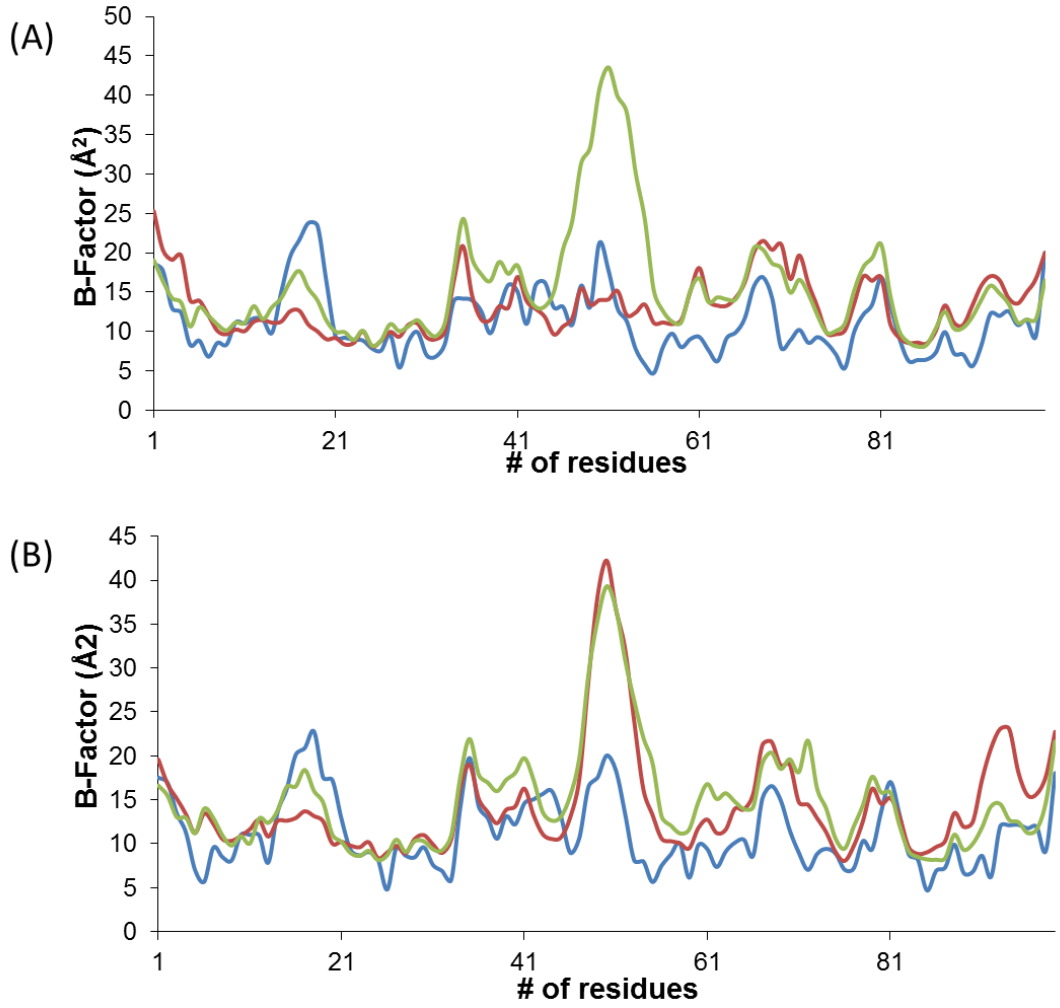


Figure 4-2

B values averaged over the main chain atoms of each residue for $PR20_{D25N}$open (blue), $PR20_{D25N}$twist (red) and $PR20_{D25N}$tuck (green). Upper panel: residues from subunit A. Lower panel: residues from subunit B.

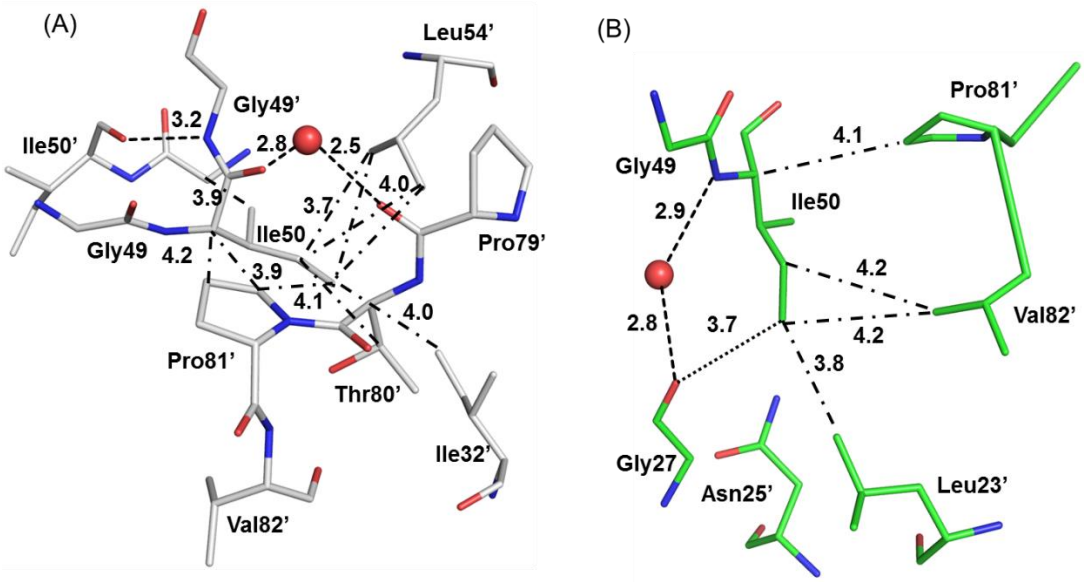


Figure 4-3 Interactions of Ile50 with PR20 residues.

(A) The closed conformation observed in PR20/GRL02031 (4J55). (B) The tucked conformation of $PR20_{D25N}$tuck.

and Pro81/Pro81' are 4.7 Å and 4.0 Å, respectively.

The $PR20_{D25N}$twist dimer has the most widely open flaps reported in any crystal structure to date. In $PR20_{D25N}$twist, the flap tips lie nearly in the same plane but are further from the catalytic Asp25/25', and almost perpendicular to their arrangement in the closed conformation. The shortest distance observed between the two flaps is 9.6 Å between the carbonyl oxygen atoms of Ile50 and Ile50', and the intersubunit separations between Ile50 and Pro81' and Ile50' and Pro81 are 8.5 Å and 5.1 Å, respectively. The wide open flaps of $PR20_{D25N}$open and $PR20_{D25N}$twist show few direct interactions with the residues near Asp25/25' at the base of the active site cavity.

Several water molecules occupy the active site cavity of the ligand-free dimers instead of an inhibitor or substrate. In the closed conformation dimer, the water mediated hydrogen bond interactions connecting the carboxyl oxygen of Pro79'/Pro79 to the carboxyl oxygen of Ile50 and the amide nitrogen of Gly51' have been proposed to stabilize the flaps (Figure 4-3 A) [171]. In the structures of $PR20_{D25N}$open and $PR20_{D25N}$twist, the carbonyl oxygen of Pro79'/Pro79 retains a hydrogen bond with water, however, the structures lack the interactions with Ile50 / Gly51' because of the greater separation of the tips of the flaps.

The unliganded $PR20_{D25N}$open and $PR20_{D25N}$twist structures can be compared to the open conformation of wild-type PR (2PC0) complex with a magnesium ion on top of the catalytic aspartates (Figure 4-4 A) [192]. The dimers of $PR20_{D25N}$open and 2PC0 superimposed with an rmsd value of 0.9 Å for the Cα atoms. The main differences of $PR20_{D25N}$open structure are observed at hinge loop, flap, residues 16 - 17 and 66 - 71 in both subunits. The tips of the flaps of $PR20_{D25N}$open are 1.8 - 2.0 Å higher than the tips of the flaps of 2PC0. Comparison of the $PR20_{D25N}$twist and 2PC0 structures gives an overall rmsd value of 1.3 Å for the Cα atoms. Larger variations are observed at hinge loop, flap, 80` loop, residues 15 - 19 and 66 - 71 for both monomers. The tips of the flaps of $PR20_{D25N}$twist are almost perpendicular to the tips of the flaps of 2PC0.

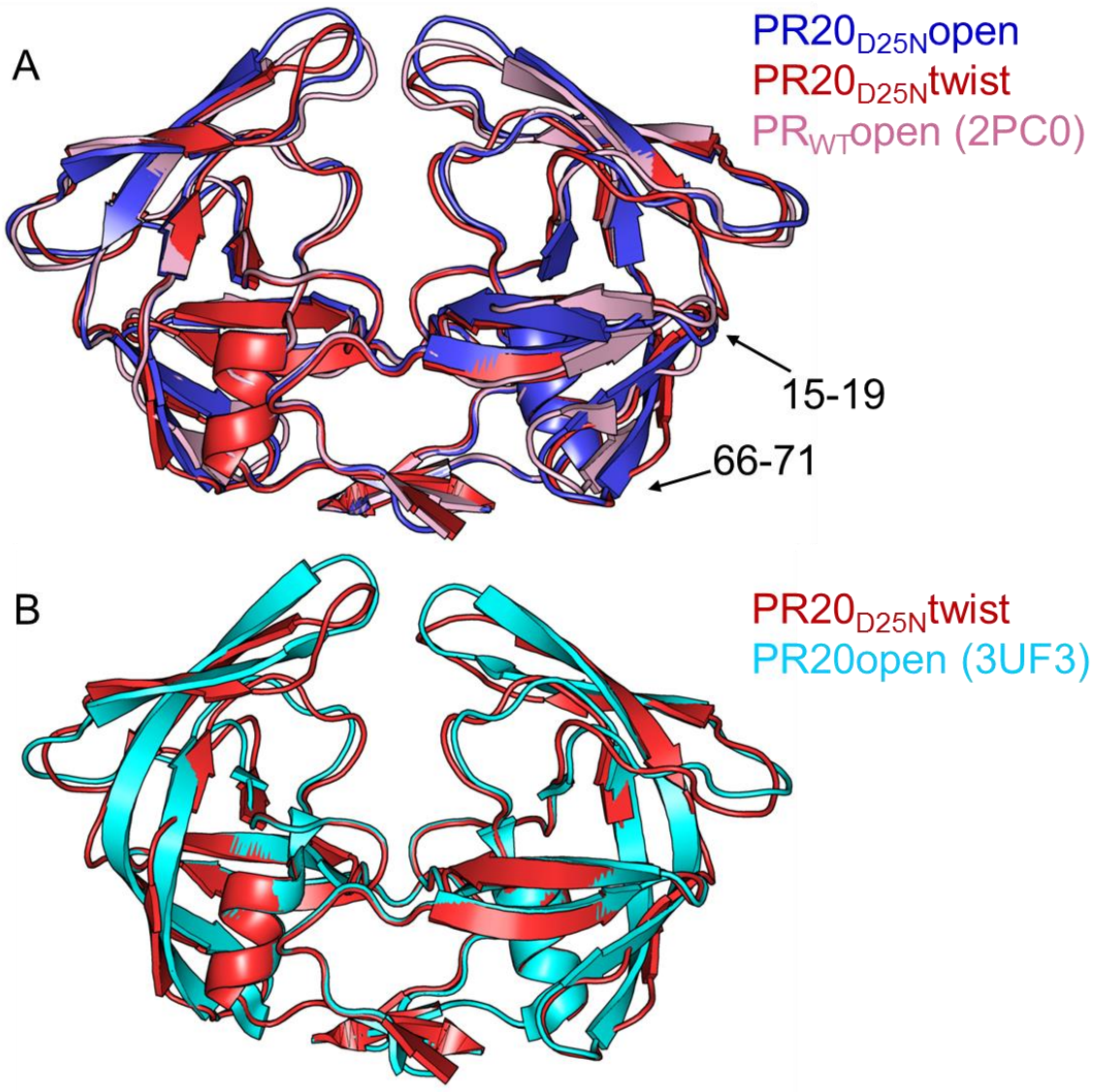


Figure 4-4 Dimer conformations of $PR20_{D25N}$, PR_{WT}open (2PC0) and PR20open(3UF3).

(A) Comparison of $PR20_{D25N}$open (blue), $PR20_{D25N}$twist(red) and PR_{WT}open(pink) shown in cartoon. (B) superposition of $PR20_{D25N}$twist(red) and PR20open (cyan)

The structures of $PR20_{D25N}$open and $PR20_{D25N}$twist are compared to the wide open structure of PR20 (3UF3) with yttrium bound near the catalytic aspartates (Figure 4-4 B) [201]. Although $PR20_{D25N}$open and PR20open dimer structures were solved in two different space groups, no significant difference was apparent in the main chain conformation as indicated by the low rmsd value of 0.3 Å for Cα atoms. The $PR20_{D25N}$twist dimer is more different from the PR20open dimer, as shown by the 1.2 Å RMS value. The major differences occur for residues in the surface loops: hinge loop, 80's loop and flap regions.

4.3.3 Unusual tucked flap conformation

The flap conformation of the $PR20_{D25N}$tuck structure does not belong to the categories defined as closed, semi-open, and open [25]. Instead, flap A of $PR20_{D25N}$tuck exhibits an unusual curled or tucked conformation (Figure 4-1). The tip of the flap tucks or inserts into the active site cavity so that Ile50 lies next to the two catalytic Asp25/25`, while the other flap is directed away from the catalytic residues. In the typical closed conformation, the side chain atoms of Ile50 interact with the side chain atoms of Val32', Val54', Thr80' and Pro81' (Figure 4-3 A). In $PR20_{D25N}$tuck, Ile50 at the tip of the tucked flap A forms completely different interactions with the carbonyl oxygen of Gly27, the side chain atoms of Leu23', Asn25', Pro81', and Val82', while water mediated hydrogen bond interactions connect the amide nitrogen of Ile50 and the carbonyl oxygen of Gly27 (Figure 4-3 B). The typical water-mediated interactions linking the 80's loop residues and the tips of the flaps are missing in the $PR20_{D25N}$tuck structure, consistent with lower stability of the flaps. The lack of stabilizing interactions is consistent with the open conformation exhibited by the flap B, however, the two flap tips show van der Waals contact of 4.2 Å between Cγ1 of Ile50' and the carbonyl oxygen of Gly51.

4.3.4 MD simulations

MD simulations were performed to assess the conformational dynamics of PR20 and wild type PR dimers in the absence of bound inhibitor. Simulations were run on the active enzyme forms with Asp25 and 25' (PR_{20}MD and PR_{WT}MD), and a second set incorporated the active site mutation D25N (PR_{20D25N}MD

and PR_{WTD25N}MD) for comparison with the new crystal structures reported here. The simulations started from the closed conformation, and ended after 10 ns. The trajectories of the RMS deviations from the initial model are shown in Figure 4-5. Simulations equilibrate quickly within the first 500 ps. PR_{WT}MD and PR_{WTD25N}MD equilibrate to similar RMS values of 1.7 ± 0.2 Å and 1.8 ± 0.3 Å, respectively. The PR_{20}MD simulation equilibrated to the highest RMS deviation of 2.4 ± 0.3 Å, while PR_{20D25N}MD gave the lowest RMS value of 1.5 ± 0.1 Å. The scale of the fluctuations is similar to the range described in previous simulations of HIV-1 protease by other groups [212, 213].

4.3.5 Cluster Analysis of simulations

Clustering is a data-mining technique that partitions geometrically closer conformations into disjoint sets. Thus, conformational information from relatively long trajectories is simplified into a small conformational space and the comparison between clusters provides insight on protein flexibility. The 1000 frames from the 10 ns simulations of PR_{WT}MD, PR_{20}MD and PR_{20D25N}MD were partitioned into three clusters, and PR_{WTD25N}MD had four clusters. Similar conclusions for the number of clusters can be obtained from a visual inspection of the dendrogram drawn from hierarchical cluster analysis.

The averaged structures of each cluster were overlaid and displayed in Figure 4-6. Generally, the main chain structures are similar to their corresponding starting models, with the exception of the flaps and other surface loops, which display a larger range of conformations. The features of the tertiary structure of the dimer were preserved over the simulations with good agreement for the elements of regular secondary structure, such as β-strands between the surface loops of residues 9-15, 33-42, and 65-72 and the alpha helix at residues 86-94. The tips of the flaps move toward the catalytic dyad in all the simulations.

The conserved antiparallel β-sheet structure of the flaps and the ordered hinge loop (residues 34−43) are well preserved in all clusters of the simulations for PR_{WT}MD, PR_{WTD25N}MD and PR_{20D25N}MD. In the PR_{20}MD simulation, the β-sheet at the tip of one flap is lost during the simulation, as shown in clusters

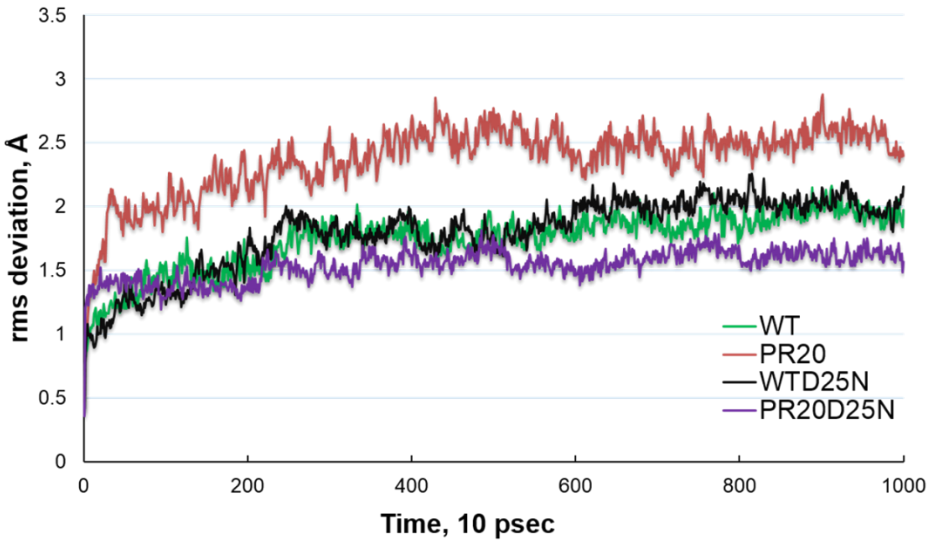


Figure 4-5 Trajectories of the MD simulations.

The time course is plotted for the RMSD calculated by superposing Cα atoms of the dimer at each time point with the corresponding atoms in the starting crystal structure. The simulations are shown for $PR_{WT}MD$ (green), $PR_{20}MD$ (red), inactive $PR_{WTD25N}MD$ (black) and $PR_{20D25N}MD$ (purple).

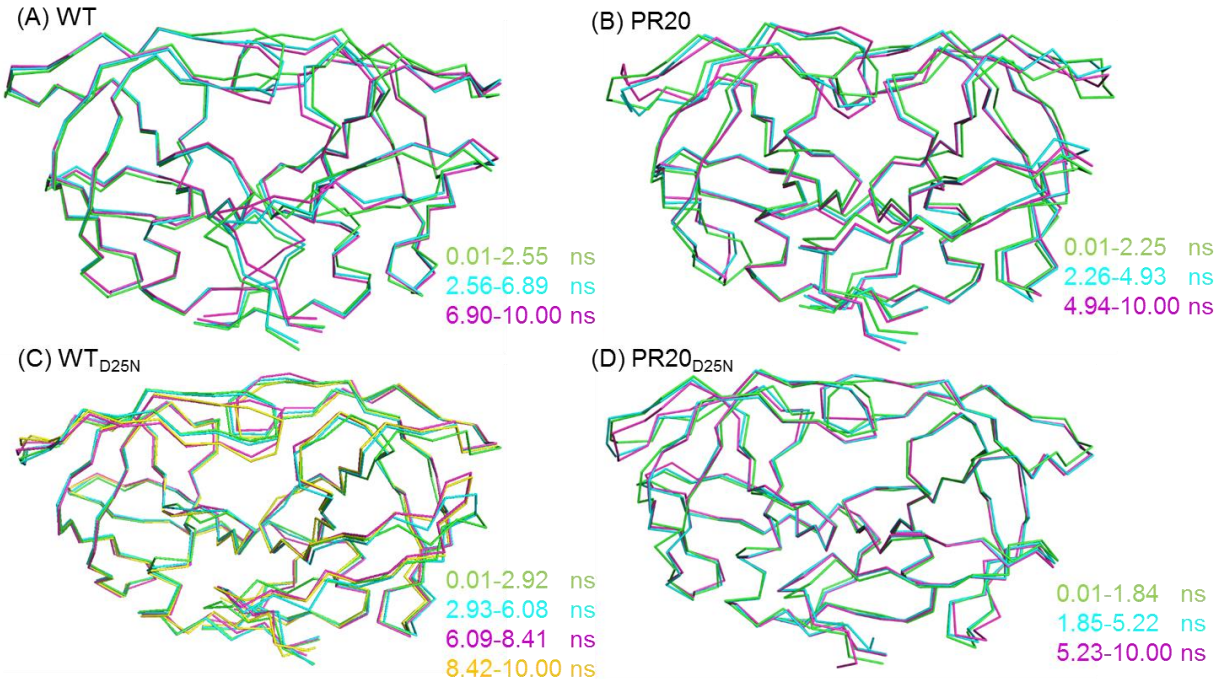


Figure 4-6 Superposition of averaged structures of each cluster calculated for simulations: (A) $PR_{WT}MD$, (B) $PR_{20}MD$, (C) $PR_{WTD25N}MD$ and (D) $PR_{20D25N}MD$.

2 and 3 (Figure 4-6). The hinge loop and residues 14'-17' of PR_{20}MD also display conformational variation. Differences between the clusters occur for the residues at the N- and C- termini of PR_{WT}MD, PR_{WTD25N}MD and PR_{20D25N}MD.

In summary, the wild type PR and PR20 simulations cluster into similar conformational spaces. The tips of the flaps move from the closed conformation toward the catalytic aspartates during our 10 ns simulations. In addition, one flap tip in the dimer of PR_{20}MD shows an extreme conformation, which presumably arises from the 20 mutations relative to the wild type PR.

4.3.6 Correlated motions extracted from MD simulations

Dynamic fluctuations in the protein conformation are important in regulating protein function, and mutations can alter the dynamics [214, 215]. The correlation coefficients of the atomic fluctuations were analyzed for each residue over time in order to identify the correlated motions and the change between wild-type PRMD and PR20MD simulations. Dynamical cross-correlation maps (DCCM) were drawn for each simulation (Figure 4-7). Correlation coefficients of higher than 0.25 or lower than -0.25 are shown in the maps [211], and peaks identified for correlated motions are labeled in the maps. The peaks of correlated motions identified in the MD simulations of wild type PR and PR20 (Figure 4-7 a and Figure 4-7 b) are notably larger than for the corresponding simulations of the inactive D25N mutants (Figure 4-7 c and Figure 4-7 d). Our MD simulations agree with the NMR study showing that the single mutation of D25N altered the dynamic properties of the PR [179]. However, introducing the D25N mutation had no significant effect on the atomic positions or hydrogen bond interactions when comparing the PR_{WT} and PR_{WT}D25N dimers [179].

Correlated motions identified in quadrants I and III of DCCM represent intra-subunit relationships and are consistent with previous MD simulations [216, 217]. The secondary structure patterns and the interface contacts have mainly positive correlations within each monomer, while anti-correlations exist between these regions. The peaks of correlated motions observed in the two monomers are similar but

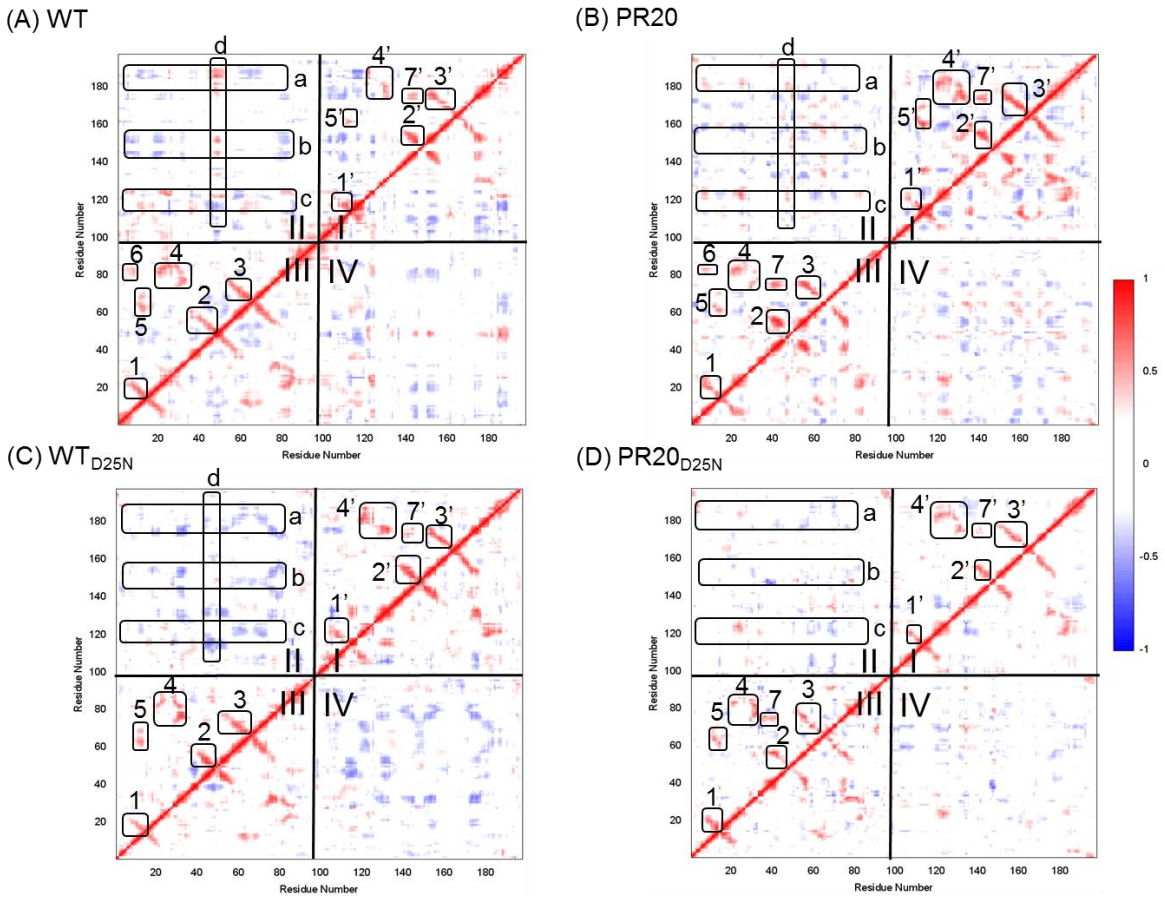


Figure 4-7 Dynamic Cross Correlation Maps show correlated motions of:

(A) $PR_{WT}MD$, (B) $PR_{20}MD$, (C) $PR_{WTD25N}MD$ and (D) $PR_{20D25N}MD$. Cross-correlation coefficients C_{ij} larger than 0.25 and smaller than -0.25 are shown in maps with the intensity represented as follows: red squares $0.25 < C_{ij} < 1$, blue squares $-1 < C_{ij} < -0.25$. The four quadrants are labeled. Quadrants I and II show intra-subunit correlated motions. Quadrant II and the symmetry related quadrant IV show inter-subunit correlated motions. Peaks of positive correlated correlation are labeled with numbers for each monomer (quadrant I and III), and groups of peaks are labeled with letters in quadrant II.

have different intensities, suggesting that each monomer has a similar pattern of dynamic motions. Correlated inter-subunit motions were identified in quadrants II and IV of DCCM. Anti-correlations are mostly seen in the inter-subunit quadrants, showing that motions between two monomers have opposing directions. Correlated motions of the flaps with the active site cavity and peripheral residues have been reported in different MD simulations of wild-type PR [217, 218]. When the DCCMs are compared for the four MD simulations, the correlation coefficients in the inter-subunit quadrant are generally weaker in PR_{20}MD and PR_{20D25N}MD than in PR_{WT}MD and PR_{WTD25N}MD. The anti-correlated motions of the flaps (residues 45-48 and 53-58) in the inter-subunit quadrant are weaker or disappear in PR_{20}MD and PR_{20D25N}MD, respectively. These motions suggest that the two flaps tend to fluctuate more independently in the PR_{20} mutant than in the wild-type enzyme. In summary, correlated motions are well preserved in each monomer, however, the cross communications between the two monomers are impaired in PR_{20} relative to wild type PR dimers.

4.4 Discussion

The variant PR_{20} from a clinical isolate is highly resistant to the tested clinical inhibitors, which show several orders of magnitude worse affinity for PR_{20} compared to the wild type enzyme. Moreover, in contrast to the wild type precursor, saquinavir and darunavir do not inhibit autoprocessing of the precursor comprising TFR-PR_{20}. Previous crystallographic studies of the PR_{20} dimer showed a large variation in the conformation of the flaps in the ligand-free structures. Here, we report two new conformations, designated twisted and tucked flap, observed in the dimer of $PR20_{D25N}$. Formerly, three highly drug resistant variants PR_{P51}, MDR769 and PR20 were observed to have widely separated flaps in their dimer structures, suggesting a common mechanism for resistance to inhibitors [105, 107, 201], and new inhibitors have been designed to target the wide open flaps[219]. Among the new structures reported here, $PR20_{D25N}$twist has the highest separation of the flap tips in the dimer. The rotation of the tips of the flaps in $PR20_{D25N}$twist might further contribute to reduce the affinity for DRV and SQV.

$PR20_{D25N}$tuck exhibits a unique flap conformation, which has not been described previously for crystal structures of HIV protease. The tucked flap enters the active site cavity, and this conformation prevents the binding of substrates or inhibitors. Again, this tucked flap provides a novel mechanism to lower the affinity for inhibitors.

The conformational dynamics of proteins contributes to their biological function, and correlated motions of domains regulate biological function [220, 221]. In the PR dimer, the substrate binding cavity is constructed by two identical monomers, and the cooperative opening and reclosing of the two flaps is critical during proteolysis of the natural substrate and binding of an inhibitor. The flaps exhibit diverse conformations in the crystal structures of ligand-free PR20 [201]. The new crystal structures reported here for ligand-free dimers of $PR20_{D25N}$ show additional conformational variation, including an unusual tucked form, where one flap is tucked inside the active site cavity. Cluster analysis of our MD simulations for PR_{20} and wild type PR shows that the mutations influence the conformational dynamics of the flaps. Isothermal titration calorimetry shows that mutations in PR_{20} increase the stability of the monomer compared to wild type PR [200]. In the MD simulations, the individual monomers in the PR_{20} and wild type enzyme dimers exhibit similar conformational dynamics. However, the correlated motions between the two subunits of the dimer differ in the PR_{20} mutant and wild type enzyme. PR_{20} lacks correlated motions between the flap and the other subunit in the dimer, which will tend to destabilize the flaps, consistent with the diverse conformations observed in crystal structures of this mutant.

In summary, subtle rearrangements in the conformational ensemble of the flaps induced by the mutations impair the dynamics, and consequently the proteolytic activity and inhibitor affinity of the mutant enzyme. These changes in dynamics will contribute to the high level resistance of PR20. Importantly, the discovery of new flap conformations in crystal structures and MD simulations may hint at designs for novel antiviral inhibitors that capture the variant flap conformations of the resistant mutants.

4.5 Acknowledgment

Data were collected at Southeast Regional Collaborative Access Team (SER-CAT) beamline 22ID at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at http://www.ser-cat.org/members.html. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract W-31-109-Eng-38.

5 A HYBRID IMPLEMENTATION OF PARALLEL AMORTIZED FAST MULTIPOLE ALGORITHM FOR THE MOLECULAR MODELING PROGRAM: AMMP

(In-preparation: Chen-Hsiang Shen and Robert W. Harrison)

5.1 Abstract

Molecular dynamics simulation is a tool used to study the molecular basis of chemical and biological problems. The efficient calculation on electrostatic and long range forces, which is the $\theta(N^2)$ summation by naive algorithm, is the central problem on using molecular dynamics simulation. An OpenMP-CUDA implementation of parallel molecular dynamic program, AMMP, is presented for protein simulation. The hybrid implementation provides high efficiency of molecular dynamic simulation. The new parallel implementation of AMMP is capable of simulating molecular systems with more than half million atoms with excellent acceleration and parallel efficiency. The combination of OpenMP-CUDA can accelerate the simulation about 20 fold faster than the 8-threaded CPU based AMMP.

5.2 Introduction

Molecular dynamics simulation is a powerful computational tool for modeling the behavior of proteins and protein-ligand complexes. In the classical limit, molecular dynamics finds a numerical solution for the motion of the molecules in a molecular mechanics potential or force-field. A molecular mechanics potential can be divided in terms of its computational complexity into two sorts of terms: covalent geometry terms that define covalent chemical bonds, angles, torsional effects and chirality and non-bonded terms like van der Waals and electrostatic terms. The number of covalent geometry terms scales asymptotically as O(N) where N is the number of atoms, while the number of non-bonded terms scales as the number of unique pairs of atoms or $O(N^2)$.

The naïve way to calculate non-bonded terms is to loop over each pair of atoms. Therefore significant effort has been spent on accelerating the calculations for the non-bonded terms. One of the first

approximations was to use a cutoff radius, which converts the non-bonded calculations to O(N). However, this approximation violates the conservation of energy and momentum with predictably dire consequences [110, 111, 118]. Recent work has focused developing algorithms that can calculate the non-bonded terms without using an explicit loop over all the pairs of atoms. Two general classes of algorithms are used for this, the Ewald method which uses a Fourier transform[112] and the Fast Multipole Method (FMM) which uses power series expansion and tree based data structure[113, 222]. This paper describes accelerating the implementation of the FMM used in the program AMMP on multi-core and GPU architectures [118, 223].

Even though both the Ewald and FMM methods are effective at speeding up the calculation of the long-range terms, the calculation is still expensive. Further economies are possible by using an Amortization algorithm [118, 223], where a small amount of additional effort translates into significant savings. Amortization is an example of a Multi-Time step algorithm [224]. The force fields used in molecular dynamics are split as fast evolving short-range terms and slow moving long-range terms. Energy terms that evolve quickly in time, namely the covalent geometry terms and non-bonded interactions between nearby atoms are explicitly treated on a fast time scale. Energy terms that evolve more slowly are treated with a slow time scale. The key difference between a classical multi-time step algorithm and an amortized algorithm is that the amortized algorithm maintains a local expansion of the forces and energies which it updates when the atoms have moved sufficiently and the multi-time step expansion simply updates the long-range terms less frequently than the covalent and local non-bonded terms. The advantage of using an update that depends on the distance the atoms have moved is that it automatically adapts the update frequency to the speed of the atoms.

5.3 Specification of Algorithms

The pseudo code for the amortized multi-time step algorithm used in AMMP is shown below.

5.3.1 Algorithm for update:

Inputs: atom's velocities, charges, positions, a list of atoms that are bonded to each other, a radius that differentiates between the local expansion and the non-local expansion, and a parameter that states when the atoms have moved enough to update the non-local expansion.

Initialization: allocate memory for the current position, the partial potential, first and second derivatives of the potential and a list of atoms that will be within the local radius.

For i = 0 to n-1 in atom's vector do

For j = 1 to n in atom vector do

Save the current position.

Calculate the distance r between $atom_{ij}$.

If r > local radius

The power series expansion:

Partial potential += $\frac{kqq}{r} - \frac{aa}{r^6} + \frac{bb}{r^{12}}$, where q,a,b are the charge, Van der Waals attractive and repulsive terms of the energy.

In addition calculate the derivatives of the potential energy around the current atomic positions in a multipole expansion.

$\frac{dV}{dx}$ +=

$$\frac{-kqq}{r^3} * \vec{r_x} + 6 * \frac{7}{2} * \frac{aa}{r^8} * \vec{r_x} - 12 * \frac{13}{2} * \frac{bb}{r^{14}} * \vec{r_x}$$

similar terms for y,z

$\frac{d^2V}{dx^2}$ +=

$$\frac{3kqq}{r5 * \vec{r_x} * \vec{r_x}} - 6 * 7 * \frac{8}{2} * \frac{aa}{r^{10}} * \vec{r_x} * \vec{r_x} - 12 * 13 * \frac{14}{2} * \frac{bb}{r^{16}} * \vec{r_x} * \vec{r_x}$$

similar terms for y^2, z^2, and multiplied by

2 for xy,xz,yz

else

add the atom to the local list.

End if

end for

end for

5.3.2 Algorithm for non-update:

Inputs: The current atomic positions and a user-specified maximum distance before update.

Initialization: the update algorithm has been invoked to build the power series, and the local interaction list.

For i = 0 to n in atom vector **do**

calculate the distance r from the saved position.

If r > specified distance

call the update algorithm for the entire model.

End if

For i = 0 to n in atom vector **do**

Calculate the difference from the saved position and use that to estimate the long range potential or its derivatives for use as a force term.

For J = 0 to m in the list of close atom **do**

calculate local interactions and add in either the potential or force terms.

End for

End for.

5.3.3 Amortized Scalar FMM Algorithm:

The amortized FMM uses the FMM and an adaptive multi-step algorithm to compute the change in the potential at each atom. The FMM divides 3D space into small boxes and uses a polynomial series to approximate long range terms. The terms from the atoms that are inside each box are grouped and treated them as a single source with a multipole expansion. The multipole expansion from the boxes is propagated to both current potential and force terms and the update expansion used for amortization. Thus the system can evolve in time without requiring an additional FMM calculation. This amortizes the computational cost of the FMM over several molecular dynamics steps. Using the amortized FMM algorithm to integrate the non-bonded terms provides compatible results with the standard non-update calculation [223]. The amortized algorithm starts from checking the location of atoms. If an atom has moved more than a pre-defined distance, it will apply update algorithm to calculate a new list of local interactions and the power series. If no atom moves more than a specific distance, the direct sum over a list of local interactions and the power series will be used to calculate the long range interactions. AMMP uses an amortized FMM algorithm integrates non-bonded terms with the asymptotic Θ(N)[223]. Below is the pseudo code for Amortized FMM algorithm:

Inputs: atom's vector, charges, positions, box size, list of atoms contributes in bond terms, a radius that differentiates close atoms.

Required: The maximum and minimum coordinate.

A three-dimensional array of boxes as MMNODEs. Each MMNODE represents a set of atoms that are expanded with a multipole expansion.

An array of MMATOM for a reverse indexed list that will store the relation of each atom with a unique MMNODE.

For i = 0 to n in atom vector **do**

decide which MMNODE $atom_i$ belongs to.

Store that in the MMATOM.

For i = 0 to m in list of MMNODE **do**

For j = 0 to n in atom vector **do**

Select the atoms in MMNODE using MMATOM

Integrate the terms of the rectangular expansion of the $atom_j$ to the $MMNODE_i$.

Except the sign of odd derivatives is different, the expansion used here is the same as the one used in local expansion.

End for

End for

For i = 0 to n in atom vector **do**

$atom_i$ looks up the MMNODE it belongs to and calculates the MMNODEs belong to $atom_i$'s neighbor.

For j = 0 to m in list of MMNODE **do**

If the MMNODEs are not belong to the atom or its neighbors

propagate and sum the distal terms from the series for all these MMNODEs

else

all of the atoms in these MMNODEs are invoked with the update algorithm for the standard amortized algorithm. Thus, the list of close atoms is generated for non-update algorithm.

End if

End for

End for

5.4 Parallelization of AMMP

The non-bonded terms integrated though amortized FMM empirically uses as much as 80% of the runtime cost in sequential version of AMMP, thus the amortized FMM is the main target for parallelization. Preparation of MMNODE data-structures and calculation of pair wise forces are the most expensive parts of amortized FMM algorithm. Previous work on parallelization of amortized FMM is implemented by Parallel Virtual Machine and gain speedup [223]; however, the parallel efficiency was poor because of the high cost of communication. Porting the amortized FMM to shared memory and *GPU accelerated* environments is different approach to improve performance efficient and scalable parallelization.

5.4.1 The OpenMP AMMP

The OpenMP library was used for the first implementation of parallelizing bond and non-bonded terms. This version will be used to compare the performance gain by GPU accelerated computing. Using OpenMP to parallelize short range terms is done quickly though adding directives to guide compiler and it is easy to split loop into different threads to avoid race condition without rewriting sequential code. Threads are created and executed on each core on OpenMP based parallelization. The bond, angle, torsion terms have fixed relation to atoms, thus each thread is assigned a unique start and end point to loop though all the bond, angle and torsion vectors.

Each MMNODE data structure is built independently, thus partitioning the construction of the MMNODE data structures onto individual threads gives reasonable speedup without race conditions or other issues. However, the calculation of pairwise forces between atoms or between atoms and MMNODE requires random access of an array of atom data structures. Therefore it is difficult to avoid race conditions when calculating the pairwise forces with a naive approach. Protecting of the critical region with mutex locks or atomic operations hampers the performance and makes the parallel algorithm more complex and dependent upon individual properties of specific CPU's. Because the AMMP molecular

dynamics program is not memory intensive software, race conditions are avoided by creating duplicated arrays of atom data structures for each thread and these duplicated vectors are summed at the end of amortized FMM function.

5.4.2 The OpenMP-CUDA AMMP

The implementation of Amortized FMM on CUDA is different from the CPU based parallelization. Implementation of GPU accelerated Amortized FMM function starts from OpenMP version because of some light weight loops already parallelized, moving these loops to GPU might be trivial due to frequently moving memory between device and host memory. Only complicated loops such as the preparation of MMNODEs and calculation of pair wise forces, in Amortized FMM function are moved to the GPU. The programming model on GPU has to create many threads to compensate for the high memory latency cost. Thus as many threads are created as there are atoms and these threads are executed by many of CUDA cores concurrently. Because the hardware design of CUDA cores is different to CPU, the atomic function is faster on CUDA cores and is used to protect critical regions.

5.5 Results

5.5.1 Sample preparation for simulations

To perform tests for the overall speed of our implementation as well as the effectiveness of parallelized AMMP, three MD systems were built from PDBs. An ATPase system (PDB ID: 1E79)[225] obtained from X-ray diffraction was prepared for performance test by removing non-protein molecules and filling with 100,000 random water molecules; thus there are 351274 atoms and 51274 atoms belong to protein. Another system with 665197 atoms was prepared from HIV-1 protease structure (PDB ID: 3UHL)[201]. Non-protein atoms were removed from PDB file, 34 dimeric proteases, total 105468 atoms, were prepared based on symmetric molecules, and 186243 water molecules were added randomly. Two

hundred steps were executed, and elapsed times are record for total execution time and run time of amortized FMM function.

To investigate the stability of parallelized AMMP on the structure and dynamics of the protein, drug resistant mutant PR20 (PDB ID: 3UCB) was prepared by removing co-complexed inhibitor DRV[201]. The crystallographic water molecules were kept in simulation. Fifty sodium ion, fifty chloride ions and about nine thousands water molecules were randomly added. Prior to the MD simulation, the random water molecules were equilibrium for 1 ns. Sequential AMMP and OpenMP-CUDA AMMP were used to run the simulation. The RMSD value was calculated from the 5 ns simulation.

5.5.2 Run-time Analyses

Tests were performed on 1. DELL PRECISION T5500 with two Intel® Xeon E5607 CPU, 8 cores available, and an EVGA Geforce GTX 780 graphic card running Ubuntu 10.04. 2. IBM System x3850 X5 with four Intel® Xeon E7-4850 CPU, 40 cores available, running Red Hat Enterprise Linux version. ATPase and HIV PRs were benchmarked with solvated water in two different machines. Parallel efficiency of each core is calculated by the equation: (run time of 1 core / number of cores) / actual run time with multi cores. For example execution times are 20s obtained by 1 core and 5.5s obtained by 4 cores, and the parallel efficiency is 90%.

The run time analyses of molecular dynamics done by OpenMP-AMMP are presented in Figure 5-1 and Figure 5-2. Basically, the elapsed times on simulations of ATPase and HIV PR decrease while the number of cores included in calculation. The implementation of OpenMP-AMMP produces good speedups on DELL PRECISION T5500; up to 8 cores, there is no significant performance drop observed in both simulations. The parallel efficiency maintains higher than 95% in DELL system. Similar results observed in IBM System x3850 X5 when using 8 cores for parallel computation; however, as the figure showed, the parallel efficiencies drop to below 80% in the parallel computation with 40 cores. This might be caused by the

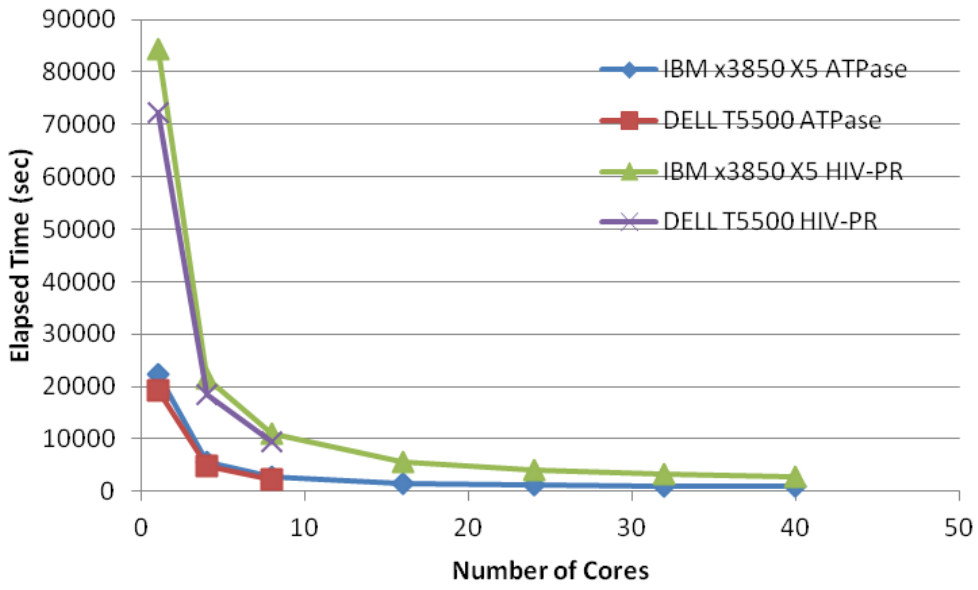


Figure 5-1 The elapsed time of HIV protease and ATPase simulated on DELL workstation and IBM x3850.

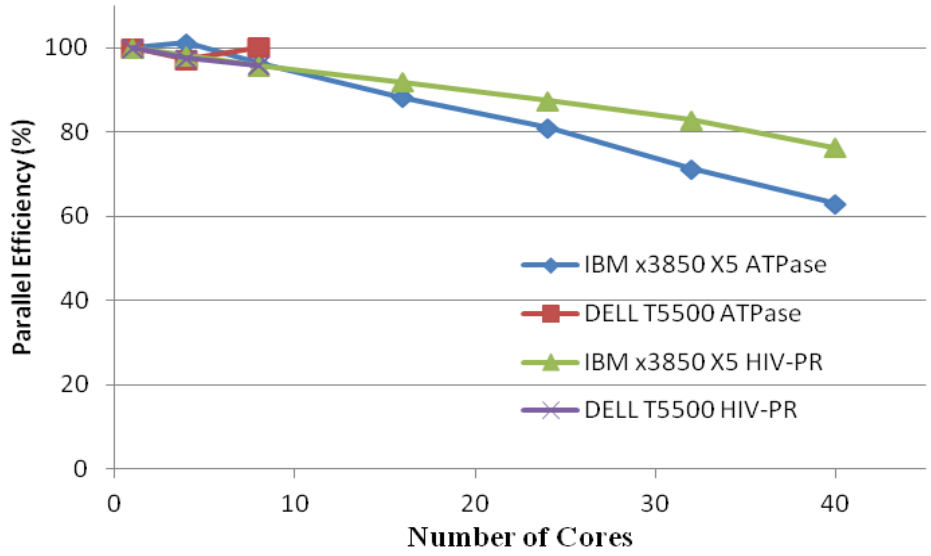


Figure 5-2 The parallel efficiencies are calculated from the simulations with various numbers of cores included in computation.

limitation on memory bandwidth. Clearly, the implementation of OpenMP-AMMP obtains reasonable improvement on performance

The OpenMP-CUDA-AMMP is tested on DELL T5500, the elapsed times from OpenMP-CUDA-AMMP and OpenMP-AMMP are compared in Table 5-1. The elapsed time form hybrid implementation of AMMP reduced 20 folds than OpenMP-AMMP. Taken together, passing computation of Amortized FMM to CUDA cores significantly improves performance.

5.5.3 Stability of simulation by parallelized AMMP

The most direct way to estimate the stability of the simulation is calculating the root mean square deviation (RMSD) that generated by superposition the dynamic trajectory to the starting crystallographic structure. The RMSD on the main chain atoms from the initial model are shown in Figure 5-3. The average RMSD for sequential AMMP and OpenMP-CUDA-AMMP are 2.4 ± 0.3 Å and 2.3 ± 0.3 Å, respectively. The scale of the fluctuations is similar to the range described in other simulations of HIV-1 protease [212, 213]. It is clear that the newly hybrid implementation of AMMP displays comparable stability to the sequential AMMP.

5.6 Discussion

Currently, the hybrid implementation of AMMP by OpenMP and CUDA provides compatible results without cut-off and more than 10 fold faster than sequential version of AMMP. Two possible improvement of AMMP are: including more functions in GPU and designing new parallel algorithm to split calculation to several nodes. Based on current architectures, the GPU significantly can do more floating point operations (FLOPs) per unit time than CPU. More speedups can be expected and reached while there are more routine functions implemented on GPU. Now, only amortized FMM function ports to GPU, more force field functions implemented on GPU might further improve performance. Currently, AMMP executes in single node. The memory bandwidth is saturated when included more cores in computation,

Table 5-1 Run time analysis of ATPase and HIV protease of two different implementations.

Box(Å)		OpenMP-AMMPa (Sec)	CUDA-AMMPb (Sec)	Speed-up (Fold Change)
10	ATPase	1052.4	53.2	19.8
	HIV PR	3951.5	195.1	20.3
9	ATPase	1264.2	63.9	19.8
	HIV PR	5021.9	229.8	21.8
8	ATPase	1636.9	79.7	20.5
	HIV PR	6362.4	289.2	22.0
7	ATPase	2324.8	118.6	19.6
	HIV PR	9793.4	473.9	20.7
6	ATPase	3881.4	189.0	20.5
	HIV PR	13718.0	701.3	19.6

[a] Simulation using 8 cores of CPU

[b] Simulation running on 8 cores of CPU and one GPU

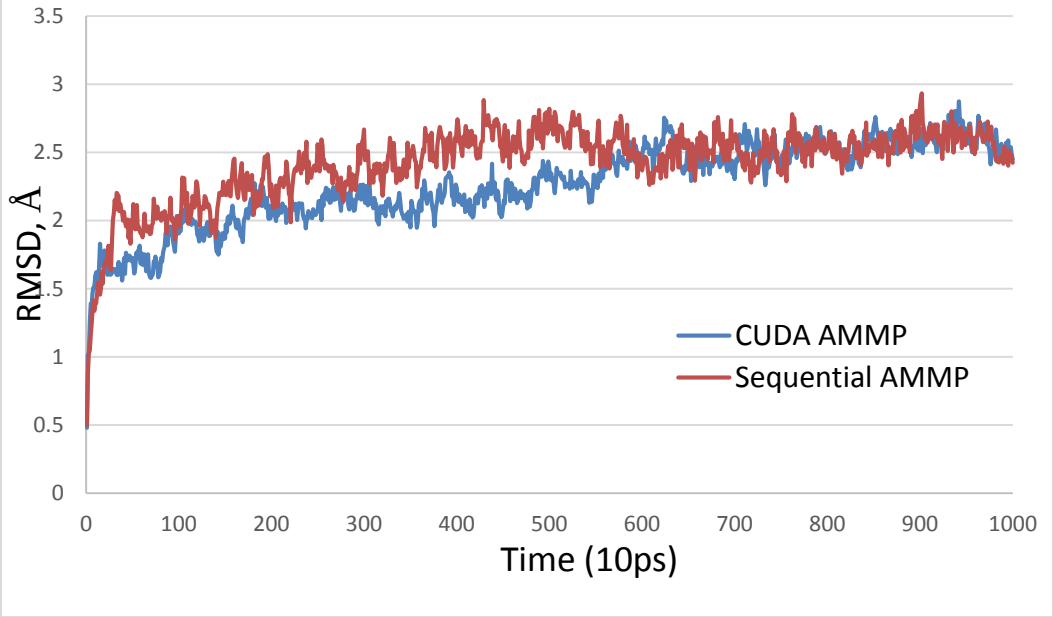


Figure 5-3 Time course plot of the RMSD draws by the superposition main chain atoms of the whole protein to the initial crystallographic structure.

which lower parallel efficiency. The parallel algorithm using distributed memory system, such as spatial decomposition algorithm that partition atoms to different nodes, might improve performance.

5.7 Acknowledgment

This work is partially supported by NIH U01-GM062920. Chen-Hsiang Shen was supported by a Molecular Basis of Disease Fellowship from Georgia State University. IBM System x3850 X5 was supported by research computing at Georgia State University.

6 OVERALL SUMMARY AND DISCUSSION

HIV-1 PR is an indispensable enzyme in the viral life cycle, which makes it an effective target in anti-retroviral therapy. Since 1995, mortality of AIDS patients has been greatly reduced due to using PIs in anti-retroviral therapy. However, the challenge to the long-term efficacy of PIs is the rapid emergence of drug-resistant PR variants. The extensive studies of HIV-1 PR and resistant variants for their structures and catalytic activities will help better understand the molecular basis of drug resistance; thus, various techniques, X-ray crystallography, enzyme kinetic assay or molecular dynamic simulation, were used in one or more of the following structural studies: I. PR_{WT} and drug resistant mutants, PR_{V32I}, PR_{I50V}, PR_{I54V}, PR_{I54M}, PR_{I84V}, and PR_{L90M} complexed with APV. II. PR_{WT} and drug resistant mutants, PR_{V32I} and PR_{I47V}. III. highly resistant $PR20_{D25N}$.

The flap dynamics, stability of PR and the binding of inhibitors are potentially altered by mutations within the hydrophobic cluster. In first study, wild type PR and PR variants (PR_{V32I}, PR_{I50V}, PR_{I54V}, PR_{I54M}, PR_{I84V}, and PR_{L90M}) bearing single mutations were investigated by crystallographic and kinetic analyses. The mutants PR_{V32I}, PR_{I54M} and PR_{L90M} showed substitution of a larger side chain that introduced new hydrophobic interactions with flap residues, Pro79 and Thr80 on 80's loop and the catalytic Asp25, respectively. PR_{I50V} and PR_{I54V} have the substitution of a smaller side chain that decreased internal hydrophobic contacts. The worse inhibition by APV of 10-, and 30- fold~~s~~ is consistent with the observed structural alternations of PR_{V32I} and PR_{I50V}, respectively. Structural comparisons between PR mutants complexed with APV or SQV indicate that the dimeric structure and activity of PR can be maintained by rearrangement of side chains in the hydrophobic cluster. Furthermore, rearrangements of the PR residues are observed on flaps and 80's loop to accommodate two different shapes of PIs. Thus, the different drug resistant patterns of resistant mutations for the two PIs are produced by the modification in the loops and in the hydrophobic cores of PR.

Three distinct reaction intermediates were captured in crystals of wild-type PR and PR mutants with substitutions of I47V and V32I. Two tetrahedral intermediates aligned in opposite orientations were captured in the PR_{WT} structure. The metastable reaction intermediate, gem-diol, was observed in substrate YDQIIxIEIA where x indicates as the cleavage site between P1 Ile and P1' Ile residues in PR sequence. The N- and C-terminal products of autoproteolytic substrate co-exist in the substrate binding cavity of PR_{I47V} structure. The scissile bond between amide and carboxyl carbon has lengthened to 2.2Å indicating the incomplete cleavage of peptide bond before the release of products, which represents the stage immediately after hydrolysis of the peptide bond. The C-terminal product occupies S1-S5 subsites of PR_{V32I} structure. Two water molecules were observed in S2' subsite and had hydrogen bond interactions with D29' and D30'. In addition, one water molecule was refined with partial occupancy in catalytic site. It forms hydrogen bond interactions between the carboxylate group of C- terminal product and carboxyl oxygen of Gly27', which illustrates the possible role of Gly27 in recycling the enzyme. A short hydrogen bond interaction, with distance between 2.3 and 2.4 Å, was observed in three crystal structures connecting the outer oxygen of aspartate group of Asp25' and reaction intermediates. Overall, the reaction intermediates observed in the crystal structures represent consecutive steps in the catalytic reaction.

Generally, multiple drug resistant mutations accumulate in clinically isolated PR variants. It is more difficult directly to apply the drug resistant mechanisms observed from PR with a single mutation to multiple mutations. Clinically isolated highly resistant mutant, PR_{20}, harboring 20 mutations significantly reduces binding affinity to clinical inhibitors. Two ligand-free crystal structures of the PR20 mutant with the inactivating D25N mutation, $PR20_{D25N}$, revealed various flap conformations. Different to the majority of PR structures, which show a closed flap in the presence of ligand, various conformations of flaps were observed in structures of $PR20_{D25N}$ and categorized as open, tucked and twisted conformations. The open conformation, $PR20_{D25N}$open, has been reported in previous studies on wild type PR and MDR769 [102,

105, 192]. The unusual tucked conformation, $PR20_{D25N}$tuck, shows that the Ile50 at the tip of the flap inserted into substrate binding cavity and stabilized by hydrophobic interactions with side chain of Val32', Val54', Thr80' and Pro81'. For the twisted conformation, $PR20_{D25N}$twist, the tips of the flaps were raised from the closed conformation and the tips of the flaps were rotated to lie nearly in the same plane. Thus, the tips of the flaps were almost perpendicular to the arrangement in the closed conformation. The MD simulations show that mutations in PR20 alter in the dimer conformations of flap and affect correlated interactions between two monomers in the dimer. In conclusion, combining the results of structural analysis and MD simulations, unusual flap conformations lacking correlated inter-subunit motions may contribute to the high resistance of PR20.

HIV-1 PR is one of the pharmacologically important targets in AIDS therapy. One challenge in long term AIDS therapy is drug resistance. The high mutation rate of the virus means that new patterns of mutation lead to susceptibility to the PIs. Mutations selected by PI treatment are observed in PR sequence as well as in the protease cleavage sites. The relation of specific mutations in the PR sequence and protease cleavage sites might confer decreased susceptibility to the PIs and maintain biological function. Thus, continuous efforts on understanding the molecular mechanisms that responsible for clinically isolated multi drug resistance and the coevolution of protease and its substrate are necessary for developing new generation PIs.

The flaps are one of the critical elements forming the substrate binding cavity, where inhibitors and substrates compete for the same site. Structural analyses of multi drug resistant mutants indicate that various conformations of the flaps, causing loss of interactions with PIs, are associated with the low binding affinity of inhibitors. However, the structures of substrates are more flexible than the rigid inhibitors to adapt to the various conformations of flaps caused by the mutations. Targeting the flexible flap region may be not enough to combat drug resistance. An allosteric site in the hinge loop has been proposed as another drug target site for regulating protease activity [226]. The hinge loop is the region

preceding the flap. Molecular dynamics simulations show that the distance between the hinge loop and residues 15-17 is correlated with the opening of the flaps [212]. Introducing a new disulfide bond between residues G16 and L38 in wild-type PR and using monoclonal antibody to bind residues 36-46 significantly lowered proteolytic activities, which suggests that the activities of flexible flap can be regulated by the hinge loop [227, 228]. APV, DRV and small molecules have been observed in the binding site formed by the hinge residues Trp42, Pro44, Met46 Lys55, Val56 and Arg57 [69, 229-231]. In sum, the two solvent-accessible hinge loops might be a promising drug target for designing a non-competitive inhibitor, and studies of the structural alternations of hinge loop will improve the understanding of the molecular basis of flap opening /closing.

Overall, our crystallographic studies increase the fundamental knowledge of PR catalytic mechanism and the effects of drug resistant mutations on PR structure and inhibition. These will direct us to a better understanding of drug-resistant variants and benefit the design of new inhibitors to combat drug resistance to AIDS.

In addition to crystallographic study, a computational study was done on improving the efficiency of molecular dynamics program, AMMP. Two different parallelization models were implemented in AMMP, which are OpenMP-AMMP and hybrid OpenMP-CUDA-AMMP. Currently, the force fields that are used in MD simulation are parallelized in OpenMP-AMMP. The performance of OpenMP-AMMP was assessed in machine with 40 cores and improved as expected. Also, the OpenMP-AMMP provides the foundation for using Intel® Many Integrated Core (Intel® MIC) Architecture, which uses OpenMP library to improve efficiency. In OpenMP-CUDA-AMMP, the Amortized Fast Multipole Algorithm that integrates long range terms is parallelized by using CPU and CUDA libraries, which gives about 20X speed-up compared to OpenMP-AMMP in our test simulations. Overall, the development of OpenMP-CUDA-AMMP would be very useful for researches and applications using all atom MD simulations.

7 REFERENCES

1. Gallo, R.C., et al., *Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS.* Science, 1984. **224**(4648): p. 500-3.
2. UNAIDS, *UNAIDS report on the global AIDS epidemic 2013.* 2013: p. 1-152.
3. Gilbert, P.B., et al., *Comparison of HIV-1 and HIV-2 infectivity from a prospective cohort study in Senegal.* Stat Med, 2003. **22**(4): p. 573-93.
4. Maddon, P.J., et al., *The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain.* Cell, 1986. **47**(3): p. 333-48.
5. Sattentau, Q.J. and J.P. Moore, *Conformational changes induced in the human immunodeficiency virus envelope glycoprotein by soluble CD4 binding.* J Exp Med, 1991. **174**(2): p. 407-15.
6. Goff, S.P., *Retroviral reverse transcriptase: synthesis, structure, and function.* J Acquir Immune Defic Syndr, 1990. **3**(8): p. 817-31.
7. Miller, M.D., C.M. Farnet, and F.D. Bushman, *Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition.* J Virol, 1997. **71**(7): p. 5382-90.
8. Nie, Z., et al., *The putative alpha helix 2 of human immunodeficiency virus type 1 Vpr contains a determinant which is responsible for the nuclear translocation of proviral DNA in growth-arrested cells.* J Virol, 1998. **72**(5): p. 4104-15.
9. Cassan, M., et al., *Translational frameshifting at the gag-pol junction of human immunodeficiency virus type 1 is not increased in infected T-lymphoid cells.* J Virol, 1994. **68**(3): p. 1501-8.
10. Klimas, N., A.O. Koneru, and M.A. Fletcher, *Overview of HIV.* Psychosom Med, 2008. **70**(5): p. 523-30.
11. Louis, J.M., G.M. Clore, and A.M. Gronenborn, *Autoprocessing of HIV-1 protease is tightly coupled to protein folding.* Nat. Struct. Biol., 1999. **6**: p. 868-875.
12. Louis, J.M., et al., *HIV-1 protease: maturation, enzyme specificity, and drug resistance.* Adv. Pharmacol., 2000. **49**: p. 111-146.
13. Erickson-Viitanen, S., et al., *Cleavage of HIV-1 gag polyprotein synthesized in vitro: sequential cleavage by the viral protease.* AIDS Res Hum Retroviruses, 1989. **5**(6): p. 577-91.
14. Gottlinger, H.G., J.G. Sodroski, and W.A. Haseltine, *Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1.* Proc Natl Acad Sci U S A, 1989. **86**(15): p. 5781-5.
15. Louis, J.M., et al., *HIV-1 Protease: Structure, Dynamics, and Inhibition.* Adv Pharmacol, 2007. **55**: p. 261-98.
16. Frankel, A.D. and J.A. Young, *HIV-1: fifteen proteins and an RNA.* Annu Rev Biochem, 1998. **67**: p. 1-25.
17. Turner, B.G. and M.F. Summers, *Structural biology of HIV.* J Mol Biol, 1999. **285**(1): p. 1-32.
18. Darke, P.L., et al., *HIV-1 protease specificity of peptide cleavage is sufficient for processing of gag and pol polyproteins.* Biochem Biophys Res Commun, 1988. **156**(1): p. 297-303.
19. Gustchina, A. and I.T. Weber, *Comparison of inhibitor binding in HIV-1 protease and in non-viral aspartic proteases: the role of the flap.* FEBS Lett, 1990. **269**(1): p. 269-72.
20. Weber, I. and J. Agniswamy, *HIV-1 Protease: Structural Perspectives on Drug Resistance.* Viruses, 2009. **1**(3): p. 1110-1136.
21. Tie, Y., et al., *Molecular basis for substrate recognition and drug resistance from 1.1 to 1.6 angstroms resolution crystal structures of HIV-1 protease mutants with substrate analogs.* Febs J, 2005. **272**(20): p. 5265-77.

22. Wlodawer, A. and J. Vondrasek, *Inhibitors of HIV-1 protease: a major success of structure-assisted drug design.* Annu Rev Biophys Biomol Struct, 1998. **27**: p. 249-84.
23. Shao, W., et al., *Sequence requirements of the HIV-1 protease flap region determined by saturation mutagenesis and kinetic analysis of flap mutants.* Proc Natl Acad Sci U S A, 1997. **94**(6): p. 2243-8.
24. Kear, J.L., et al., *Subtype polymorphisms among HIV-1 protease variants confer altered flap conformations and flexibility.* J Am Chem Soc, 2009. **131**(41): p. 14650-1.
25. Robbins, A.H., et al., *Structure of the unbound form of HIV-1 subtype A protease: comparison with unbound forms of proteases from other HIV subtypes.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 3): p. 233-42.
26. Krausslich, H.G., *Human immunodeficiency virus proteinase dimer as component of the viral polyprotein prevents particle assembly and viral infectivity.* Proc Natl Acad Sci U S A, 1991. **88**(8): p. 3213-7.
27. Wlodawer, A. and A. Gustchina, *Structural and biochemical studies of retroviral proteases.* Biochim Biophys Acta, 2000. **1477**(1-2): p. 16-34.
28. Weber, I.T., *Comparison of the crystal structures and intersubunit interactions of human immunodeficiency and Rous sarcoma virus proteases.* J Biol Chem, 1990. **265**(18): p. 10492-6.
29. Todd, M.J., N. Semo, and E. Freire, *The structural stability of the HIV-1 protease.* J Mol Biol, 1998. **283**(2): p. 475-88.
30. Naicker, P., et al., *F99 is critical for dimerization and activation of South African HIV-1 subtype C protease.* Protein J, 2013. **32**(7): p. 560-7.
31. Johnson, V.A., et al., *Update of the drug resistance mutations in HIV-1: March 2013.* Top Antivir Med, 2013. **21**(1): p. 6-14.
32. Ceccherini-Silberstein, F., et al., *Identification of the minimal conserved structure of HIV-1 protease in the presence and absence of drug pressure.* AIDS, 2004. **18**(12): p. F11-9.
33. Schramm, H.J., et al., *Lipopeptides as dimerization inhibitors of HIV-1 protease.* Biol Chem, 1999. **380**(5): p. 593-6.
34. Bannwarth, L., et al., *Dimer disruption and monomer sequestration by alkyl tripeptides are successful strategies for inhibiting wild-type and multidrug-resistant mutated HIV-1 proteases.* Biochemistry, 2009. **48**(2): p. 379-87.
35. FDA. *HIV and AIDS Activities - Antiretroviral drugs used in the treatment of HIV infection*. [WebContent] 2014; Available from: http://www.fda.gov/ForConsumers/byAudience/ForPatientAdvocates/HIVandAIDSActivities/ucm118915.htm.
36. Hansen, J., et al., *Partial purification and substrate analysis of bacterially expressed HIV protease by means of monoclonal antibody.* EMBO J, 1988. **7**(6): p. 1785-91.
37. Seelmeier, S., et al., *Human immunodeficiency virus has an aspartic-type protease that can be inhibited by pepstatin A.* Proc Natl Acad Sci U S A, 1988. **85**(18): p. 6612-6.
38. Kohl, N.E., et al., *Active human immunodeficiency virus protease is required for viral infectivity.* Proc. Natl. Acad. Sci. U.S.A., 1988. **85**(13): p. 4686-4690.
39. Hyland, L.J., et al., *Human immunodeficiency virus-1 protease. 1. Initial velocity studies and kinetic characterization of reaction intermediates by 18O isotope exchange.* Biochemistry, 1991. **30**(34): p. 8441-53.
40. Smith, R., et al., *Ionization states of the catalytic residues in HIV-1 protease.* Nat Struct Biol, 1996. **3**(11): p. 946-50.
41. Hyland, L.J., T.A. Tomaszek, Jr., and T.D. Meek, *Human immunodeficiency virus-1 protease. 2. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism.* Biochemistry, 1991. **30**(34): p. 8454-63.

42. Adachi, M., et al., *Structure of HIV-1 protease in complex with potent inhibitor KNI-272 determined by high-resolution X-ray and neutron crystallography.* Proc Natl Acad Sci U S A, 2009. **106**(12): p. 4641-6.
43. Weber, I.T., et al., *Joint X-ray/Neutron Crystallographic Study of HIV-1 Protease with Clinical Inhibitor Amprenavir: Insights for Drug Design.* J Med Chem, 2013. **56**(13): p. 5631-5.
44. Suguna, K., et al., *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, 1987. **84**(20): p. 7009-13.
45. Brik, A. and C.H. Wong, *HIV-1 protease: mechanism and drug discovery.* Org Biomol Chem, 2003. **1**(1): p. 5-14.
46. Oroszlan, S. and R.B. Luftig, *Retroviral proteinases.* Curr Top Microbiol Immunol, 1990. **157**: p. 153-85.
47. Wlodawer, A. and J. Vondrasek, *Inhibitors of HIV-1 protease: a major success of structure-assisted drug design.* Annu. Rev. Biophys. Biomol. Struct., 1998. **27**: p. 249-284.
48. Wensing, A.M., N.M. van Maarseveen, and M. Nijhuis, *Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance.* Antiviral Res, 2010. **85**(1): p. 59-74.
49. Supuran, C.T., et al., *Antiviral sulfonamide derivatives.* Mini Rev Med Chem, 2004. **4**(2): p. 189-200.
50. Miller, M., et al., *Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 A resolution.* Science, 1989. **246**(4934): p. 1149-1152.
51. Wlodawer, A., et al., *Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease.* Science, 1989. **245**(4918): p. 616-21.
52. Weber, I.T.K., Andrey Y.; Harrison, Robert W., *Frontiers in Drug Design & Discovery*. Structures of HIV Protease Guide Inhibitor Design to Overcome Drug Resistance, ed. I.C. Atta-ur-Rahman, M.; Caldwell, Gary W.; Player, Mark R. Vol. 3. 2007: Bentham Science Publishers. p 45-62.
53. Appelt, K., et al., *Design of enzyme inhibitors using iterative protein crystallographic analysis.* J Med Chem, 1991. **34**(7): p. 1925-34.
54. Gustchina, A. and I.T. Weber, *Comparison of inhibitor binding in HIV-1 protease and in non-viral aspartic proteases: the role of the flap.* FEBS Lett., 1990. **269**(1): p. 269-272.
55. Tie, Y., et al., *Molecular basis for substrate recognition and drug resistance from 1.1 to 1.6 A resolution crystal structures of HIV-1 protease mutants with substrate analogs.* FEBS J., 2005. **272**(20): p. 5265-5277.
56. Tie, Y., et al., *Atomic resolution crystal structures of HIV-1 protease and mutants V82A and I84V with saquinavir.* Proteins, 2007. **67**(1): p. 232-242.
57. Krohn, A., et al., *Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (R)-hydroxyethylamine isostere.* J. Med. Chem., 1991. **34**(11): p. 3340-3342.
58. Kozisek, M., et al., *Molecular analysis of the HIV-1 resistance development: enzymatic activities, crystal structures, and thermodynamics of nelfinavir-resistant HIV protease mutants.* J Mol Biol, 2007. **374**(4): p. 1005-16.
59. Liu, Z., et al., *Insights into the mechanism of drug resistance: X-ray structure analysis of multi-drug resistant HIV-1 protease ritonavir complex.* Biochem Biophys Res Commun, 2013. **431**(2): p. 232-8.
60. Mahalingam, B., et al., *Crystal structures of HIV protease V82A and L90M mutants reveal changes in the indinavir-binding site.* Eur J Biochem, 2004. **271**(8): p. 1516-24.
61. Ghosh, A.K., et al., *Design of HIV protease inhibitors targeting protein backbone: an effective strategy for combating drug resistance.* Acc Chem Res, 2008. **41**(1): p. 78-86.
62. Foulkes, J.E., et al., *Role of invariant Thr80 in human immunodeficiency virus type 1 protease structure, function, and viral infectivity.* J Virol, 2006. **80**(14): p. 6906-16.

63. Ogden, R.C. and C.W. Flexner, *Protease inhibitors in AIDS therapy*. Infectious disease and therapy. 2001, New York: Marcel Dekker. xi, 310 p.
64. Ghosh, A.K., et al., *Enhancing protein backbone binding--a fruitful concept for combating drug-resistant HIV.* Angew Chem Int Ed Engl, 2012. **51**(8): p. 1778-802.
65. King, N.M., et al., *Structural and thermodynamic basis for the binding of TMC114, a next-generation human immunodeficiency virus type 1 protease inhibitor.* J Virol, 2004. **78**(21): p. 12012-21.
66. MacArthur, R.D., *Darunavir: promising initial results.* Lancet, 2007. **369**(9568): p. 1143-4.
67. James, J.S., *Prezista (darunavir, TMC-114) approved; may be important treatment advance.* AIDS Treat News, 2006(418): p. 2-3.
68. Tie, Y., et al., *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, 2004. **338**(2): p. 341-52.
69. Kovalevsky, A.Y., et al., *Ultra-high resolution crystal structure of HIV-1 protease mutant reveals two binding sites for clinical inhibitor TMC114.* J Mol Biol, 2006. **363**(1): p. 161-73.
70. Liu, F., et al., *Effect of flap mutations on structure of HIV-1 protease and inhibition by saquinavir and darunavir.* J Mol Biol, 2008. **381**(1): p. 102-15.
71. Condra, J.H., et al., *In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors.* Nature, 1995. **374**(6522): p. 569-71.
72. Najera, I., et al., *Pol gene quasispecies of human immunodeficiency virus: mutations associated with drug resistance in virus from patients undergoing no drug therapy.* J Virol, 1995. **69**(1): p. 23-31.
73. Kozal, M.J., et al., *Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays.* Nat Med, 1996. **2**(7): p. 753-9.
74. Velazquez-Campoy, A., et al., *Catalytic efficiency and vitality of HIV-1 proteases from African viral subtypes.* Proc Natl Acad Sci U S A, 2001. **98**(11): p. 6062-7.
75. Robertson, D.L., et al., *Recombination in HIV-1.* Nature, 1995. **374**(6518): p. 124-6.
76. Kuwata, T., et al., *The rapid spread of recombinants during a natural in vitro infection with two human immunodeficiency virus type 1 strains.* J Virol, 1997. **71**(9): p. 7088-91.
77. Zhang, Y.M., et al., *Drug resistance during indinavir therapy is caused by mutations in the protease gene and in its Gag substrate cleavage sites.* J. Virol., 1997. **71**(9): p. 6662-6670.
78. Cote, H.C., Z.L. Brumme, and P.R. Harrigan, *Human immunodeficiency virus type 1 protease cleavage site mutations associated with protease inhibitor cross-resistance selected by indinavir, ritonavir, and/or saquinavir.* J. Virol., 2001. **75**(2): p. 589-594.
79. Rhee, S.Y., et al., *Human immunodeficiency virus reverse transcriptase and protease sequence database.* Nucleic Acids Res., 2003. **31**(1): p. 298-303.
80. Shafer, R.W., et al., *Identification of biased amino acid substitution patterns in human immunodeficiency virus type 1 isolates from patients treated with protease inhibitors.* J. Virol., 1999. **73**(7): p. 6197-202.
81. Feher, A., et al., *Effect of sequence polymorphism and drug resistance on two HIV-1 Gag processing sites.* Eur J Biochem, 2002. **269**(16): p. 4114-20.
82. Tamiya, S., et al., *Amino acid insertions near Gag cleavage sites restore the otherwise compromised replication of human immunodeficiency virus type 1 variants resistant to protease inhibitors.* J Virol, 2004. **78**(21): p. 12030-40.
83. Pereira-Vaz, J., et al., *Detection of the protease codon 35 amino acid insertion in sequences from treatment-naive HIV-1 subtype C infected individuals in the Central Region of Portugal.* J Clin Virol, 2009. **46**(2): p. 169-72.

84. Jordan, P.S., et al., *A novel codon insert in protease of clade B HIV type 1.* AIDS Res Hum Retroviruses, 2009. **25**(5): p. 547-50.
85. Kozisek, M., et al., *Ninety-nine is not enough: molecular characterization of inhibitor-resistant human immunodeficiency virus type 1 protease mutants with insertions in the flap region.* J Virol, 2008. **82**(12): p. 5869-78.
86. Walker, B.D. and D.R. Burton, *Toward an AIDS vaccine.* Science, 2008. **320**(5877): p. 760-4.
87. Sakai, K. and M. Takiguchi, *Toward an effective AIDS vaccine development.* Eur J Immunol, 2013. **43**(12): p. 3087-9.
88. Brenner, B.G., D. Turner, and M.A. Wainberg, *HIV-1 drug resistance: can we overcome?* Expert Opin Biol Ther, 2002. **2**(7): p. 751-61.
89. Tozser, J., *HIV inhibitors: problems and reality.* Ann N Y Acad Sci, 2001. **946**: p. 145-59.
90. Miller, V., *International perspectives on antiretroviral resistance. Resistance to protease inhibitors.* J Acquir Immune Defic Syndr, 2001. **26 Suppl 1**: p. S34-50.
91. Pieniazek, D., et al., *Protease sequences from HIV-1 group M subtypes A-H reveal distinct amino acid mutation patterns associated with protease resistance in protease inhibitor-naive individuals worldwide. HIV Variant Working Group.* AIDS, 2000. **14**(11): p. 1489-95.
92. Rhee, S.Y., et al., *Human immunodeficiency virus reverse transcriptase and protease sequence database.* Nucleic Acids Res, 2003. **31**(1): p. 298-303.
93. Shafer, R.W., *Genotypic testing for human immunodeficiency virus type 1 drug resistance.* Clin Microbiol Rev, 2002. **15**(2): p. 247-77.
94. Johnson, V.A., et al., *2011 update of the drug resistance mutations in HIV-1.* Top Antivir Med, 2011. **19**(4): p. 156-64.
95. Tie, Y., et al., *Atomic resolution crystal structures of HIV-1 protease and mutants V82A and I84V with saquinavir.* Proteins, 2007. **67**(1): p. 232-42.
96. Liu, F., et al., *Mechanism of drug resistance revealed by the crystal structure of the unliganded HIV-1 protease with F53L mutation.* J Mol Biol, 2006. **358**(5): p. 1191-9.
97. Mahalingam, B., et al., *Structural implications of drug-resistant mutants of HIV-1 protease: high-resolution crystal structures of the mutant protease/substrate analogue complexes.* Proteins, 2001. **43**(4): p. 455-64.
98. Mahalingam, B., et al., *Structural and kinetic analysis of drug resistant mutants of HIV-1 protease.* Eur J Biochem, 1999. **263**(1): p. 238-45.
99. Tozser, J., et al., *Comparison of the HIV-1 and HIV-2 proteinases using oligopeptide substrates representing cleavage sites in Gag and Gag-Pol polyproteins.* FEBS Lett, 1991. **281**(1-2): p. 77-80.
100. Mahalingam, B., et al., *Combining mutations in HIV-1 protease to understand mechanisms of resistance.* Proteins, 2002. **48**(1): p. 107-16.
101. Johnson, V.A., et al., *Update of the drug resistance mutations in HIV-1: March 2013.* Top. Antivir. Med., 2013. **21**: p. 4-12.
102. Logsdon, B.C., et al., *Crystal structures of a multidrug-resistant human immunodeficiency virus type 1 protease reveal an expanded active-site cavity.* J Virol, 2004. **78**(6): p. 3123-32.
103. Dierynck, I., et al., *Binding kinetics of darunavir to human immunodeficiency virus type 1 protease explain the potent antiviral activity and high genetic barrier.* J Virol, 2007. **81**(24): p. 13845-51.
104. Louis, J.M., et al., *Inhibition of autoprocessing of natural variants and multidrug resistant mutant precursors of HIV-1 protease by clinical inhibitors.* Proc. Natl. Acad. Sci. U.S.A., 2011. **108**(22): p. 9072-9077.
105. Martin, P., et al., *"Wide-open" 1.3 A structure of a multidrug-resistant HIV-1 protease as a drug target.* Structure, 2005. **13**(12): p. 1887-95.

106. Wang, Y., et al., *The higher barrier of darunavir and tipranavir resistance for HIV-1 protease.* Biochem Biophys Res Commun, 2011. **412**(4): p. 737-42.
107. Zhang, Y., et al., *Structures of Darunavir-Resistant HIV-1 Protease Mutant Reveal Atypical Binding of Darunavir to Wide Open Flaps.* ACS Chem Biol, 2014.
108. McCammon, J.A., B.R. Gelin, and M. Karplus, *Dynamics of folded proteins.* Nature, 1977. **267**(5612): p. 585-90.
109. Born, M. and R. Oppenheimer, *Zur Quantentheorie der Molekeln.* Annalen der Physik, 1927. **389**(20): p. 457-484.
110. Norberto de Souza, O. and R.L. Ornstein, *Molecular dynamics simulations of a protein-protein dimer: particle-mesh Ewald electrostatic model yields far superior results to standard cutoff model.* J Biomol Struct Dyn, 1999. **16**(6): p. 1205-18.
111. Harrison, R.W., *Stiffness and energy conservation in molecular dynamics: An improved integrator.* Journal of Computational Chemistry, 1993. **14**(9): p. 1112-1122.
112. Darden, T., D. York, and L. Pedersen, *Particle mesh Ewald: An N ·log(N) method for Ewald sums in large systems.* The Journal of Chemical Physics, 1993. **98**(12): p. 10089-10092.
113. Greengard, L.F., *The rapid evaluation of potential fields in particle systems*. 1987, Yale University. p. 101.
114. Toth, G. and A. Borics, *Flap opening mechanism of HIV-1 protease.* J Mol Graph Model, 2006. **24**(6): p. 465-74.
115. Hamelberg, D. and J.A. McCammon, *Fast peptidyl cis-trans isomerization within the flexible Gly-rich flaps of HIV-1 protease.* J Am Chem Soc, 2005. **127**(40): p. 13778-9.
116. Cai, Y., et al., *Differential Flap Dynamics in Wild-type and a Drug Resistant Variant of HIV-1 Protease Revealed by Molecular Dynamics and NMR Relaxation.* J Chem Theory Comput, 2012. **8**(10): p. 3452-3462.
117. Scott, W.R. and C.A. Schiffer, *Curling of flap tips in HIV-1 protease as a mechanism for substrate entry and tolerance of drug resistance.* Structure, 2000. **8**(12): p. 1259-65.
118. Harrison, R.W. and I.T. Weber, *Molecular dynamics simulations of HIV-1 protease with peptide substrate.* Protein Eng, 1994. **7**(11): p. 1353-63.
119. Sadiq, S.K., F. Noe, and G. De Fabritiis, *Kinetic characterization of the critical step in HIV-1 protease maturation.* Proc Natl Acad Sci U S A, 2012. **109**(50): p. 20449-54.
120. Kunze, J., et al., *Targeting dynamic pockets of HIV-1 protease by structure-based computational screening for allosteric inhibitors.* J Chem Inf Model, 2014. **54**(3): p. 987-91.
121. Antunes, D.A., et al., *New insights into the in silico prediction of HIV protease resistance to nelfinavir.* PLoS One, 2014. **9**(1): p. e87520.
122. Rao, B.G., E.E. Kim, and M.A. Murcko, *Calculation of solvation and binding free energy differences between VX-478 and its analogs by free energy perturbation and AMSOL methods.* J Comput Aided Mol Des, 1996. **10**(1): p. 23-30.
123. Chen, X., I.T. Weber, and R.W. Harrison, *Molecular dynamics simulations of 14 HIV protease mutants in complexes with indinavir.* J Mol Model, 2004. **10**(5-6): p. 373-81.
124. Bandyopadhyay, P. and B.R. Meher, *Drug resistance of HIV-1 protease against JE-2147: I47V mutation investigated by molecular dynamics simulation.* Chem Biol Drug Des, 2006. **67**(2): p. 155-61.
125. Gotz, A.W., et al., *Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born.* J Chem Theory Comput, 2012. **8**(5): p. 1542-1555.
126. Tarmyshov, K.B. and F. Muller-Plathe, *Parallelizing a molecular dynamics algorithm on a multiprocessor workstation using OpenMP.* J Chem Inf Model, 2005. **45**(6): p. 1943-52.
127. Zhu, Y.L., et al., *GALAMOST: GPU-accelerated large-scale molecular simulation toolkit.* J Comput Chem, 2013. **34**(25): p. 2197-211.

128. Lupo, J.A., et al., *A large scale molecular dynamics simulation code using the fast multipole algorithm (FMD): performance and application.* J Mol Graph Model, 2002. **21**(2): p. 89-99.
129. Guoxing, F., A. Sabnis, and R.W. Harrison. *A deterministic-stochastic crossover algorithm for simulation of complex biochemical systems*. in *Computational Advances in Bio and Medical Sciences (ICCABS), 2013 IEEE 3rd International Conference on*. 2013.
130. Gropp, W., E. Lusk, and R. Thakur, *Using MPI-2 : advanced features of the message-passing interface*. Scientific and engineering computation. 1999, Cambridge, Mass.: MIT Press. xxi, 382 p.
131. Chapman, B., *Using OpenMP : portable shared memory parallel programming*. 2008, Cambridge, Mass.: MIT Press. xxii, 353 p.
132. Johnson, V.A., et al., *Update of the Drug Resistance Mutations in HIV-1: December 2010.* Top HIV Med, 2010. **18**(5): p. 156-63.
133. Pearson, A.R. and R.L. Owen, *Combining X-ray crystallography and single-crystal spectroscopy to probe enzyme mechanisms.* Biochem Soc Trans, 2009. **37**(Pt 2): p. 378-81.
134. Kovalevsky, A.Y., et al., *Caught in the Act: the 1.5 A resolution crystal structures of the HIV-1 protease and the I54V mutant reveal a tetrahedral reaction intermediate.* Biochemistry, 2007. **46**(51): p. 14854-64.
135. Das, A., et al., *X-ray snapshot of HIV-1 protease in action: observation of tetrahedral intermediate and short ionic hydrogen bond SIHB with catalytic aspartate.* J Am Chem Soc, 2010. **132**(18): p. 6366-73.
136. Ishima, R., et al., *Flap opening and dimer-interface flexibility in the free and inhibitor-bound HIV protease, and their implications for function.* Structure, 1999. **7**(9): p. 1047-55.
137. Brooks, C.L., M. Karplus, and B.M. Pettitt, *Proteins: a theoretical perspective of dynamics, structure, and thermodynamics*. Advances in chemical physics. 1988, New York: John Wiley. xiii, 259 p.
138. UNAIDS, *2008 Report on the Global AIDS Epidemic Report on the Global AIDS Epidemic UNAIDS Publication Series*. 2009: World Health Organization. 357.
139. Kim, E.E., et al., *Crystal structure of HIV-1 protease in complex with VX-478, a potent and orally bioavailable inhibitor of the enzyme.* J. Am. Chem. Soc., 1995. **117**(3): p. 1181-1182.
140. Williams, G.C. and P.J. Sinko, *Oral absorption of the HIV protease inhibitors: a current update.* Adv Drug Deliv Rev, 1999. **39**(1-3): p. 211-238.
141. Surleraux, D.L., et al., *Discovery and selection of TMC114, a next generation HIV-1 protease inhibitor.* J Med Chem, 2005. **48**(6): p. 1813-22.
142. Krohn, A., et al., *Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (R)-hydroxyethylamine isostere.* J Med Chem, 1991. **34**(11): p. 3340-2.
143. Johnson, V.A., et al., *Update of the Drug Resistance Mutations in HIV-1.* Top HIV Med, 2008. **16**(5): p. 138-45.
144. Wu, T.D., et al., *Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments.* J Virol, 2003. **77**(8): p. 4836-47.
145. Pazhanisamy, S., et al., *Kinetic characterization of human immunodeficiency virus type-1 protease-resistant variants.* J Biol Chem, 1996. **271**(30): p. 17979-85.
146. Liu, F., et al., *Kinetic, stability, and structural changes in high-resolution crystal structures of HIV-1 protease with drug-resistant mutations L24I, I50V, and G73S.* J Mol Biol, 2005. **354**(4): p. 789-800.
147. Kovalevsky, A.Y., et al., *Effectiveness of nonpeptide clinical inhibitor TMC-114 on HIV-1 protease with highly drug resistant mutations D30N, I50V, and L90M.* J Med Chem, 2006. **49**(4): p. 1379-87.

148. Murphy, M.D., G.I. Marousek, and S. Chou, *HIV protease mutations associated with amprenavir resistance during salvage therapy: importance of I54M.* J Clin Virol, 2004. **30**(1): p. 62-7.
149. Roberts, N.A., J.C. Craig, and J. Sheldon, *Resistance and cross-resistance with saquinavir and other HIV protease inhibitors: theory and practice.* Aids, 1998. **12**(5): p. 453-60.
150. Hoffman, N.G., C.A. Schiffer, and R. Swanstrom, *Covariation of amino acid positions in HIV-1 protease.* Virology, 2003. **314**(2): p. 536-48.
151. Maguire, M., et al., *Emergence of resistance to protease inhibitor amprenavir in human immunodeficiency virus type 1-infected patients: selection of four alternative viral protease genotypes and influence of viral susceptibility to coadministered reverse transcriptase nucleoside inhibitors.* Antimicrob Agents Chemother, 2002. **46**(3): p. 731-8.
152. Ishima, R., J.M. Louis, and D.A. Torchia, *Characterization of two hydrophobic methyl clusters in HIV-1 protease by NMR spin relaxation in solution.* J Mol Biol, 2001. **305**(3): p. 515-21.
153. Christian B. Anfinsen, J.T.E., Frederic M. Richards, *Advances in Protein Chemistry*. Vol. 47. 1995: Academic Press. 458.
154. Kellis, J.T., Jr., et al., *Contribution of hydrophobic interactions to protein stability.* Nature, 1988. **333**(6175): p. 784-6.
155. Wondrak, E.M. and J.M. Louis, *Influence of flanking sequences on the dimer stability of human immunodeficiency virus type 1 protease.* Biochemistry, 1996. **35**(39): p. 12957-62.
156. Otwinowski, Z. and W. Minor, *Processing of X-ray diffraction data collected in oscillation mode.* Methods Enzymol, 1997. **267**: p. 307-326.
157. McCoy, A.J., et al., *Likelihood-enhanced fast translation functions.* Acta Crystallogr D Biol Crystallogr, 2005. **61**(Pt 4): p. 458-64.
158. Wang, Y.F., et al., *Potent new antiviral compound shows similar inhibition and structural interactions with drug resistant mutants and wild type HIV-1 protease.* J Med Chem, 2007. **50**(18): p. 4509-15.
159. Vagin, A. and A. Teplyakov, *MOLREP: an Automated Program for Molecular Replacement.* Journal of Applied Crystallography, 1997. **30**(6): p. 1022-1025.
160. Sheldrick, G.M. and T.R. Schneider, *SHELXL: high-resolution refinement.* Methods Enzymol, 1997. **277**: p. 319-43.
161. Murshudov, G.N., A.A. Vagin, and E.J. Dodson, *Refinement of macromolecular structures by the maximum-likelihood method.* Acta Crystallogr D Biol Crystallogr, 1997. **53**(Pt 3): p. 240-55.
162. Emsley, P. and K. Cowtan, *Coot: Model-Building Tools for Molecular Graphics.* Acta Crystallographica Section D - Biological Crystallography, 2004. **60**.
163. DeLano, W.L., *The PyMOL Molecular Graphics System.* DeLano Scientific: San Carlos, CA, 2002.
164. Tie, Y., *Crystallographic Analysis and Kinetic Studies of HIV-1 Protease and Drug-Resistant Mutants*, in *Biology*. 2006, Georgia State University: Atlanta. p. 131-141.
165. Fioravanti, E., et al., *Specific radiation damage to acidic residues and its relation to their chemical and structural environment.* J Synchrotron Radiat, 2007. **14**(Pt 1): p. 84-91.
166. Vaguine, A.A., J. Richelle, and S.J. Wodak, *SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model.* Acta Crystallographica Section D, 1999. **55**(1): p. 191-205.
167. Ghosh, A.K., et al., *Flexible cyclic ethers/polyethers as novel P2-ligands for HIV-1 protease inhibitors: design, synthesis, biological evaluation, and protein-ligand X-ray studies.* J Med Chem, 2008. **51**(19): p. 6021-33.
168. Gustchina, A., et al., *Energy calculations and analysis of HIV-1 protease-inhibitor crystal structures.* Protein Eng, 1994. **7**(3): p. 309-17.
169. Spinelli, S., et al., *The three-dimensional structure of the aspartyl protease from the HIV-1 isolate BRU.* Biochimie, 1991. **73**(11): p. 1391-6.

170. Hong, L., et al., *Crystal structure of an in vivo HIV-1 protease mutant in complex with saquinavir: insights into the mechanisms of drug resistance.* Protein Sci, 2000. **9**(10): p. 1898-904.
171. Prabu-Jeyabalan, M., E. Nalivaika, and C.A. Schiffer, *How does a symmetric dimer recognize an asymmetric substrate? A substrate complex of HIV-1 protease.* J Mol Biol, 2000. **301**(5): p. 1207-20.
172. Short, G.F., 3rd, et al., *Probing the S1/S1' substrate binding pocket geometry of HIV-1 protease with modified aspartic acid analogues.* Biochemistry, 2000. **39**(30): p. 8768-81.
173. Stebbins, J., et al., *The 80's loop (residues 78 to 85) is important for the differential activity of retroviral proteases.* J Mol Biol, 1997. **267**(3): p. 467-75.
174. Foulkes-Murzycki, J.E., W.R. Scott, and C.A. Schiffer, *Hydrophobic sliding: a possible mechanism for drug resistance in human immunodeficiency virus type 1 protease.* Structure, 2007. **15**(2): p. 225-33.
175. Heaslet, H., et al., *Structural insights into the mechanisms of drug resistance in HIV-1 protease NL4-3.* J Mol Biol, 2006. **356**(4): p. 967-81.
176. Saskova, K.G., et al., *Molecular characterization of clinical isolates of human immunodeficiency virus resistant to the protease inhibitor darunavir.* J Virol, 2009. **83**(17): p. 8810-8.
177. Koh, Y., et al., *Novel bis-tetrahydrofuranylurethane-containing nonpeptidic protease inhibitor (PI) UIC-94017 (TMC114) with potent activity against multi-PI-resistant human immunodeficiency virus in vitro.* Antimicrob Agents Chemother, 2003. **47**(10): p. 3123-9.
178. Randolph, J.T. and D.A. DeGoey, *Peptidomimetic inhibitors of HIV protease.* Curr Top Med Chem, 2004. **4**(10): p. 1079-95.
179. Sayer, J.M., et al., *Effect of the active site D25N mutation on the structure, stability, and ligand binding of the mature HIV-1 protease.* J Biol Chem, 2008. **283**(19): p. 13459-70.
180. Cigler, P., et al., *From nonpeptide toward noncarbon protease inhibitors: metallacarboranes as specific and potent inhibitors of HIV protease.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15394-9.
181. Damm, K.L., et al., *A poke in the eye: inhibiting HIV-1 protease through its flap-recognition pocket.* Biopolymers, 2008. **89**(8): p. 643-52.
182. Veerapandian, B., et al., *Direct observation by X-ray analysis of the tetrahedral "intermediate" of aspartic proteinases.* Protein Sci, 1992. **1**(3): p. 322-8.
183. Coates, L., et al., *The catalytic mechanism of an aspartic proteinase explored with neutron and X-ray diffraction.* J Am Chem Soc, 2008. **130**(23): p. 7235-7.
184. Rose, R.B., et al., *Three-dimensional structures of HIV-1 and SIV protease product complexes.* Biochemistry, 1996. **35**(39): p. 12933-44.
185. Kumar, M., et al., *Observation of a tetrahedral reaction intermediate in the HIV-1 protease-substrate complex.* Biochem J, 2005. **389**(Pt 2): p. 365-71.
186. Mildner, A.M., et al., *The HIV-1 protease as enzyme and substrate: mutagenesis of autolysis sites and generation of a stable mutant with retained kinetic properties.* Biochemistry, 1994. **33**(32): p. 9405-13.
187. Shen, C.H., et al., *Amprenavir complexes with HIV-1 protease and its drug-resistant mutants altering hydrophobic clusters.* Febs J, 2010. **277**(18): p. 3699-714.
188. Rowley, R.L. and T. Pakkanen, *Determination of a methane intermolecular potential model for use in molecular simulations from ab initio calculations.* Journal of Chemical Physics, 1999. **110**(7): p. 3368-3377.
189. Yamazaki, T., et al., *Three-dimensional solution structure of the HIV-1 protease complexed with DMP323, a novel cyclic urea-type inhibitor, determined by nuclear magnetic resonance spectroscopy.* Protein science : a publication of the Protein Society, 1996. **5**(3): p. 495-506.

190. Weber, I.T., et al., *Reaction intermediates discovered in crystal structures of enzymes.* Advances in protein chemistry and structural biology, 2012. **87**: p. 57-86.
191. Kipp, D.R., et al., *Transition states of native and drug-resistant HIV-1 protease are the same.* Proc Natl Acad Sci U S A, 2012. **109**(17): p. 6543-8.
192. Heaslet, H., et al., *Conformational flexibility in the flap domains of ligand-free HIV protease.* Acta Crystallogr D Biol Crystallogr, 2007. **63**(Pt 8): p. 866-75.
193. Piana, S., et al., *Ab initio molecular dynamics-based assignment of the protonation state of pepstatin A/HIV-1 protease cleavage site.* J Am Chem Soc, 2001. **123**(36): p. 8730-7.
194. Titanji, B.K., et al., *Protease inhibitors effectively block cell-to-cell spread of HIV-1 between T cells.* Retrovirology, 2013. **10**(1): p. 161.
195. Domingo, E., et al., *Basic concepts in RNA virus evolution.* FASEB J, 1996. **10**(8): p. 859-64.
196. Harrigan, P.R., et al., *Predictors of HIV drug-resistance mutations in a large antiretroviral-naive cohort initiating triple antiretroviral therapy.* Journal of Infectious Diseases, 2005. **191**(3): p. 339-47.
197. Weiser, S.D., et al., *Higher rates of viral suppression with nonnucleoside reverse transcriptase inhibitors compared to single protease inhibitors are not explained by better adherence.* HIV Clin Trials, 2004. **5**(5): p. 278-87.
198. Wensing, A.M., et al., *2014 Update of the drug resistance mutations in HIV-1.* Top Antivir Med, 2014. **22**(3): p. 642-50.
199. Louis, J.M., et al., *Inhibition of autoprocessing of natural variants and multidrug resistant mutant precursors of HIV-1 protease by clinical inhibitors.* Proc Natl Acad Sci U S A, 2011. **108**(22): p. 9072-7.
200. Louis, J.M., et al., *Enhanced Stability of Monomer Fold Correlates with Extreme Drug Resistance of HIV-1 Protease.* Biochemistry, 2013. **52**(43): p. 7678-88.
201. Agniswamy, J., et al., *HIV-1 protease with 20 mutations exhibits extreme resistance to clinical inhibitors through coordinated structural rearrangements.* Biochemistry, 2012. **51**(13): p. 2819-28.
202. McCoy, A.J., et al., *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.
203. Winn, M.D., G.N. Murshudov, and M.Z. Papiz, *Macromolecular TLS refinement in REFMAC at moderate resolutions.* Methods in enzymology, 2003. **374**: p. 300-21.
204. Brunger, A.T., *Version 1.2 of the Crystallography and NMR system.* Nature protocols, 2007. **2**(11): p. 2728-33.
205. Krissinel, E. and K. Henrick, *Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2256-68.
206. Fang, B., et al., *Caspase-3 binds diverse P4 residues in peptides as revealed by crystallography and structural modeling.* Apoptosis : an international journal on programmed cell death, 2009. **14**(5): p. 741-52.
207. Chen, J., et al., *Insights into the functional role of protonation states in the HIV-1 protease-BEA369 complex: molecular dynamics simulations and free energy calculations.* J Mol Model, 2009. **15**(10): p. 1245-52.
208. Kar, P. and V. Knecht, *Energetic basis for drug resistance of HIV-1 protease mutants against amprenavir.* Journal of computer-aided molecular design, 2012. **26**(2): p. 215-32.
209. Everitt, B. and T. Hothorn, *A handbook of statistical analyses using R*. 2nd ed. 2010, Boca Raton: CRC Press. 355 p.
210. Team, R.C., *R: A language and environment for statistical computing.* 2014, R Foundation for Statistical Computing.

211. Harte, W.E., Jr., et al., *Domain communication in the dynamical structure of human immunodeficiency virus 1 protease.* Proc Natl Acad Sci U S A, 1990. **87**(22): p. 8864-8.
212. Perryman, A.L., J.H. Lin, and J.A. McCammon, *HIV-1 protease molecular dynamics of a wild-type and of the V82F/I84V mutant: possible contributions to drug resistance and a potential new target site for drugs.* Protein Sci, 2004. **13**(4): p. 1108-23.
213. Lepsik, M., Z. Kriz, and Z. Havlas, *Efficiency of a second-generation HIV-1 protease inhibitor studied by molecular dynamics and absolute binding free energy calculations.* Proteins, 2004. **57**(2): p. 279-93.
214. Karplus, M. and J. Kuriyan, *Molecular dynamics and protein function.* Proc Natl Acad Sci U S A, 2005. **102**(19): p. 6679-85.
215. Kokkinidis, M., N.M. Glykos, and V.E. Fadouloglou, *Protein flexibility and enzymatic catalysis.* Adv Protein Chem Struct Biol, 2012. **87**: p. 181-218.
216. Trylska, J., et al., *HIV-1 protease substrate binding and product release pathways explored with coarse-grained molecular dynamics.* Biophys J, 2007. **92**(12): p. 4179-87.
217. Piana, S., P. Carloni, and M. Parrinello, *Role of conformational fluctuations in the enzymatic reaction of HIV-1 protease.* J Mol Biol, 2002. **319**(2): p. 567-83.
218. Kurt, N., et al., *Cooperative fluctuations of unliganded and substrate-bound HIV-1 protease: a structure-based analysis on a variety of conformations from crystallography and molecular dynamics simulations.* Proteins, 2003. **51**(3): p. 409-22.
219. Dewdney, T.G., et al., *Ligand modifications to reduce the relative resistance of multi-drug resistant HIV-1 protease.* Bioorg Med Chem, 2013. **21**(23): p. 7430-4.
220. Bahar, I., et al., *Global dynamics of proteins: bridging between structure and function.* Annu Rev Biophys, 2010. **39**: p. 23-42.
221. Fenwick, R.B., et al., *Correlated motions are a fundamental property of beta-sheets.* Nat Commun, 2014. **5**: p. 4070.
222. Board Jr, J.A., et al., *Accelerated molecular dynamics simulation with the parallel fast multipole algorithm.* Chemical Physics Letters, 1992. **198**(1–2): p. 89-94.
223. Harrison, R.W., *Amortized Fast Mutipole Algorithm for Molecular Modeling*. 2003: ICCSA-2003 proceedings. 77-81.
224. García-Archilla, B., J. Sanz-Serna, and R. Skeel, *Long-Time-Step Methods for Oscillatory Differential Equations.* SIAM Journal on Scientific Computing, 1998. **20**(3): p. 930-963.
225. Gibbons, C., et al., *The structure of the central stalk in bovine F(1)-ATPase at 2.4 A resolution.* Nat Struct Biol, 2000. **7**(11): p. 1055-61.
226. Pokorna, J., et al., *Current and Novel Inhibitors of HIV Protease.* Viruses, 2009. **1**(3): p. 1209-39.
227. Mittal, S., et al., *Hydrophobic core flexibility modulates enzyme activity in HIV-1 protease.* J Am Chem Soc, 2012. **134**(9): p. 4163-8.
228. Lescar, J., et al., *Three-dimensional structure of an Fab-peptide complex: structural basis of HIV-1 protease inhibition by a monoclonal antibody.* J Mol Biol, 1997. **267**(5): p. 1207-22.
229. Agniswamy, J., et al., *Extreme multidrug resistant HIV-1 protease with 20 mutations is resistant to novel protease inhibitors with P1'-pyrrolidinone or P2-tris-tetrahydrofuran.* J Med Chem, 2013. **56**(10): p. 4017-27.
230. Perryman, A.L., et al., *Fragment-based screen against HIV protease.* Chem Biol Drug Des, 2010. **75**(3): p. 257-68.
231. Tiefenbrunn, T., et al., *Crystallographic fragment-based drug discovery: use of a brominated fragment library targeting HIV protease.* Chem Biol Drug Des, 2014. **83**(2): p. 141-8.

8 APPENDICES

Appendix A

Protein	ligand	PDB accession code	Resolution (Å)
PR_{I50V}	APV	3NU5	1.29
PR_{I54M}	APV	3NU6	1.16
PR_{I84V}	APV	3NU9	1.85
PR_{I54V}	APV	3NUJ	1.50
PR_{L90M}	APV	3NUO	1.35
PR20	SQV	3UFN	1.45
PR20	APV	4J5J	1.80
PR_{WT}^{*}	gem-diol	4FL8	1.20
PR_{I47V}^{*}	products	4FLG	1.31
PR_{V32I}	product	4FM6	1.40
PR_{I50V}	GRL-0519	4HDP	1.22
PR_{I54M}	GRL-0519	4HE9	1.06
$PR20_{D25N}open^{*}$	-		1.75
$PR20_{D25N}twist/tuck^{*}$	-		1.45

Appendix B

BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

HIV-1 Protease with 20 Mutations Exhibits Extreme Resistance to Clinical Inhibitors through Coordinated Structural Rearrangements

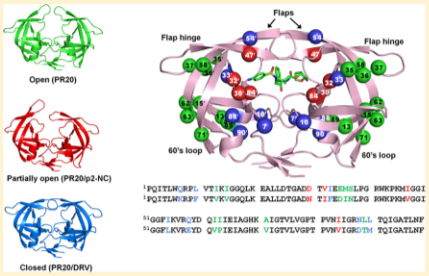
Johnson Agniswamy,[†] Chen-Hsiang Shen,[†] Annie Aniana,[‡] Jane M. Sayer,[‡] John M. Louis,[‡] and Irene T. Weber*[,†]

[†]Department of Biology, Molecular Basis of Disease Program, Georgia State University, Atlanta, Georgia 30303, United States

[‡]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, DHHS, Bethesda, Maryland 20892-0520, United States

Supporting Information

ABSTRACT: The escape mutant of HIV-1 protease (PR) containing 20 mutations (PR20) undergoes efficient polyprotein processing even in the presence of clinical protease inhibitors (PIs). PR20 shows >3 orders of magnitude decreased affinity for PIs darunavir (DRV) and saquinavir (SQV) relative to PR. Crystal structures of PR20 crystallized with yttrium, substrate analogue p2-NC, DRV, and SQV reveal three distinct conformations of the flexible flaps and diminished interactions with inhibitors through the combination of multiple mutations. PR20 with yttrium at the active site exhibits widely separated flaps lacking the usual intersubunit contacts seen in other inhibitor-free dimers. Mutations of residues 35−37 in the hinge loop eliminate interactions and perturb the flap conformation. Crystals of PR20/p2-NC contain one uninhibited dimer with one very open flap and one closed flap and a second inhibitor-bound dimer in the closed form showing six fewer hydrogen bonds with the substrate analogue relative to wild-type PR. PR20 complexes with PIs exhibit expanded S2/S2′ pockets and fewer PI interactions arising from coordinated effects of mutations throughout the structure, in agreement with the strikingly reduced affinity. In particular, insertion of the large aromatic side chains of L10F and L33F alters intersubunit interactions and widens the PI binding site through a network of hydrophobic contacts. The two very open conformations of PR20 as well as the expanded binding site of the inhibitor-bound closed form suggest possible approaches for modifying inhibitors to target extreme drug-resistant HIV.



Multidrug resistance poses a serious challenge in long-term therapy for human immunodeficiency virus (HIV) infection. The HIV protease (PR) plays a crucial role in replication by generating mature infectious viral particles through cleavage of Gag and Gag-Pol polyprotein precursors, and consequently, PR is a highly successful therapeutic target for HIV/AIDS.[1] Therapeutic intervention with the first protease inhibitor (PI) greatly improved the survival of patients infected with HIV but was followed by rapid emergence of drug resistance.[2,3] Drug resistance develops by selection of "major" mutations within the viral PR gene that decrease binding of the PIs, accompanied by decreased binding of natural substrates and reduced viral replication.[4] The replication of viruses containing major mutations is improved by the compensating effect of "minor" resistance mutations in naturally variable regions.[5] To date, 15 sites for major mutations and 19 for minor mutations have been identified for all nine FDA approved PIs.[6] In addition, drug resistance can develop due to insertion of 1−6 amino acids at various sites in the viral PR sequence[7,8] or by mutations in the PR cleavage sites in the Gag precursor.[9]

Mature PR is released by autoproteolysis of the Gag-Pol precursor. It is active as a homodimer of 99-residue subunits in which each subunit contributes one of the two aspartates required for catalysis. Substrate binding is accompanied by a conformational change from an open form that permits substrate entry into the active site cavity to a closed form in which the two flexible flaps (residues 44−57) close down to bind the substrate.[10] Current clinical inhibitors were designed to bind PR with the closed conformation of the flaps. Hence, mutations that affect flap conformation are selected frequently in resistance to PIs and can alter both inhibitor binding and catalytic activity of the enzyme.[6,10]

We have recently characterized a clinically derived multidrug-resistant protease (PR20)[11] bearing 20 mutations (Figure 1),[12] of which 15 are classified as either major or minor drug resistance mutations.[6] Mature PR20 exhibits a dimer dissociation constant (K_d) of ∼30 nM, which is >3-fold higher than for PR, and is catalytically competent with a similar turnover rate (k_{cat}) and an ∼13-fold higher K_m for a synthetic substrate relative to PR (Table S1 in ref 11). Relative to PR, PR20 shows a drastically lower affinity for PIs by >3 orders of

Received: December 12, 2011
Revised: February 24, 2012
Published: March 10, 2012

dx.doi.org/10.1021/bi2018317 | Biochemistry 2012, 51, 2819−2828

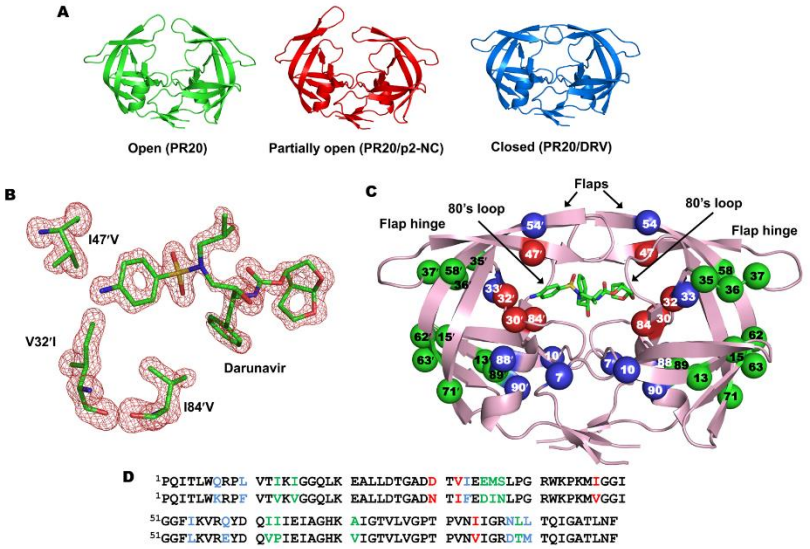


Figure 1. Crystal structures of PR20 exhibit three different conformations, open, partially open, and closed, in the free PR20 and inhibited structures of PR20/p2-NC and PR20/DRV (A). (B) $F_o - F_c$ omit map of resistant mutations V32′I, I47′V, I84′ and the major conformation of DRV bound at the active site contoured at 3σ level. (C) Sites of the 20 multidrug-resistant mutations are mapped on HIV-1 PR (pink cartoon representation) with the bound inhibitor DRV colored with green sticks for carbon atoms. The mutations with direct interaction to inhibitors are colored as red spheres, while second shell mutations are shown as blue spheres, and more distal mutations are shown as green spheres. (D) Amino acid sequence of HIV-1 PR (upper line) and PR20 (lower line). The residues are colored similar to panel C. Note that the PR20 sequence includes Q7K, and PR used for comparison in this study includes Q7K, L33I, and L63I to prevent autoproteolysis and both include C67A and C95A to eliminate cysteine-induced aggregation.

magnitude. Inhibitor-dissociation constants (K_I) for DRV and SQV binding to PR20 are 41 and 930 nM, respectively, relative to the corresponding K_I values for PR with DRV (0.005−0.01 nM) and SQV (0.4 nM) (Table 1 in ref 11). Even though the thermal stability of uninhibited PR20 is significantly greater than that of PR, as shown by a 6 °C higher T_m on DSC, consistent with their weak binding, PIs stabilize the ternary complexes of PR20 (dimer + PI) to a significantly lesser extent than when bound to PR. Thus, values of ΔT_m (inhibitor bound minus unbound) are markedly lower for PR20 at 5.3 and 3.1 °C for DRV and SQV, respectively,[11] than for PR (22.4 and 19.3 °C).[13]

Autocatalytic cleavage (autoprocessing) of the PR from the viral Gag-Pol precursor polyprotein, particularly at its N-terminus, is crucial for its release, viral maturation, and propagation. A PR20 precursor analogue consisting of PR20 fused at its N-terminus to the S6-amino acid transframe region (TFR), when expressed in *E. coli*, undergoes efficient autoprocessing at the TFR/PR20 site to release mature, catalytically active PR20. Importantly, autoprocessing of TFR-PR20 is unresponsive to inhibition by all clinical PIs in current use.[11] Inhibition is not observed even in the presence of 150−250 μM SQV or DRV, which far exceeds the estimated plasma or intracellular concentration on administration of these drugs in human subjects.[11] In contrast, the IC_{50} for inhibition of wild-type TFR-PR autoprocessing by DRV in *E. coli* is 1−2 μM.

These observations indicate that PR20 is a highly evolved drug-resistant mutant and is likely to be clinically unresponsive to all currently available PIs. To examine the structural basis for this extreme drug resistance, we determined the crystal structures of PR20 alone, and bound to DRV and SQV, which fail to block autoprocessing of TFR-PR20, although they are the most effective PIs for inhibition of the wild-type TFR-PR precursor.[11] We also determined the structure of PR20 bound to a substrate analogue that mimics the p2-NC natural cleavage site in the Gag-Pol polyprotein in order to assess differences in substrate binding. Three distinct dimeric structures were obtained: a wide open conformation, a semiopen conformation, and a closed conformation showing significantly diminished interactions with inhibitors and substrate analogue. Comparison of PR20 with wild-type PR structures reveals the evolving mechanisms in HIV to evade PIs and thus may lead to improved strategies for targeting extreme multidrug-resistant mutants.

■ MATERIALS AND METHODS

Construction, Expression, and Purification of PR20. A synthetic gene encoding 99 amino acids of the protease derived from the sequence of a clinical isolate,[12] termed PR20, was cloned between the Nde1 and BamH1 sites of pET11a vector (Novagen, San Diego, CA) and transformed into *E. coli* BL-21 (DE3; Stratagene). Protein expression, purification, and folding were carried out as described.[14]

Crystallization. PR20 at 0.25 mM concentration was complexed with 1.25 mM of DRV, SQV, or substrate analogue p2-NC at a 1:5 molar ratio and incubated for 30 min. All crystallization trials were performed using hanging drop vapor diffusion technique at room temperature. Crystals of the PR20/

Table 1. Crystallographic Data Collection and Refinement Statistics

	PR20/DRV	PR20/p2-NC	PR20	PR20/SQV
space group	$P2_12_12_1$	$P4_32_12$	$P2_12_12_1$	$P2_12_12_1$
cell dimensions				
a (Å)	28.64	63.79	44.34	28.83
b (Å)	65.67	63.79	45.85	66.26
c (Å)	94.03	213.66	103.99	93.17
resolution range	50.0−1.38	50.0−2.2	50.0−1.65	50.0−1.45
total observations	213 451	156 472	166 031	175 446
unique reflections	35 869	23 454	25 386	31 978
redundancy	6.0 (3.2)	6.7 (5.0)	6.5 (3.2)	
completeness	95.8 (72.0)[a]	99.6 (98.2)	93.9 (61.7)	99.0 (95.4)
$\langle I/\sigma(I)\rangle$	27.3 (2.6)	24.7 (5.6)	31.9 (2.3)	9.9 (3.2)
R_{sym} (%)	5.4 (38.0)	7.1 (28.2)	5.1 (41.8)	11.6 (33.8)
refinement resolution range	10−1.38	47.5−2.2	10.0−1.63	10−1.45
R_{cryst} (%)	16.2	22.4	18.8	17.6
R_{free} (%)	22.3	29.3	22.2	22.5
number of solvent	108	106	166	130
mean *B*-factor (Å^2)	27.9	46.1	27.3	20.9
mean *B*-factor of the ligand (Å^2)	17.5	54.6		12.5
rms deviations from ideality				
bond lengths (Å)	0.01	0.02	0.03	0.01
angles	0.03 (Å)[b]	1.9 (deg)[c]	2.1 (deg)[c]	0.03 (Å)[b]

[a]Values in parentheses are given for the highest resolution shell. [b]The angle rmsd in SHELX97 is indicated by distance in Å. [c]The angle rmsd in REFMAC 5.2 is indicated by angle in degrees.

DRV complex were grown by mixing 1 μL of protein (5 mg/mL) and 1 μL of reservoir solution (1.6 M sodium chloride and 0.1 M sodium acetate buffer at pH 4.6). Crystals of PR20/SQV were grown with a reservoir solution containing 0.93 M sodium chloride and 0.03 M citrate−phosphate buffer at pH 3.8. The well solution used for growing PR20/p2-NC crystals was 0.9 M ammonium sulfate and 0.1 M sodium citrate buffer at pH 4.5. The PR20 by itself was crystallized with 0.9 M sodium chloride, 0.1 M yttrium chloride, and 0.1 M sodium acetate buffer at pH 5.5. The crystals were frozen with cryoprotectant containing the respective mother liquor together with 30% glycerol.

X-ray Data Collection and Structure Determination. Diffraction data were collected at 100 K on beamline 22-ID of the Southeast Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source, Argonne National Laboratory. The data were integrated and scaled with HKL2000.[15] The PR20/DRV and PR20 crystal structures were solved by molecular replacement with the wild-type HIV-1 PR in complex with DRV (2IEN[16]) as the initial model using PHASER.[17,18] The PR20/SQV crystal structure was solved by molecular replacement with the structure of PR mutant I50V in complex with DRV (2F8G[19]) using MOLREP. The wild-type structure in complex with p2-NC (2AOD) was used as the starting model to solve PR20/p2-NC. The 20 mutated residues were pruned to alanine during molecular replacement, and the correct amino acid side chain was added during refinement. For PR20/p2-NC structure, the flaps in the model were pruned for successful structure solution and rebuilt during refinement. The PR20/DRV and PR20/SQV structures were refined using SHELX-97,[20] while PR20 was refined with REFMAC 5.2,[22] and the model building was carried out in COOT.[21] The lower resolution PR20/p2-NC structure was subjected to several rounds of refinement in REFMAC 5.2[22] with the two dimers as a single TLC group and model building with COOT.[21] The inhibitors were fitted into unambiguous electron density in the three complex structures. Solvent molecules were inserted at stereochemically reasonable positions using $2F_o - F_c$ and $F_o - F_c$ maps at 1 and 3 sigma levels, respectively. Molecular figures were prepared with Molscript, Raster3D, and PyMOL.

RESULTS AND DISCUSSION

Crystal Structures of PR20 Exhibit Diverse Flap Conformations. The four crystal structures of PR20 reveal three distinct conformations with remarkable variation in the flaps and shed light on the biophysical and biochemical properties of this resistant mutant (Figure 1A). The diverse flap conformations are well ordered without unusually high *B*-factors (Figure S1A). The crystal structures were refined to resolutions of ∼1.4 to 2.2 Å and *R*-factors from 16 to 22.5% (Table 1). The free PR20 ($PR20_{open}$), PR20/DRV, and PR20/SQV complexes were crystallized with one dimer (residues numbered 1−99 and 1′−99′) in the asymmetric unit. The PR20 crystallized with the p2-NC peptide analogue had two dimers per asymmetric unit exhibiting different flap conformations: $PR20/p2\text{-}NC_{open}$ has no bound peptide, whereas $PR20/p2\text{-}NC_{closed}$ contains the p2-NC analogue bound in the active site (Figure S1B). The PR20/p2-NC structure has a high mean *B*-factor of 46 Å^2 compared to values of 21−28 Å^2 for the other PR20 structures, probably due to the lower resolution and increased flexibility of the two dimers in the asymmetric unit. The unusual existence of two dimer conformations in the same crystal structure, as well as possible disorder in the individual conformers, may result from weak binding of this substrate analogue.

All 20 drug-resistant mutations, inhibitors, and the p2-NC substrate analogue were unambiguously modeled in the structures (see examples in Figure 1B and Figure S1C,D). The mutations are distributed broadly around the inhibitor binding site, flaps, flap hinge, dimer interface, and regions distal to the active site (Figure 1C,D). Four "first shell" mutations D30N, V32I, I47V, and I84V alter residues making direct contacts with DRV and SQV inhibitors. Six other mutations

dx.doi.org/10.1021/bi2018317 | *Biochemistry* 2012, 51, 2819−2828

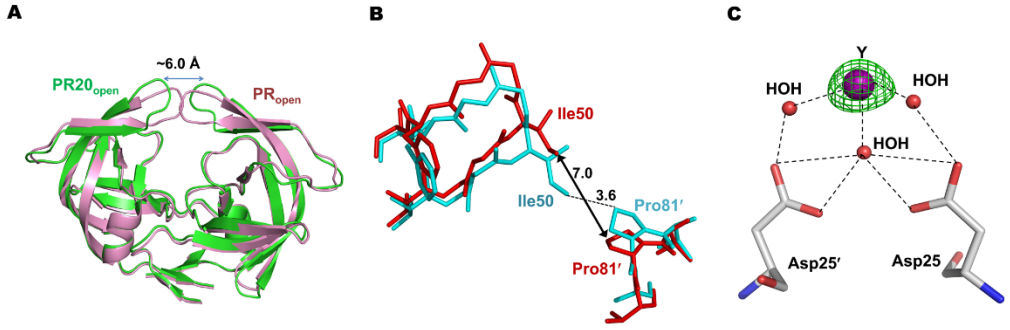


Figure 2. PR20 has an open conformation in the absence of inhibitor. (A) Superposition of free $PR20_{open}$ dimer in cartoon (green) and wild-type PR_{open} dimer (pink) showing a 6.0 Å separation between the flaps of $PR20_{open}$. (B) Comparison of flap A interaction with 80's loop in monomer B between $PR20_{open}$ (red) and MDR 769 (cyan). $PR20_{open}$ has no flap contacts to Monomer B. (C) $F_o - F_c$ omit map (green contours) of yttrium (magenta) bound at the active site of $PR20_{open}$ contoured at the 12σ level. The hydrogen bond interactions are indicated by dashed lines.

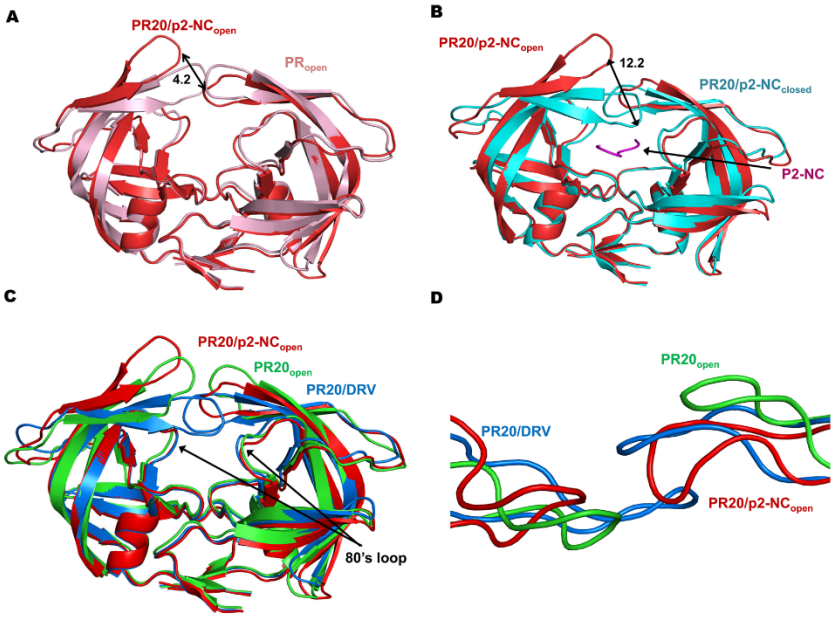


Figure 3. Conformation of the flaps varies in the PR20 structures. (A) Comparison of $PR20/p2\text{-}NC_{open}$ (red) with wild-type PR_{open} (pink). The closest distance between the two flaps of $PR20/p2\text{-}NC_{open}$ is 4.0 Å. (B) Comparison $PR20/p2\text{-}NC_{open}$ (red) with $PR20/p2\text{-}NC_{closed}$ (cyan) with the bound p2-NC molecule in $PR20/p2\text{-}NC_{closed}$ complex shown in magenta. The flap A position of $PR20/p2\text{-}NC_{open}$ and $PR20/p2\text{-}NC_{closed}$ deviates by a maximum of 12.2 Å. (C) Superimposition of $PR20_{open}$ (green), $PR20/p2\text{-}NC_{open}$ (red) and PR20/DRV (blue) shown in cartoon. (D) Close-up of flaps of $PR20_{open}$ (green), $PR20/p2\text{-}NC_{open}$ (red), and PR20/DRV (blue) orthogonal to the view in (C). The flap B of $PR20/p2\text{-}NC_{open}$ is curled at the tip in relation to the closed flap B of PR20/DRV.

Q7K, L10F, L33F, I54L, N88D, and L90 M are located in the second shell with direct influence on the inhibitor-interacting residues.

PR20 Crystallized with Yttrium Reveals a Wide Open Conformation with No Contact between Flap Tips. PR20 crystallizes in an open conformation in the absence of inhibitor, designated $PR20_{open}$, in the presence of the metal ion yttrium. The inhibitor-free $PR20_{open}$ shows large structural changes in comparison with the free wild-type PR (1HHP) as measured by an rmsd value of 1.73 Å for 198 Cα atoms (Figure 2A). The greatest disparity occurs at the flap tips with maximum deviations of >7 Å occurring at Gly49/49′. In the open form of wild-type PR, the tips of the two flaps have hydrophobic and hydrophilic interactions with each other. In contrast, the two flaps are completely separated in $PR20_{open}$, exposing a channel where the narrowest constriction of 6.1 Å occurs between the

side chain of Ile50 and the carbonyl oxygen of Ile50′. The flap residues 46−54 and 46′−54′ differ in conformation from those of the inhibitor-free wild-type PR. This conformation falls into the "open-flap" category described previously.[23] Notably, $PR20_{open}$ most closely resembles the free multidrug-resistant mutant named MDR769 (1TW7) containing 9 mutations and the inactivating D25N mutation to prevent self-proteolysis,[24] as indicated by the lower rmsd value of 0.61 Å for 198 Cα atoms. The major difference still resides in the flap region with the maximum deviation for Ile50 at the tip of the flap (Figure 2B). The flaps in MDR769 retain intersubunit contacts; the side chains of Ile50/50′ at the tip of both flaps have van der Waals contacts with Pro81′/81 from the opposite monomer, despite the presence of D25N that by itself increases the K_d of wild-type PR by 2 orders of magnitude.[25] In $PR20_{open}$, however, the flap of monomer A has no contact with monomer B since the shortest separation is about 7 Å between Ile50 and Pro81′ (Figure 2B).

Yttrium was identified by very high electron density visible at the active site of the $PR20_{open}$ structure (Figure 2C). The yttrium ion coordinates with three water molecules that form hydrogen bonds with the catalytic Asp25 and 25′. The central water molecule bound to yttrium mimics the central hydroxyl group common to PIs. Various metal ions have been reported to inhibit HIV PR,[26] and other crystal structures have metal ions bound at the active site.[27] We considered the possibility that binding of yttrium influences the flap conformation. However, the central yttrium has no direct or water-mediated interaction with flap residues, although a surface layer of water molecules is visible extending from the catalytic aspartate toward flap residue I54L. This water-coordinated yttrium ion resembles the magnesium ion and coordinated water observed in two independent open conformation apo HIV PR structures (2pc0 and 2hb4).[23] These three structures with magnesium or yttrium ions have three significantly different open flap conformations, suggesting that the metal binding does not influence the flap conformation in the open form. Since the clinical inhibitors were designed to bind the closed conformation of PR, the water-coordinated yttrium in the open conformation may suggest a novel framework for the design of compounds to combat resistance by targeting the widely separated flaps of multidrug-resistant PRs.

Two Unusual Flap Conformations Occur in PR20/p2-NC_{open} Dimer. The open conformation dimer (PR20/p2-NC_{open}) has no inhibitor bound at the active site, and the flaps have different conformations; flap A is very open, while flap B has a twisted closed form. Similar to $PR20_{open}$, the two flaps in the PR20/p2-NC_{open} dimer have minimal contact with each other with the closest distance of ~4.2 Å occurring between the side chain of Ile50 and Gly51′ (Figure 3A). Moreover, the flaps are 12.2 and 5.4 Å distant from Pro81′/81, respectively, and hence completely lack the intersubunit contacts seen in wild-type PR and most mutants. The PR20/p2-NC_{open} dimer differs from the free wild-type PR and $PR20_{open}$ showing overall rmsd values of 1.9 and 1.8 Å, respectively, with the largest deviations of over 7 Å occurring at Ile50 and Ile50′. The flap A of PR20/p2-NC_{open} exhibits the most widely open conformation reported for crystal structures of free HIV PRs since the tip shifts by ~4.4 Å relative to $PR20_{open}$ and ~12.2 Å relative to the inhibitor-bound PR20/p2-NC_{closed} (Figure 3B). In contrast, the flap B of PR20/p2-NC_{open} is more similar to the closed form flap in PR20/p2-NC_{closed} (Figure 3B), although the tip of the flap is twisted by ~72° due to the interaction with the flap of a symmetry-related molecule in the crystal lattice. The molecules intertwine so that the two flaps of one dimer occupy the active site cavity of the symmetry-related dimer (Figure S2A).

Superimposition of the $PR20_{open}$, PR20/p2-NC_{open}, and PR20/DRV structures reveals a striking variation in flap conformation (Figure 3C,D). A combination of lateral and vertical movements separates the flap tips as seen for the $PR20_{open}$ structure relative to the closed form of PR20/DRV. The very open PR20/p2-NC_{open} and $PR20_{open}$ structures show a vertical rise of the flaps away from the catalytic site, which eliminates intersubunit contacts between the flap tips and the 80's loop (Figure 3D). Also, the two flaps separate from each other in the horizontal plane. This variation in flap conformation is presumed to interfere mainly with binding of inhibitors as the PR20 precursor is fully competent for its release[11] and mediating polyprotein processing.[11] Similarly, a recent double electron−electron resonance study showed that polymorphism in HIV-1 PR subtypes can confer altered flap conformation and flexibility.[28]

PR20/p2-NC_{closed} Lacks Significant Interactions with Substrate Analogue. PR20/p2-NC_{closed} dimer shows the closed conformation of flaps and diminished interactions with the reduced peptide analogue p2-NC. Comparison of PR20/p2-NC_{closed} with the corresponding wild-type PR/p2-NC structure[29] reveals significant differences as indicated by an rmsd of 1.1 Å with the maximum deviation of 4.3 Å at the E35D mutation in the hinge loop. PR20 lacks several hydrogen bonds with p2-NC relative to those seen in the wild-type complex (Figure 4A,B). Notably, the PR20 flaps are missing 3 out of 4 hydrogen bonds connecting the main chain of the substrate analogue with Gly48 and 48′ in wild-type PR. In contrast, the MDR769/p2-NC structure shows open flaps lacking all hydrogen bonds with peptide substrate.[30] PR20 has no hydrogen bonds with the side chain of P2′ Gln unlike the three seen in the wild-type complex. Also, the side chain of P3′ Arg in PR20/p2-NC_{closed} forms an ion pair with Asp29 instead of the intersubunit ion pair between Asp29 and Arg8′ in the wild-type PR. In PR20, the absence of six hydrogen bonds with the substrate analogue is consistent with the 10-fold lower catalytic activity (~13× higher K_m) of the mutant and with the presence of a free dimer in the same asymmetric unit. Despite the markedly diminished interactions of the mature PR20 with this substrate analogue, its precursor undergoes autocatalytic processing even in the presence of inhibitors to produce viable virion.[11,12]

Flap Hinge Mutations Perturb Flap Conformation. All the PR20 structures show the same large structural change in the flap hinge region (residues 34−43) relative to the corresponding wild-type PR structures, which is attributed in part to the E35D, M36I, and S37N mutations (Figure 5A). In fact, residues 31−37 were shown to be critical for the specificity differences between HIV-1 and Simian immunodeficiency virus PRs.[31] In the PR20 structures, the shorter aspartate side chain introduced by the E35D mutation eliminates the ion pair observed between Glu35 and Arg57 at the base of the flap in the wild-type PR, which may affect the monomer stability and increase the variability of PR20 flaps. Further, in the wild-type structures, Met36 interacts with Ile15 and Ile33. All three residues are mutated in PR20 where the M36I mutation causes the loop to rearrange so that the shorter Ile36 side chain maintains hydrophobic contacts with mutated I15V and I33F. This rearrangement is in agreement with the molecular dynamics studies that predict increased flexibility of the flaps

dx.doi.org/10.1021/bi2018317 | Biochemistry 2012, 51, 2819−2828

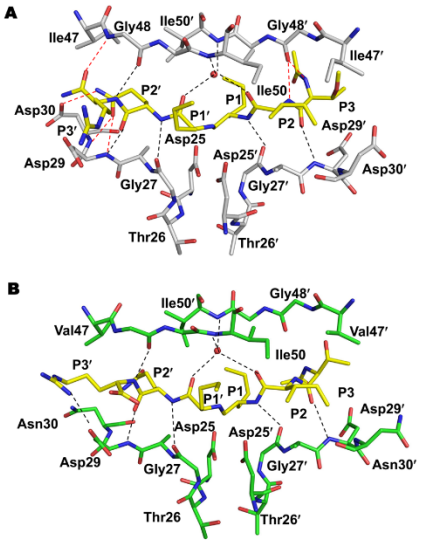


Figure 4. Hydrogen bond interactions of p2-NC substrate analogue with (A) PR and (B) PR20. PR is shown with gray carbons, PR20 with green carbons, and p2-NC is colored with yellow carbons. The p2-NC analogue comprises residues P3-P3′, which lie in subsites S3-S3′ of the substrate-binding cavity, with the nonhydrolyzable reduced peptide bond between norleucine at P1 and P1′. The hydrogen bonds are represented by dashed lines, and the red dashed lines indicate the hydrogen bonds in PR/p2-NC that are absent in PR20/p2-NC$_{closed}$.

due to E35D and M36I mutations.[32,33] Similar twisting of the hinge loop is seen in the crystal structures of HIV-1 PRs from subtype B, subtype F, and group N that bear the M36I substitution, despite their different space groups and inhibitors.[14,34,35] Thus, the hinge loop mutations break the ion pair anchors at the base of the flaps and propagate structural changes to the flap tips. The more variable flaps are expected to influence the monomer−dimer equilibrium and alter the binding of inhibitors as well as the substrates.

PR20/DRV and PR20/SQV Reveal Coordinated Effects of First and Second Shell Mutations. Our high resolution (1.38 and 1.45 Å) structures of PR20 with clinical inhibitors DRV and SQV reveal the coordinated effects of the mutations on the inhibitor binding site. Overall, the two inhibitor complexes are very similar to each other with rmsd value of 0.40 Å for 198 Cα atoms. In contrast, PR20 dimers show larger differences of 1.06 and 0.93 Å, respectively, relative to the corresponding wild-type PR complexes with DRV and SQV (2IEN,[16] 3OXC[36]). Similar to the PR20/p2-NC$_{closed}$ complex, the largest deviation of more than 4 Å occurs at the hinge loop. The active site cavity of the PR20/DRV complex is occupied by DRV in two alternate orientations related by a 180° rotation with occupancies of 0.7 and 0.3, while SQV binds in a single conformation. A second inhibitor molecule is seen bound to different sites in the PR20/DRV and PR20/SQV dimers (see Supporting Information and Figure S2B), as reported previously for two other PR mutant structures with the same inhibitors.[37,38] The PR20/SQV complex was crystallized at a slightly lower pH (3.8) than the other PR20 complexes (pH of 4.5); however, the conformation of the catalytic aspartates is essentially identical in all PR20 structures. Similarly, the previously published HIV-1 protease single mutant complex of M46L/DRV was crystallized at pH 3.6 with no major conformational change relative to PR/DRV.[37]

Unlike the PR20/p2-NC$_{closed}$ complex, PR20 retains the hydrogen bond interactions with SQV and DRV seen for the wild-type complexes with minor exceptions. SQV shows a shorter hydrogen bond of 3.1 Å with the carbonyl oxygen of Gly27 compared to a 3.6 Å long interaction in the wild-type PR, similar to that described recently in the SQV complex with the L76V single mutant.[39] In the wild-type PR, the side chain of Asp30 forms either a direct or a water-mediated hydrogen bond with the P2′ aniline in the major and minor conformations of DRV. In PR20/DRV, however, the Asn side chain of D30N forms a water-mediated hydrogen bond with the major conformation of DRV and introduces a new hydrogen bond with the side chain of the N88D mutation in the second shell (Figure 5B). The minor conformation of DRV exhibits a direct hydrogen bond to D30N, and no hydrogen bond forms between the side chains of D30N and N88D. Variations in the interaction of D30N with DRV and N88D were reported for a DRV-resistant clinical isolate.[40] The D30N single mutant does not form a hydrogen bond with Asn88 and resembles the wild-type structure, while D30N in the D30N, N88D double mutant forms a water-mediated hydrogen bond with DRV and a direct hydrogen bond with N88D similar to those in the PR20/DRV complex.[41] On the other hand, the conformation of the D30N side chain in PR20/SQV resembles that of the PR/SQV complex and lacks a hydrogen bond with N88D.[6] D30N is the only mutation of a charged first shell residue in PR20, and second shell mutation N88D may compensate for the altered charge. Analysis of mutation patterns in 8060 virus isolates reveals that N88D is positively associated with D30N and facilitates the occurrence of major resistance mutations D30N and L90M, resulting in multidrug resistance.[42]

The second shell mutation of L90M exhibits a similar effect in PR20 and single mutant structures.[19,43] The longer side chain of Met90 forms a shorter C−H···O interaction of 3.4 Å with the carbonyl oxygen of catalytic Asp25 in both monomers compared to the longer van der Waals contact in the wild-type PR. These short interactions of L90 M disturb the active site at the dimer interface and may contribute to the increased dimer dissociation observed for PR20 relative to PR. The L90M single mutation significantly reduces the dimer stability of PR[43] and increases the catalytic activity.[19,44] Also, *in vitro* studies show that combining the two mutations D30N and L90M results in PR that exhibits high levels of resistance to nelfinavir.[45] This inhibitor independent mechanism may explain why L90M has been identified as a resistance mutation for all clinical inhibitors except for DRV and TPV.[6]

Mutation L10F acts to perturb the intersubunit ion pairs of Arg8-Asp29′ and Arg8′-Asp29 at the outer edges of the substrate binding site in the closed forms of PR20. Conserved residues Arg8/8′ and Asp29′/29 contribute to the S2/S2′ pockets and form critical interactions with inhibitors. Mutations that eliminate this ion pair produce a temperature-sensitive phenotype with altered enzyme activity and thermal stability.[46,47] In the structures of PR20$_{open}$ and the monomer A of PR20/p2-NC$_{open}$ with the open flap, the intersubunit ion pairs of Arg8/8′ and Asp29′/29 exist as usually seen in the open and closed conformations of PR. This ion pair cannot form in PR20/p2-NC$_{closed}$ and in monomer B of PR20/p2-NC$_{open}$ with

dx.doi.org/10.1021/bi2018317 | *Biochemistry* 2012, 51, 2819−2828

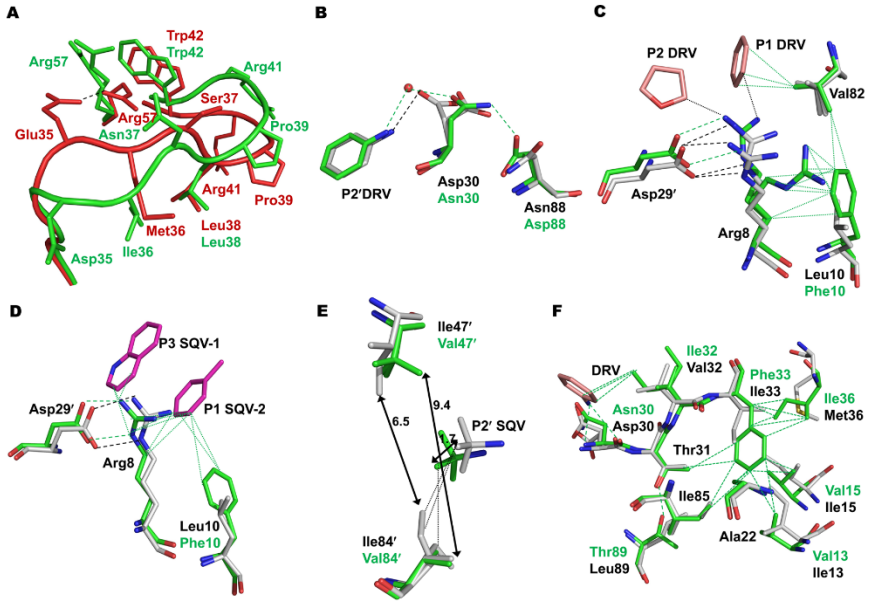


Figure 5. Comparison of PR20 and wild-type PR reveals structural rearrangements associated with mutations. (A) Structural changes in the flap hinge region of PR20/DRV (green) complex in comparison to PR/DRV complex (red). (B) The direct hydrogen bond between Asp30 and DRV in the wild-type complex (gray carbons) is replaced in PR20/DRV (green carbons) by a water-mediated interaction between D30N and the major conformation of DRV and a hydrogen bond interaction with N88D. (C) L10F mutation disrupts the intersubunit ion pair between the major conformation of Arg8 and Asp29′ in both the monomers of PR20/DRV (green) due to rotation of the guanidinium group of Arg8. PR/DRV is shown in gray carbons and DRV groups in red bonds. (D) In PR20/SQV (green carbons), the P1 group of SQV-2 (magenta) inserts between Arg8 and L10F and prevents the movement of Arg8. (E) I47′V and I84′V mutations in PR20/SQV (green carbons) widen the S2′ pocket by more than 2 Å. (F) The L33F (green carbons) mutation connects the active site with the hydrophobic core by hydrophobic interactions indicated by dotted lines. The L89T mutation forms a new hydrogen bond with the hydroxyl of Thr31 which lies between the inhibitor interacting mutations D30N and I32V.

the closed flap, however, due to rotation of the Arg8/8′ side chain toward mutation Phe10/10′. In the PR20/DRV monomers, the Arg8 side chains exhibit two alternate conformations (0.6 and 0.4 occupancy). The major conformation of the Arg8 side chain rotates to form multiple van der Waals contacts with the edge of the aromatic Phe side chain of L10F, which abolishes the ion pair with Asp29′ and the hydrophobic contacts with DRV (Figure 5C). This rotation of Arg8 away from Asp29′ is prevented in the PR20/SQV complex, since the P2 group of SQV-2 bound at the second site inserts between the side chains of Arg8 and L10F mutation (Figure 5D). Similar interactions of a symmetry-related SQV with Arg8′ and Phe10′ occur in the other monomer. Thus, the open flap conformation of PR20 retains the strong intersubunit ion pairs between Arg8/8′ and Asp29′/29, but engagement of substrate or inhibitor can lock down the flaps and eliminate these intersubunit ion pairs, which weakens Arg8/8′ interactions with inhibitors. L10F has a critical role in this reorganization, which is consistent with the identification of L10F as an accessory mutation associated with resistance to several PIs.[6] By serendipity, the PR20/SQV structure reveals a possible mechanism to prevent the L10F-mediated elimination of the intersubunit ion pair by introducing a large hydrophobic group designed to insert between L10F and Arg8.

Multiple Mutations Combine To Expand the S2/S2′ Subsites. The most significant structural changes in the active site cavity occur as an expansion of the S2/S2′ binding sites of the three PR20 inhibitor complexes due to coordinated effects of multiple mutations. First, mutation of four (D30N, V32I, I47V, and I84V) of the seven residues forming the S2/S2′ subsites alters their size, shape, and charge. In particular, mutation of Ile47 and Ile84 to smaller residues increases space within the subsites. The distances between the side chain of residues 47 and 84 on opposite sides of the S2/S2′ subsites in the PR20/DRV complex increase by about 2 Å over the corresponding distances in the wild-type complex. Similarly, the distances between residues 47 and 84 in PR20/SQV (Figure 5E) and PR20/p2-NC$_{closed}$ complexes expand by ∼2−4 Å relative to the wild-type structures revealing a wider S2/S2′ pocket in all PR20 inhibitor complexes.

The expanded S2 and S2′ pockets of PR20/p2-NC$_{closed}$ eliminate two hydrogen bonds of Gly48 and 48′ with p2-NC and permit rotation of the P2′Gln side chain removing three hydrogen bonds with Asp29′ and 30′ seen in wild-type PR/p2-NC (Figure 4). In the PR20/SQV complex, the P2′ group of SQV shifts more than 1.5 Å toward the shorter I84′V to compensate for the larger S2′ subsite and maintain interactions with D30N, V32I, I47V, and I84V (Figure 5E). The I84V mutation in PR20/SQV eliminates hydrophobic contacts with the P1′ and P1 groups of SQV, similar to the effects of I84V in PR20/DRV and in the single I84V mutant complex.[16] Mutation I47V shows different effects with DRV and SQV. The shorter

dx.doi.org/10.1021/bi2018317 | *Biochemistry* 2012, 51, 2819−2828

side chain of Val47 eliminates van der Waals contacts with the P2 group of DRV and with Ile50 that occur in the wild-type complex. In the PR20/SQV structure, however, I47V retains the contacts with SQV seen in the PR/SQV. Mutation V32I on the opposite side of the S2/S2′ subsite restores hydrophobic contact with the side chain of I47V, while showing little effect on interactions with inhibitors.

The expanded binding site of PR20 is coupled with the networks of mutations around the inhibitor-binding residues. The interactions formed by second shell residues Thr31 and L33F and 6 distal mutations maintain the main chain of residues 30−32 further from the inhibitor as shown by ∼1 Å increased separation between the Cα atoms of residues 32 and 50 in the PR20 inhibitor complexes compared to the equivalent wild-type structures. Notably, the large aromatic side chain of L33F protrudes into the hydrophobic core of the protein enhancing hydrophobic contacts to distal mutations of I13V, I15V, M36I, and second shell residues Ala22 and Ile85 (Figure 5F). Also, the hydroxyl side chain of Thr31 forms a new hydrogen bond with the hydroxyl side chain introduced by distal mutation L89T (Figure 5F). Mutations of Leu89 are associated with resistance to DRV and TPV.[6] Similarly, the larger L10F side chain adds contacts with Val82 in PR20 (Figure 5C). The hydrophobic contacts of Val82 and Ile85 with L10F and L33F may contribute to the decreased interactions of the 80's loop with inhibitor and the flaps observed in PR20 complexes. Mutations of remote residues like I13V, I15V, M36I, and L89T act to accommodate the larger side chains of mutated Phe10 and Phe33 in the hydrophobic core and coordinate with second and first shell mutations to expand the inhibitor binding site and decrease hydrophobic contacts with inhibitors.

Implications for Drug Resistance. The mutations and structural changes in PR20 act to increase its dimer dissociation constant (K_d) by a factor ≥30 and reduce the k_{cat}/K_m by ∼10-fold, due mainly to a 13-fold increase in K_m relative to wild-type PR.[11] These observations are consistent with the enlarged binding site cavity and increased flap mobility in PR20, which may influence binding of substrates and inhibitors. Notably, however, the effects of these factors on inhibitor binding are much larger than the 10-fold decrease in catalytic efficiency for substrate hydrolysis. For example, the binding of PR20 to DRV is 4000−8000 times weaker (larger dissociation constant) than that of wild-type PR; their precursors also exhibit a similar >250-fold difference in binding affinity to DRV.[11] Consequently, PR20 is not compromised in precursor processing even in the presence of clinical inhibitors and hence is highly resistant to inhibition by these drugs.

The additional mutations in PR20 compared with the nine in MDR769[24] appear to enhance its resistance to drugs. In comparison with the wild-type enzyme, clinical PIs exhibit 1.8−590-fold weaker inhibition of MDR769[48] and even weaker binding to PR20.[11] Although PR20 and MDR769[24] exhibit similar twisting of the hinge loop due to the M36I mutation, PR20 bears other mutations in the hinge loop, such as I33F and E35D, that can alter the flexibility of its flaps. Similarly, mutations in the active site cavity of PR20 (D30N, V32I, I47V, and I84V) alter the charge and increase the distance between the flaps and the base of the cavity relative to the values in MDR769 (V82A and I84V). Furthermore, PR20 possesses a network of internal mutations (L10F, I13V, I15V, L33F, M36I, N88D, and L89T) that restrain the conformation of residues in the inhibitor binding site. It is of interest that responsiveness of PR20 to DRV inhibition is not restored by reverting 1 or 2 mutant residues to wild type, suggesting that its extreme drug resistance requires the simultaneous selection of multiple mutations.[11]

One approach to overcome multidrug resistance is the design of PIs to better accommodate the changes induced by mutations, especially the widely separated flaps and expanded S2/S2′ pocket revealed in PR20. The recently reported potent PR inhibitor GRL-0519 with a larger tris-THF moiety as the P2 group compared to the bis-THF group in DRV may provide a good example for further development.[49] Also, drugs with novel interactions with the flaps may be effective to neutralize the changes due to hinge mutations and flap variability. Moreover, the structure of $PR20_{open}$ with yttrium at the active site provides insights for design of drugs that bind the wide open conformation. For example, compounds designed to target the open PR conformation include metallacarboranes and pyrrolidine-based inhibitors.[50,51]

■ ASSOCIATED CONTENT

Supporting Information

A second inhibitor is seen bound to PR20 in PR20/DRV and PR20/SQV complexes in addition to the inhibitors bound at the active site. This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

The structure coordinates and factors have been deposited in the Protein Data Bank with access codes 3UCB for PR20/DRV, 3UFN for PR20/SQV, 3UHL for PR20/p2-NC, and 3UF3 for PR20.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: iweber@gsu.edu; phone: 404 413-5411; fax: 404 413-5301.

Funding

This research was supported, in whole or in part, by the Intramural Research Program of the NIDDK, National Institutes of Health (NIH), Intramural AIDS-Targeted Antiviral Program of the Office of the Director, NIH, and Grant GM062920 from the NIH. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract W-31-109-Eng-38.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Chen-Hsiang Shen was supported in part by the Georgia State University Molecular Basis of Disease Fellowship and the Georgia State University Research Program Enhancement Award in Bioinformatics. This research was authored, in whole or in part, by National Institutes of Health staff. Data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) beamline 22ID at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at www.ser-cat.org/members.html. DRV and SQV were obtained through the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH.

■ ABBREVIATIONS

HIV-1, human immunodeficiency virus type 1; PR, mature HIV-1 protease; PR20, mature HIV-1 protease with 20 mutations; TFR, transframe region; PI, clinical inhibitor of PR; DRV, darunavir; SQV, saquinavir; TPV, tipranavir; DSC, differential scanning calorimetry; rmsd, root-mean-square deviation.

■ REFERENCES

(1) Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, R. A., Scolnick, E. M., and Sigal, I. S. (1988) Active human immunodeficiency virus protease is required for viral infectivity. *Proc. Natl. Acad. Sci. U. S. A. 85*, 4686−4690.

(2) Ridky, T., and Leis, J. (1995) Development of drug resistance to HIV-1 protease inhibitors. *J. Biol. Chem. 270*, 29621−29623.

(3) Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M., and Markowitz, M. (1995) Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature 373*, 123−126.

(4) Croteau, G., Doyon, L., Thibeault, D., McKercher, G., Pilote, L., and Lamarre, D. (1997) Impaired fitness of human immunodeficiency virus type 1 variants with high-level resistance to protease inhibitors. *J. Virol. 71*, 1089−1096.

(5) Mammano, F., Trouplin, V., Zennou, V., and Clavel, F. (2000) Retracing the evolutionary pathways of human immunodeficiency virus type 1 resistance to protease inhibitors: virus fitness in the absence and in the presence of drug. *J. Virol. 74*, 8524−8531.

(6) Johnson, V. A., Brun-Vezinet, F., Clotet, B., Gunthard, H. F., Kuritzkes, D. R., Pillay, D., Schapiro, J. M., and Richman, D. D. (2010) Update of the Drug Resistance Mutations in HIV-1: December 2010. *Top. HIV Med. 18*, 156−163.

(7) Winters, M. A., and Merigan, T. C. (2005) Insertions in the human immunodeficiency virus type 1 protease and reverse transcriptase genes: clinical impact and molecular mechanisms. *Antimicrob. Agents Chemother. 49*, 2575−2582.

(8) Kozisek, M., Saskova, K. G., Rezacova, P., Brynda, J., van Maarseveen, N. M., De Jong, D., Boucher, C. A., Kagan, R. M., Nijhuis, M., and Konvalinka, J. (2008) Ninety-nine is not enough: molecular characterization of inhibitor-resistant human immunodeficiency virus type 1 protease mutants with insertions in the flap region. *J. Virol. 82*, 5869−5878.

(9) Maguire, M. F., Guinea, R., Griffin, P., Macmanus, S., Elston, R. C., Wolfram, J., Richards, N., Hanlon, M. H., Porter, D. J., Wrin, T., Parkin, N., Tisdale, M., Furfine, E., Petropoulos, C., Snowden, B. W., and Kleim, J. P. (2002) Changes in human immunodeficiency virus type 1 Gag at positions L449 and P453 are linked to I50V protease mutants in vivo and cause reduction of sensitivity to amprenavir and improved viral fitness in vitro. *J. Virol. 76*, 7398−7406.

(10) Agniswamy, J., and Weber, I. T. (2009) HIV-1 protease: structural prespectives on drug resistance. *Viruses 1*, 1110−1136.

(11) Louis, J. M., Aniana, A., Weber, I. T., and Sayer, J. M. (2011) Inhibition of autoprocessing of natural variants and multidrug resistant mutant precursors of HIV-1 protease by clinical inhibitors. *Proc. Natl. Acad. Sci. U. S. A. 108*, 9072−9077.

(12) Dierynck, I., De Wit, M., Gustin, E., Keuleers, I., Vandersmissen, J., Hallenberger, S., and Hertogs, K. (2007) Binding kinetics of darunavir to human immunodeficiency virus type 1 protease explain the potent antiviral activity and high genetic barrier. *J. Virol. 81*, 13845−13851.

(13) Sayer, J. M., and Louis, J. M. (2009) Interactions of different inhibitors with active-site aspartyl residues of HIV-1 protease and possible relevance to pepsin. *Proteins 75*, 556−568.

(14) Sayer, J. M., Agniswamy, J., Weber, I. T., and Louis, J. M. (2010) Autocatalytic maturation, physical/chemical properties, and crystal structure of group N HIV-1 protease: relevance to drug resistance. *Protein Sci. 19*, 2055−2072.

(15) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307−326.

(16) 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.

(17) McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 61*, 458−464.

(18) Storoni, L. C., McCoy, A. J., and Read, R. J. (2004) Likelihood-enhanced fast rotation functions. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 432−438.

(19) Kovalevsky, A. Y., Tie, Y., Liu, F., Boross, P. I., Wang, Y. F., Leshchenko, S., Ghosh, A. K., Harrison, R. W., and Weber, I. T. (2006) Effectiveness of nonpeptide clinical inhibitor TMC-114 on HIV-1 protease with highly drug resistant mutations D30N, I50V, and L90M. *J. Med. Chem. 49*, 1379−1387.

(20) Sheldrick, G. M., and Schneider, T. R. (1997) SHELXL: high-resolution refinement. *Methods Enzymol. 277*, 319−343.

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

(22) Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 53*, 240−255.

(23) 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., Sect. D: Biol. Crystallogr. 63*, 866−875.

(24) 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.

(25) Sayer, J. M., Liu, F., Ishima, R., Weber, I. T., and Louis, J. M. (2008) Effect of the active site D25N mutation on the structure, stability, and ligand binding of the mature HIV-1 protease. *J. Biol. Chem. 283*, 13459−13470.

(26) Woon, T. C., Brinkworth, R. I., and Fairlie, D. P. (1992) Inhibition of HIV-1 proteinase by metal ions. *Int. J. Biochem. 24*, 911−914.

(27) Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J., and Kent, S. B. (1989) Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science 245*, 616−621.

(28) Kear, J. L., Blackburn, M. E., Veloro, A. M., Dunn, B. M., and Fanucci, G. E. (2009) Subtype polymorphisms among HIV-1 protease variants confer altered flap conformations and flexibility. *J. Am. Chem. Soc. 131*, 14650−14651.

(29) Tie, Y., Boross, P. I., Wang, Y. F., Gaddis, L., Liu, F., Chen, X., Tozser, J., Harrison, R. W., and Weber, I. T. (2005) Molecular basis for substrate recognition and drug resistance from 1.1 to 1.6 angstroms resolution crystal structures of HIV-1 protease mutants with substrate analogs. *FEBS J. 272*, 5265−5277.

(30) 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.

(31) Swairjo, M. A., Towler, E. M., Debouck, C., and Abdel-Meguid, S. S. (1998) Structural role of the 30's loop in determining the ligand specificity of the human immunodeficiency virus protease. *Biochemistry 37*, 10928−10936.

(32) Meiselbach, H., Horn, A. H., Harrer, T., and Sticht, H. (2007) Insights into amprenavir resistance in E35D HIV-1 protease mutation from molecular dynamics and binding free-energy calculations. *J. Mol. Model. 13*, 297−304.

(33) Ode, H., Matsuyama, S., Hata, M., Neya, S., Kakizawa, J., Sugiura, W., and Hoshino, T. (2007) Computational characterization of structural role of the non-active site mutation M36I of human immunodeficiency virus type 1 protease. *J. Mol. Biol. 370*, 598−607.

(34) Sanches, M., Krauchenco, S., Martins, N. H., Gustchina, A., Wlodawer, A., and Polikarpov, I. (2007) Structural characterization of

B and non-B subtypes of HIV-protease: insights into the natural susceptibility to drug resistance development. *J. Mol. Biol. 369*, 1029−1040.
(35) Clemente, J. C., Moose, R. E., Hemrajani, R., Whitford, L. R., Govindasamy, L., Reutzel, R., McKenna, R., Agbandje-McKenna, M., Goodenow, M. M., and Dunn, B. M. (2004) Comparing the accumulation of active- and nonactive-site mutations in the HIV-1 protease. *Biochemistry 43*, 12141−12151.
(36) Tie, Y., Kovalevsky, A. Y., Boross, P., Wang, Y. F., Ghosh, A. K., Tozser, J., Harrison, R. W., and Weber, I. T. (2007) Atomic resolution crystal structures of HIV-1 protease and mutants V82A and I84V with saquinavir. *Proteins 67*, 232−242.
(37) Kovalevsky, A. Y., Liu, F., Leshchenko, S., Ghosh, A. K., Louis, J. M., Harrison, R. W., and Weber, I. T. (2006) Ultra-high resolution crystal structure of HIV-1 protease mutant reveals two binding sites for clinical inhibitor TMC114. *J. Mol. Biol. 363*, 161−173.
(38) Tie, Y., Wang, Y. F., Boross, P. I., Chiu, T. Y., Ghosh, A. K., Tozser, J., Louis, J. M., Harrison, R. W., and Weber, I. T. (2011) Critical differences in HIV-1 and HIV-2 protease specificity for clinical inhibitors. *Protein Sci. 21*, 339−350.
(39) Louis, J. M., Zhang, Y., Sayer, J. M., Wang, Y. F., Harrison, R. W., and Weber, I. T. (2011) The L76V Drug Resistance Mutation Decreases the Dimer Stability and Rate of Autoprocessing of HIV-1 Protease by Reducing Internal Hydrophobic Contacts. *Biochemistry 50*, 4786−4795.
(40) Saskova, K. G., Kozisek, M., Rezacova, P., Brynda, J., Yashina, T., Kagan, R. M., and Konvalinka, J. (2009) Molecular characterization of clinical isolates of human immunodeficiency virus resistant to the protease inhibitor darunavir. *J Virol. 83*, 8810−8818.
(41) Bandaranayake, R. M., Kolli, M., King, N. M., Nalivaika, E. A., Heroux, A., Kakizawa, J., Sugiura, W., and Schiffer, C. A. (2010) The effect of clade-specific sequence polymorphisms on HIV-1 protease activity and inhibitor resistance pathways. *J. Virol. 84*, 9995−10003.
(42) Mitsuya, Y., Winters, M. A., Fessel, W. J., Rhee, S. Y., Hurley, L., Horberg, M., Schiffer, C. A., Zolopa, A. R., and Shafer, R. W. (2006) N88D facilitates the co-occurrence of D30N and L90M and the development of multidrug resistance in HIV type 1 protease following nelfinavir treatment failure. *AIDS Res. Hum. Retroviruses 22*, 1300−1305.
(43) Mahalingam, B., Wang, Y. F., Boross, P. I., Tozser, J., Louis, J. M., Harrison, R. W., and Weber, I. T. (2004) Crystal structures of HIV protease V82A and L90M mutants reveal changes in the indinavir-binding site. *Eur. J. Biochem. 271*, 1516−1524.
(44) Shen, C. H., Wang, Y. F., Kovalevsky, A. Y., Harrison, R. W., and Weber, I. T. (2010) Amprenavir complexes with HIV-1 protease and its drug-resistant mutants altering hydrophobic clusters. *FEBS J. 277*, 3699−3714.
(45) Kozisek, M., Bray, J., Rezacova, P., Saskova, K., Brynda, J., Pokorna, J., Mammano, F., Rulisek, L., and Konvalinka, J. (2007) Molecular analysis of the HIV-1 resistance development: enzymatic activities, crystal structures, and thermodynamics of nelfinavir-resistant HIV protease mutants. *J. Mol. Biol. 374*, 1005−1016.
(46) Mahalingam, B., Louis, J. M., Reed, C. C., Adomat, J. M., Krouse, J., Wang, Y. F., Harrison, R. W., and Weber, I. T. (1999) Structural and kinetic analysis of drug resistant mutants of HIV-1 protease. *Eur. J. Biochem. 263*, 238−245.
(47) Manchester, M., Everitt, L., Loeb, D. D., Hutchison, C. A. III, and Swanstrom, R. (1994) Identification of temperature-sensitive mutants of the human immunodeficiency virus type 1 protease through saturation mutagenesis. Amino acid side chain requirements for temperature sensitivity. *J. Biol. Chem. 269*, 7689−7695.
(48) 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.
(49) Ghosh, A. K., Xu, C. X., Rao, K. V., Baldridge, A., Agniswamy, J., Wang, Y. F., Weber, I. T., Aoki, M., Miguel, S. G., Amano, M., and Mitsuya, H. (2010) Probing multidrug-resistance and protein-ligand interactions with oxatricyclic designed ligands in HIV-1 protease inhibitors. *ChemMedChem 5*, 1850−1854.
(50) Kozisek, M., Cigler, P., Lepsik, M., Fanfrlik, J., Rezacova, P., Brynda, J., Pokorna, J., Plesek, J., Gruner, B., Grantz Saskova, K., Vaclavikova, J., Kral, V., and Konvalinka, J. (2008) Inorganic polyhedral metallacarborane inhibitors of HIV protease: a new approach to overcoming antiviral resistance. *J. Med. Chem. 51*, 4839−4843.
(51) 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.

Appendix C

Journal of Medicinal Chemistry

Article

pubs.acs.org/jmc

Novel P2 Tris-tetrahydrofuran Group in Antiviral Compound 1 (GRL-0519) Fills the S2 Binding Pocket of Selected Mutants of HIV-1 Protease

Hongmei Zhang,[†] Yuan-Fang Wang,[†] Chen-Hsiang Shen,[†] Johnson Agniswamy,[†] Kalapala Venkateswara Rao,[∥] Chun-Xiao Xu,[∥] Arun K. Ghosh,[∥] Robert W. Harrison,[†,‡] and Irene T. Weber*[,†,§]

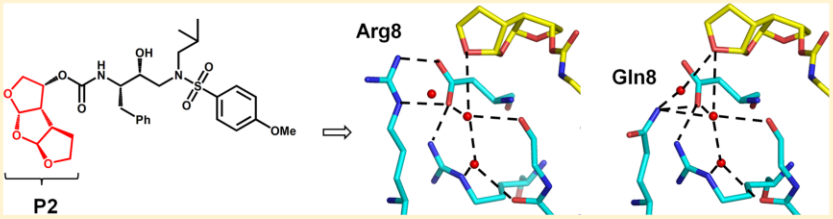
[†]Department of Biology, Georgia State University, Atlanta, Georgia 30303, United States

[‡]Department of Computer Science, Molecular Basis of Disease Program, Georgia State University, Atlanta, Georgia 30303, United States

[§]Department of Chemistry, Molecular Basis of Disease Program, Georgia State University, Atlanta, Georgia 30303, United States

[∥]Department of Chemistry and Department of Medicinal Chemistry, Purdue University, West Lafayette, Indiana 47907, United States

Supporting Information



ABSTRACT: GRL-0519 (**1**) is a potent antiviral inhibitor of HIV-1 protease (PR) possessing tris-tetrahydrofuran (tris-THF) at P2. The high resolution X-ray crystal structures of inhibitor **1** in complexes with single substitution mutants PR_{R8Q}, PR_{D30N}, PR_{I50V}, PR_{I54M}, and PR_{V82A} were analyzed in relation to kinetic data. The smaller valine side chain in PR_{I50V} eliminated hydrophobic interactions with inhibitor and the other subunit consistent with 60-fold worse inhibition. Asn30 in PR_{D30N} showed altered interactions with neighboring residues and 18-fold worse inhibition. Mutations V82A and I54M showed compensating structural changes consistent with 6−7-fold lower inhibition. Gln8 in PR_{R8Q} replaced the ionic interactions of wild type Arg8 with hydrogen bond interactions without changing the inhibition significantly. The carbonyl oxygen of Gly48 showed two alternative conformations in all structures likely due to the snug fit of the large tris-THF group in the S2 subsite in agreement with high antiviral efficacy of **1** on resistant virus.

■ INTRODUCTION

HIV/AIDS is a life-threatening disease that interferes with the immune system, and approximately 1.8 million people died of AIDS-related illnesses in 2010.[1] Although there is no effective vaccine available,[2] highly active antiretroviral therapy (HAART), which employs a combination of different antiretroviral drugs, improves the lives of AIDS patients.[3] HIV-1 protease (PR) is critical for viral particle maturation because it cleaves the viral precursor polypeptides Gag and Gag-Pol into the mature structural and enzymatic proteins.[4,5] Thus PR is an effective target for antiviral drugs, however, the most severe challenge to the long-term efficacy of protease inhibitors (PIs) in HAART is the emergence of drug-resistant mutants of PR.[6]

HIV-1 PR is catalytically active as a homodimer. Structural regions critical for PR activity and stability are the dimer interface including the catalytic Asp25 from each subunit and the flexible flaps comprising residues 45−55.[7,8] To date, there are nine approved clinical PIs. The first clinical inhibitors, such as saquinavir (SQV), were designed to bind tightly in the active site cavity of the wild-type enzyme; however, their binding affinity can be readily lowered by mutations. Analysis of the structural and biochemical properties of PR mutants suggests that resistant mutations act by multiple mechanisms, including mutations in the binding site that directly lower inhibitor affinity, mutations at the dimer interface that destabilize the catalytically active dimer, and flap mutations that alter the conformational flexibility.[7] Drug resistant PR mutants exhibit decreased binding affinity for inhibitors while maintaining the critical PR function in viral replication.[9] Two clinical PIs, darunavir (DRV) and amprenavir (APV), contain tetrahydro-

Received: October 18, 2012
Published: January 8, 2013

furan (THF) in the P2 group; APV has a single THF, while DRV incorporates bis-THF.

The bis-THF of DRV introduces more hydrogen bonds with PR main chain atoms, and DRV has demonstrated high potency and clinical efficacy on resistant viral infections.[10,11] Recently, a third THF ring was added to enlarge P2 and fit better in the S2 binding pocket of PR, leading to the novel PI called GRL-0519 (**1**) (Figure 1A). The incorporation of the third ring endows inhibitor **1** with excellent antiviral activity on drug resistant virus.[12]

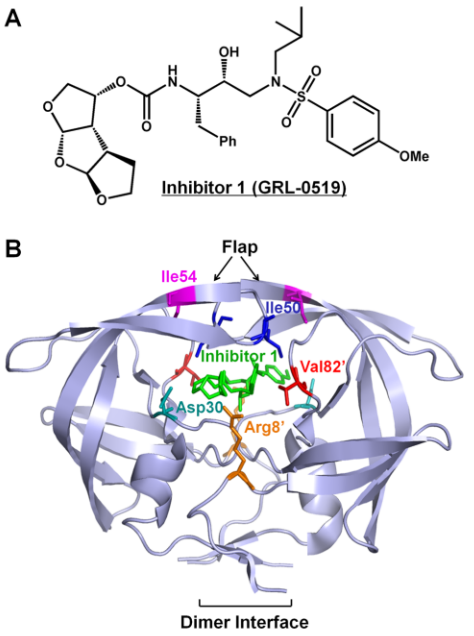


Figure 1. (A) Chemical structures of protease inhibitor compound **1**. (B) Structure of HIV-1 PR_{WT}/inhibitor **1**. The HIV-1 protease dimer is shown in light-blue cartoon representation. The inhibitor **1** and wild-type residues at the mutation sites are indicated by differently colored sticks. The same residues on the two subunits are shown in the same color with only one of them labeled.

The crystal structure of inhibitor **1** complexed with wild-type PR (PR_{WT}) was reported previously.[12] The third THF ring showed new water molecule-mediated hydrogen bonds with conserved PR residues Gly27, Asp29, and Arg8′. To study the molecular basis for the potency of inhibitor **1** against drug resistant viral strains, crystal structures of inhibitor **1** complexes with PR mutants bearing single substitutions of R8Q, D30N, I50V, I54M, and V82A (PR_{R8Q}, PR_{D30N}, PR_{I50V}, PR_{I54M}, and PR_{V82A}) were analyzed. The location of these mutations in the PR dimer is indicated in Figure 1B. These mutations, with the exception of R8Q, are common in drug resistant clinical isolates.[13] R8Q was one of the first resistant mutants identified in the laboratory for an investigational inhibitor.[14] In the wild-type enzyme, Arg8 forms an ionic interaction with Asp29′ in the other subunit as an important component of the dimer interface.[15,16] This intersubunit ionic interaction was eliminated in the mutant with the single substitution of R8Q.[17] Moreover, in the PR_{WT}–inhibitor **1** complex, the side chain of Arg8 forms a water molecule-mediated hydrogen bond with the third THF of inhibitor **1**.[12] Therefore, it is of particular interest to test how the R8Q mutation affects the binding of inhibitor **1**. D30N is a major mutation that is associated with resistance to nelfinavir (NFV).[18] Asp30 forms hydrogen bond interactions with the bis-THF of DRV, thus mutation of this residue may alter the inhibitor binding. Mutations of the flap residues such as Ile50 and Ile54 can alter the conformational dynamics of this region, thereby affecting the binding affinity for inhibitors.[19−22] Ile50 is located at the tip of the flap where its side chain forms hydrophobic interactions with inhibitors. Mutation of I50V to a shorter side chain is expected to reduce the binding affinity for inhibitors. Indeed, PR with I50V mutation exhibits significantly reduced inhibition by indinavir, SQV, and DRV.[21−23] I50V also has a significant effect in destabilizing the PR dimer.[21] Mutations of Val82 are found frequently in resistant virus.[13] The mutation V82A in the active site cavity can eliminate interactions with inhibitor and also exhibits a shift of its main chain atoms to adapt to inhibitor.[24−26] Here, the inhibitor **1** complexes with PR mutants PR_{R8Q}, PR_{D30N}, PR_{I50V}, PR_{I54M}, and PR_{V82A} are analyzed in relation to the PR_{WT}–inhibitor **1** complex and the inhibition values.

RESULTS

Compound 1 Inhibition of PR_{WT} and Selected Mutants. The kinetic parameters and inhibition constants (K_i) of compound **1** for PR_{WT} and selected mutants PR_{R8Q}, PR_{D30N}, PR_{I50V}, PR_{I54M}, and PR_{V82A} are shown in Table 1. The catalytic efficiency (k_{cat}/K_m) of PR_{V82A} showed 2.7-fold increase, PR_{I54M} was essentially identical, and PR_{R8Q} had a slight decrease relative to PR_{WT}. The lowest catalytic efficiency of 10% of PR_{WT} was observed for PR_{D30N} and PR_{I50V}. The relative catalytic efficiency of PR_{D30N} measured here is consistent with previous reports using a different substrate.[23] The relative activities of PR_{I50V} and PR_{V82A} are slightly different from those reported earlier; however, the trends are identical, with decreased k_{cat}/K_m for PR_{I50V} and increased k_{cat}/K_m for

Table 1. Kinetic Parameters for Substrate Hydrolysis and Inhibition of Compound 1

protease	K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m ($min^{-1}\cdot\mu M^{-1}$)	relative k_{cat}/K_m	K_i (nM)	relative K_i
PR_{WT}[a]	30 ± 5	190 ± 20	6.5 ± 1.30	1.0	0.50 ± 0.06	1
PR_{R8Q}	38 ± 2	161 ± 4	4.2 ± 0.25	0.6	1.20 ± 0.08	2
PR_{D30N}	76 ± 8	48 ± 2	0.6 ± 0.72	0.1	8.9 ± 0.4	18
PR_{I50V}[a]	109 ± 8	68 ± 5	0.6 ± 0.03	0.1	30.9 ± 0.6	60
PR_{I54M}[a]	41 ± 5	300 ± 40	7.3 ± 0.80	1.1	3.5 ± 0.2	7
PR_{V82A}[b]	29 ± 3	512 ± 26	17.7 ± 2.00	2.7	3.2 ± 0.2	6

[a] K_m and k_{cat} values previously reported in ref 22. [b] K_m and k_{cat} values previously reported in ref 41.

dx.doi.org/10.1021/jm301519z | J. Med. Chem. 2013, 56, 1074−1083

Table 2. Crystallographic Statistics for PR Mutants in Complex with Inhibitor 1

inhibitor **1** complex	PR_{R8Q}	PR_{D30N}	PR_{I50V}	PR_{I54M}	PR_{V82A}
space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
unit cell dimensions (Å)					
a	58.37	58.72	58.59	58.78	58.61
b	85.96	86.24	85.95	85.84	86.15
c	46.18	45.90	46.07	46.02	46.33
resolution range (Å)	50−1.46	50−1.49	50−1.22	50−1.06	50−1.29
unique reflections	40949	37721	64041	103379	59015
R_{merge} (%) overall (final shell)	5.3 (49.7)	6.2 (42.1)	5.3(47.4)	6.3 (56.8)	5.6 (40.7)
$I/\sigma(I)$ overall (final shell)	30.8 (2.6)	25.5 (2.1)	31(2.1)	26.7 (2.1)	18.9 (2.2)
completeness (%) overall (final shell)	99.2 (94.8)	97.0 (76.3)	91.5(65.4)	97.4 (85.6)	99.0 (92.7)
data range for refinement (Å)	10−1.46	10−1.49	10−1.22	10−1.06	10−1.29
R (%)	15.8	15.6	16.5	14.9	16.3
R_{free} (%)	21.0	21.4	20.3	17.8	20.9
no. of solvent atoms (total occupancies)	175(170.5)	189 (144.2)	251(237.6)	248(234.5)	182(179.0)
RMS deviation from ideality					
bonds (Å)	0.010	0.010	0.014	0.016	0.012
angle distance (deg)	0.029	0.030	0.034	0.038	0.031
average *B*-factors (Å^2)					
main chain atoms	15.0	17.5	12.9	12.6	13.0
side chain atoms	21.1	24.1	19.6	19.1	19.8
solvent	31.6	33.7	30.8	28.8	30.5
inhibitor	22.6	25.3	19.5	17.1	19.3
relative occupancy of **1**	0.55/0.45	0.55/0.45	0.7/0.3	0.53/0.47	0.6/0.4

PR_{V82A}. Compound **1** showed a range of inhibition of up to 60-fold for the various mutants relative to the wild-type enzyme. The inhibition constant of PR_{R8Q} for compound **1** was not significantly different from the value of 0.5 nM for PR_{WT}. The PR_{I54M} and PR_{V82A} showed 7- and 6-fold increased K_i for compound **1** relative to PR_{WT}. Again, the mutants PR_{D30N} and PR_{I50V} showed the most significant changes with 18- and 60-fold worse inhibition, respectively.

Crystallographic Analysis of Inhibitor 1 Complexes. Crystal structures of PR mutants PR_{R8Q}, PR_{D30N}, PR_{I50V}, PR_{I54M}, and PR_{V82A} were obtained in complex with inhibitor **1**, and the structures were refined at resolutions of 1.06−1.49 Å. The crystallographic data collection and refinement statistics are summarized in Table 2. All five complexes were determined in space group $P2_12_12$ and refined with anisotropic *B*-factors, solvent molecules, and hydrogen atoms to the *R*-factors of 13.8−16.3. The asymmetric unit contains one PR dimer of residues labeled 1−99 and 1′−99′ for each monomer. Well-defined electron density was observed for the PR residues, as illustrated for the mutated residues in Figure 2. The mutated residues, except for Ala82, show alternative conformations. The side chain of mutated residue Asn30/Asn30′ in both subunits of PR_{D30N}−inhibitor **1** structure has two alternative conformations with relative occupancies of 0.55/0.45. Alternative conformations of the side chains of mutated Met54 and Gln8 with relative occupancies of 0.58/0.42 and 0.55/0.45, respectively, were observed in one subunit in the PR_{I54M} and PR_{R8Q} complexes. Mutated residue Val50 showed alternative conformations for both main chain and side chain atoms in both subunits of PR_{I50V}−inhibitor **1** with relative occupancies of 0.7/0.3. Inhibitor **1** has two alternative conformations related by 180° in all five complexes with relative occupancies listed in Table 2.

The residues 48−51/52 at the tip of the flaps show alternative conformations in all the complexes, which was associated with approximately 180° rotation of the peptide

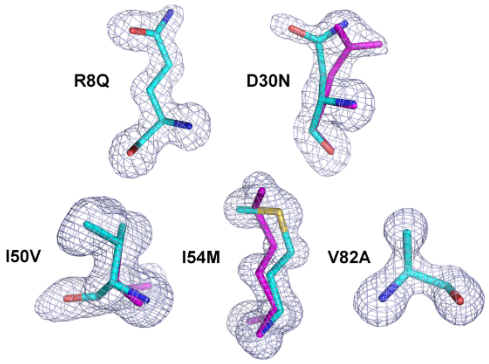


Figure 2. $F_o - F_c$ omit maps (contoured at 3.0 σ) for the mutated residues. The magenta sticks indicate the alternative conformations of Asn30, Val50, and Met54.

bond between Ile50 and Gly51 and alternative conformations of the carbonyl group of Gly48 in both subunits. The electron density map of residues Gly48−Gly52 from the atomic resolution (1.06 Å) structure of PR_{I54M}−inhibitor **1** illustrates clearly the alternative conformations (Figure 3A). Corresponding regions of the other mutant structures are shown in Supporting Information Figure S1. Similar disorder in the flaps was observed in the wild-type complex with inhibitor **1**. However, alternative conformations are unusual for Gly48 in complexes with other inhibitors. Here, in each conformation of the flap residues, the carbonyl oxygen of Gly48 forms several C−H···O interactions with the tris-THF P2 group and Gly48′ forms a C−H···O interaction with the P2′ aniline in the corresponding conformation of inhibitor **1** (Figure 3B). The

dx.doi.org/10.1021/jm301519z | J. Med. Chem. 2013, 56, 1074−1083

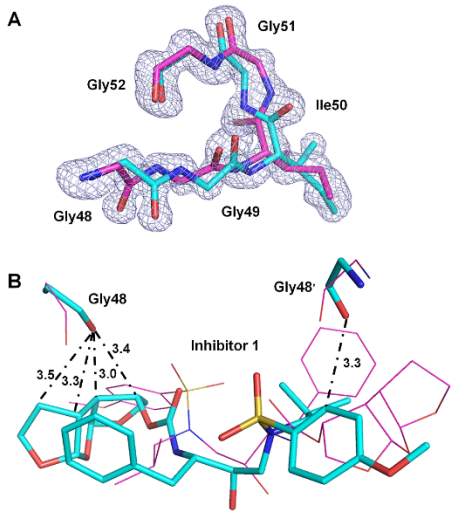


Figure 3. (A) $F_o - F_c$ omit map (contoured at 3.0 σ) for the flap residues Gly48–Gly52 in complex PR_{I54M}–inhibitor **1** (1.06 Å). Carbon atoms are colored cyan and magenta for the major and minor conformations, respectively. (B) Interactions of Gly48/48′ with inhibitor **1** in complex PR_{I50V}–inhibitor **1** (0.7/0.3 occupancy). The major and minor conformations of Gly48/48′ and the inhibitor **1** are represented by sticks with carbon atoms colored cyan and lines with carbon atoms colored magenta, respectively. C–H···O interactions between the major conformations are indicated by dash–dotted lines with distances in Å. C–H···O interactions between the minor conformations are the same but not shown here for clarity.

two conformations of Gly48 may arise from steric hindrance due to the tight fit of the tris-THF group in the S2 subsite.

The mutant dimers superimpose on the PR_{WT}–inhibitor **1** complex (PDB ID: 3OK9) with pairwise root-mean-square deviations (rmsd) of 0.13–0.24 Å on 198 C_α atoms. Both alternative inhibitor conformations are considered, except for the minor conformation of inhibitor in PR_{I50V}–inhibitor **1**, which has the low relative occupancy of 0.3. Hydrogen bond interactions of inhibitor **1** with all five mutants are similar except for variation in interactions of residue 30 (Supporting Information Table S1). The side chain of Asn30/30′ in PR_{D30N} forms a hydrogen bond interaction with the first THF oxygen of inhibitor **1** as observed for the major conformation of inhibitor **1** in the PR_{WT} complex; however, this interaction was not observed for Asp30/Asp30′ in the other four mutants. Interatomic distances for hydrogen bond interactions (O–H···O, N–H···O, etc.) are considered in the range of 2.4–3.4 Å, C–H···O interactions at 3.0–3.7 Å separation, C–H···π interactions for distances of <4.0 Å, and van der Waals contacts when distances are 3.8–4.2 Å, as described previously.[27] Structural changes are described for each mutant separately or in related pairs in the following sections.

Influence of D30N on PR Structure and Its Interaction with Inhibitor 1. Residue 30 is located at one end of the active site cavity and interacts with the inhibitor, although the side chain frequently has alternate conformations. In PR_{WT}–inhibitor **1**, the side chain and main chain amide of the major conformation of Asp30′ form hydrogen bond interactions with the first THF ring of inhibitor **1** (Figure 4A). The main chain of residue 30′ lies in essentially the same position in PR_{D30N}–

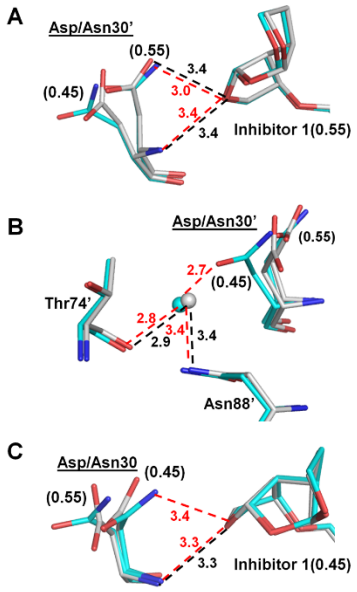


Figure 4. Structural changes of PR_{D30N}. Carbon atoms are colored gray for PR_{WT} –inhibitor **1** and cyan for PR_{D30N}–inhibitor **1**. Hydrogen bond interactions in PR_{WT}–inhibitor **1** and PR_{D30N}–inhibitor **1** are represented by black and red dashed lines, respectively, with distances in Å. Water molecules are shown as gray and cyan spheres for PR_{WT}–inhibitor **1** and PR_{D30N}–inhibitor **1**, respectively. Occupancies for alternative conformations are labeled in parentheses for Asp/Asn30′. (A) Interactions of Asp/Asn30′ with the major conformation of inhibitor **1**. (B) Interactions of Asp/Asn30′ with the neighboring residues. (C) Interactions of Asp/Asn30 with the minor conformation of inhibitor **1**.

inhibitor **1** and PR_{WT}–inhibitor **1**, thus the hydrogen bond interaction of the main chain amide of Asp30′ with O26 of inhibitor **1** is preserved in the mutant (Figure 4A). In addition, the major conformation of the side chain (occupancy of 0.55) of residue 30′ was similar in both structures and retained the hydrogen bond interaction with inhibitor. However, the minor conformation (occupancy of 0.45) of the side chain differs in the mutant and the wild-type enzyme, especially the position of Oδ1 atom, due to the Cγ movement of approximately 0.9 Å as well as rotation of the side chain (approximately 90°) of Asn30′ compared to Asp30′ of PR_{WT}. As a result, the side chain of Asn30′ in the mutant forms a hydrogen bond with a water molecule, which further interacts with the side chain of Asn88′ as well as the main chain of Thr74′ (Figure 4B). This water molecule also forms hydrogen bond interactions with the main chain and side chain of Thr31′ in both wild-type and mutant (not shown in Figure 4B). In the other subunit (1–99), the Asn30 side chain in PR_{D30N} forms a hydrogen bond with the minor conformation of inhibitor **1**, which was not observed in PR_{WT} (Figure 4C). The interactions of Asn30 with neighboring residues are similar to those shown in Figure 4C for Asn30′. It

dx.doi.org/10.1021/jm301519z | J. Med. Chem. 2013, 56, 1074–1083

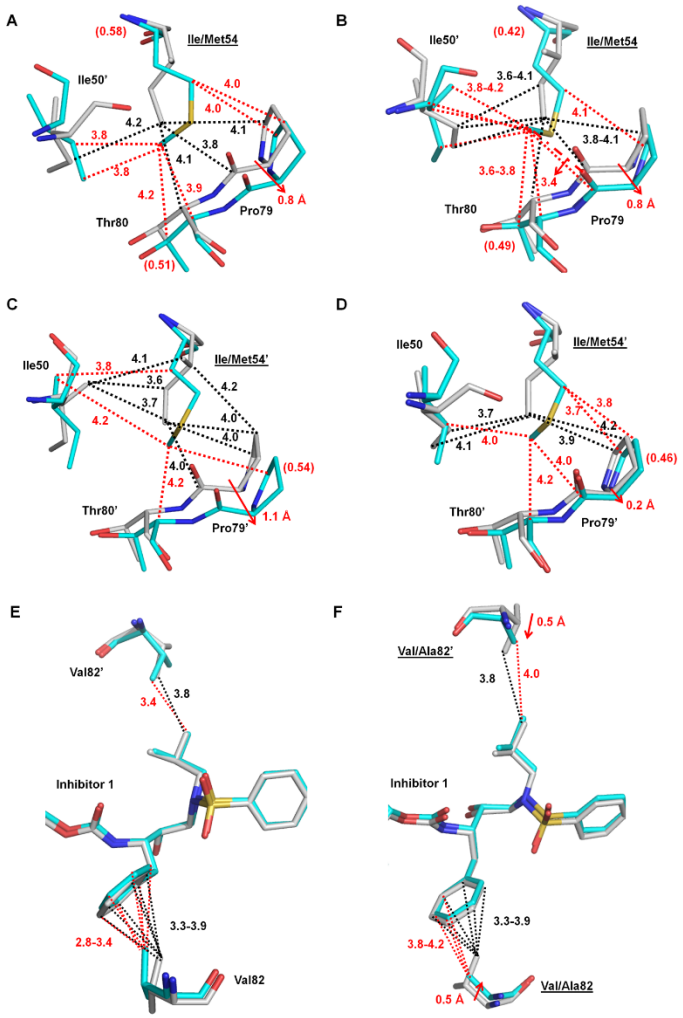


Figure 5. Structural changes of PR_{I54M} and PR_{V82A}. Carbon atoms are colored gray for PR_{WT}–inhibitor **1** and cyan for PR_{I54M}–inhibitor **1** or PR_{V82A}–inhibitor **1**. Interactions are indicated by black lines for PR_{WT}–inhibitor **1**, and red lines for the mutants. van der Waals and C–H···π interactions are indicated by dotted lines. (A,B) Interactions of the major (A) and minor (B) conformations of Ile/Met54 with Ile50′ and 80s loop residues in subunit 1–99. (C,D) Interactions of Ile/Met54′ with Ile50 and the major (C) and minor (D) conformations of 80′s loop residues in subunit 1′–99′. Relative occupancies for alternative conformations of residues are shown in parentheses. Red arrows indicate the shifts of the mutants relative to PR_{WT}. (E) Interactions of Val82/82′ with inhibitor **1** in PR_{WT} and PR_{I54M}. (F) Interactions of Val/Ala82/82′ with inhibitor **1** in PR_{WT} and PR_{V82A}.

seems that the observed 18-fold increase in the inhibition constant of the inhibitor **1** for PR_{D30N} does not arise from changes in the direct interactions between residue 30 and the inhibitor. Instead, the elimination of the negative charge of Asp30 and alterations within an internal network of interactions may provide indirect effects on inhibition. Similar changes in interactions of residues 30 and 88 have been reported in mutants containing D30N and N88D and also proposed to effect inhibition.[28,29]

Mutations I54M and V82A Produce Compensating Shifts in the 80s Loop Residues. Ile54, which is located in the flap region, does not contact inhibitors directly, but it forms many hydrophobic interactions with surrounding residues, including residues in the 80s loop (residues 78–82) and Ile50′. In the subunit comprising residues 1–99, Met54 had two alternative conformations as shown in Figure 2. Compared with Ile54 in PR_{WT}, both the major (Figure 5A) and minor (Figure 5B) conformations of Met54 in PR_{I54M} formed one more

dx.doi.org/10.1021/jm301519z | J. Med. Chem. 2013, 56, 1074−1083

intersubunit van der Waals contact with the major and minor conformations of Ile50′. The 80s loop in PR_{I54M} has shifted to accommodate the larger side chain of Met54 compared with Ile54 in PR_{WT}. Pro79 has shifted away from residue 54 by about 0.8 Å at the Cα atom to maintain their van der Waals interactions. The major conformation of Met54 (relative occupancy of 0.58) formed the same number of van der Waals contacts with Pro79 (Figure 5A). In addition, it formed one new van der Waals interaction with the major conformation of Thr80. The minor conformation of Met54 (relative occupancy of 0.42) formed the same number of van der Waals contacts with the proline ring (Figure 5B); however, it gained a new C−H···O interaction with the main chain carbonyl oxygen atom of Pro79 and one more van der Waals interaction with the minor conformation of Thr80 compared with Ile54 in PR_{WT}. In the other subunit 1′-99′, Met54′ had a single conformation and formed one less intersubunit van der Waals contact with both the major (Figure 5C) and minor (Figure 5D) conformations of Ile50 compared with Ile54′ in PR_{WT}. Unlike subunit 1−99, the 80′s loop residue Pro79′ in subunit 1′−99′ exhibited two alternative conformations. A larger shift of approximately 1.1 Å was observed for the major conformation of Pro79′, whereas the minor conformation showed an insignificant change of 0.2 Å. Although Met54′ had less van der Waals interactions with the major conformation of Pro79′, it gained a new one with Thr80′ (Figure5C). The interactions of residue 54′ with the minor conformations of proline were identical for PR_{WT} and the mutant (Figure 5D). However, it gained new interactions with the C atoms of Pro79′ and Thr80′. Taken together, there are either two conformations of Met54 or the 80s loop residues to accommodate the large side chain of the mutated Met54. Ile50 exhibits two alternative conformations in all structures as mentioned previously. Met54 in PR_{I54M} has a similar number of intersubunit van der Waals contacts with Ile50′ as does Ile54 in PR_{WT} (Figures 5A−D). The separation of Pro79 and residue 54 is increased in the mutant so that their van der Waals interaction is retained. The new interactions between Met54 and Thr80 may help maintain the conformation of the rest of the 80s loop residues despite the shift of Pro79/79′. As shown in Figure 5E, the hydrophobic interactions of Val82 and 82′ with the P1 and P1′ groups of inhibitor **1** are retained in PR_{WT} and PR_{I54M}. Therefore, the similar catalytic efficiency of PR_{I54M} and PR_{WT} may reflect the similar internal contacts of residue 54. Moreover, the hydrogen bond between the major conformation of inhibitor **1** and Asp30 in PR_{WT} was not observed in PR_{I54M}, which may contribute to the 7-fold lower inhibition for PR_{I54M}.

Val82/82′ has extensive C−H···π interactions with the P1 phenyl group and a van der Waals contact with the P1′ group of inhibitor **1** in PR_{WT} (Figure 5F). Mutation V82A introduces the smaller Ala side chain, leading to the loss of C−H···π interactions with the P1 phenyl group of inhibitor **1** compared with those in the wild-type complex. However, the main chain of Ala82′ has shifted toward the inhibitor by around 0.5 Å compared to the wild-type complex, which preserves the hydrophobic interaction with the P1′ of inhibitor **1** (Figure 5F). The interactions of Val82/82′ with the major and minor conformations of the inhibitor were identical, thus only those with the major conformation of the inhibitor **1** are shown. Similar to PR_{I54M}, the hydrogen bond between the major conformation of inhibitor **1** and Asp30 in PR_{WT} was also not observed in PR_{V82A}. The absence of a hydrogen bond plus fewer interactions of Ala82 with the P1 phenyl of inhibitor **1** may be responsible for the decreased inhibitory activity of inhibitor **1** for PR_{V82A}. However, the compensation mechanisms observed in PR_{I54M} and PR_{V82A} may explain why there was only 6−7-fold decrease of the inhibition constants for these two mutants relative to PR_{WT}.

Mutations I50V and R8Q Alter the Dimer Interface and Interactions with Inhibitor 1. The side chain of Ile50′ in PR_{WT} had extensive interactions with inhibitor **1**, including C−H···O and C−H···π interactions (Figure 6A). However, the

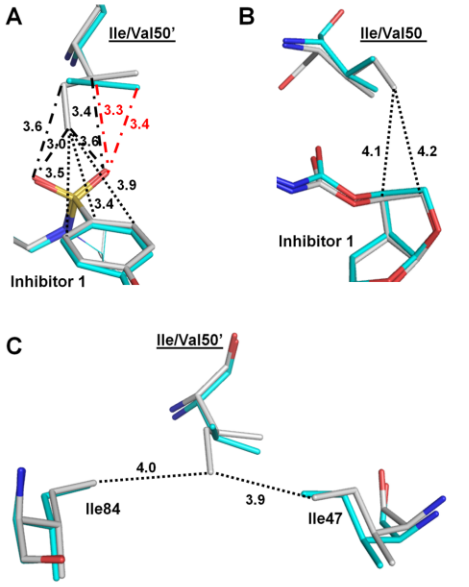


Figure 6. Structural changes of PR_{I50V}. (A) Interactions of Ile/Val50′ with inhibitor **1**. (B) Interactions of Ile/Val50 with inhibitor **1**. (C) Intersubunit interactions of Ile/Val50′ with surrounding residues. Carbon atoms are colored gray for PR_{WT}−inhibitor **1** and cyan for PR_{I50V}−inhibitor **1**. Interactions in PR_{WT}−inhibitor **1** and PR_{I50V}−inhibitor **1** are represented by black lines and red lines, respectively. C−H···O and van der Waals/C−H···π interactions are indicated by dash−dotted lines and dotted lines, respectively, with distances in Å.

interactions with the P2′ group of inhibitor were eliminated for the small Val50′ side chain in the mutant PR_{I50V} and only two C−H···O interactions with the sulfonamide oxygen were maintained. Similarly, the two van der Waals interactions between Ile50 and the P2 tris-THF group of inhibitor **1** were lost in the other subunit of PR_{I50V} (Figure 6B). Similar to PR_{I54M} and PR_{V82A}, the hydrogen bond between the major conformation of inhibitor **1** and Asp30 in PR_{WT} was also absent in PR_{I50V}. Taken together, the considerably reduced interactions of mutant PR_{I50V} with inhibitor **1** are consistent with the significantly decreased inhibition (60-fold) for PR_{I50V} compared to the PR_{WT}.

The mutation of Ile50 to Val50 also affects the intersubunit interactions. In PR_{WT}, the Ile50′ side chain had hydrophobic interactions with the side chains of Ile84 and Ile47, and the same intersubunit interactions were also observed for Ile50.

dx.doi.org/10.1021/jm301519z | J. Med. Chem. 2013, 56, 1074−1083

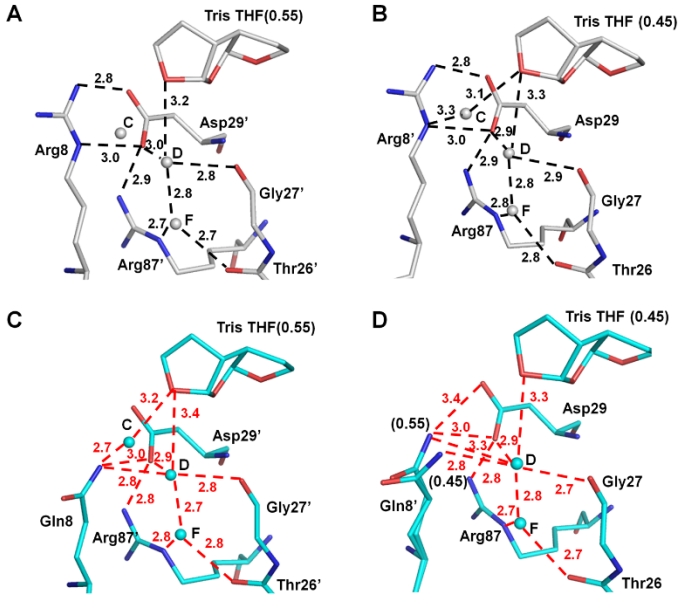


Figure 7. Structural changes of PR_{R8Q}. (A,B) Interactions of Arg8/8′ in PR_{WT}–inhibitor **1**. (C,D) Interactions of Gln8/8′ in PR_{R8Q}–inhibitor **1**. Carbon atoms are colored gray for PR_{WT}–inhibitor **1** and cyan for PR_{R8Q}–inhibitor **1**. Water molecules are shown as gray and cyan spheres for PR_{WT}–inhibitor **1** and PR_{R8Q}–inhibitor **1**, respectively. The hydrogen bond interactions are indicated by the dashed lines with distances in Å. Occupancies for alternative conformations of inhibitor **1** and Gln8′ are indicated in parentheses. Hydrogen bond interactions in R_{R8Q}–inhibitor **1** are colored red. Interactions between water molecules C and D in all figures (A–D) are omitted for clarity.

However, the interactions of Val50′ with both Ile84 and Ile47 are eliminated in the PR_{I50V} mutant as illustrated in Figure 6C. The loss of the interaction between Val50 and Ile47′ was also observed in the other subunit. These findings are consistent with our previous analysis of PR_{I50V} structures,[21] and the overall decreased interaction between subunits may account for the 10-fold decrease of its catalytic efficiency as well as likely contributing to the significantly worse inhibition relative to PR_{WT}.

In contrast to mutant PR_{I50V}, mutation of R8Q has altered the type of interaction rather than eliminated interactions. In the wild-type PR, Arg8′ forms an ionic interaction with Asp29 in both subunits (Figure 7A,B). However, this intersubunit ionic interaction is eliminated in PR_{R8Q} (Figure 7C,D). In one subunit of PR_{R8Q}, a direct hydrogen bond is formed between the side chains of Gln8 and Asp29′ (Figure 7C). The side chain of Gln8 in PR_{R8Q} has shifted away from Arg8 in PR_{WT} by approximately 0.9 Å, thus gaining new hydrogen bond interactions with conserved water molecules C and D and a water molecule-mediated interaction with the tris-THF of inhibitor. In the other subunit, Gln8′ has two alternative conformations with relative occupancies of 0.55/0.45 (Figure 7D). The side chain of Gln8′ in the major conformation forms two direct and one water molecule-mediated hydrogen bond with Asp29, whereas that of the minor conformation forms one water molecule-mediated hydrogen bond with Asp29. Overall, direct as well as water molecule-mediated intersubunit hydrogen bonds between Gln8 and Asp29′ were observed in both subunits, suggesting that these new interactions might compensate for the loss of the intersubunit ionic interactions in the wild-type enzyme (and many other mutants).

The conserved water molecule D also has a hydrogen bond interaction with the oxygen of the third THF ring of inhibitor **1** in both PR_{WT} and PR_{R8Q} (Figure 7). This water molecule mediates a hydrogen bond interaction between the inhibitor **1** and Gln8/8′, which is not seen for Arg8/8′ in PR_{WT} (Figure 7C,D). Moreover, conserved water molecule C mediates an additional hydrogen bond between the major conformation of inhibitor **1** and the side chain of Gln8, which resembles the interaction between the minor conformation of inhibitor **1** and Arg8′ in PR_{WT} (Figure 7B).

In summary, although PR_{R8Q} has lost the strong intersubunit ionic interaction between Arg8 and Asp29′, compensation is provided by new direct and water molecule-mediated hydrogen bonds. Meanwhile, even though the hydrogen bond between the major conformation of inhibitor **1** and Asp30 in PR_{WT} was not seen in PR_{R8Q}, the water molecule-mediated hydrogen bond interaction between inhibitor **1** and Gln8/8′ was enhanced to some extent. Therefore, there is little absolute change in the interaction of PR_{R8Q} with inhibitor **1**. In contrast, PR_{I50V} showed significant loss of interactions with inhibitor **1**. These structural changes explain why inhibitor **1** has 60-fold worse inhibition for PR_{I50V}, while its inhibition for PR_{R8Q} was not affected significantly.

■ DISCUSSION AND CONCLUSIONS

Inhibitor **1** is a novel inhibitor incorporating a unique tris-THF group at P2 to target drug resistant HIV-1 PR mutants.

dx.doi.org/10.1021/jm301519z | J. Med. Chem. 2013, 56, 1074−1083

Previously, inhibitor **1** was reported to possess potent antiviral activity against wild-type HIV-1 and multidrug-resistant strains.[12] Here, structural and kinetic analyses are described for inhibitor **1** with selected PR mutants. The single mutants showed various effects. Substitution of the smaller side chains in PR_{I50V} and PR_{V82A} resulted in the loss of their interactions with inhibitor; however, mutation V82A caused a shift of its main chain atoms to compensate for the loss. Mutation to larger side chain in PR_{I54M} pushed the Pro79 away from residue 54 maintaining their interactions; moreover, new interactions were formed between Met54 and Thr80. The elimination of the interactions caused by the change of the residue size in PR_{I50V} and PR_{V82A} may be related to their worse inhibition (K_i) by inhibitor **1**. In addition, the mechanisms PR_{V82A} and PR_{I54M} adopted to compensate for the loss of the interactions may account for the moderate decrease (6−7-fold) in K_i for PR_{V82A} and PR_{I54M} compared to significant reduction (60-fold) in K_i for PR_{I50V}. Moreover, intersubunit interactions were lost in PR_{I50V}, which is expected to contribute to the worse inhibition. Mutations in PR_{D30N} and PR_{R8Q} eliminated the negative and positive charges of Asp30 and Arg8, respectively. In PR_{R8Q}, new hydrogen bond interactions of Gln8 replaced the intersubunit ionic interaction of Arg8 with Asp29′. Meanwhile, Gln8 formed new water molecule-mediated hydrogen bonds with inhibitor **1**, which may explain why inhibitor **1** has similar inhibition for PR_{R8Q} and PR_{WT}. In terms of hydrogen bonds with inhibitor **1**, PR_{WT} and all five mutants are similar except for variation in interactions with residue 30 due to the mobility of its side chain. The variation in Asp30/30′ and its interaction with inhibitor did not appear to be a major factor in the K_i for these mutants.

PR_{D30N} differs from the other mutants because the major changes are loss of negative charge for Asp30 and altered internal interactions of residue 30/30′ to account for 18-fold decrease of its K_i value for inhibitor **1**. D30N mutation is the major mutation associated with resistance to NFV,[13] which has been suggested to arise from the change of conformational entropy upon inhibitor binding.[29] Our structural analysis also showed that mutation D30N introduced changes in its interactions with neighboring residues Thr74 and Asn88. In PR_{D30N}, the side chain of Asn30 formed water molecule-mediated hydrogen bond interactions with both Thr74 and Asn88, whereas these interactions were absent in PR_{WT}−inhibitor **1**. Moreover, the proteolytic activity is sensitive to mutations of residues Gly86−Arg87−Asn88.[16,30] Considering the positive association of mutation D30N and N88D,[31] it would be interesting to study how N88D influences the structure of D30N in complex with this new inhibitor **1**.

Compound **1** is an inhibitor designed to fill the S2 binding pocket of PR by incorporating a third THF ring in the P2 group relative to DRV (Figure 1A). As analyzed in its complex with PR_{WT}[12] (Figure 7A,B), the oxygen of the third THF ring of inhibitor **1** gains a hydrogen bond interaction with a conserved water molecule (D in Figure 7A) and helps stabilize the intersubunit interaction of Arg8′ with Asp29 and other surrounding residues. In the mutant PR_{R8Q}, the interaction of the newly incorporated ring of inhibitor **1** with this conserved network was maintained, which is consistent with almost identical inhibition for PR_{WT} and PR_{R8Q}. Mutation R8Q, however, has rarely been found in resistant clinical isolates. DRV has a K_i value of 6.6 nM for mutant PR_{D30N} and up to 18 nM for PR_{I50V}, consistent with resistance to virus with these mutations.[22,23] Similarly, our study suggests that viral strains containing D30N or I50V are likely to show resistance to inhibitor **1** given their K_i values of 8.9 and 30.9 nM for PR_{D30N} and PR_{I50V}, respectively. Inhibitor **1** shows moderate effects on the other two mutants, PR_{V82A} and PR_{I54M}, similar to their effects with DRV.[22,26] Therefore, mutations D30N, I50V, I54M, and V82A exhibit similar effects on both inhibitors, although inhibitor **1** is less effective than DRV on PR_{I50V}. Mutation R8Q exhibits higher sensitivity to compound **1** than other mutations studied here, suggesting that compound **1** may be a good inhibitor against viral strains bearing R8Q mutation. Drug resistant clinical isolates generally have more than one mutation, however, and previous studies of single and double mutations indicate that the changes in structure and other properties of the respective single mutants may not be preserved in the double mutants.[29,32] More importantly, the influence of drug resistant mutants on the viral infectivity depends not only on the inhibitor sensitivity, but also on the protease activity, especially for those mutations lacking direct interactions with the inhibitors.[33] Their study showed that mutant PR_{I50V} had the most negative effect on viral infectivity. Our previous studies suggested that the loss of dimer stability due to I50V mutation was a major contributor to the diminished catalytic activity and inhibition of PR_{I50V}.[21] Mutant PR_{D30N} also showed ∼50% loss of infectivity,[33] in agreement with our previous observation of altered activity on different precursor cleavage sites.[17] Therefore, a single mutation may have a variety of effects on inhibition, dimer stability, catalytic activity, and viral replication. Finally, because inhibitor **1** possesses the large tris-THF P2 group, our structural analysis suggests it may fit better with the mutants that have an expanded S2/S2′ pocket, such as the extreme multiple mutant PR20.[28]

■ EXPERIMENTAL SECTION

General. Inhibitor **1** has shown analytical purity of >98% by HPLC.[12] The structure was confirmed by ^{1}H and ^{13}C NMR spectral analysis and high resolution mass spectrometry. HRMS (ESI) $[M + Na]^+$ calcd for $C_{30}H_{40}N_2O_9SNa$, 627.2352; found, 627.2359.

Protein Expression and Purification. To prevent PR autoproteolysis, a PR clone (Genbank HIVHXB2CG) engineered with five mutations (Q7K, L33I, L63I, C67A, and C95A) was used as a template. Mutants (R8Q, D30N, I50V, I54M, and V82A) were generated by using the Quick-Change mutagenesis kit (Stratagene, La Jolla, CA). The expression, purification, and refolding were performed as previously described.[34,35]

Enzyme Kinetics. Kinetic parameters were determined by a fluorescence assay. The fluorogenic substrate was Abz-Thr-Ile-Nle-*p*-nitro-Phe-Gln-Arg-NH_2, where Abz is anthranilic acid and Nle is norleucine (Bachem, King of Prussia, PA, USA), with the sequence derived from the p2/NC cleavage site of the Gag polyprotein. Then 10 μL of protease (final concentration of 70−120 nM) diluted in 98 μL of reaction buffer (100 mM Mes, pH 5.6, 400 mM sodium chloride, 1 mM EDTA and 5% glycerol) and 2 μL of dimethylsulfoxide or inhibitor (dissolved in dimethylsulfoxide) were incubated at 37 °C for 5 min. The reaction was initialized by adding 90 μL of substrate. The reaction was monitored over 5 min in the POLARstar OPTIMA microplate reader at wavelengths of 340 and 420 nm for excitation and emission. Data analysis was performed using the program sigmaplot 9.0 (SPSS Inc., Chicago, IL, USA). K_m and k_{cat} values were obtained by standard data fitting with the Michaelis−Menten equation. The K_i value was obtained from the IC_{50} values estimated from an inhibitor dose−response curve using the equation $K_i = (IC_{50})\ [E]/2)/(1 + [S]/K_m)$, where [E] and [S] are the PR and substrate concentrations.

Crystallographic Analysis. Crystals were grown by the hanging drop vapor diffusion method using protein solutions preincubated with inhibitor, which was dissolved in dimethylsulfoxide, in a molar ratio of

dx.doi.org/10.1021/jm301519z | J. Med. Chem. 2013, 56, 1074−1083

1:5−10. The final crystallization drop was 1.6−2 μL with reservoir solution and protein in a 1:1 ratio by volume. Crystallization conditions for different complexes were as follows: 22−24% saturated ammonium sulfate, 130−135 mM sodium phosphate, 0.05 M sodium citrate, pH 6.1 for PR_{R8Q}; 1.46 M NaCl, 0.1 M sodium citrate, pH 5.0 for PR_{D30N}; 1.26−1.46 M NaCl, 0.06 M sodium acetate, pH 5.0−5.4 for PR_{I50V}; 0.6−0.93 M NaCl, 0.06 M sodium acetate, pH 4.6−5.0 for PR_{I54M}; and 10% ammonium sulfate, 0.05 M citrate-phosphate, pH 5.6 for PR_{V82A}. The crystals for X-ray data collection were soaked in the reservoir solution with 20−30% glycerol as cryoprotectant for ∼1 min and frozen immediately in liquid nitrogen. X-ray data for all the complexes were collected on the SER-CAT beamline at the Advanced Photon Source, Argonne National Laboratory. Data were processed using HKL2000.[38] The structures were solved by molecular replacement using PHASER[37] in the CCP4i suite of programs.[38] Crystal structures were refined with SHELX97.[39] Manual adjustments and rebuilding were performed using the molecular graphics program COOT.[40,41] Figures of the structures were prepared with PYMOL (http://www.pymol.org).

■ ASSOCIATED CONTENT

Supporting Information

Hydrogen bonds between protease mutants and inhibitor **1** and figures of electron density for alternate conformations of residues 48−52 for all mutants. This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

Crystallographic data are available for inhibitor **1** complexes with: PR_{R8Q} (PDB ID: 4HEG), PR_{D30N} (PDB ID: 4HDB), PR_{I50V} (PDB ID: 4HDP), PR_{I54M} (PDB ID: 4HE9), and PR_{V82A} (PDB ID: 4HDF).

■ AUTHOR INFORMATION

Corresponding Author

*Phone: (404) 413-5411. Fax: (404) 413-5301. E-mail: iweber@gsu.edu.

Notes

The authors declare no competing financial interests.

■ ACKNOWLEDGMENTS

This work was supported in part by the Georgia State University Molecular Basis of Disease Fellowships (H.Z. and C.-H.S.), the Georgia State University Research Program Enhancement Award in Bioinformatics (C.-H.S.), and the National Institutes of Health grants GM062920 (I.W.) and GM053386 (A.K.G.). X-ray data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) beamlines 22ID (R8Q, D30N, I50V, I54M) and 22BM (V82A) at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at www.ser-cat.org/members.html. Use of the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science, under contract no. W-31-109-Eng-38.

■ ABBREVIATIONS USED

HAART, highly active antiretroviral therapy; PR_{WT}, mature HIV-1 protease; PI, clinical inhibitor of PR; DRV, darunavir; NFV, nelfinavir; tris-THF, tris-tetrahydrofuran; PR_{R8Q}, PR with R8Q mutation; PR_{D30N}, PR with D30N mutation; PR_{I50V}, PR with I50V mutation; PR_{I54M}, PR with I54M mutation; PR_{V82A}, PR with V82A mutation

■ REFERENCES

(1) *UNAIDS Report on the Global AIDS Epidemic*; UNAIDS: Geneva, 2010.

(2) Girard, M. P.; Bansal, G. P. HIV/AIDS vaccines: a need for new concepts? *Int. Rev. Immunol.* **2008**, *27*, 447−471.

(3) Brenner, B. G.; Turner, D.; Wainberg, M. A. HIV-1 drug resistance: can we overcome? *Expert Opin. Biol. Ther.* **2002**, *2*, 751−761.

(4) Darke, P. L.; Nutt, R. F.; Brady, S. F.; Garsky, V. M.; Ciccarone, T. M.; Leu, C. T.; Lumma, P. K.; Freidinger, R. M.; Veber, D. F.; Sigal, I. S. HIV-1 protease specificity of peptide cleavage is sufficient for processing of gag and pol polyproteins. *Biochem. Biophys. Res. Commun.* **1988**, *156*, 297−303.

(5) Oroszlan, S.; Luftig, R. B. Retroviral proteinases. *Curr. Top. Microbiol. Immunol.* **1990**, *157*, 153−185.

(6) Richman, D. D.; Morton, S. C.; Wrin, T.; Hellmann, N.; Berry, S.; Shapiro, M. F.; Bozzette, S. A. The prevalence of antiretroviral drug resistance in the United States. *AIDS* **2004**, *18*, 1393−1401.

(7) Weber, I. T.; Agniswamy, J. HIV-1 Protease: Structural Perspectives on Drug Resistance. *Viruses (Basel)* **2009**, *1*, 1110−1136.

(8) Weber, I. T.; Wang, Y.-F. HIV-1 Protease: Role in Viral Replication, Protein−Ligand X-Ray Crystal Structures and Inhibitor Design. In *Aspartic Acid Proteases as Therapeutic Targets*; Wiley-VCH: New York, 2010; pp 107−137.

(9) Miller, V. International perspectives on antiretroviral resistance. Resistance to protease inhibitors. *JAIDS, J. Acquired Immune Defic. Syndr.* **2001**, *26* (Suppl 1), S34−S50.

(10) Ghosh, A. K.; Kincaid, J. F.; Cho, W.; Walters, D. E.; Krishnan, K.; Hussain, K. A.; Koo, Y.; Cho, H.; Rudall, C.; Holland, L.; Buthod, J. Potent HIV protease inhibitors incorporating high-affinity P2-ligands and (*R*)-(hydroxyethylamino)sulfonamide isostere. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 687−690.

(11) Surleraux, D. L.; Tahri, A.; Verschueren, W. G.; Pille, G. M.; de Kock, H. A.; Jonckers, T. H.; Peeters, A.; De Meyer, S.; Azijn, H.; Pauwels, R.; de Bethune, M. P.; King, N. M.; Prabu-Jeyabalan, M.; Schiffer, C. A.; Wigerinck, P. B. Discovery and selection of TMC114, a next generation HIV-1 protease inhibitor. *J. Med. Chem.* **2005**, *48*, 1813−1822.

(12) Ghosh, A. K.; Xu, C. X.; Rao, K. V.; Baldridge, A.; Agniswamy, J.; Wang, Y. F.; Weber, I. T.; Aoki, M.; Miguel, S. G.; Amano, M.; Mitsuya, H. Probing multidrug-resistance and protein−ligand interactions with oxatricyclic designed ligands in HIV-1 protease inhibitors. *ChemMedChem* **2010**, *5*, 1850−1854.

(13) Johnson, V. A.; Calvez, V.; Gunthard, H. F.; Paredes, R.; Pillay, D.; Shafer, R.; Wensing, A. M.; Richman, D. D. 2011 update of the drug resistance mutations in HIV-1. *Top. Antiviral Med.* **2011**, No. 19, 156−164.

(14) Ho, D. D.; Toyoshima, T.; Mo, H.; Kempf, D. J.; Norbeck, D.; Chen, C. M.; Wideburg, N. E.; Burt, S. K.; Erickson, J. W.; Singh, M. K. Characterization of human immunodeficiency virus type 1 variants with increased resistance to a *C*2-symmetric protease inhibitor. *J. Virol.* **1994**, *68*, 2016−2020.

(15) Weber, I. T. Comparison of the crystal structures and intersubunit interactions of human immunodeficiency and Rous sarcoma virus proteases. *J. Biol. Chem.* **1990**, *265*, 10492−10496.

(16) Ishima, R.; Gong, Q.; Tie, Y.; Weber, I. T.; Louis, J. M. Highly conserved glycine 86 and arginine 87 residues contribute differently to the structure and activity of the mature HIV-1 protease. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1015−1025.

(17) Mahalingam, B.; Louis, J. M.; Reed, C. C.; Adomat, J. M.; Krouse, J.; Wang, Y. F.; Harrison, R. W.; Weber, I. T. Structural and kinetic analysis of drug resistant mutants of HIV-1 protease. *Eur. J. Biochem.* **1999**, *263*, 238−245.

(18) Jarvis, B.; Faulds, D. Nelfinavir. A review of its therapeutic efficacy in HIV infection. *Drugs* **1998**, *56*, 147−167.

(19) Liu, F.; Kovalevsky, A. Y.; Louis, J. M.; Boross, P. I.; Wang, Y. F.; Harrison, R. W.; Weber, I. T. Mechanism of drug resistance revealed by the crystal structure of the unliganded HIV-1 protease with F53L mutation. *J. Mol. Biol.* **2006**, *358*, 1191−1199.

(20) Pazhanisamy, S.; Stuver, C. M.; Cullinan, A. B.; Margolin, N.; Rao, B. G.; Livingston, D. J. Kinetic characterization of human immunodeficiency virus type-1 protease-resistant variants. *J. Biol. Chem.* **1996**, *271*, 17979−17985.
(21) Liu, F.; Boross, P. I.; Wang, Y. F.; Tozser, J.; Louis, J. M.; Harrison, R. W.; Weber, I. T. Kinetic, stability, and structural changes in high-resolution crystal structures of HIV-1 protease with drug-resistant mutations L24I, I50V, and G73S. *J. Mol. Biol.* **2005**, *354*, 789−800.
(22) Liu, F.; Kovalevsky, A. Y.; Tie, Y.; Ghosh, A. K.; Harrison, R. W.; Weber, I. T. Effect of flap mutations on structure of HIV-1 protease and inhibition by saquinavir and darunavir. *J. Mol. Biol.* **2008**, *381*, 102−115.
(23) Kovalevsky, A. Y.; Tie, Y.; Liu, F.; Boross, P. I.; Wang, Y. F.; Leshchenko, S.; Ghosh, A. K.; Harrison, R. W.; Weber, I. T. Effectiveness of nonpeptide clinical inhibitor TMC-114 on HIV-1 protease with highly drug resistant mutations D30N, I50V, and L90M. *J. Med. Chem.* **2006**, *49*, 1379−1387.
(24) Mahalingam, B.; Wang, Y. F.; Boross, P. I.; Tozser, J.; Louis, J. M.; Harrison, R. W.; Weber, I. T. Crystal structures of HIV protease V82A and L90M mutants reveal changes in the indinavir-binding site. *Eur. J. Biochem.* **2004**, *271*, 1516−1524.
(25) Tie, Y.; Kovalevsky, A. Y.; Boross, P.; Wang, Y. F.; Ghosh, A. K.; Tozser, J.; Harrison, R. W.; Weber, I. T. Atomic resolution crystal structures of HIV-1 protease and mutants V82A and I84V with saquinavir. *Proteins: Struct., Funct., Bioinf.* **2007**, *67*, 232−242.
(26) Tie, Y.; Boross, P. I.; Wang, Y. F.; Gaddis, L.; Hussain, A. K.; Leshchenko, S.; Ghosh, A. K.; Louis, J. M.; Harrison, R. W.; Weber, I. T. 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.* **2004**, *338*, 341−352.
(27) Kovalevsky, A. Y.; Liu, F.; Leshchenko, S.; Ghosh, A. K.; Louis, J. M.; Harrison, R. W.; Weber, I. T. Ultra-high resolution crystal structure of HIV-1 protease mutant reveals two binding sites for clinical inhibitor TMC114. *J. Mol. Biol.* **2006**, *363*, 161−173.
(28) Agniswamy, J.; Shen, C. H.; Aniana, A.; Sayer, J. M.; Louis, J. M.; Weber, I. T. HIV-1 protease with 20 mutations exhibits extreme resistance to clinical inhibitors through coordinated structural rearrangements. *Biochemistry* **2012**, *51*, 2819−2828.
(29) Kozisek, M.; Bray, J.; Rezacova, P.; Saskova, K.; Brynda, J.; Pokorna, J.; Mammano, F.; Rulisek, L.; Konvalinka, J. Molecular analysis of the HIV-1 resistance development: enzymatic activities, crystal structures, and thermodynamics of nelfinavir-resistant HIV protease mutants. *J. Mol. Biol.* **2007**, *374*, 1005−1016.
(30) Ishima, R.; Ghirlando, R.; Tozser, J.; Gronenborn, A. M.; Torchia, D. A.; Louis, J. M. Folded monomer of HIV-1 protease. *J. Biol. Chem.* **2001**, *276*, 49110−4916.
(31) Mitsuya, Y.; Winters, M. A.; Fessel, W. J.; Rhee, S. Y.; Hurley, L.; Horberg, M.; Schiffer, C. A.; Zolopa, A. R.; Shafer, R. W. N88D facilitates the co-occurrence of D30N and L90M and the development of multidrug resistance in HIV type 1 protease following nelfinavir treatment failure. *AIDS Res. Hum. Retroviruses* **2006**, *22*, 1300−1305.
(32) Mahalingam, B.; Boross, P.; Wang, Y. F.; Louis, J. M.; Fischer, C. C.; Tozser, J.; Harrison, R. W.; Weber, I. T. Combining mutations in HIV-1 protease to understand mechanisms of resistance. *Proteins: Struct., Funct., Bioinf.* **2002**, *48*, 107−116.
(33) Henderson, G. J.; Lee, S. K.; Irlbeck, D. M.; Harris, J.; Kline, M.; Pollom, E.; Parkin, N.; Swanstrom, R. Interplay between single resistance-associated mutations in the HIV-1 protease and viral infectivity, protease activity, and inhibitor sensitivity. *Antimicrob. Agents Chemother.* **2012**, *56*, 623−633.
(34) Wondrak, E. M.; Louis, J. M. Influence of flanking sequences on the dimer stability of human immunodeficiency virus type 1 protease. *Biochemistry* **1996**, *35*, 12957−12962.
(35) Mahalingam, B.; Louis, J. M.; Hung, J.; Harrison, R. W.; Weber, I. T. Structural implications of drug-resistant mutants of HIV-1 protease: high-resolution crystal structures of the mutant protease/substrate analogue complexes. *Proteins: Struct., Funct., Bioinf.* **2001**, *43*, 455−464.
(36) Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology, Volume 276: Macromolecular Crystallography, Part A*; Academic Press: New York, **1997**; Vol. *276*, pp 307−326.
(37) McCoy, A. J.; Grosse-Kunstleve, R. W.; Storoni, L. C.; Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2005**, *61*, 458−464.
(38) Potterton, E.; Briggs, P.; Turkenburg, M.; Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003**, *59*, 1131−1137.
(39) Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **2008**, *64*, 112−122.
(40) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126−2132.
(41) Chang, Y. C.; Yu, X.; Zhang, Y.; Tie, Y.; Wang, Y. F.; Yashchuk, S.; Ghosh, A. K.; Harrison, R. W.; Weber, I. T. Potent antiviral HIV-1 protease inhibitor GRL-02031 adapts to the structures of drug resistant mutants with its P1′-pyrrolidinone ring. *J. Med. Chem.* **2012**, *55*, 3387−3397.

Appendix D

Journal of Medicinal Chemistry

Article

pubs.acs.org/jmc

Extreme Multidrug Resistant HIV-1 Protease with 20 Mutations Is Resistant to Novel Protease Inhibitors with P1′-Pyrrolidinone or P2-Tris-tetrahydrofuran

Johnson Agniswamy,[†] Chen-Hsiang Shen,[†] Yuan-Fang Wang,[†] Arun K. Ghosh,[‡] Kalapala Venkateswara Rao,[‡] Chun-Xiao Xu,[‡] Jane M. Sayer,[§] John M. Louis,[§] and Irene T. Weber*[,†]

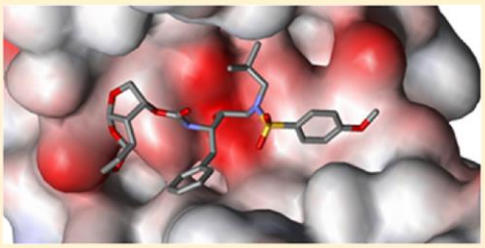
[†]Department of Biology, Molecular Basis of Disease Program, Georgia State University, Atlanta, Georgia 30303, United States

[‡]Department of Chemistry and Medicinal Chemistry, Purdue University, West Lafayette, Indiana 47907, United States

[§]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, DHHS, Bethesda, Maryland 20892-0520, United States

Supporting Information

ABSTRACT: Extreme drug resistant mutant of HIV-1 protease (PR) bearing 20 mutations (PR20) has been studied with the clinical inhibitor amprenavir (**1**) and two potent antiviral investigational inhibitors GRL-02031 (**2**) and GRL-0519 (**3**). Clinical inhibitors are >1000-fold less active on PR20 than on wild-type enzyme, which is consistent with dissociation constants (K_L) from isothermal titration calorimetry of 40 nM for **3**, 178 nM for amprenavir, and 960 nM for **2**. High resolution crystal structures of PR20−inhibitor complexes revealed altered interactions compared with the corresponding wild-type PR complexes in agreement with relative inhibition. Amprenavir lacks interactions due to PR20 mutations in the S2/S2′ subsites relative to PR. Inhibitors **2** and **3** lose interactions with Arg8′ in PR20 relative to the wild-type enzyme because Arg8′ shifts to interact with mutated L10F side chain. Overall, inhibitor **3** compares favorably with darunavir in affinity for PR20 and shows promise for further development.



INTRODUCTION

The introduction of combination therapy, also known as highly active antiretroviral therapy (HAART), has greatly improved the outcome of HIV/AIDS therapy.[1] However, the major challenge for the success of current treatment is the emergence of drug resistant strains due to the high mutational rate caused by the infidelity of HIV reverse transcriptase.[2−4] Greater transmission by drug resistant HIV strains has been recorded in several regions around the world, underscoring the therapeutic challenge.[5−11] Resistance testing is now recommended for patients in therapy.[12] With the Treatment 2.0 initiative by UNAIDS secretariat and the World Health Organization, millions of infected people will initiate or be maintained on antiretroviral therapy.[11] Furthermore, prevention strategies using antiviral drugs have the potential to increase the prevalence of HIV drug resistant strains.[13,14] More than ever, there is an urgent demand for renewed efforts to design potent inhibitors targeting the resistant viral strains.

Recently, we characterized a clinically derived HIV-1 protease (PR20) bearing 20 mutations [Q7K, L10F, I13V, I15V, D30N, V32I, L33F, E35D, M36I, S37N, I47V, I54L, Q58E, I62V, L63P, A71V, I84V, N88D, L89T, and L90M] and extremely resistant to all clinical protease inhibitors (PIs).[15,16] The inhibitor dissociation constants (K_L) of PR20 for darunavir, which was approved in 2006 for patients infected by drug resistant strains, and for saquinavir, were increased by more than 8000- and 2000-fold, respectively.[15] Thermal stability studies by DSC showed that the PIs stabilize the ternary complex of PR20 (dimer + PI) to a significantly lesser extent than for PR. In addition, the critical autocatalytic cleavage of a PR20 precursor, which is required to release fully active PR, was unresponsive to all clinical PI's, whereas the autoprocessing of wild-type precursor is inhibited by darunavir and saquinavir. Recent crystallographic studies of PR20 by itself and in complexes with darunavir, saquinavir, and a substrate analogue that mimics the p2-NC natural cleavage site in the Gag-Pol polyprotein revealed distinct conformations for its highly flexible flaps.[16] PR20 forms six fewer hydrogen bonds with the substrate analogue relative to the wild-type PR, emphasizing its weak affinity. Further, PR20 mutations result in an expanded S2/S2′ binding pocket and diminished interactions with inhibitors, in agreement with the lower affinity relative to wild-type PR.[15,16] These observations demonstrate that PR20 is a highly evolved drug resistant mutant and may

Received: February 14, 2013
Published: April 16, 2013

   dx.doi.org/10.1021/jm400231v | J. Med. Chem. 2013, 56, 4017−4027

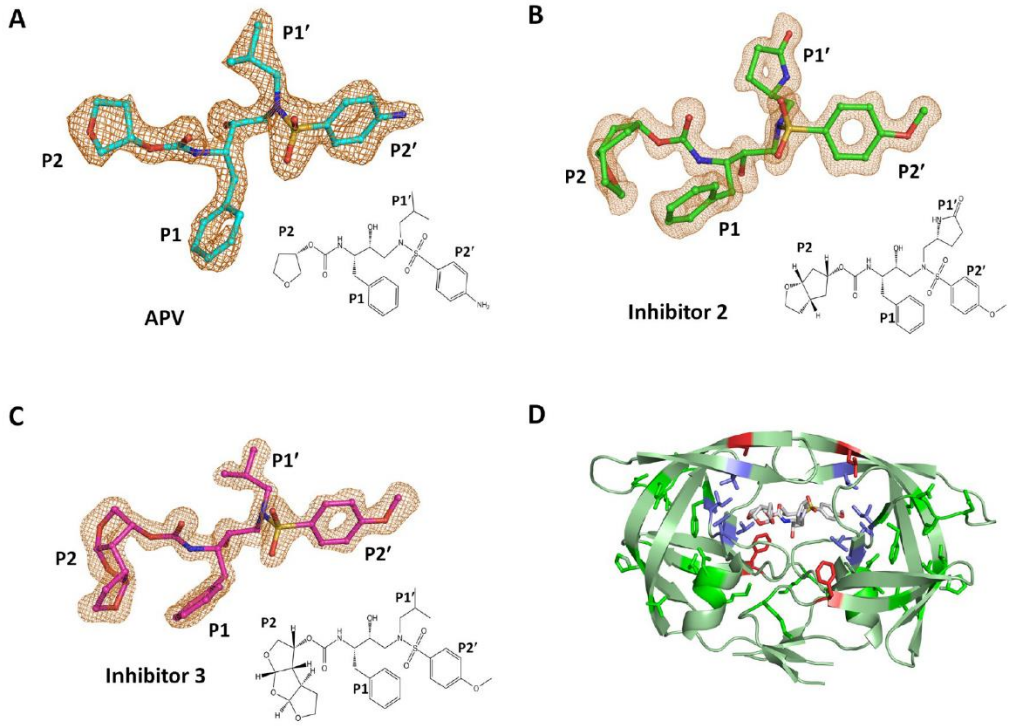


Figure 1. $F_o - F_c$ omit maps contoured at 3σ level for (A) amprenavir (APV), (B) inhibitor **2**, and (C) inhibitor **3**. The chemical structure of the inhibitors are shown in the inserts of panels (A), (B), and (C). (D) Sites of the 20 mutations of highly drug resistant PR20 are mapped on HIV-1 PR dimer (green cartoon representation), with the bound inhibitor **3** colored by atom type in stick representation. The mutations with direct interaction to inhibitors are shown as blue sticks, while the distal mutations are shown as green sticks and the critical L10F and I54L mutations are shown as red sticks.

Table 1. ITC Data for Inhibitor Binding to PR20[a]

protein	inhibitor	K_A (M^{-1})	K_L (nM)	ΔH (kcal/mol)	ΔS (cal/mol/deg)	$-T\Delta S$[b] (kcal/mol)
PR20						
	Darunavir	$24.5 \pm 5.9 \times 10^6$	41[c]	-7.61 ± 0.11	8.53	−2.57[d]
	Amprenavir	$5.61 \pm 1.69 \times 10^6$	178	-5.76 ± 0.17	11.8	−3.55[d]
	2	$1.04 \pm 0.07 \times 10^6$	960	-5.62 ± 0.08	8.86	−2.67[d]
	3	$25.3 \pm 5.3 \times 10^6$	39.5	-6.50 ± 0.07	12.3	−3.70[d]
PR[e]	Darunavir	3.2×10^{11}	0.0045	−12.1		−3.1[f]
	Amprenavir	2.5×10^9	0.39	−7.3		−5.3[f]

[a] K_A, ligand association constant; K_L ligand dissociation constant ($1/K_A$). [b] $-T\Delta S$ is given in the same units to indicate the relative contribution of entropy vs enthalpy to the overall free energy ($\Delta G = \Delta H - T\Delta S$) of ligand binding. [c] Previously reported in ref 15. [d] At 28 °C. [e] Data from ref 33. [f] At 20 °C.

prove invaluable as a model system in the quest to design potent inhibitors against multidrug resistant viral strains.

Amprenavir (inhibitor **1**; Figure 1A) is a nonpeptide *N,N*-disubstituted amino-sulfonamide protease inhibitor approved for HIV/AIDS therapy in 1999.[17] Amprenavir shares a hydroxyethylamine core with other inhibitors such as saquinavir, but the sulfonamide group increases its solubility over saquinavir which enhances oral bioavailability.[18] Recently, two new potent antiviral inhibitors of HIV PR derived on the darunavir scaffold were reported: GRL-02031 (**2**; Figure 1B) and GRL-0519 (**3**; Figure 1C).[19,20] These inhibitors were designed to maximize the protease−inhibitor interaction and promote extensive hydrogen bond interactions with PR backbone in the structure guided strategy pioneered for darunavir.[21] Inhibitor **2** was designed with an enhanced functional group at P1′ and demonstrated to form new interactions with Gly27′ and Arg8 in the S1′ subsite.[19] Inhibitor **2** was as potent as all FDA approved drugs against wild-type PR, except for darunavir, which is nearly 10-fold more potent. Moreover, **2** maintained near full potency against multidrug

dx.doi.org/10.1021/jm400231v | J. Med. Chem. 2013, 56, 4017−4027

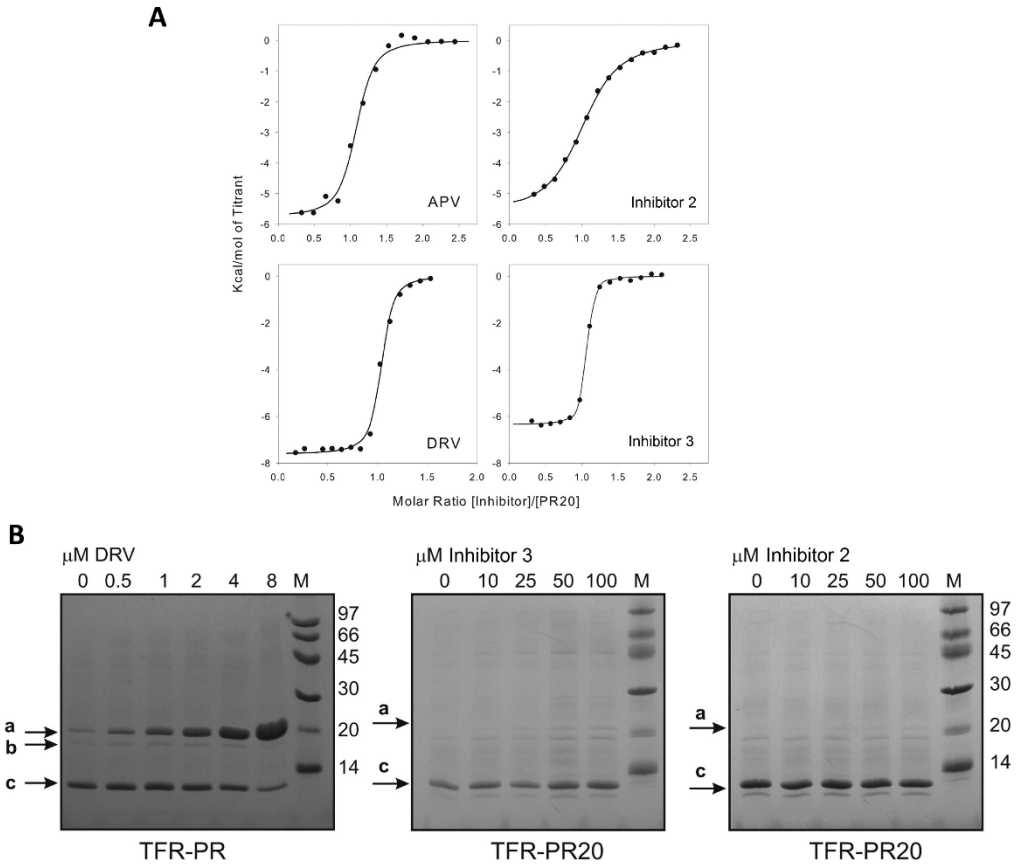


Figure 2. (A) Calorimetric titration of PR20 with inhibitors. For conditions, see Experimental Section. Quantitative results are tabulated in Table 1. (B) Left panel, dose–response for darunavir (DRV) inhibition of autocatalytic processing of wild-type TFR-PR in the *E. coli* expression system as described in ref 15. Band a, full length TFR-PR; b, cleavage product between F8/L9 within the TFR; c, mature protease. Note that virtually complete inhibition by 8 μM darunavir is accompanied by extensive buildup of the precursor. No inhibition is observed by added **3** (center panel) or **2** (right panel) on processing of the precursor TFR-PR20 to mature protease.

resistant HIV-1 isolates, while other clinical drugs, except for darunavir, showed 5–10-fold loss in potency. Inhibitor **3** was designed with a third tetrahydrofuran (THF) ring fused to the bis-THF ring at the P2 position of darunavir.[20] Inhibitor **3** displayed a 10-fold increase in potency over darunavir against multidrug resistant clinical HIV-1 strains.[20] Interestingly, both inhibitors introduce new interactions with the conserved Arg8 and Gly27 residues of PR. To elucidate the molecular basis for the excellent potency of these inhibitors against drug resistant strains, we studied the binding affinity and the crystal structures of PR20 in complexes with **2**, **3**, and the clinical drug amprenavir. The crystal structures showed several conserved interactions between the PR20 complexes and corresponding wild-type complexes but also revealed diminished interactions due to mutations in PR20. The acquired knowledge will be helpful to improve strategies for attacking extremely resistant mutants.

RESULTS

Calorimetric Analysis of Inhibitor Binding to PR20. The binding constants for the inhibitors (Table 1) were measured using isothermal titration calorimetry (ITC) by titrating PR20 with inhibitors at ~10-fold of PR20 concentration (Figure 2A and Table 1). Despite the presence of a third THF ring, the ligand dissociation constant (K_L) of **3** is similar to that reported for darunavir.[15] The K_L of amprenavir is 4.5-fold higher than that for **3**. Inhibitor **2** is the weakest binder of the set with a K_L value 5-fold higher than for amprenavir and 24-fold higher than for **3** as measured by ITC.

Inhibition Studies on Autoprocessing of TFR-PR20 Precursor by Compounds 2 and 3. We have reported previously that the autocatalytic cleavage (autoprocessing) of wild-type precursor TFR-PR to give mature PR is inhibited by saquinavir, darunavir, and atazanavir, but the autoprocessing of PR20 precursor TFR-PR20 is unresponsive to all the clinical

dx.doi.org/10.1021/jm400231v | J. Med. Chem. 2013, 56, 4017–4027

inhibitors.[15] We have further studied the effect of added **2** and **3** in the culture medium on processing of the PR20 precursor (band a) to mature PR20 (band c) as shown in Figure 2B, center and right panels, respectively. For comparison the effect of darunavir on processing of the wild-type precursor, TFR-PR, is shown in the left panel. We had previously reported that even 100−150 μM darunavir or amprenavir fails to inhibit processing of TFR-PR20 in the *E. coli* test system. Similarly, no inhibition of processing of TFR-PR20 is observed in the presence of **2** or **3** even at concentrations up to 100 μM, and thus neither of these inhibitors provides any advantage over darunavir for inhibition of PR20 precursor maturation.

The Overall Crystal Structure. PR20 was crystallized in complex with amprenavir and the new potent antiviral inhibitors **2** and **3**. The crystallographic data collection and refinement statistics are summarized in Table 2. The crystals,

Table 2. Crystallographic Data Collection and Refinement Statistics

	PR20 complexes		
	PR20/amprenavir	PR20/**2**	PR20/**3**
space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
cell dimensions			
a (Å)	28.72	28.64	28.57
b (Å)	65.80	65.97	65.79
c (Å)	94.56	93.71	92.82
resolution range	50.0−1.8	50.0−1.31	50.0−1.55
unique reflections	15764	39582	24299
redundancy	4.5 (3.1)	5.5 (2.3)	6.1 (1.8)
completeness	97.6 (80.0)[a]	90.8 (50.1)	92.5 (50.2)
$\langle I/\sigma(I)\rangle$	19.3 (2.0)	16.2 (3.0)	22.6 (2.5)
R_{sym} (%)	7.3 (46.8)	7.6 (26.3)	6.9 (26.5)
refinement resolution range	10 − 1.8	10.0 − 1.31	10.0−1.55
R_{cryst} (%)	19.8	16.8	20.0
R_{free} (%)	25.0	22.9	26.0
number of solvent molecules	102	116	114
average *B*-factor (Å^2)			
main chain	21.4	17.0	16.4
side chain	23.4	23.7	23.8
inhibitor	32.8	14.9	11.1
solvent	28.6	25.9	22.4
RMS deviations from ideality			
bond length (Å)	0.02	0.01	0.01
angles	1.8 (deg)[b]	0.03 (Å)	0.03 (Å)[c]

[a]Values in parentheses are given for the highest resolution shell. [b]The angle rmsd in REFMAC 5.2 is indicated by angle in degrees. [c]The angle rmsd in SHELX97 is indicated by distance in Å.

which are in the orthorhombic space group $P2_12_12_1$, diffracted to high resolutions of 1.3−1.8 Å and refined to *R*-factors of 16.8−20%. One PR20 homodimer with residues numbered 1-99 and 1′-99′ is in the asymmetric unit of each crystal. The three inhibitors and the 20 mutations were unambiguously visible in the electron density maps, as shown by examples in Figure 1. The 20 mutations are distributed throughout the PR tertiary structure (Figure 1D). The salient structural features of PR20 identified in the PR20/darunavir and PR20/saquinavir inhibitor complexes[16] and attributed to mutations in PR20 also appear in the new PR20 complex structures. For example, the L10F mediated breakage of intermonomer ion pair between Arg8-Asp29′ and Arg8′-Asp29[16] is observed in all three complexes with the exception of a minor conformation of Arg8 in PR20/**2** retaining an ion pair with Asp29′.

Second Binding Site for Amprenavir. Apart from the inhibitor bound at the active site, a second inhibitor molecule of 0.7 occupancy was visibly bound to one flap in the PR20/amprenavir complex. This is the first crystal structure showing amprenavir bound in a second site resembling the flap location of the second darunavir reported in the PR20/darunavir (3UCB) and PR_{V32I}/darunavir (2HS1) complexes[16,22] and as suggested by kinetic data.[23] The backbone and P2′ side chain of amprenavir and darunavir share very similar conformations in their PR20 complexes, and most of the conformational variation occurs for the P1′, P1, and P2 groups (Figure 3A). All the polar contacts with protein are conserved for darunavir and amprenavir bound in the second site of PR20. The P2′ aniline group has a van der Waals contact with the carbonyl of Val56′, while the P1′ forms van der Waals interactions with the carbonyl oxygen atoms of Lys55′ and Val77′.) The second amprenavir observed at one flap of PR20 is surrounded by three symmetry related PR20 molecules (Figure 3B). The P2 and P2′ groups of amprenavir form extensive van der Waals contacts with one of the symmetry related PR20 molecules, while the P1 residue has van der Waals contacts with a second symmetry related PR20. The third symmetry related molecule has one van der Waals contact with the P2 group of amprenavir. Raw ITC data for APV gave an *N* value of ∼0.8, consistent with titration by ≤1 equivalent of APV, possibly due to underestimation of its concentration. The existence of crystal lattice contacts taken together with the close to 1:1 binding stoichiometry observed in ITC indicates that the amprenavir binds only weakly to the second flap binding site and may be an artifact due to crystal packing.

Expanded S2/S2′ Pocket of PR20 Associated with Lower Inhibition by Amprenavir. The PR20/amprenavir complex superimposes on the wild-type PR/amprenavir structure (3NU3) with the relatively high root-mean-square deviation (rmsd) of 1.07 Å for 198 equivalent Cα atoms. In contrast to the wild-type complex, amprenavir binds PR20 in a single orientation, which coincides with the minor conformation in the wild-type PR structure. All the ionic interactions observed between amprenavir and the wild-type PR are retained in the PR20/amprenavir complex, except that the direct hydrogen bond of the minor conformation of the P2′ group with the side chain of Asp30′ in the PR/amprenavir structure is replaced by a water-molecule-mediated interaction in the PR20/amprenavir complex. Comparison with the PR/amprenavir structure reveals that mutations D30N, V32I, I47V, and I84V in PR20 alter the size, shape, and charge of the S2/S2′ pocket in the PR20/amprenavir complex as described previously for the PR20/darunavir and PR20/saquinavir (3UFN) complexes.[16] The expanded pockets are demonstrated by the increased distance of ∼1.5/1.2 Å between the two mutated residues I47/47′V and I84/84′V on either side of the S2/S2′ pocket in the two monomers of PR20/amprenavir relative to the wild-type PR structure (Supporting Information Figures S1A, S1B). Also, mutations in the 80s loop have been shown to alter the potency of protease inhibitors significantly.[24]

The P2′ aniline of amprenavir in the PR20 complex shifts by ∼0.6 Å to form hydrophobic contacts with the shorter valine side chain of mutation I84′V in comparison to both major and minor P2′ conformations in the wild-type PR/amprenavir complex (Figure 4A). I84′V forms two hydrophobic interactions with the aniline in the PR20/amprenavir instead

dx.doi.org/10.1021/jm400231v | J. Med. Chem. 2013, 56, 4017−4027

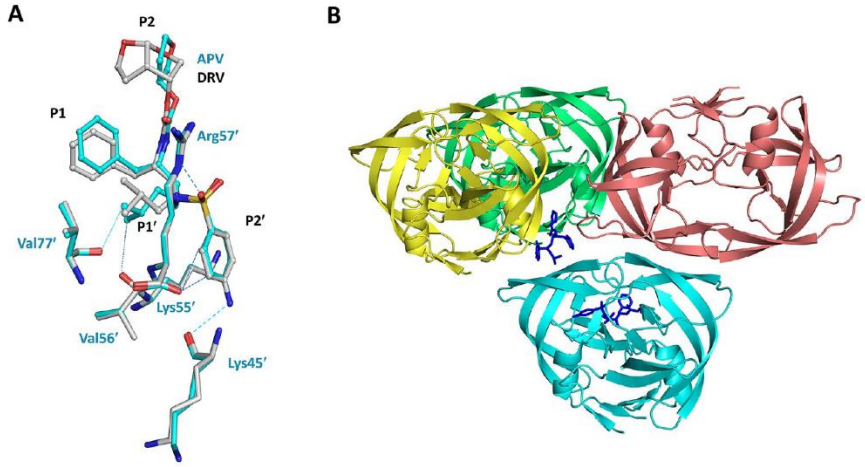


Figure 3. (A) The second amprenavir is observed bound in the PR20/amprenavir complex (cyan carbons) at the same location as the second darunavir site seen in the PR20/darunavir complex (gray carbons). The hydrogen bonds are indicated by broken lines and van der Waals contacts as dotted lines. (B) The second amprenavir (blue sticks) binds near one flap of PR20 (cyan), forming two hydrogen bond interactions and several van der Waals interactions, and also has hydrophobic contacts with three symmetry related PR20 dimers. The P2 and P2′ groups of amprenavir form extensive van der Waals contacts with the symmetry molecule 1 (yellow), while the P1 group has van der Waals contacts with the symmetry molecule 2 (pink). The symmetry molecule 3 (green) has a single van der Waals contact with the P2 group of amprenavir.

of one in the wild-type complex. Also, the P2 THF ring of amprenavir in the PR20 complex is shifted by ~0.5 Å to form van der Waals interactions with the mutated Val84 in comparison to the major conformation of P2 THF in wild-type PR/amprenavir (Figure 4B). In contrast, the THF ring in the PR20 complex is shifted toward the I47V mutation in comparison with the position in the wild-type PR complex (Figure 4C). In addition, amprenavir in its PR20 complex lacks 1−3 van der Waals contacts observed between Ile84/84′ and the P1/P1′ group of amprenavir in the wild-type PR complex (Figure 4A,B). Unlike PR20/amprenavir, the single mutant PR_{I84V}/amprenavir complex retained van der Waals contacts to P1/P1′ similar to wild-type PR/amprenavir.[25] Thus, the shifts in the P2 and P2′ residues of amprenavir due to the expanded S2/ S2′ pocket of PR20 add an additional interaction to I84/84′ at the expense of eliminating interactions between I84/84′V and P1′/P1 residues. The shorter side chain of I47V mutation in the PR20/amprenavir complex removes interactions of the side chain with the P2 THF ring of amprenavir (Figure 4C), but its contact with the P2′ aniline group is retained. In addition, the smaller I84V mutation results in the loss of the van der Waals contact seen between the sulfonyl oxygen of amprenavir and Ile 84′ of wild-type PR. The loss of several interactions between amprenavir and the S2/S2′ mutations in the PR20 complex shows that the P2 and P2′ groups fit loosely in the expanded S2/S2′ pocket. In the S2′ subsite, the mutated Asn30′ of PR20 forms a hydrogen bond to the mutated side chain of Asp88′ in addition to a water-molecule-mediated hydrogen bond to P2′ aniline of amprenavir (Figure 4D,E). Thus, mutation N88D may compensate for the change due to D30N mutation in the active site cavity of PR20. The wild-type PR/amprenavir complex lacks a hydrogen bond interaction between Asp30 and Asn88 in both major (Figure 4D) and minor (Figure 4E) conformations.

The amprenavir complex of PR20 can be compared with the corresponding complex with the HIV-2 protease (PR2/amprenavir) (3S45).[26] The wild-type enzymes from HIV-1 and HIV-2 share 39−48% sequence identity; however, several of the mutations in PR20 exist in the wild-type PR2 sequence including V32I and I47V in the S2/S2′ subsites. The PR20 and PR2 complexes share similar polar interactions with amprenavir except for those of the aniline group. In PR2/amprenavir, the aniline group lacks the hydrogen bonds to the main chain amide and carbonyl group of Asp 30′ but gains a hydrogen bond with the Asp30′ side chain consistent with its 15-fold poorer inhibition by amprenavir in comparison to wild-type HIV-1 PR.[26] However, PR20/amprenavir does not show such dramatic change in interactions with the aniline group and is very similar to wild-type HIV-1 PR/amprenavir complex except for the shifts at the P2 and P2′ groups and the changes in hydrophobic contacts between amprenavir and PR20 (Figure 4F).

Amprenavir and darunavir differ only in the number of THF rings at the P2 position. Darunavir with its two THF rings forms more hydrogen bond interactions with the wild-type PR and binds tighter to wild-type PR than amprenavir with a single THF ring. The PR20/ amprenavir complex superimposes well with the previously determined PR20/darunavir complex (3UCB) with a rmsd of 0.4 Å. The amprenavir complex lacks the two additional hydrogen bond interactions with Asp29 together with new van der Waals interactions formed by the second THF ring of darunavir. The increase in size and the loss of charge in the S2 pocket of PR20 causes the THF ring of amprenavir to shift toward I84V resulting in an elongated C-H···O interaction (from 3.3 to 3.6 Å) with the flap residue Gly48 compared to the wild-type PR. Thus, the single THF group of amprenavir does not fill the expanded S2/S2′ pocket of PR20, thereby reducing the susceptibility of the enzyme to this clinical inhibitor.

dx.doi.org/10.1021/jm400231v | J. Med. Chem. 2013, 56, 4017−4027

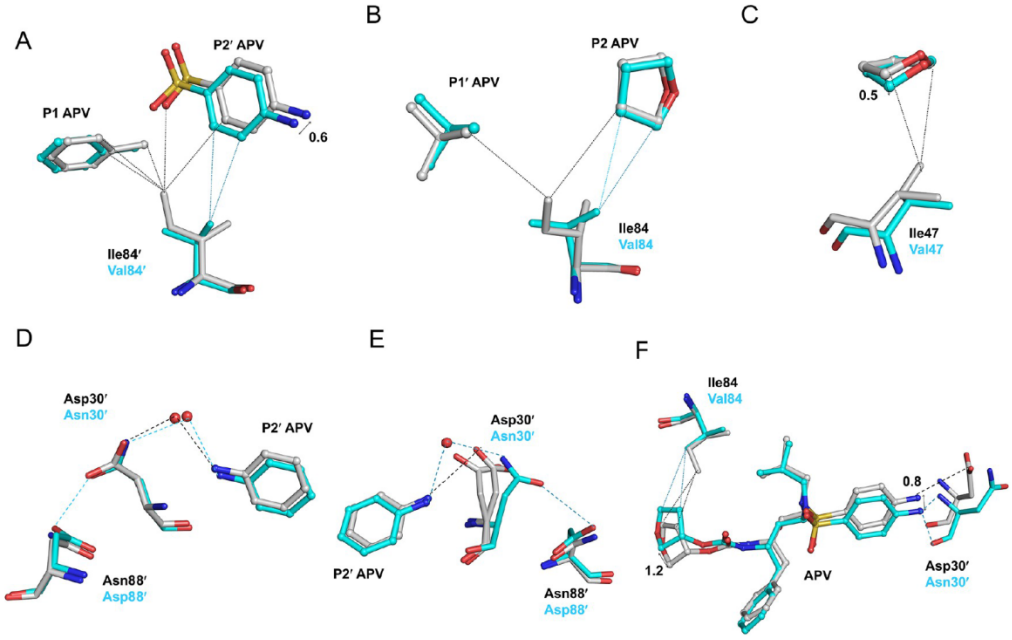


Figure 4. Structural comparison of amprenavir (APV) complexes with PR20 (cyan carbons) and wild-type PR (gray carbons). The protein is represented in sticks, while the inhibitor atoms are in ball and stick format. The hydrogen bonds are indicated by broken lines, and the van der Waals contacts are in dotted lines colored cyan or black according to the complex. (A) The P2′ group of amprenavir is shifted by ~0.6 Å toward the I84V mutation in PR20 in comparison to the major conformation of amprenavir in its complex with wild-type PR (gray carbon), which also eliminates PR20 interactions with the P1 group. (B) The P2 THF group in PR20/amprenavir forms an additional van der Waals contact with the I84V side chain, while the interaction with the P1′ group is eliminated in comparison to the major conformation of amprenavir in the wild-type PR complex. (C) The P2 THF group loses van der Waals interactions with the mutated Val47 side chain in PR20/amprenavir in comparison to the major conformation of amprenavir in its wild-type PR complex. (D) The mutated Asn30 forms a water-molecule-mediated interaction with P2′ and hydrogen bond interaction with the N88D mutation in PR20/amprenavir in comparison to the major conformation of amprenavir in the wild-type PR complex. (E). The direct hydrogen bond between P2′ aniline and Asp30′ in the minor conformation of wild-type PR/amprenavir is lost in PR20/amprenavir. However, the side chain of D30′N forms a hydrogen bond with N88′D in PR20/amprenavir. (F) The P2 and P2′ groups of PR20/amprenavir are shifted toward I84/84′V mutation in comparison to the PR2/amprenavir complex (gray carbons). In contrast to PR2/amprenavir, the P2′ aniline of PR20/amprenavir retains the hydrogen bonds to the main chain amide and the carbonyl of D30N similar to wild-type PR/amprenavir but lacks the hydrogen bond to the side chain of Asp30.

P1′-Pyrrolidinone Has Fewer Interactions with PR20 Than with PR. The HIV-1 PR inhibitor **2** incorporates methyl-2-pyrrolidinone as the P1′ ligand, which was shown to interact with the Gly27′ carbonyl oxygen and Arg8 side chain in the S1′ subsite of wild-type PR.[19] Inhibitor **2** has potent activity against various drug resistant clinical isolates.[27] It retains inhibition of PR mutants bearing single substitutions associated with drug resistance.[28] The PR20/**2** complex superimposes with the previously determined wild-type PR/**2** complex[19] (3H5B) with a relatively large rmsd of 1.14 Å for 198 Cα atoms. The overall structure of PR20/**2** is more similar to the other structures of PR20/amprenavir and PR20/**3** with rmsd values of 0.51 and 0.56 Å, respectively. Unlike the wild-type PR structure where inhibitor **2** binds in two conformations, the inhibitor binds to PR20 in a single conformation resembling the major conformation in the wild-type PR complex.

When compared with amprenavir and darunavir, the P2′ position of **2** differs in the substitution of the oxymethyl group for the amino group in the aniline moiety (Figure 1). The P2′ methoxy oxygen of **2** forms hydrogen bond interactions with main chain amide groups of Asp30′ and Asn30′ in wild-type PR and PR20 complexes, respectively (Figure 5A). The methoxy methyl of the inhibitor also forms a van der Waals interaction with the side chain oxygen of Asp30′ and Asn30′ in both complexes. Thus, the methoxy substitution in **2** retains similar interactions with both Asp30′ and mutated Asn30′. Also, the PR20 structure does not show the dynamic side chain positions of Asn30 observed in various D30N single[29] and double mutants like D30N/N88D and D30N/N88S.[30] The side chain of Asn30′ in PR20/**2** complex forms a polar interaction with the side chain of mutated Asp88′ as described for other PR20 complexes (Figure 5A). The P2 group of **2** makes similar interactions with both PR20 and wild-type PR. However, unlike in the S2′ subsite, the Asn30 side chain in the S2 pocket of PR20/**2** does not form a hydrogen bond interaction with Asp88 (Supporting Information Figure S1C).

The P1′ group incorporates the major design change in **2**. In the wild-type PR/**2** structure, the P1′-pyrrolidine ring is visible

dx.doi.org/10.1021/jm400231v | J. Med. Chem. 2013, 56, 4017−4027

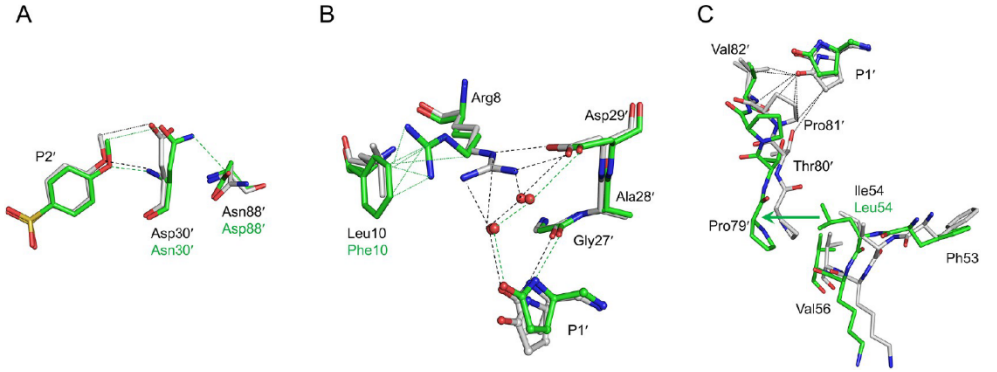


Figure 5. Comparison of inhibitor **2** complexes with PR20 (green carbons) and wild-type PR (gray carbons). The protein is represented in sticks, while the inhibitor atoms are in ball and stick format. The hydrogen bonds are indicated by broken lines, and the van der Waals contacts are in dotted lines colored according to the complex. (A) The P2′ of **2** forms hydrogen bond interactions and C-H···O interaction with main chain amide and the side chain oxygen of Asp30′ and Asn30′ in wild-type PR and PR20 complexes. The D30′N mutation in PR20/**2** complex forms hydrogen bond interactions with N88′D. (B) The L10F mutation eliminates the ion pair between Arg8 and Asp29′. The critical water-molecule-mediated interactions between P1′and Arg′ are also lost in the PR20/**2** complex. (C). The displacement of 80s loop by the I54L flap mutation in PR20/**2** prevents the interactions observed between Pro81′, Val82′, and the second conformation of P1′ in the wild-type PR/**2** complex.

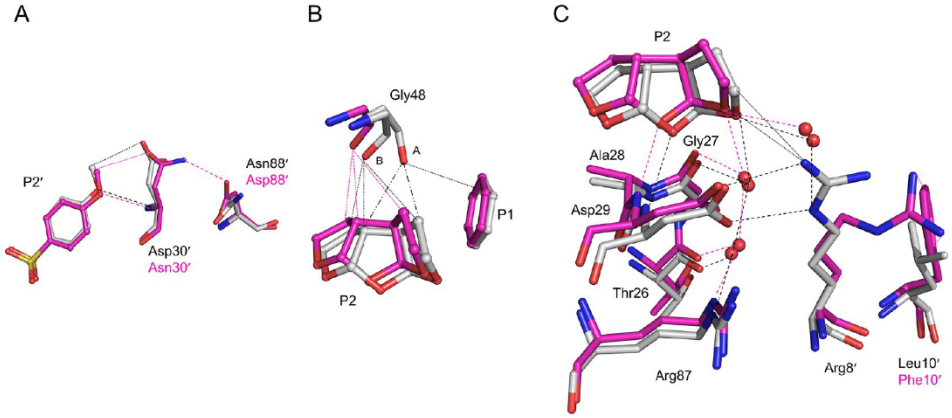


Figure 6. Comparison of inhibitor **3** complexes with PR20 (magenta carbons) and wild-type PR (gray carbons). The protein is represented in sticks, while the inhibitor atoms are in ball and stick format. The hydrogen bonds are indicated by broken lines, the van der Waals contacts are in dotted lines, and the C-H···O interactions are shown as (··—··—··—) and colored according to the complex. (A) The residue 30 forms a hydrogen bond and a C-H···O interaction with the P2′ oxymethyl group of **3** in both wild-type PR and PR20. The mutated Asn30 side chain also forms a hydrogen bond with mutated Asp88 in PR20 complex in contrast to their lack of interaction in the wild-type PR complex. (B) The flap residue Gly48 showed two alternative conformations (A and B) and different interactions with inhibitor in the wild-type PR/**3** complex. In PR20/**3**, Gly48 adopts the conformation B observed in the wild-type PR complex, which results in the loss of stronger C-H···O interactions with the tris-THF rings of **3**. In addition, the van der Waals contact with the P1 group of **3** is also lost. (C) The L10′F mediated buckling of Arg8′ in PR20 results in the loss of the ion pair with Arg29 and the loss of water-molecule-mediated hydrogen bonds connecting the third THF ring and Arg8′. The van der Waals contacts between the third THF ring and Arg8′ are also not observed in the PR20/**3** complex.

in two alternative conformations with equal occupancy and related by 18° rotation around the C12−C13 bond (Figure 5B).[19] In the first conformation, P1′-pyrrolidinone has a hydrogen bond between the NH and the carbonyl oxygen of Gly27. A water-molecule-mediated hydrogen bond is formed between the pyrrolidinone carbonyl and the side chain of Arg8. The second conformation of P1′-pyrrolidinone has hydrophobic and C-H···O interactions with Pro81′ and Val82′. In comparison, the P1′ pyrrolidinone of PR20/**2** occurs in a single conformation similar to the first conformation seen in the wild-type PR structure and forms a hydrogen bond with the carbonyl of Gly27′. However, the P1′-pyrrolidinone in PR20 does not form the water-molecule-mediated hydrogen bond with the side chain of Arg8, presumably due to the L10F mutation (Figure 5B). The L10F mutation facilitates flipping of the Arg8 side chain to form van der Waals contacts with the

dx.doi.org/10.1021/jm400231v | J. Med. Chem. 2013, 56, 4017−4027

bulky Phe10 side chain, which breaks the intersubunit ion pair between Arg8 and Asp29′. The second conformation of P1′-pyrrolidinone is likely not favored in the PR20/**2** structure due to the I54L mutation, which displaces Pro81′ by ∼1.6 Å, eliminating its potential interactions with the second P1′ conformation found in the wild-type PR complex (Figure 5C). Thus, changes due to PR20 mutations L10F and I54L may correlate with the higher K_L value for inhibitor **2** relative to the related PIs lacking the P1′-pyrrolidinone.

The P2 Tris-THF in 3 Adds New Interactions with PR20. The potent HIV-1 PR inhibitor was developed by structure based design.[20] The major difference from darunavir is the presence of a third THF ring in **3** instead of the bis-THF group in the P2 position of darunavir. This larger P2 group was designed to provide additional interactions and fill the S2 pocket in wild-type PR more effectively.[20] Also, similar to **2**, the P2′ amino group in the aniline moiety is substituted with an oxymethyl group. Biological testing showed that **3** had 10-fold higher antiviral potency than darunavir against a set of drug resistant strains.[20] Studies of PR with single mutations, however, suggested **3** would be less effective on variants with the I50V mutation.[31] Superposition of the PR20/**3** complex with the wild-type PR/**3** (3OK9) dimer gave a significant rmsd of 1.0 Å for 198 equivalent Cα atoms. Overall, the PR20/**3** complex is very similar to the PR20/amprenavir and PR20/**2** complexes and can be superimposed with lower rmsd values of 0.31 and 0.56 Å, respectively, for all Cα atoms. Inhibitor **3** binds in the active site of wild-type PR in two orientations related by 180° with relative occupancies of 0.55/0.45.[20] In the PR20/**3** complex, the inhibitor binds in a single orientation similar to the major orientation in the wild-type PR complex. The interactions of the P1 and P1′ groups of the inhibitor are similar in both the wild-type and PR20 complexes. Similar to inhibitor **2**, the P2′ position of **3** has the longer oxymethyl group substituted for the amino group in the aniline moiety of amprenavir (Figure 1). Also, the P2′ methoxy oxygen of **3** forms hydrogen bond interactions with main chain amides of Asp30′ and Asn30′ in wild-type PR and PR20 complexes, respectively (Figure 6A). The methoxy methyl of the inhibitor forms a van der Waals interaction with the side chain oxygen Asn30′ in the PR20 complex similar to the van der Waals contact with the side chain of Asp30′ observed in the wild-type PR complex. The Asn30′ side chain in the PR20 complex also forms a hydrogen bond interaction with the mutated side chain of Asp88′, which does not occur in wild-type PR. Thus, similar to inhibitor **2**, the extended oxymethyl group in **3** maintains Asn30′ in a single conformation that interacts with the inhibitor and Asp88′ mutated side chain.

The critical difference between amprenavir, darunavir, and **3** is the presence of one, two, and three THF rings at the P2 positions, respectively. The conformation and interactions of the first two THF rings of **3** are very similar in its complexes with wild-type PR and PR20. The first THF oxygen forms a hydrogen bond interaction with Asp30 and Asn30 in the wild-type PR and PR20 complexes, respectively. The Asp30 in wild-type PR exhibits two alternate conformations, with the major conformation interacting with the first THF group. The majority of differences in inhibitor interactions between PR20/**3** and the wild-type PR complex are confined to the interactions of the third THF ring. In the wild-type PR complex, the presence of the third THF ring drives the carbonyl of the flap residue Gly48 into two alternate conformations, with the major conformation forming a C-H···O interaction with the third THF ring in addition to the van der Waals contact to the P1 phenyl group (Figure 6B). The minor conformation of Gly48 forms van der Waals interactions with all three THF rings. Gly48 carbonyl has a single conformation in the PR20 complex, probably due to the expanded S2 subsite, and retains interactions similar to the minor conformation in the wild-type PR structure. In the wild-type PR complex, the oxygen of the third THF forms a water mediated hydrogen bond with the guanidinium group of Arg8′ and also van der Waals contacts with Arg8′. In the PR20 complex, the third THF ring has lost these interactions because the Arg8′ side chain moves to form hydrophobic interactions with the mutated side chain of L10′F (Figure 6C). However, the third THF ring in the PR20 complex retains all the van der Waals interactions, with Asp29 observed in the wild-type complex. Also, the water-molecule-mediated hydrogen bond interactions between the oxygen of the third THF ring and main chain carbonyl oxygens of Thr26 and Gly27 in the characteristic catalytic triplet (Asp25-Thr26-Gly27) of aspartic proteases and the water-molecule-mediated interactions with the side chains of Asp29 and Arg87 are conserved in both the wild-type and PR20 complexes (Figure 6C). The two water molecules involved in these interactions are internal structural water molecules conserved in the majority of PR structures.[32] Further, the third THF ring in the PR20 complex packs neatly against the P1 phenyl group of the inhibitor as described for the wild-type complex and retains good internal hydrophobic contacts that cannot occur for the single or bis-THF group in amprenavir and darunavir, respectively. Thus, despite the new interactions of the P2 group of inhibitor **3**, some of the interactions with Arg8 and to the flap are lost in the PR20 complex, resulting in binding affinity similar to that of darunavir for PR20.

■ DISCUSSION AND CONCLUSIONS

We propose that PR20 provides a useful model for evaluation and development of new inhibitors targeting emerging strains of multidrug resistant virus. Recently, we characterized the structural changes in the clinically derived extreme drug resistant protease PR20 bearing 20 mutations.[16] PR20 showed substantially lower affinity for PIs by >3 orders of magnitude relative to wild-type PR.[15] Analysis of ITC data shows that the markedly lower affinity of PR20 for darunavir of 10^4-fold arises primarily from a ΔH that is significantly less favorable by 4.5 kcal/mol accompanied by a small (∼0.5 kcal/mol) unfavorable decrease in the entropic effect (less negative $-T\Delta S$) relative to the corresponding wild-type values. Amprenavir exhibits a ∼500-fold lower affinity for PR20 relative to wild-type PR, due to unfavorable changes of 1.5 and 1.7 kcal/mol in ΔH and $-T\Delta S$, respectively.[33] Overall, energy contributions to binding of darunavir, **1**, and structurally related inhibitors to PR20 are predominantly enthalpic, unlike the large change in entropy reported recently for binding to different mutants.[34]

In addition, the autocatalytic cleavage of PR20 precursor to release mature protease is not inhibited significantly by the clinical PIs, although the autoprocessing of wild-type PR precursor is inhibited effectively by darunavir and saquinavir.[15,35,36] Our studies show that **2** and **3**, like the clinical inhibitors, do not inhibit the autocatalytic cleavage of PR20 precursor.

The new crystal structures of PR20 in complex with amprenavir, **2**, and **3** reveal the molecular mechanisms for evasion of these inhibitors. Amprenavir has the smallest P2

dx.doi.org/10.1021/jm400231v | J. Med. Chem. 2013, 56, 4017−4027

group among these three inhibitors, which appears to fit poorly in the expanded S2 pocket of PR20 consistent with 4-fold weaker affinity relative to darunavir. Inhibitor **2** has the weakest affinity for PR20, probably due to poorer interaction of the P1′ moiety in PR20 compared with wild-type PR. The highest affinity for PR20 was shown by **3** because its bulky tris-THF rings fill the expanded hydrophobic S2 pocket of PR20 more effectively than do the smaller P2 groups in the other two inhibitors. However, despite the addition of a third THF ring, the K_L values of **3** and darunavir are almost identical. The favorable affinity of **3** for PR20 agrees with its high antiviral potency.[20] In fact, **3** exhibits 10-fold higher potency than darunavir on tested multidrug resistant viral strains, which may reflect different effects in the cell compared to the isolated protein. Overall, **3** is a promising inhibitor for further development against extreme drug resistant variants exemplified by PR20. Structural analysis suggests inhibitors may be improved by adding more polar interactions with the flexible flaps as well as increasing the size of P2/P2′ groups. Preventing the L10F mediated breakage of the ion pair between Arg8/8′ and Asp29′/29 may be important for stronger inhibition of PR20. The insertion between L10F and Arg8 of the P2 group of saquinavir from the second binding site in the PR20/saquinavir complex[16] may suggest a possible approach to overcome the effects of the L10F mutation.

■ EXPERIMENTAL SECTION

General. Inhibitor **2** has shown analytical purity >99% by HPLC analysis,[28] and inhibitor **3** has shown analytical purity of >98% by HPLC.[31] The structures were confirmed by ^{1}H and ^{13}C NMR spectral analysis and high resolution mass spectrometry. For inhibitor **2**, HRMS (m/z) calcd for C30H40N3O8S $[M + H]^+$ 602.2536, found 602.2536. For inhibitor **3**, HRMS (ESI) $[M + Na]^+$ calcd for C30H40N2O9SNa 627.2352, found 627.2359.

Construction, Expression, and Purification of PR20. A 99 amino acid synthetic gene termed PR20 derived from the protease sequence of a clinical isolate[37] was cloned between the Nde1 and BamH1 sites of pET11a vector (Novagen, San Diego, CA) and transformed into *E. coli* BL-21 (DE3; Stratagene). Protein expression, purification, and folding were carried out as described previously.[38]

In Vivo Autoprocessing of TFR-PR20 in *E. coli*. Precursor construct TFR-PR20 comprising the multidrug resistant PR20 sequence fused to the full length TFR was cloned and expressed in *E. coli*, and the effect of added inhibitors to inhibit autoprocessing was assessed by growing the cultures in the presence of increasing concentrations of **3** and **2** in the medium. The expressed protein was partially purified and analyzed by SDS-PAGE as described and compared to the wild-type TFR-PR construct inhibited by darunavir.[15]

Isothermal Titration Calorimetry. PR20 (53 μL of a 1.8−2.15 mg/mL solution in 12 mM HCl) was folded by addition of 5 mM sodium acetate buffer, pH 6 (buffer A), to give 175 μL, followed immediately by 175 μL of 100 mM sodium acetate buffer, pH 5 (buffer B), and centrifuged for 2−4 min at 16,000g. The supernatant was transferred to the calorimeter cell (MicroCal high-precision iTC$_{200}$ microcalorimeter, GE Healthcare) and titrated as described,[15] with inhibitors at ∼10-fold the PR20 concentration in a buffer of the same composition. Data were analyzed by use of the Origin software provided with the instrument. Because it was not possible to determine accurately the concentrations of the **2** and **3** inhibitors because of the limited amounts available, inhibitor concentrations were estimated from the observed N values at known PR20 concentration and an active-site binding stoichiometry of 1:1 inhibitor:PR20, to give corrected N values of 1.0. For consistency, the same correction procedure was also used for darunavir and amprenavir.

Crystallization, X-ray Data Collection, and Structure Determination. The inhibitors amprenavir, **2**, and **3** were mixed with PR20 at a 5:1 molar ratio and incubated on ice for 30 min prior to crystallization trials. The complexes were crystallized by the hanging drop vapor diffusion technique at room temperature. The PR20/amprenavir crystals were grown by mixing 1 μL of protein complex (5 mg/ml) and 1 μL of reservoir solution well solution containing 1.67 M sodium chloride and 67 mM citrate-phosphate buffer at pH 4.2. The reservoir used for growing PR20/**2** was 1.2 M lithium chloride and 0.1 M sodium acetate at pH 4.8. The PR20/**3** complex crystals were obtained with well solution containing 0.25 M potassium iodide and 0.1 M sodium acetate buffer at pH 4.8. The cryoprotectant for crystals was composed of 30% glycerol and the respective mother liquor. Diffraction data at 100 K were collected on beamline 22-ID (SER-CAT) at the Advance Photon Source, Argonne National Laboratory (Argonne, IL, USA). All data were integrated and scaled with HKL2000.[39]

The crystal structures were solved by molecular replacement using Phaser.[40,41] The PR dimer from the crystal structure of PR$_{I50V}$/darunavir (2F8G)[42] was used as the search model for the PR20/amprenavir complex. For PR20/**2** and PR20/**3** complexes, the search model was the PR dimer from PR$_{M46L}$/darunavir (2HS2).[22] The atomic models were refined by iterative rounds of model building into electron density maps and refinement using COOT,[43] REFMAC[44] (PR20/amprenavir), and SHELX-97[45] (PR20/**2** and PR20/**3**). The surface loops, including 10s, 30s and 60s loops with ambiguous densities, were pruned during early stages of refinement and successfully rebuilt. The correct amino acid side chain corresponding to the 20 mutations were adding during the refinement. The three inhibitors were fitted into unambiguous electron densities of the respective structures. Solvent molecules were inserted at stereochemically reasonable positions using $2F_o - F_c$ and $F_o - F_c$ maps contoured at 1 and 3 σ levels, respectively. The hydrogen bond interaction is defined by a distance in the range of 2.6−3.5 Å between hydrogen donor and acceptor atoms. The C-H···O interaction has a C···O distance of 3.0−3.7 Å.[47] The van der Waals interaction has C-H···H-C distance of 3.8−4.2 Å.[46] Molecular figures were prepared with PyMOL (http://www.pymol.org).

■ ASSOCIATED CONTENT

Supporting Information

Figure comparing PR20 and wild-type PR structures. This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

Crystallographic data are available for inhibitor complexes: PR20-APV (PDB ID: 4J5J), PR20-GRL0519 (PDB ID: 4J54), PR20-GRL02031 (PDB ID: 4J55).

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 404 413-5411. Fax: 404 413-5301. E-Mail: iweber@gsu.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The research was supported in part by the National Institutes of Health grants GM062920 (I.T.W.) and GM053386 (A.K.G.), by the Georgia State University Molecular Basis of Disease Fellowship and the Georgia State University Research Program Enhancement Award in Bioinformatics (H.-C.S.), and by the Intramural Research Program of the NIDDK and the Intramural AIDS-Targeted Antiviral Program of the Office of the Director, NIH. This research was authored, in whole or in part, by National Institutes of Health staff. Data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) beamline 22ID at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at

www.ser-cat.org/members.html. Amprenavir and darunavir were obtained through the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH.

■ ABBREVIATIONS USED

APV, amprenavir; DRV, darunavir; HIV-1, human immunodeficiency virus type 1; HAART, highly active antiretroviral therapy; ITC, isothermal titration calorimetry; PR, mature HIV-1 protease; PI, clinical inhibitor of PR; tris-THF, tris-tetrahydrofuran; PDB, Protein Data Bank; rmsd, root-mean-square deviation

■ REFERENCES

(1) Sepkowitz, K. A. AIDS—the first 20 years. *N. Engl. J. Med.* **2001**, *344*, 1764−1772.
(2) Condra, J. H.; Schleif, W. A.; Blahy, O. M.; Gabryelski, L. J.; Graham, D. J.; Quintero, J. C.; Rhodes, A.; Robbins, H. L.; Roth, E.; Shivaprakash, M.; et al. In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors. *Nature* **1995**, *374*, 569−571.
(3) Mehellou, Y.; De Clercq, E. Twenty-six years of anti-HIV drug discovery: where do we stand and where do we go? *J. Med. Chem.* **2010**, *53*, 521−538.
(4) Menendez-Arias, L. Molecular basis of human immunodeficiency virus drug resistance: an update. *Antiviral Res.* **2010**, *85*, 210−231.
(5) Grubb, J. R.; Singhatiraj, E.; Mondy, K.; Powderly, W. G.; Overton, E. T. Patterns of primary antiretroviral drug resistance in antiretroviral-naive HIV-1-infected individuals in a midwest university clinic. *AIDS* **2006**, *20*, 2115−2116.
(6) Liao, L.; Xing, H.; Dong, Y.; Qin, G.; Ma, Y.; Lu, H.; Chen, L.; Zhang, L.; Osborne, C.; Seguy, N.; Wei, D.; Sun, F.; Yang, J.; Ruan, Y.; Shao, Y. Surveys of Transmitted HIV Drug Resistance in 7 Geographic Regions in China, 2008−2009. *Clin. Infect. Dis.* **2012**, *54* (Suppl 4), S320−S323.
(7) Hunt, G. M.; Ledwaba, J.; Basson, A. E.; Moyes, J.; Cohen, C.; Singh, B.; Bertagnolio, S.; Jordan, M. R.; Puren, A.; Morris, L. Surveillance of Transmitted HIV-1 Drug Resistance in Gauteng and KwaZulu-Natal Provinces, South Africa, 2005−2009. *Clin. Infect. Dis.* **2012**, *54* (Suppl 4), S334−S338.
(8) Duc, N. B.; Hien, B. T.; Wagar, N.; Tram, T. H.; Giang le, T.; Yang, C.; Wolfe, M. L.; Hien, N. T.; Tuan, N. A. Surveillance of Transmitted HIV Drug Resistance Using Matched Plasma and Dried Blood Spot Specimens From Voluntary Counseling and Testing Sites in Ho Chi Minh City, Vietnam, 2007−2008. *Clin. Infect. Dis.* **2012**, *54* (Suppl 4), S343−S347.
(9) Hamers, R. L.; Wallis, C. L.; Kityo, C.; Siwale, M.; Mandaliya, K.; Conradie, F.; Botes, M. E.; Wellington, M.; Osibogun, A.; Sigaloff, K. C.; Nankya, I.; Schuurman, R.; Wit, F. W.; Stevens, W. S.; van Vugt, M.; de Wit, T. F. HIV-1 drug resistance in antiretroviral-naive individuals in sub-Saharan Africa after rollout of antiretroviral therapy: a multicentre observational study. *Lancet Infect. Dis.* **2011**, *11*, 750−759.
(10) Sigaloff, K. C.; Mandaliya, K.; Hamers, R. L.; Otieno, F.; Jao, I. M.; Lyagoba, F.; Magambo, B.; Kapaata, A.; Ndembi, N.; Rinke de Wit, T. F. High Prevalence of Transmitted Antiretroviral Drug Resistance Among Newly HIV Type 1 Diagnosed Adults in Mombasa, Kenya. *AIDS Res. Hum. Retroviruses* **2012**, *28*, 1033−1037.
(11) Jordan, M. R.; Bennett, D. E.; Wainberg, M. A.; Havlir, D.; Hammer, S.; Yang, C.; Morris, L.; Peeters, M.; Wensing, A. M.; Parkin, N.; Nachega, J. B.; Phillips, A.; De Luca, A.; Geng, E.; Calmy, A.; Raizes, E.; Sandstrom, P.; Archibald, C. P.; Perriens, J.; McClure, C. M.; Hong, S. Y.; McMahon, J. H.; Dedes, N.; Sutherland, D.; Bertagnolio, S. Update on World Health Organization HIV Drug Resistance Prevention and Assessment Strategy: 2004−2011. *Clin. Infect. Dis.* **2012**, *54* (Suppl 4), S245−S249.
(12) Hirsch, M. S.; Gunthard, H. F.; Schapiro, J. M.; Brun-Vezinet, F.; Clotet, B.; Hammer, S. M.; Johnson, V. A.; Kuritzkes, D. R.; Mellors, J. W.; Pillay, D.; Yeni, P. G.; Jacobsen, D. M.; Richman, D. D. Antiretroviral drug resistance testing in adult HIV-1 infection: 2008 recommendations of an International AIDS Society−USA panel. *Clin. Infect. Dis.* **2008**, *47*, 266−285.
(13) Heneine, W.; Kashuba, A. HIV Prevention by Oral Preexposure Prophylaxis. *Cold Spring Harb. Perspect. Med.* **2012**, *2*, a007419.
(14) Supervie, V.; Garcia-Lerma, J. G.; Heneine, W.; Blower, S. HIV, transmitted drug resistance, and the paradox of preexposure prophylaxis. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 12381−12386.
(15) Louis, J. M.; Aniana, A.; Weber, I. T.; Sayer, J. M. Inhibition of autoprocessing of natural variants and multidrug resistant mutant precursors of HIV-1 protease by clinical inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 9072−9077.
(16) Agniswamy, J.; Shen, C. H.; Aniana, A.; Sayer, J. M.; Louis, J. M.; Weber, I. T. HIV-1 Protease with 20 Mutations Exhibits Extreme Resistance to Clinical Inhibitors through Coordinated Structural Rearrangements. *Biochemistry* **2012**, *51*, 2819−2828.
(17) Flexner, C. HIV drug development: the next 25 years. *Nature Rev. Drug Discovery* **2007**, *6*, 959−966.
(18) Williams, G. C.; Sinko, P. J. Oral absorption of the HIV protease inhibitors: a current update. *Adv. Drug Delivery Rev.* **1999**, *39*, 211−238.
(19) Ghosh, A. K.; Leshchenko-Yashchuk, S.; Anderson, D. D.; Baldridge, A.; Noetzel, M.; Miller, H. B.; Tie, Y.; Wang, Y. F.; Koh, Y.; Weber, I. T.; Mitsuya, H. Design of HIV-1 protease inhibitors with pyrrolidinones and oxazolidinones as novel P1′-ligands to enhance backbone-binding interactions with protease: synthesis, biological evaluation, and protein−ligand X-ray studies. *J. Med. Chem.* **2009**, *52*, 3902−3914.
(20) Ghosh, A. K.; Xu, C. X.; Rao, K. V.; Baldridge, A.; Agniswamy, J.; Wang, Y. F.; Weber, I. T.; Aoki, M.; Miguel, S. G.; Amano, M.; Mitsuya, H. Probing multidrug-resistance and protein−ligand interactions with oxatricyclic designed ligands in HIV-1 protease inhibitors. *ChemMedChem* **2010**, *5*, 1850−1854.
(21) Ghosh, A. K.; Anderson, D. D.; Weber, I. T.; Mitsuya, H. Enhancing protein backbone binding—a fruitful concept for combating drug-resistant HIV. *Angew. Chem., Int. Ed. Engl.* **2012**, *51*, 1778−1802.
(22) Kovalevsky, A. Y.; Liu, F.; Leshchenko, S.; Ghosh, A. K.; Louis, J. M.; Harrison, R. W.; Weber, I. T. Ultra-high resolution crystal structure of HIV-1 protease mutant reveals two binding sites for clinical inhibitor TMC114. *J. Mol. Biol.* **2006**, *363*, 161−173.
(23) Kovalevsky, A. Y.; Ghosh, A. K.; Weber, I. T. Solution kinetics measurements suggest HIV-1 protease has two binding sites for darunavir and amprenavir. *J. Med. Chem.* **2008**, *51*, 6599−6603.
(24) Mahalingam, A. K.; Axelsson, L.; Ekegren, J. K.; Wannberg, J.; Kihlstrom, J.; Unge, T.; Wallberg, H.; Samuelsson, B.; Larhed, M.; Hallberg, A. HIV-1 protease inhibitors with a transition-state mimic comprising a tertiary alcohol: improved antiviral activity in cells. *J. Med. Chem.* **2010**, *53*, 607−615.
(25) Shen, C. H.; Wang, Y. F.; Kovalevsky, A. Y.; Harrison, R. W.; Weber, I. T. Amprenavir complexes with HIV-1 protease and its drug-resistant mutants altering hydrophobic clusters. *FEBS J.* **2010**, *277*, 3699−3714.
(26) Tie, Y.; Wang, Y. F.; Boross, P. I.; Chiu, T. Y.; Ghosh, A. K.; Tozser, J.; Louis, J. M.; Harrison, R. W.; Weber, I. T. Critical differences in HIV-1 and HIV-2 protease specificity for clinical inhibitors. *Protein Sci.* **2012**, *21*, 339−350.
(27) Koh, Y.; Das, D.; Leschenko, S.; Nakata, H.; Ogata-Aoki, H.; Amano, M.; Nakayama, M.; Ghosh, A. K.; Mitsuya, H. GRL-02031, a novel nonpeptidic protease inhibitor (PI) containing a stereochemically defined fused cyclopentanyltetrahydrofuran potent against multi-PI-resistant human immunodeficiency virus type 1 in vitro. *Antimicrob. Agents Chemother.* **2009**, *53*, 997−1006.
(28) Chang, Y. C.; Yu, X.; Zhang, Y.; Tie, Y.; Wang, Y. F.; Yashchuk, S.; Ghosh, A. K.; Harrison, R. W.; Weber, I. T. Potent antiviral HIV-1 protease inhibitor GRL-02031 adapts to the structures of drug resistant mutants with its P1′-pyrrolidinone ring. *J. Med. Chem.* **2012**, *55*, 3387−3397.

(29) Kozisek, M.; Bray, J.; Rezacova, P.; Saskova, K.; Brynda, J.; Pokorna, J.; Mammano, F.; Rulisek, L.; Konvalinka, J. Molecular analysis of the HIV-1 resistance development: enzymatic activities, crystal structures, and thermodynamics of nelfinavir-resistant HIV protease mutants. *J. Mol. Biol.* **2007**, *374*, 1005−1016.
(30) Bihani, S. C.; Das, A.; Prashar, V.; Ferrer, J. L.; Hosur, M. V. Resistance mechanism revealed by crystal structures of unliganded nelfinavir-resistant HIV-1 protease non-active site mutants N88D and N88S. *Biochem. Biophys. Res. Commun.* **2009**, *389*, 295−300.
(31) Zhang, H.; Wang, Y. F.; Shen, C. H.; Agniswamy, J.; Rao, K. V.; Ghosh, A. K.; Harrison, R. W.; Weber, I. T. Novel P2 tris-tetrahydrofuran group in antiviral compound 1 (GRL-0519) fills the S2 binding pocket of selected mutants of HIV-1 protease. *J. Med. Chem.* **2013**, *56*, 1074−1083.
(32) Mahalingam, B.; Louis, J. M.; Hung, J.; Harrison, R. W.; Weber, I. T. Structural implications of drug-resistant mutants of HIV-1 protease: high-resolution crystal structures of the mutant protease/substrate analogue complexes. *Proteins* **2001**, *43*, 455−464.
(33) King, N. M.; Prabu-Jeyabalan, M.; Nalivaika, E. A.; Wigerinck, P.; de Bethune, M. P.; Schiffer, C. A. Structural and thermodynamic basis for the binding of TMC114, a next-generation human immunodeficiency virus type 1 protease inhibitor. *J. Virol.* **2004**, *78*, 12012−12021.
(34) King, N. M.; Prabu-Jeyabalan, M.; Bandaranayake, R. M.; Nalam, M. N.; Nalivaika, E. A.; Ozen, A.; Haliloglu, T.; Yilmaz, N. K.; Schiffer, C. A. Extreme Entropy−Enthalpy Compensation in a Drug-Resistant Variant of HIV-1 Protease. *ACS Chem. Biol.* **2012**, *7*, 1536−1546.
(35) Huang, L.; Li, Y.; Chen, C. Flexible catalytic site conformations implicated in modulation of HIV-1 protease autoprocessing reactions. *Retrovirology* **2011**, *8*, 79.
(36) Davis, D. A.; Soule, E. E.; Davidoff, K. S.; Daniels, S. I.; Naiman, N. E.; Yarchoan, R. Activity of Human Immunodeficiency Virus Type 1 Protease Inhibitors Against the Initial Autocleavage in Gag-Pol Polyprotein Processing. *Antimicrob. Agents Chemother.* **2012**, *56*, 3620−3628.
(37) Dierynck, I.; De Wit, M.; Gustin, E.; Keuleers, I.; Vandersmissen, J.; Hallenberger, S.; Hertogs, K. Binding kinetics of darunavir to human immunodeficiency virus type 1 protease explain the potent antiviral activity and high genetic barrier. *J. Virol.* **2007**, *81*, 13845−13851.
(38) Sayer, J. M.; Agniswamy, J.; Weber, I. T.; Louis, J. M. Autocatalytic maturation, physical/chemical properties, and crystal structure of group N HIV-1 protease: relevance to drug resistance. *Protein Sci.* **2010**, *19*, 2055−2072.
(39) Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **1997**, *276*, 307−326.
(40) Storoni, L. C.; McCoy, A. J.; Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 432−438.
(41) McCoy, A. J.; Grosse-Kunstleve, R. W.; Storoni, L. C.; Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2005**, *61*, 458−464.
(42) Kovalevsky, A. Y.; Tie, Y.; Liu, F.; Boross, P. I.; Wang, Y. F.; Leshchenko, S.; Ghosh, A. K.; Harrison, R. W.; Weber, I. T. Effectiveness of nonpeptide clinical inhibitor TMC-114 on HIV-1 protease with highly drug resistant mutations D30N, I50V, and L90M. *J. Med. Chem.* **2006**, *49*, 1379−1387.
(43) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126−2132.
(44) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53*, 240−255.
(45) Sheldrick, G. M.; Schneider, T. R. SHELXL: high-resolution refinement. *Methods Enzymol.* **1997**, *277*, 319−343.
(46) Weber, I. T., Wang, Y,-F. HIV protease: role in viral replication, protein−ligand X-ray crystal structures and inhibitor design. In *Aspartic Protease as Therapeutic Targets*; Gosh, A. K., Ed.; Wiley-VCH Verlag GmbH & Co.: New York, 2010; Vol. *45*, pp 109−137.
(47) Panigrahi, S. K.; Desiraju, G. R. Strong and weak hydrogen bonds in the protein-ligand interface. *Proteins* **2007**, *67*, 128−141.

Appendix E

PEER-REVIEWED PRIMARY RESEARCH PAPERS

1. **Shen CH**, Chang, Y.-C, Agniswamy J, Harrison R.W. and Weber, I.T. Dynamic Variation in the Flaps of an Extreme Drug Resistant HIV Protease Variant, in preparation.
2. Agniswamy J, **Shen CH,** Wang YF, Ghosh AK, Rao KV, Xu CX, Sayer JM, Louis JM, Weber IT. Extreme multidrug Resistant HIV-1 protease with 20 mutations is resistant to novel protease inhibitors with P1'-pyrrolidinone or P2-Tris-tetrahydrofuran. J Med Chem 2013 56(10):4017-27
3. Zhang H, Wang YF, **Shen CH**, Agniswamy J, Rao KV, Xu CX, Ghosh AK, Harrison RW, Weber IT. Novel P2 Tris-tetrahydrofuran group in antiviral compound 1 (GRL-0519) fills the S2 binding pocket of selected mutants of HIV-1 protease. J Med Chem 2013 56(3):1074-83.
4. **Shen CH**, Tie Y, Yu X, Wang YF, Kovalevsky AY, Harrison RW, Weber IT. Capturing the reaction pathway in near-atomic-resolution crystal structures of HIV-1 protease. Biochemistry. 2012 51(39): 7726-32.
5. J Agniswamy**, CH Shen**, A Aniana, JM Sayer, JM Louis, IT Weber. 2012 HIV-1 protease with 20 mutations exhibits extreme resistance to clinical inhibitors through coordinated structural rearrangements. Biochemistry 2012 51(13), 2819-2828.
6. **Shen CH**, Wang YF, Kovalevsky AY, Harrison RW, Weber IT. 2010. Amprenavir complexes with HIV-1 protease and its drug-resistant mutants altering hydrophobiclusters. FEBS J. 2010 Sep;277(18):3699-714.

BOOK CHAPTER

Weber, I.T., Agniswamy, J., Fu, G., **Shen, C.H**., Harrison, R.W. Reaction Intermediates Discovered in Crystal Structures of Enzymes. In "Structural and Mechanistic Enzymology: Bringing Computations and Experiments Together", Eds., Christo Christov and Tatyana Karabencheva, Advances in Protein Chemistry and Structural Biology, Elsevier, Inc. (2012) (invited review) 87, 57-86.