8-10-2021

Crystallographic and enzyme kinetics studies of multidrug resistant HIV-1 protease mutants PR20 and PRS17.

Shelley Burnaman

CRYSTALLOGRAPHIC AND ENZYME KINETICS STUDIES OF MULTIDRUG RESISTANT HIV-1 PROTEASE MUTANTS PR20 AND PRS17

by

SHELLEY HINKLE BURNAMAN

Under the Direction of Irene Weber, PhD

ABSTRACT

HIV/AIDS continues to be a public health threat, with about 1 million people newly infected each year and nearly 40 million deaths from HIV/AIDS since 1980. HIV-1 protease (PR) is an important drug target for HIV antiretroviral therapy (ART). PR catalyzes the final step in the HIV-1 life cycle and is necessary for HIV to become infectious. Nine inhibitors of PR have been FDA approved for use in ART since 1995, but resistance to PR inhibitors is a growing problem. Several drug resistant mutants of PR have previously been studied to determine mechanisms of resistance and inform design of novel PR inhibitors. This research investigates two highly drug resistant

mutants of PR, PR20 and PRS17, that are resistant to all clinical PR inhibitors but have different drug resistance mechanisms. PR20 evades PR inhibitors via a cluster of mutations in the active site that alter inhibitor binding. PRS17 has two mutations in the active site (V82S and G48V) and a unique curled flap conformation that starts at V48 and continues through G52. A revertant mutant of PRS17 (V48G) was characterized to examine the contribution of the G48V mutation to drug resistance.

Novel antiviral inhibitors GRL-0489A and GRL-0739A show similar inhibition to darunavir. Analysis of active site hydrogen bonding in PR20 structures showed no major differences versus PR20/DRV.

Enzymes kinetic assays showed PR inhibitors were more effective for revertant mutant $PRS17_{V48G}$ than for PRS17. Crystal structures were solved for the revertant with and without inhibitor. The curled flap conformation seen in PRS17 was reversed in $PRS17_{V48G}$. Molecular dynamics (MD) simulations of open flap inhibitor-less $PRS17_{V48G}$ revealed the flaps were less dynamic than in PRS17. Structural analysis, MD simulations and enzyme kinetics assays show that $PRS17_{V48G}$ is more stable and more susceptible to inhibitors than PRS17. These data can be used to inform design of novel protease inhibitors to target PR mutants that are not resistant due to active site mutations.

INDEX WORDS: Human immunodeficiency virus, HIV-1, HIV-1 protease, Protease inhibitors, Drug resistance, X-ray crystallography, Crystal structure, Enzyme kinetics, Protease mutant, PRS17, PR20, Amprenavir

CRYSTALLOGRAPHIC AND ENZYME KINETICS STUDIES OF MULTIDRUG RESISTANT HIV-1 PROTEASE MUTANTS PR20 AND PRS17

by

SHELLEY HINKLE BURNAMAN

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

2021

CRYSTALLOGRAPHIC AND ENZYME KINETICS STUDIES OF MULTIDRUG RESISTANT HIV-1 PROTEASE MUTANTS PR20 AND PRS17

by

SHELLEY HINKLE BURNAMAN

Committee Chair: Irene Weber

Committee: Robert Harrison

William Walthall

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

August 2021

DEDICATION

I dedicate this thesis to everyone that helped me to get where I am today. To my family, particularly my mom, husband, and son, for their endless support and understanding every time I needed space to work odd hours or be up all night writing. Without their love and patience, I could not have completed this dissertation.

Thank you to my committee members: my PI, Dr. Irene Weber, for keeping me going in the right direction and supporting me through all the ups and down of grad school. To Dr. Bill Walthall for giving me my first chance to work in a biology lab and for being my first mentor in science. To Dr. Robert Harrison for your invaluable guidance through this process.

Thank you to Yuan-Fang Wang for spending hours teaching me protein structure refinement and for all of the lovely conversations about crystallography. Thank you to Dr. Johnny Agniswamy, Dr. Daniel Kneller, and Dr. Andres Wong, for all the help and guidance with protein purification and enzyme kinetics. Thank you to all of the teachers I have had over the years for pushing me to always produce my best work.

Thank you to the GSU biology department and the Molecular Basis of Disease Fellowship for all the support and guidance over the years. Thank you to Asia Bright for being my best friend, lunch buddy, and sounding board for all these years as a PhD student. Extra special thanks to LaTesha Warren for knowing exactly what I needed to do in every situation I got myself into, and especially for being a listening ear when things were going wrong and for being the best cheerleader when things were going right. You are truly missed.

ACKNOWLEDGEMENTS

I would like to acknowledge the GSU Molecular Basis of Disease program for financially supporting me throughout my PhD years, the Southeast Regional Collaborative Access Team (SER-CAT) for granting access to beamlines for remote data collection, our collaborators at Oak Ridge National Laboratory, Purdue University, and National Institutes of Health. Thank you to the GSU Biology Core Facilities for providing equipment and training on countless occasions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

HIV-1	Human immunodeficiency virus -1
ART	Antiretroviral therapy
AIDS	Acquired immunodeficiency syndrome
HIV+	Person living with HIV
PR	Protease
APV	Amprenavir
DRV	Darunavir
WT	Wild type
WHO	World Health Organization
UNAIDS	Joint United Nations Programme on HIV/AIDS
CDC	Centers for Disease Control and Prevention
PDB	Protein Data Bank
CCP4	Collaborative Computational Program, number 4
RMSF	Root mean square fluctuation
RMSD	Root mean square difference
M	Molar
µM	micromolar
nM	Nanomolar
K_m	Michaelis constant
K_{cat}	turnover number
K_i	kinetic inhibition constant
EC_{50}	Half maximal effective concentration

IC_{50}	Half maximal inhibition concentration
CNS	Central nervous system
BBB	Blood brain barrier
FPLC	Fast protein liquid chromatography
HPLC	High performance liquid chromatography
RPC	Reverse phase chromatography
w/v	weight by volume
PEG	Polyethylene glycol
EDTA	Ethylenediaminetetraacetic acid
MES	2-ethanesulfonic acid
NaCl	Sodium chloride
[S]	substrate concentration
[E]	enzyme concentration
Å	Angstrom
VDW	van der Waals
MD	molecular dynamics
ns	nanosecond
µs	microsecond
bis-THF	*bis*-tetrahydrofuran
Cp-THF	cyclopentyl-tetrahydrofuran
LB	Luria Bertani
TE	tris-EDTA

1 INTRODUCTION

1.1 HIV continues to be a global public health threat

The year 2021 marks 40 years since the first cases of HIV/AIDS were reported. 38 million people worldwide are living with HIV as of 2020[1]. The 90-90-90 by 2020 goals were adopted by the WHO and UNAIDS in 2014[2]; this means that 90% of people living with HIV should know their status, 90% of people that are HIV+ are receiving antiretroviral therapy (ART), and 90% of people on ART have achieved viral suppression. As of 2020, ~84% of people living with HIV know their status, ~73% of people living with HIV are receiving ART, and 90% of people on ART are virally suppressed [1]. The WHO and UNAIDS adopted a new resolution on 11 June 2021 to revise these goals to 95-95-95 by 2025 and an end to the HIV/AIDS pandemic by 2030[3]. In the United States as of 2019, ~1.2 million adults and adolescents are living with HIV, and 14% of HIV+ people are unaware of their status[4, 5]. 81% of people living with HIV in the US are receiving at least some treatment for HIV, and 57% of HIV patients on ART in the US are virally suppressed[5]. The United States has adopted a program named Ending the HIV Epidemic in the US; these guidelines are similar to the resolution adopted by the WHO and UNAIDS to end the HIV pandemic by 2030[6]. This program, like the WHO and UNAIDS goals, includes widespread adoption of pre-exposure prophylaxis (PrEP) in critical at-risk populations, increased availability of rapid testing, decreased societal stigma, and increased access to medical care as key components of its strategy to defeat HIV. As of 2019, 23.4% of people in the US that are at an increased risk of exposure to HIV are taking PrEP[5].

HIV-1 causes significant weakening of the immune system by attacking CD4 T-cells and using them to replicate new infectious HIV-1 virions. HIV-1 infection presents clinically as flu-like symptoms at first, then progressing to weight loss, rash, swollen lymph nodes, oral ulcers, severe exhaustion, and generalized muscle aches[1]. If left untreated, HIV-1 will deplete the host immune system and progress to Acquired Immune Deficiency Syndrome (AIDS) in ~11 years[1]. HIV progresses to stage 3 (AIDS) when CD4 T lymphocyte counts fall below 200 [7]. HIV+ people are more susceptible to opportunistic infections that can lead to death, such as pneumonia, tuberculosis, or cancer. Although there is no cure for HIV, current ART guidelines have resulted in an extended life expectancy of people living with HIV. HIV is now considered to be a chronic, medically manageable disease.

1.2 The HIV life cycle and targets of ART

HIV-1 infection begins when virions bind the CD4 T-cell cell membrane by recognizing the CD4 surface protein and co-receptor CCR5 or CXCR4. The virion then fuses with the cell membrane using viral envelope glycoproteins, and the RNA genome and protein contents of the virion are emptied into the cytoplasm and transported into the nucleus. HIV reverse transcriptase (RT) transcribes the RNA genome to DNA, and HIV integrase (IN) inserts the viral DNA into the host genome. The virus hijacks host DNA replication machinery to make millions of copies of viral DNA. Once the viral DNA is copied, host RNA transcriptase transcribes the viral DNA to RNA, and RNA polymerase translates the RNA to Gag and Gag-Pol polyproteins. Gag, Gag-Pol, and a copy of the HIV-1 viral RNA genome assemble at the cell membrane and form a new immature virion. After the new virion is released from the cell membrane, protease (PR)

self-cleaves from Gag and Gag-Pol. PR then cleaves the other viral proteins to form a mature infectious virion.

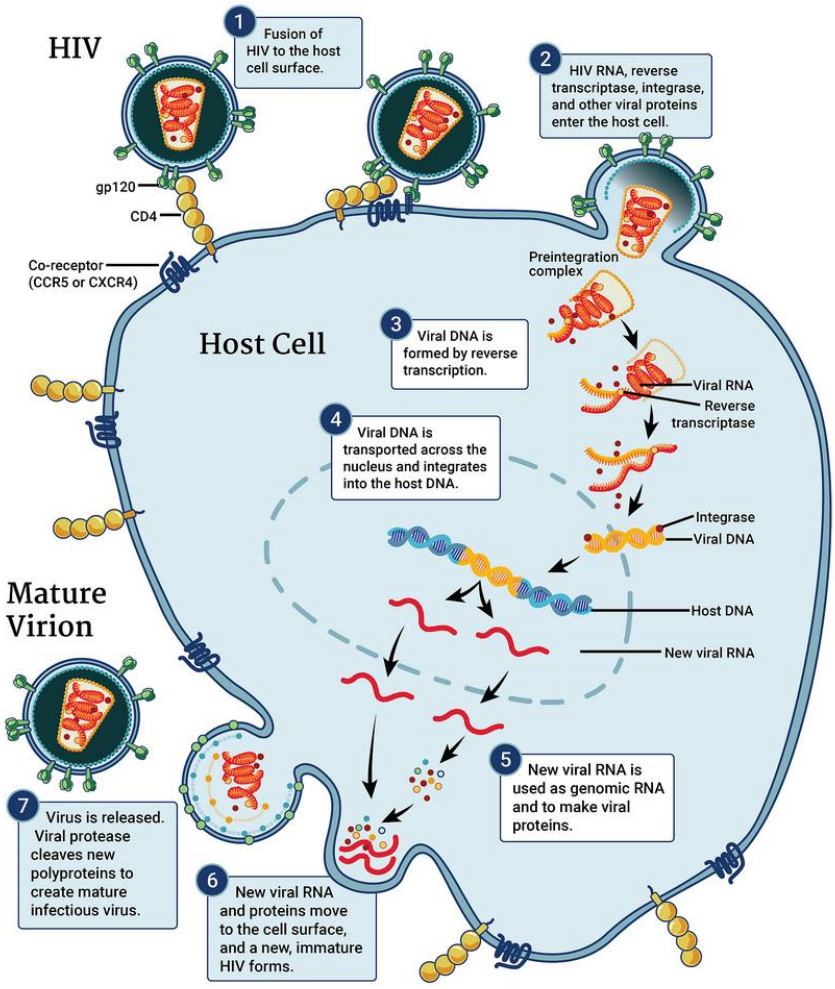

Figure 1: The Life Cycle of HIV-1 [8].

1.3 Current HIV ART guidelines

ART consists of a combination of drugs designed to target different parts of the HIV life cycle. Inhibitors of integrase, reverse transcriptase, protease, and viral binding and entry into the cell are part of the current arsenal of HIV medications. WHO recommends[9] first-line ART for adults and adolescents to include two nucleoside

reverse transcriptase inhibitors (NRTIs) plus a non-nucleoside reverse transcriptase inhibitor (NNRTI) or an integrase inhibitor (INSTI). For children, a first line regimen containing an NRTI and the protease inhibitor (PI) lopinavir (LPV). LPV is preferred for children due to its many pharmacologic formulations (powder, liquid, pill, etc.) and well-tolerated side effects. The second line treatment recommendation for adults and adolescents is two NRTIs and one protease inhibitor. LPV and atazanavir (ATV) are recommended for use first. A third PI, darunavir (DRV) is recommended as a third-line drug for patients with pre-treatment resistance to PIs or failure of treatment with LPV or ATV. Third line treatment generally combines DRV with other newer drugs like raltegravir (RAL), an integrase inhibitor, or second-generation NNRTIs. Cross-resistance with other first line NNRTIs and PIs is a problem; drugs like DRV and RAL are saved as a treatment of last resort. There is evidence that third line treatments are also failing, as pre-treatment resistance to current ART drugs is on the rise in places where adherence to ART is difficult. There is an urgent need for development of new HIV medications alongside strong HIV prevention programs. New PIs, in particular, are an important tool in combating the spread of HIV due to their ability to prevent HIV from becoming mature and infectious.

1.4 Inhibitors of HIV-1 protease and growing resistance

Nine PIs have been approved by the Food and Drug Administration (FDA) since 1995 (Figure 3): saquinavir (SQV), ritonavir (RTV), indinavir (IDV), nelfinavir (NFV), lopinavir (LPV), atazanavir (ATV), fosamprenavir (APV/FPV), tipranavir (TPV), and darunavir (DRV) (Figure 3). Ritonavir, though it has antiretroviral activity, is more commonly used as a booster for the most effective current formulations of PIs, such as

DRV/r, LPV/r, and ATV/r. The last PI to be approved was DRV in 2006. DRV has been used in salvage therapy for many years, but resistance to the last line PI is an increasing threat[10]. LPV, once used as a last line therapy, is now being used as the first choice PI for all age groups[11]. The crystal structure of HIV-1 PR, first described in 1989 [12, 13] was used to guide drug design for all the currently approved clinical inhibitors and novel protease inhibitors.

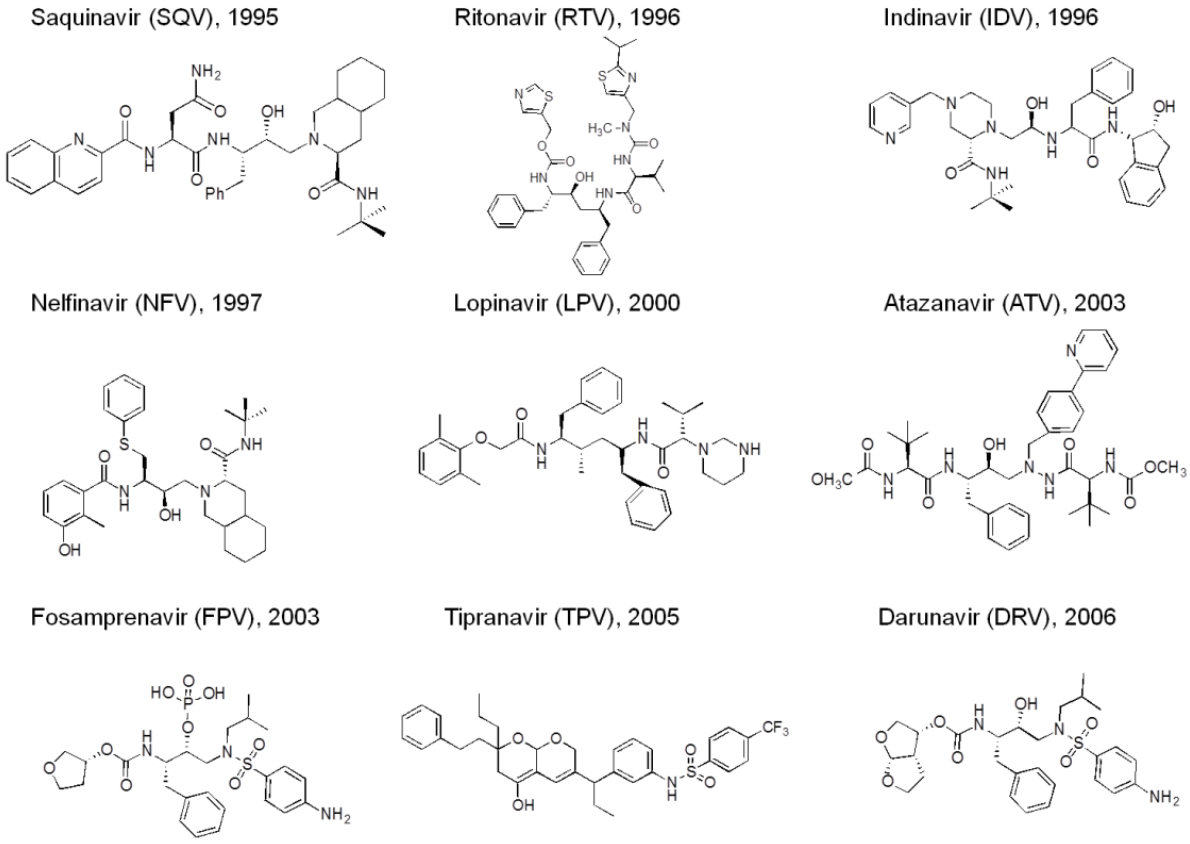

Figure 2. HIV PR inhibitors, currently FDA approved. *Fosamprenavir is also called amprenavir (APV). Inhibitors that are currently used today are lopinavir (LPV), atazanavir (ATV), and darunavir (DRV). DRV is used if treatment with other PIs has failed, or if a person has pre-treatment drug resistance to LPV or ATV [9, 11].*

Mutations in the viral PR gene cause resistance to PIs. The non-proofreading and very low fidelity aspects of HIV-1 RT results in a high rate of mutations to the HIV

genome. This allows the virus to mutate easily in response to drug pressure, selecting for new sets of mutations that allow the virus to survive in the presence of different drugs. The list of mutations that confer drug resistance (Figure 4) is periodically updated as new data become available[14]. The current research focuses on two mutant HIV-1 PRs: PR20 and PRs17. Each contains a set of mutations that allow PR to retain some catalytic activity under drug pressure while decreasing affinity for PIs. Both mutants were found as clinical isolates; they are from HIV patients that received ART.

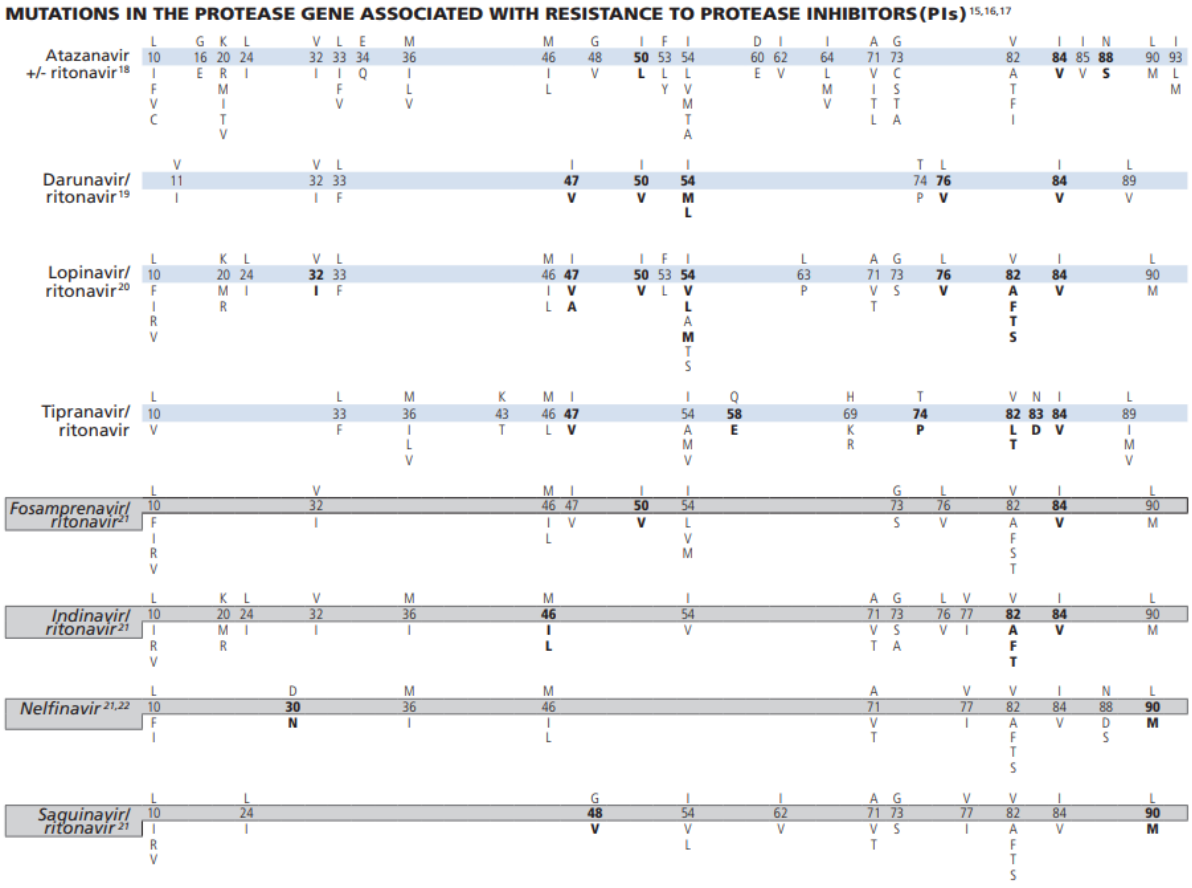

Figure 3. Drug resistance mutations for HIV-1 PR. *Major resistance mutations are highlight in bold, minor resistance mutations are shown in plain text. [14]*

1.5 Structure of HIV-1 Protease

There are currently more than 500 structures of HIV-1 protease in the Protein Data Bank (PDB; www.rcsb.org). HIV-1 protease (PR) is a homodimer consisting of two 99 amino acid subunits with amino acids number 1-99 and 1'-99' (Figure 4). Protease cleaves itself from Gag and Gag-Pol polyproteins, and then cleaves the rest of the viral proteins in the final maturation step. The homodimer forms an active site that uses two aspartic acid residues, one from each subunit, to catalyze the cleavage of viral polyproteins. Above the active site are two flexible flaps (shown in yellow in Figure 4) that open to allow substrate or inhibitor to enter the active site and close when an inhibitor or substrate is bound.

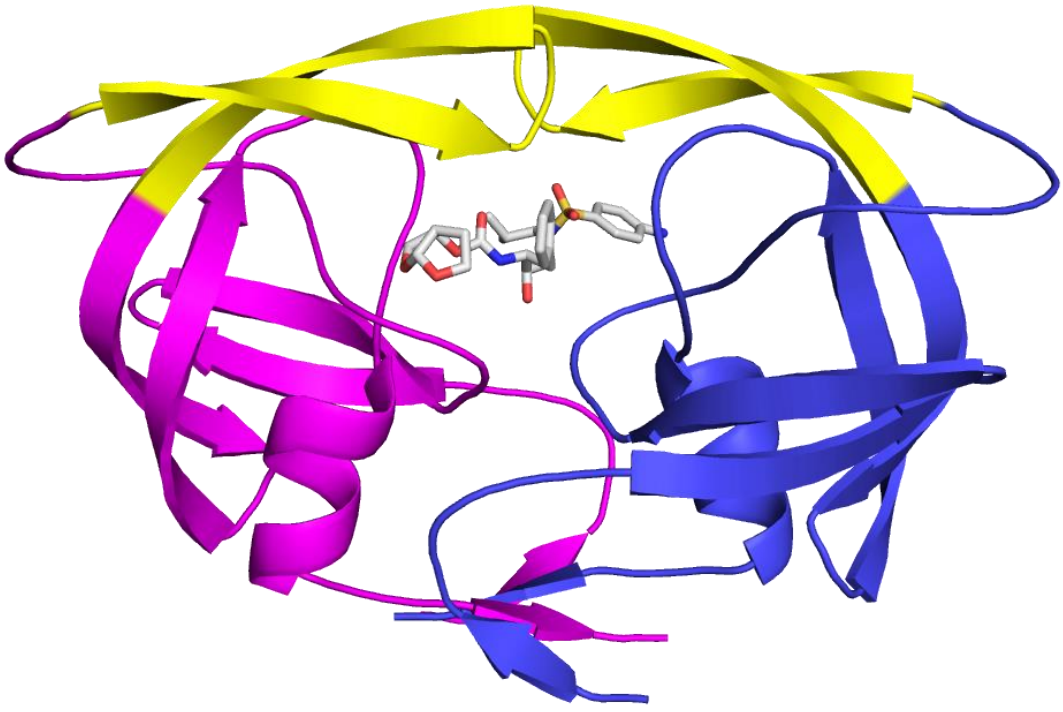

Figure 4: General structure of HIV-1 protease. *PR is made of two subunits, shown here in magenta and blue. The flaps (yellow) at the top of the protein close when substrate or inhibitor is bound. The dimer is held together by an inter-subunit β-sheet. Shown here is PDB structure 5T2Z (PRS17 bound to DRV).*

1.6 Drug-resistant mutants PR20 and PRS17

PR20 was chosen for study from [15] and PRS17 was selected using machine learning on genotype-phenotype data from the Stanford University HIV Drug Resistance Database (http://hivdb.stanford.edu). PR20 has 19 mutations that confer drug resistance and one mutation (Q7K) added to prevent autoproteolysis during purification[16]. It is resistant to all nine FDA-approved protease inhibitors and many novel inhibitors[16-18]. PR20 evades protease inhibitors via changes to the active site and the flaps. The active site mutations D30N, V32I, I47V, and I84V reduce the number of hydrogen bonds that both substrate and inhibitors make with PR, leading to increases in K_m and K_i and resistance to clinical inhibitors. This research investigates how two novel PIs interact with PR20.

PRS17 is also highly resistant to all clinical protease inhibitors, though its resistance is through a different mechanism than PR20. While PR20 evades PR inhibitors through structural rearrangements in the active site, the mutations in PRS17 are mostly found distal to the active site. Only two out of 17 mutations in PRS17 are found in the active site: G48V and V82S. Both mutations are critical for binding CA-p2 and p2-NC substrate analogs. G48V and V82S together enhance binding of substrates to PR while decreasing inhibitor binding, conferring resistance to PIs. G48V is associated with resistance to ATV and SQV[14], and it is selected for by ATV, IDV, LPV, and NFV[19-21]. G48V and V82S enhance binding of substrate analogs CA-p2 and p2-NC through structural rearrangements in the 80s loop[22]. This research further probes

the contribution to resistance of mutation G48V by studying the revertant mutation (V48G).

1.7 Rationale for studying novel inhibitors GRL-0489 and GRL-0739 bound to multidrug resistant PR20

Protease inhibitors based on the structure of darunavir are being designed to combat drug resistance by targeting the active site backbone and filling the expanded binding pocket in highly mutated proteases. PR20 has a larger binding pocket relative to WT PR, weakening chemical interactions between the inhibitor and protein, thus decreasing the binding affinity of the inhibitor[16]. The two novel inhibitors in this research were designed to fill the expanded binding pocket and to interact with the backbone of mutated PRs at Gly48. GRL-0739A [23, 24] has a *tris*-THF derivative (cyclohexyl-*bis*-THF) P2 group (cyclohexane ring fused to *bis*-THF) and GRL-0489A [25] has a cyclopentyl-THF group in place of the *bis*-THF group of DRV. Both contain a methoxy (OMe) P2' group in place of DRV's NH_2 P2'.

GRL-0739A is of particular interest due to its impressive antiviral potency (EC_{50} = 0.007 to 0.033 μM for multidrug resistant HIV-1 strains), inhibition of a variety of multidrug-resistant proteases, reduced susceptibility to drug resistance, and ability to penetrate the central nervous system (CNS)[24, 26]. Microglial cells in the CNS are the primary reservoir for HIV [27], as most current HIV drugs are not able to cross the blood brain barrier (BBB). Penetrating the BBB is an important goal for new PIs due to the high prevalence of HIV-associated neurocognitive disorder (HAND). HAND is a syndrome characterized by progressive loss of coordination, balance, memory, and decision-making, and it is a common cause of dementia in HIV patients[28]. HAND is a

problem for 40-50% of people living with HIV, regardless of whether they are receiving ART. Improving the design on PIs so that they can inhibit HIV-1 in viral reservoirs in the CNS will prevent or reduce the onset of HAND or HIV-associated dementia.

1.8 Rationale for studying revertant mutant PRS17$_{V48G}$

PRS17 is highly resistant to all clinical protease inhibitors, but there are only two mutations found in the active site. Previous work has demonstrated that changes distal to the active site are responsible for structural rearrangements that lead to worse inhibition by clinical protease inhibitors [29, 30]. The mechanism was further probed in the current research by reverting the flap mutation, G48V, back to the wild-type amino acid glycine. The mutation G48V has been previously shown to work together with mutation V82S to change how the CA-p2 substrate analog binds to PRS17[22]. G48V breaks the van der Waals (VDW) contacts made by CA-p2 P3 Arg to Phe53, causing P3 Arg to swing towards and make a water-mediated contact with PRS17 Arg8. This disrupts a crucial inter-subunit ion pair formed by Arg8’ and Asp29. Breaking this ion pair destabilizes the protease dimer. Though this work did not investigate how CA-p2 interacts with PRS17V48G revertant mutant, this could be explored in the future.

The scope of this project involves generating the revertant mutant, characterizing the inhibitor-free open flap form, and investigating how it interacts with two clinical protease inhibitors, DRV and APV.

2 REVERTANT MUTATION V48G IN HIGHLY DRUG RESISTANT PROTEASE S17 ALTERS FLAP CONFORMATION AND BINDING OF INHIBITORS

Submitted to Journal of Molecular Graphics and Modelling on 10 June 2021

Shelley H. Burnaman[1], Daniel W. Kneller[2], Yuan-Fang Wang[1], Andrey Kovalevsky[2], Irene T. Weber[1]

[1]Department of Biology, Georgia State University, Atlanta GA, USA.

[2]Neutron Scattering Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN 37831, USA; National Virtual Biotechnology Laboratory, US Department of Energy, Washington, DC, USA.

Corresponding author: Irene T. Weber, Department of Biology, P.O. Box 4010, Georgia State University, Atlanta GA, USA, 404-413-5411, iweber@gsu.edu

Highlights:

Mutation G48V in the protease flaps is common in drug resistant HIV.

Highly resistant protease with G48V shows curled flaps and extreme conformational dynamics.

Substituting wild type Gly48 alters inhibition and dynamics of mutant toward wild type values.

ABSTRACT

Drug resistance is a serious problem for controlling the HIV/AIDS pandemic. Current antiviral drugs show several orders of magnitude worse inhibition of highly resistant clinical variant PRS17 of HIV-1 protease compared with wild type protease. We have analyzed the effects of a common resistance mutation G48V in the flexible flaps of the protease by assessing the revertant $PRS17_{V48G}$ for changes in enzyme kinetics, inhibition, structure, and dynamics. Both PRS17 and the revertant showed about 10-fold poorer catalytic efficiency than wild type enzyme (0.55 and 0.39 $\mu M^{-1}min^{-1}$ compared to 6.3 $\mu M^{-1}min^{-1}$). Clinical inhibitors, amprenavir and darunavir, showed 2-fold and 8-fold better inhibition, respectively, of the revertant than of PRS17, although the inhibition constants for $PRS17_{V48G}$ were still 25 to 1,200-fold worse than for wild type protease. Crystal structures of inhibitor-free revertant and amprenavir complexes with revertant and PRS17 were solved at 1.3-1.5 Å resolution. The amprenavir complexes of $PRS17_{V48G}$ and PRS17 showed no significant differences in the interactions with inhibitor, although changes were observed in the conformation of Phe53 and the interactions of the flaps. The inhibitor-free structure of the revertant showed flaps in an open conformation, however, the flap tips do not have the unusual curled conformation seen in inhibitor-free PRS17. Molecular dynamics simulations were run for 1 μs on the two inhibitor-free mutants and wild type protease. PRS17 exhibited higher conformational fluctuations than the revertant, while the wild-type protease adopted the closed conformation and showed the least variation. The second half of the simulations captured the transition of the flaps of PRS17 from a closed to a semi-open state, whereas the flaps of $PRS17_{V48G}$ tucked into the active site and the wild type protease

retained the closed conformation. These results suggest that mutation G48V contributes to drug resistance by altering the conformational dynamics of the flaps.

2.1 Introduction

The HIV/AIDS pandemic has caused almost 33 million deaths in the past 40 years. By current estimates, about 38 million people are infected worldwide, including 1.7 million new infections as of 2019[1]. 68% of infected adults and 53% of infected children are on antiretroviral therapy (ART)[1]. ART uses a combination of drugs designed to target different aspects of the viral life cycle, such as inhibitors of the viral enzymes, integrase, reverse transcriptase, protease (PR), and viral entry into the host cell. However, the effectiveness of ART is under attack by the growing problem of drug resistance. In addition to resistance acquired during treatment, pre-treatment drug resistance is also a worsening problem worldwide. The World Health Organization (WHO) estimates that over 50% of new diagnoses in children in some parts of Sub-Saharan Africa already have resistance to non-nucleoside reverse transcriptase inhibitors (NNRTIs)[31]. Up to 30% of adults in Central and South America with new HIV infections are initially placed on second line protease inhibitor (PI)-containing ART due to resistance to first line NNRTI or NRTI regimens. WHO recommends using PI-containing ART as a first treatment in countries where >10% of new HIV infections show pre-treatment NRTI/NNRTI resistance[11]. The recent disruption caused by the COVID-19 pandemic for ART in low-income and middle-income countries is likely to be understated due to development of drug resistance[32]. HIV PR cleaves the viral Gag

and Gag-Pol precursor polyproteins in the last step of viral maturation to yield infectious viral particles[33].

HIV PR is an aspartic protease that consists of two 99-residue subunits. Nine PIs have been approved for therapeutic use, including amprenavir (APV) and darunavir (DRV). Resistance to PIs occurs primarily by mutations in the PR gene [14]. The virus accumulates different mutations that improve viral fitness and increase drug resistance. Highly resistant variants with 9-20 mutations in PR and 3 orders of magnitude worse inhibition have been observed in patients[34].

We have investigated the structures and activities of highly resistant clinical variants of PR, including PR20[16, 35], PRS17 [22, 29, 30] and PRS5B[36]. DRV has the highest affinity (5 pM) of the clinical PIs for wild type PR, however, it is unlikely to be effective on virus bearing the mutations in PR20 or PRS17, since the binding affinity for these mutants is 40 and 50 nM, respectively[16, 29, 30]. PR20 bears 20 mutations, including 4 in the inhibitor-binding site, and gains resistance to PIs through an enlarged binding site for inhibitors and altered flap dynamics [16, 35]. PRS17 was chosen by machine learning on genotype-phenotype data from the Stanford University HIV Drug Resistance Database[37, 38] to represent a variant that is extremely resistant to several PIs [39-41]. Experiments confirmed that PRS17 is poorly inhibited by all tested PIs with APV showing the best inhibition (*Ki* of 11 nM) [29, 30]. PRS17 has 17 mutations including only two mutations, G48V and V82S, in the active site. Unlike PR20, the mechanism of resistance for PR^{S17} is primarily through structural rearrangements that alter the conformational dynamics, coupled with enhanced binding to substrate analog peptides [22, 29].

The current study seeks to assess the effect of drug resistant mutation G48V in the flap of PR. The flap comprises residues 45-55, which form a pair of antiparallel beta-strands in each subunit. The tips of the two flaps interact in the PR dimer and are critical for substrate and inhibitor binding[42, 43]. Mutations of several residues in the flaps confer drug resistance[14]. G48V was first identified as a drug resistance mutation (DRM) in virus resistant to SQV[44, 45]. It is a major DRM for SQV and nelfinavir, however, it is not associated with resistance to APV or DRV [14]. In PRS17 structures, G48V is associated with shifts in the active site and differences in interactions with substrate analogs compared to the wild type protease [22, 29]. Flap mutations G48V, M46L and I54V contribute to a 144-170° twist and ~3 Å displacement in the flap, leading to a curled conformation when inhibitor is absent[29]. In order to investigate the contribution of flap mutation G48V to drug resistance of PRS17, we constructed the revertant to the glycine present in wild type enzyme. We hypothesized that the V48G revertant would alter the enzymatic properties and flap conformation of PRS17 towards those of the wild type PR. This hypothesis was assessed using enzyme kinetics experiments, crystallography, and molecular dynamics simulations on the revertant, PRS17, and the wild type PR.

2.2 Methods

2.2.1 Inhibitors

Inhibitors APV and DRV were obtained from the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH.

2.2.2 Expression and purification of PRS17 and PRS17$_{V48G}$

Genes coding for PRS17 and PRS17$_{V48G}$ were synthesized and cloned in pJ414 vector (Atum) and expressed in *E. coli* BL21 (DE3) as previously described[22], with some modifications for PRS17$_{V48G}$. Protein overexpression was induced with IPTG and incubated overnight at room temperature. Homogenized cells were incubated overnight at 4 °C with DNase (Roche) and Triton X-100 (Sigma Millipore). Cells were incubated for an additional hour at room temperature the following day, then sonicated and clarified. Protein inclusion bodies were purified using Fast Protein Liquid Chromatography and Reverse-phase High Performance Liquid Chromatography (RP-HPLC) as described in[22]. RP-HPLC fractions were dialyzed and refolded as previously described[13] and stored in 0.3 mg/mL aliquots for use in enzyme kinetics assays. RP-HPLC fractions were concentrated using Amicon centrifugal filters to 4.2 mg/mL for protein X-ray crystallography experiments.

2.2.3 Enzyme kinetics and kinetic inhibition assays

Fluorescence resonance energy transfer-based enzyme kinetics assays were performed as described[13] in reaction buffer (50 mM MES, 200 mM NaCl, 0.5 mM EDTA, and 2.5% glycerol), anthranilyl-HIV protease p2-NC cleavage site substrate analog (Ace-T-I-Nle-r-Nle-Q-R), and purified enzyme. Enzyme kinetics experiments were performed in triplicate at 37 °C in 100 uL reaction volume. K_m and k_{cat} were determined based on Michaelis-Menten kinetics using 12-96 µM substrate [S]. Enzyme concentration [E] was determined via active site titration with APV in triplicate. k_{cat} was calculated as (Vmax)/[E]. IC_{50} was obtained from dose-response curves with 60 µM [S].

K_i values were determined in triplicate measurements using equation K_i = (IC50 – [E]/2)/(1 + [S]/K_m). All kinetic calculations were performed using the enzyme kinetics module and non-linear regression tool of SigmaPlot 12.0 (Systat Software Inc., San Jose, CA) and the data solver tool for non-linear curve fitting in Microsoft Excel.

2.2.4 X-ray data collection and structure refinement

Crystals of inhibitor-free PRS17$_{V48G}$ and APV complexes of PRS17/APV and PRS17$_{V48G}$/APV were grown by hanging drop vapor diffusion at room temperature by mixing 1 µL of protein and 1 µL of well solution. Crystals of inhibitor-free PRS17$_{V48G}$ were grown from a mother liquor containing 0.2 M Na citrate tribasic dihydrate, 0.1 M HEPES pH 7.5, and 20% w/v 2-propanol. PRS17$_{V48G}$ at 4.2 mg/mL was complexed with APV at a 1:8 molar ratio, and crystals were obtained from a mother liquor containing 26% PEG 8000, 0.1 M sodium cacodylate pH 6.7, and 0.2 M sodium acetate. PRS17 at 5 mg/mL was complexed with APV at a 1:5 molar ratio and crystals were grown in 37.5% Tacsimate (Hampton Research). Crystals were frozen in liquid nitrogen and cryoprotectant made of 30% glycerol and 70% mother liquor, and X-ray diffraction data were collected on Southeast Regional Collaborative Access Team (SER-CAT) beamline 22-ID at Argonne National Lab, Chicago, IL. Structures were solved by molecular replacement using Phaser[46] and previously solved structures 2PC0[47] for the inhibitor-free PRS17$_{V48G}$ structure and 5T2E[29] for PRS17$_{V48G}$/APV and PRS17/APV. The atomic models were fit into $2F_0$-F_c and F_c-F_c maps contoured at 1 and 3 sigma levels, respectively, in iterative rounds of refinement using COOT[48] and Refmac5[49] in CCP4[50]. Two alternative conformations of APV, each at 50% occupancy, were

modeled in the active sites of PRS17/APV and PRS17$_{V48G}$/APV. Residues with ambiguous density in the PRS17/APV structure were removed in the beginning of refinement and later rebuilt into the electron density map. Extra electron density visible at the flap tips may be due to the refinement program (Refmac5) since it disappears when modeled as waters using SHELXl-2014[51]. Structural figures were made using PyMOL (http://www.pymol.org).

2.2.5 Molecular dynamics simulations

All-atom molecular dynamics (MD) simulations of inhibitor-free PRS17 (PDB: 52TE), wild-type PR (PDB: 2PC0), and the PRS17$_{G48V}$ revertant were prepared, run, and analyzed using GROMACS2020[52] with the CHARMM36m force field[53]. Each dimer was placed in a rhombic dodecahedron with a minimum of 10 Å between any protein atom and the nearest cell edge. The systems were solvated with water molecules (TIP3P model[54]) and 150 mM NaCl with additional charge-neutralizing Cl- ions. Each system was minimized to <10 kJ/mol using the steepest descent algorithm before being equilibrated to 300 K using the V-rescale thermostat[55] and 1 bar using the Berendsen[56] barostat. Following equilibration, production MD was performed for 1 µS using the leapfrog integrator with Nose-Hoover[57, 58] and Parrinello-Rahman[59] couplings with a 2 fs integration step. Periodic boundary conditions were applied. Bonds to hydrogen were constrained using the LINCS[60, 61] algorithm. Coordinates of atoms were saved every 10 ps. RMSD and RMSF calculations were performed on backbone atoms. A distinct flap closure event for PRS17$_{G48V}$ was witnessed at approximately 500

ns. This observation prompted the RMSF analysis to be split into two non-overlapping time domains consisting of the trajectories prior to and after that time point.

2.2.6 Structural Data

The coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 7N6V for PRS17$_{V48G}$/APV, 7N6X for PRS17/APV, and 7N6T for inhibitor-free PRS17$_{V48G}$.

2.3 Results

2.3.1 Kinetic parameters and inhibition of PRS17$_{V48G}$

Enzyme kinetic parameters were determined for the new mutant PRS17$_{V48G}$ and compared with values for PRS17 and wild-type PR (Table 1). The revertant mutant PRS17$_{V48G}$ has a K_m for p2-NC substrate of 80 μM or about half the value for PRS17 (K_m = 140 μM), and almost three-fold larger than the value of 30 μM for wild-type enzyme. PRS17$_{V48G}$ shows the lowest k_{cat} value of 30 min^{-1} compared to 70 min^{-1} for PRS17 and 190 min^{-1} for wild-type enzyme. The catalytic efficiency (k_{cat}/K_m) for both mutants is less than a tenth of the value for wild-type PR, while PRS17$_{V48G}$ and PRS17 have similar values of 0.39 and 0.49 $\mu M^{-1}min^{-1}$, respectively.

APV and DRV were chosen for inhibition measurements since APV is the best PI for PRS17 (K_i of 11 nM) and DRV is the most potent PI for wild-type PR (K_i of 5 pM). APV and DRV showed similar K_i values of 5-6 nM for PRS17$_{V48G}$ or about 2-fold and 8-fold better, respectively, than for PRS17. These K_i values, however, are still much worse

than the 5 pM for DRV and 0.2 nM for APV reported for wild-type enzyme, which indicates the importance of the other 16 mutations in PRS17.

Table 1: Enzyme Kinetic Data for PRS17$_{V48G}$, PRS17 and Wild-type PR.

	WT PR	PRS17	PRS17$_{V48G}$	Relative values (PRS17$_{V48G}$/PRS17)	Relative values (PRS17$_{V48G}$/WT)	Relative values (PRS17/WT)
K_m (μM)	30	140	80 +/- 13	0.57	2.7	4.7
k_{cat} (min^{-1})	190	70	30 +/- 6	0.43	0.16	0.37
k_{cat}/K_m ($\mu M^{-1} min^{-1}$)	6.3	0.49	0.39 +/- 0.01	0.78	0.06	0.08
K_i APV (nM)	0.2[62]	11	5.2 +/- 2	0.47	26.0	55.0
K_i DRV (nM)	0.005[63]	50	6.4 +/- 1	0.13	1,280	10,000

Kinetic parameters for PRS17 and wild-type PR are taken from *[16, 22]* and [64], respectively. Inhibition values for PRS17 are taken from [30], and values for wild-type PR from [62, 63]. Standard deviations are reported for new measurements.

2.3.2 Crystal structures of APV complexes with PRS17 and PRS17$_{V48G}$ have different interactions in the flap tips

Crystal structures were obtained for the mutants PRS17 and PRS17$_{V48G}$ in complex with APV. Crystallographic data collection and refinement statistics are in Table 2. The structures of PRS17$_{V48G}$/APV and PRS17/APV were solved in space group $P6_1$ at resolutions of 1.49 Å and 1.47 Å and R-factors of 16.8 and 19.9 %, respectively. Similar to the wild-type PR/APV, these structures have one dimer in the asymmetric unit, and amino acids in the subunits are numbered 1-99 and 1'-99'. In both complexes, APV was refined in two alternate orientations of 50% occupancy and related by 180°. Residues of PRS17 and PRS17$_{V48G}$ and the inhibitor APV were unambiguously

modeled into electron density maps. The dimer of PRS17/APV and the sites of mutation are illustrated in Figure 1.

Table 2: Crystallographic data collection and refinement statistics.

Statistics	**PRS17$_{V48G}$ inhibitor-free**	**PRS17/APV**	**PRS17$_{V48G}$/APV**
Space group	$P4_1$	$P6_1$	$P6_1$
Cell dimensions			
a (Å)	45.34	62.95	62.81
b (Å)	45.34	62.95	62.81
c (Å)	105.1	82.55	82.45
γ (°)	90.0	120.0	120.0
Resolution range (Å)	50-1.32	50-1.47	50-1.49
Unique reflections	46202	26478	29996
Redundancy	4.7 (3.3)	4.7 (4.9)	9.4(8.0)
Completeness	97.6 (86.1)	98.4 (98.4)	100(100)
<I/σ(I)>	21.6 (2.48)	31.9 (3.74)	24.0 (4.0)
R_{sym} (%)	5.6 (52.8)	4.1 (49.4)	8.3 (57.2)
Refinement resolution range (Å)	34.35-1.32	32.97-1.47	32.88-1.49
R_{cryst} (%)	15.9	19.9	16.8
R_{free} (%)	18.9	22.7	22.16
Number of solvent molecules	161	66	75
Average B-factor (Å^2)	25.3	22.3	25.0
RMS deviations from ideality			
Bond length (Å)	0.02	0.02	0.01
Angles (°)	2.2	2.4	1.9

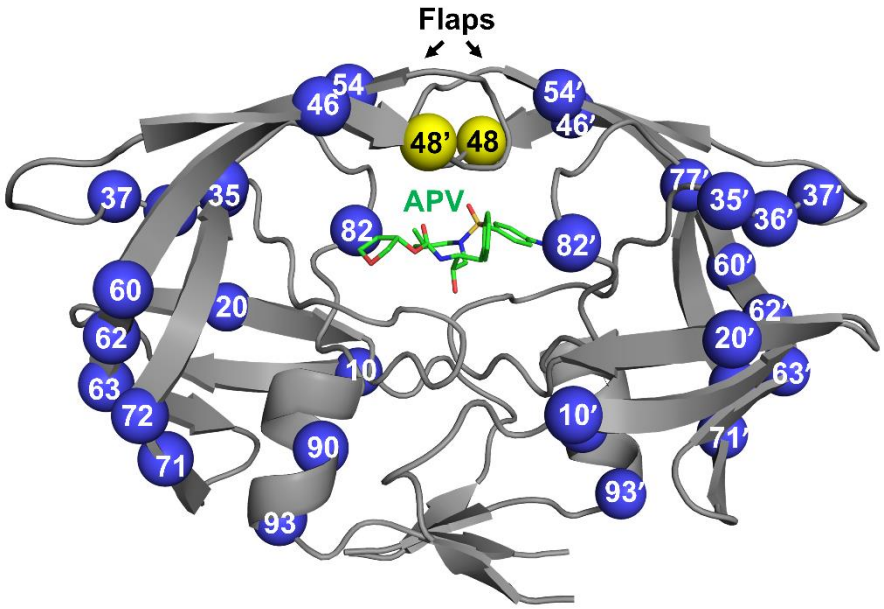

Figure 5: Sites of mutations in PRS17 dimer. The PRS17 dimer is shown in grey ribbons with inhibitor APV in green sticks. Sites of mutations are shown as yellow spheres for G48V in the flaps and blue spheres for the other 16 mutations (L10I, K20R, E35D, M36I, S37D, M46L, I54V, D60E, I62V, L63P, A71V, I72V, V77I, V82S, L90M, I93L).

The overall structures were consistent with previously solved complexes of PRS17[22, 29, 30]. The dimer of PRS17$_{V48G}$/APV superposed onto PRS17/APV with a low average root-mean square deviation (RMSD) of 0.26 Å on 198 corresponding Cα atoms, which suggests the two structures are almost identical. Since the two structures were solved in the same space group at similar resolutions, structural differences between PRS17/APV and PRS17$_{V48G}$/APV can be attributed to the V48G revertant mutation. The structures of PRS17/APV, PRS17V$_{48G}$/APV and wild-type PR/APV show conserved (<0.2 Å variation in length) hydrogen bond interactions with APV in the active sites (Figure 2).

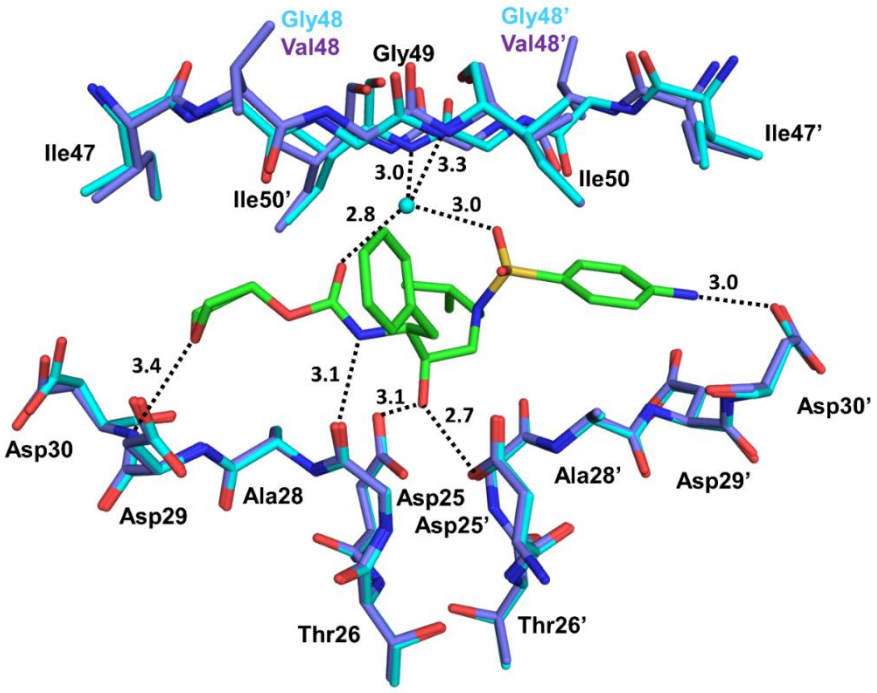

Figure 6: Hydrogen bond interactions of PRS17$_{V48G}$ and PRS17 with APV. *Hydrogen bond interactions with APV are conserved in the structures of PRS17$_{V48G}$/APV (cyan) and PRS17/APV (purple). APV is shown in green and the cyan sphere is a conserved water molecule. Alternate conformations of APV and active site residues were omitted for clarity. Distances are in angstroms (Å).*

Instead of altering the hydrogen bond interactions with APV, the structure of PRS17$_{V48G}$/APV shows changes in the hydrophobic contacts with inhibitor and around residues in the flap tips compared to PRS17/APV (Figure 3). In PRS17/APV, the side chain of Phe53 rotates away from the opposite antiparallel β-strand of the flap to accommodate the valine side chain at residue 48 (Fig 3A). This rotation of Phe53 is reversed in the A subunit of PRS17$_{V48G}$/APV with Gly48 and changes to a lesser extent in the B subunit. These structural rearrangements of Phe53 are accompanied by small changes in the hydrophobic contacts of residues 46-53 in both subunits. In the A subunit of PRS17/APV, 14 hydrophobic interactions are observed between the two

strands of the flap, while the number of contacts increases to 25 in the revertant due to the rotated Phe53 (Fig 3 B and C). Associated small shifts in Ile50 at the flap tip also increase the interactions with Pro81'. In the same subunit, Gly48 and Pro81' show seven more van der Waals contacts with P1 and P2 groups of APV in PRS17$_{V48G}$/APV compared to PRS17/APV. In the B subunit of PRS17/APV, the side chain of Phe53' forms a hydrophobic contact with Leu46' that is absent in the revertant (Fig 3 D and E). There are no significant changes, however, in the interactions with P2' of APV in the B subunits of both mutants.

Overall, the differences in hydrophobic interactions around residue 48, the rotated conformation of Phe53, and the additional van der Waals contacts with P1 of APV are likely to contribute to the increased susceptibility of PRS17$_{V48G}$ to APV. The presence of glycine or valine at residue 48 may also influence the flap dynamics, which will contribute to the better inhibition by APV and DRV seen for PRS17$_{V48G}$ compared to PRS17.

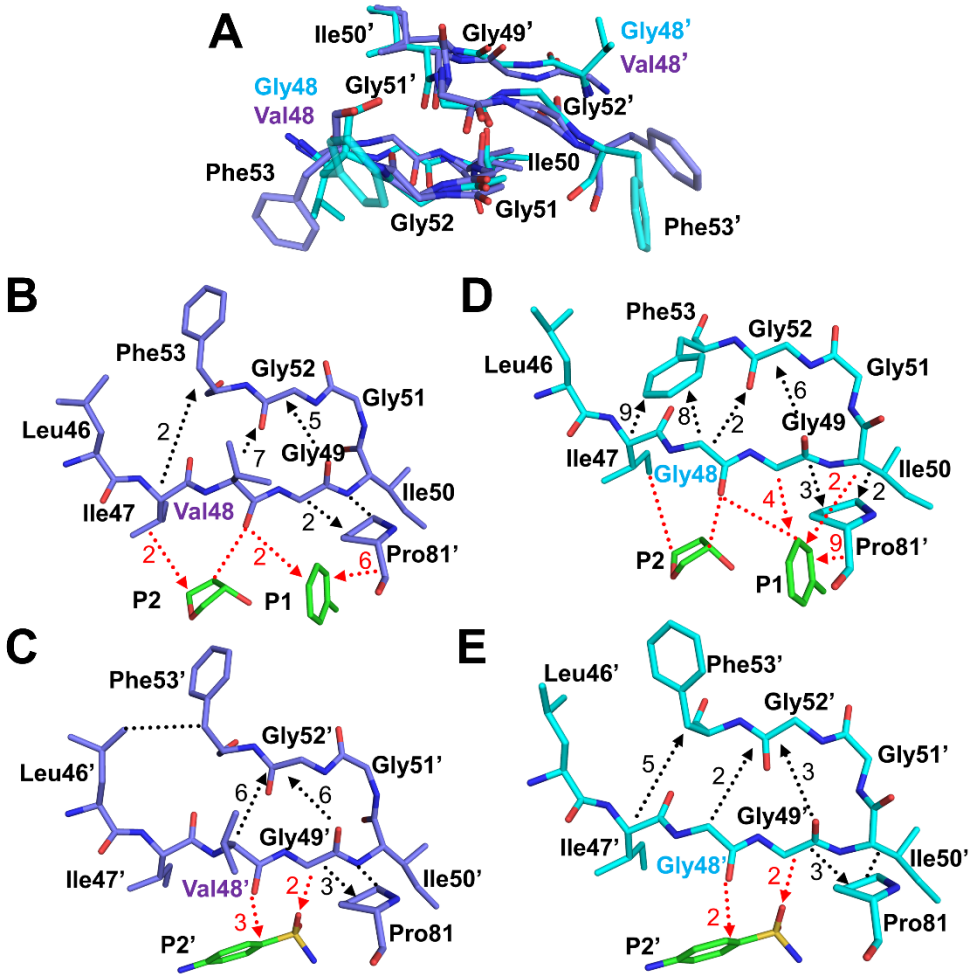

Figure 3: Hydrophobic interactions around residue 48 in PRS17$_{V48G}$ and PRS17. *PRS17 (purple) and PRS17$_{V48G}$ (cyan) around residue 48. A) Superimposed flap residues 48-53 in both subunits showing different side chain conformations of Phe53/53'. B) and C) Interactions of the A and B subunits of PRS17/APV. D) and E) Interactions of the A and B subunits of PRS17$_{V48G}$/APV. APV groups P2, P1 and P2' are in green. Van der Waals contacts are indicated by dotted lines with numbers and arrows for multiple contacts. Interactions with APV are shown in red. Residues without hydrophobic interactions were omitted for clarity.*

2.3.3 PRS17 curled flap conformation is lost in inhibitor-free PRS17V48G revertant

The ligand-free dimer of PRS17V48G is overall very similar to that of PRS17, with the most notable difference occurring in the flap region at the site of the V48G

revertant mutation. Many of the structural changes seen in PRS17 compared to WT have been rescued in the PRS17V48G revertant. PRS17 has a curl in the flaps that propagates through residues Val48, Gly49, Ile50, Gly51, and Gly52 (Figure 3). This flap curl leads to conformational changes in the position of the 80's loop, an important structure for substrate and inhibitor binding. All residue shifts were measured as the distance between Cα atoms. The carbonyl of Gly48 in PRS17V48G rotates ~140° and Ile50 shifts ~6.4 Å back to the WT position. The distance between Ile50 and Thr80 in PRS17V48G moves back to a WT-like 14.0 Å. Ile84 moves ~0.2 Å, weakening VDW contacts with APV compared to PRS17/APV. VDW interactions between Phe53 and Gly48 are restored in PRS17V48G. As expected, VDW interactions between PRS17 mutation M46L and Phe53 are not restored in PRS17V48G. In PRS17, Phe53 shifts ~1.1 Å compared to WT PR, but upon restoration of Gly48 the Phe53 shifts back ~0.4 Å closer to WT conformation in PRS17V48G. Even though the flap curl reverted closer to the WT position, the distance between Ile50 and Asp 25 (tip of flap to catalytic Asp in active site) is 19.0 Å, the same as PRS17. The distance between Ile50 and Asp25 in WT is 17.6 Å. These residue shifts together with enzyme kinetics data show that the V48G mutation in PRS17 appears to be responsible for many meaningful conformational changes to PRS17, but these data for the inhibitor-free structure do not fully explain why PRS17V48G is less resistant to PIs than PRS17.

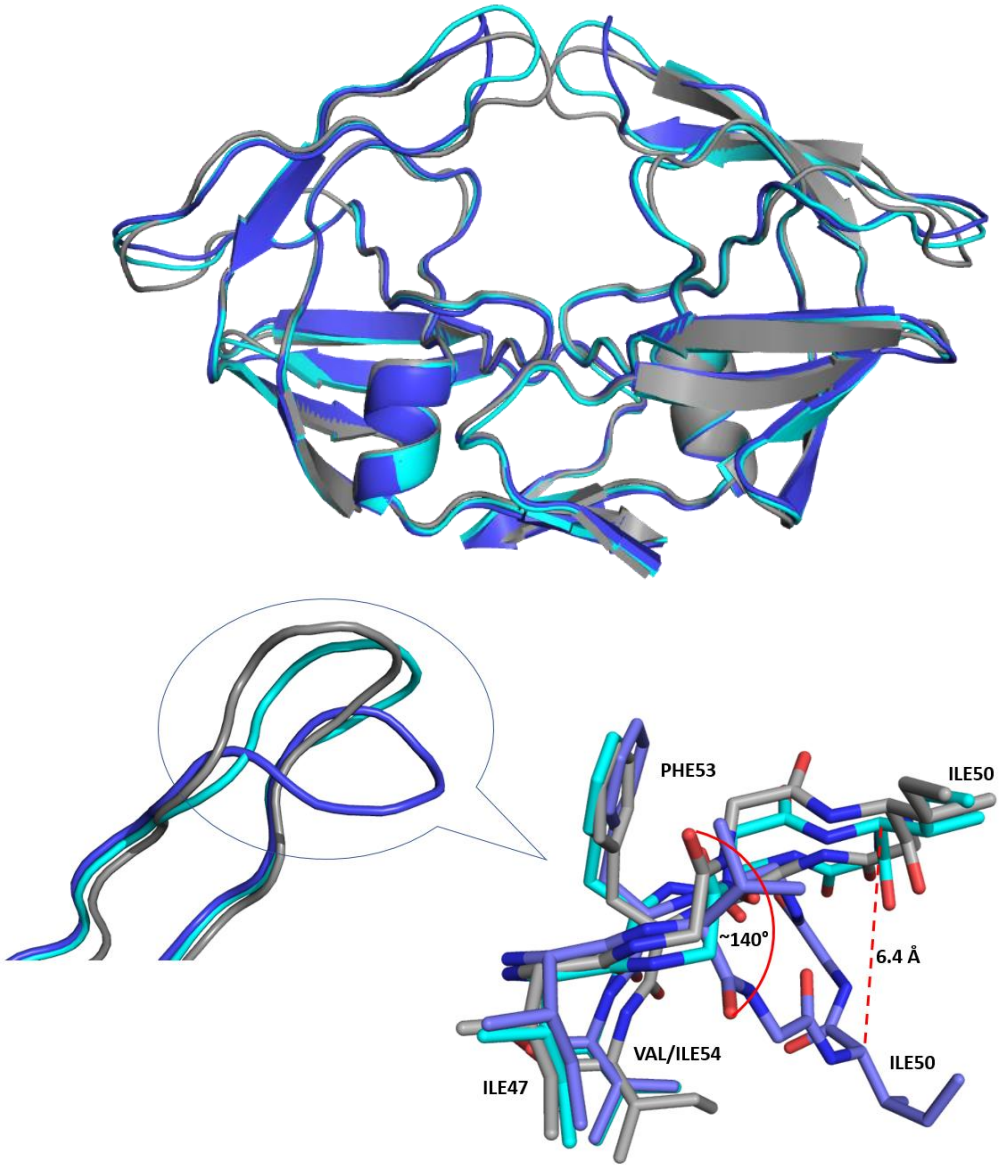

Figure 7. Comparison of PRS17V48G (cyan), PRS17 (blue), and WT protease (grey) flap conformations. *The body of each protease is largely the same, but the conformation of the flap tip differs for PRS17 vs. PRS17V48G and WT PR. A) Superposition of open flap conformation structures of PRS17V48G (cyan), PRS17 (5T2E; blue), and WT PR (2PC0; grey). The flaps of PRS17V48G adopt a more WT-like conformation, losing the curl seen in PRS17. B) PRS17V48G (cyan), PRS17 (blue), and WT PR (grey) were superposed to highlight differences between the structures. The carbonyl of G48 in PRS17V48G in flips ~140° and ILE50 (measured Cα to Cα) moves 6.4 Å.*

3.3 Crystal structure of inhibitor-free PRS17$_{V48G}$ lacks the curled flap conformation of PRS17

The crystal structure of inhibitor-free PRS17$_{V48G}$ was solved in space group $P4_1$ at 1.32 Å resolution and refined to an R-factor of 15.9 % (Table 2). This new structure of inhibitor-free revertant was compared to the most similar dimers of PRS17 and WT. Inhibitor-free PRS17 superimposed onto inhibitor-free PRS17$_{V48G}$ with an RMSD of 1.27 Å for Cα atoms (Figure 4A). The maximum deviation occurs at the flap tip (residues 49-52) with the largest difference of 7.9 Å measured between Cα atoms of Ile50 and 50' (Figure 4B). Inhibitor-free PRS17$_{V48G}$ superimposed on WT PR with an RMSD of 0.839 for Cα atoms. The maximum deviation of 3.6 Å occurred at residue Glu35. Compared to the open conformation of wild-type PR, both mutants show differences in the flap tips between residues 48-52, though ligand-free PRS17$_{V48G}$ is more similar to wild-type PR. The distance between PRS17$_{V48G}$ Ile50 and WT Ile50 was 2.1 Å, compared to 7.9 Å for PRS17$_{V48G}$ versus PRS17 and 6.9 Å for PRS17 versus WT PR [22]. The structure of inhibitor-free PRS17 has a distinctive curled conformation of the flaps that extends through residues Val48 to Gly52 (Figures 4A and B). In the revertant PRS17$_{V48G}$, the carbonyl of Gly48 rotates by ~140° which shifts the entire flap tip and moves Ile50 about 8 Å closer to the position of wild type Ile50.

The consequences of the changes in the flap tips can be assessed by examining the distances between Ile50 and the catalytic Asp25 or the 80's loop. The open conformation flaps of PRS17 and PRS17$_{V48G}$ are 1.1-1.4 Å farther from the catalytic Asp25 than in wild-type PR, (Figure 4C). The curled flap in PRS17 significantly increases the distance to residues of the 80's loop, which is an important region for

interactions with substrate and inhibitor. This change is reversed in the revertant. The flap tip of PRS17$_{V48G}$ shifts about 3 Å toward Thr80 in the 80s loop relative to the position of the curled flap of PRS17, although it is still 0.8 Å farther from Thr80 in the wild-type PR (Figure 4D). Overall, the flaps of PRS17$_{V48G}$ lose the curled tips seen in PR17 and adopt a conformation more similar to that of the wild-type PR. These structural shifts suggest that the V48G mutation in PRS17 has a major contribution to the distinctive open conformation with curled flap tips.

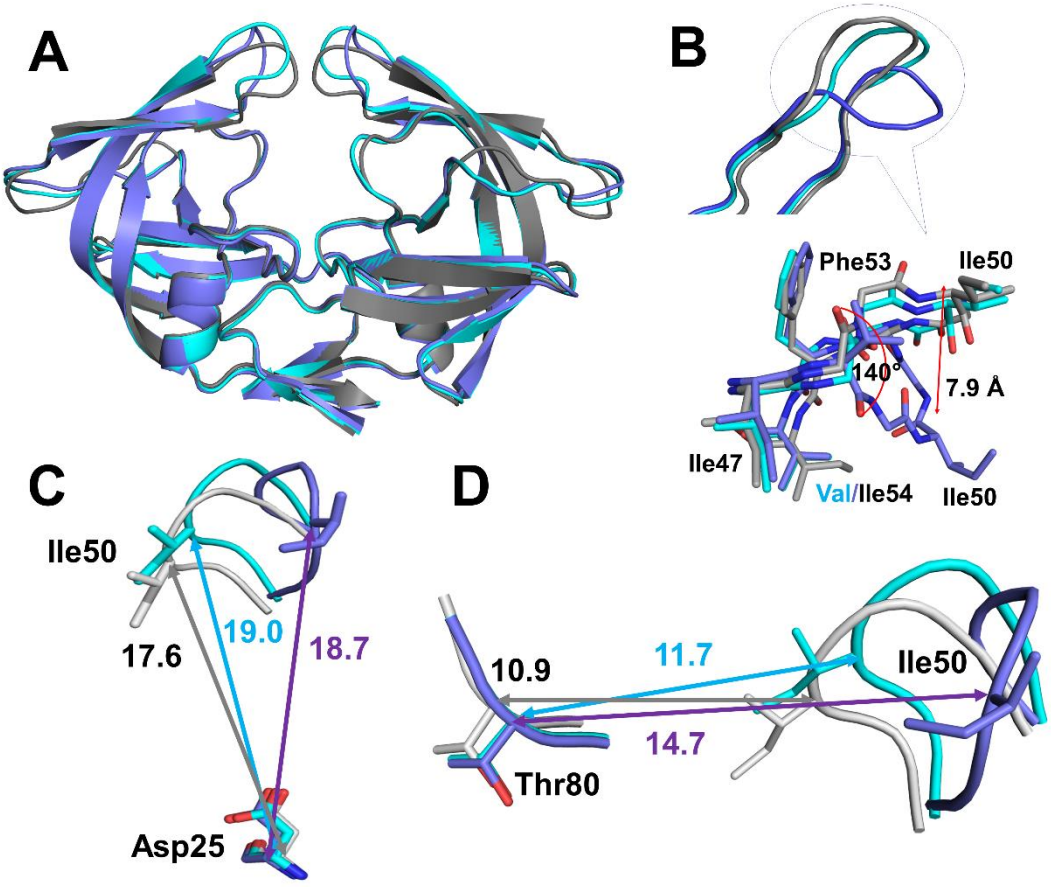

Figure 4: Change in flap tip conformation and distances between Asp25 and Thr80 due to V48G revertant mutation. *A) Superposition of open flap conformation structures of PRS17$_{V48G}$ (cyan), PRS17 (5T2E; blue), and WT PR (2PC0; grey). B) Differences in flap residues 47-54 of the structures. The carbonyl of Gly48 in PRS17$_{V48G}$ rotates by ~140° (red arc)*

and Ile50 (measured Cα to Cα) moves by 7.9 Å (red arrow). C) The distances in Å between Ile50 and Asp25 in the three structures. D) Distances between Ile50 at the flap tip and Thr80.

2.3.4 Molecular dynamics simulations exhibit greater flap dynamics for PRS17 than for PRS17$_{V48G}$

The influence of valine or glycine at position 48 on the conformational variation of the mutants was addressed in molecular dynamics (MD) simulations. MD simulations were conducted for inhibitor-free dimers of PRS17, the V48G revertant, and wild-type PR to assess how the identity of residue 48 effects the conformational variation of the flaps. To sample flap closing and opening events, 1 μs MD simulations were run starting from models of the inhibitor-free crystal structures with open flap conformations. The root mean square deviation (RMSD) of backbone atoms shows <4.5 Å difference between the coordinates in the simulation and the equilibrated starting structures for both dimers (Figure 5A). An earlier 10 ns MD simulation of the open conformation of PR20 showed <2.5 Å difference from the starting structure [35], whereas a more extensive 50 ns simulation gave up to about 5 Å deviation[65].

Visualization of the trajectories clearly indicates that PRS17 and PRS17$_{V48G}$ dimers display significant conformational variation in the flaps for the first ~500 ns followed by a distinct flap closure event where both flaps tuck into the active site. Approximately 100 ns later, the flaps of PRS17 partially open and continue to display flexibility in a semi-open conformation. In contrast, the flaps of PRS17$_{V48G}$ remain in a tucked conformation for the remainder of the trajectory. Wild-type PR quickly adopts a closed conformation and stays closed for the remainder of the simulation.

The degree of flap openness for the two subunits was quantified by calculating the distance between Cα atoms of the catalytic Asp25/25' and the flap tip at Ile50/50' as well as between the two flap tips (Figure 5B). These metrics have historically been used in MD simulations of HIV PR to describe flap conformation variability and separation over time[36, 65-67]. The separation between Asp25/25' and Ile50/50' at the flap tips fluctuates repeatedly for both mutant systems for the first ~500 ns, reaching as high as >25 Å for subunit A. PRS17 showed larger flap opening distances in subunit B than $PRS17_{V48G}$, while the opposite is shown for subunit A. The flaps of PRS17 also explored four short-lived closed flap states with distances <10 Å during the first 500 ns. Flap separation, as measured by the distance between the flap tips (Ile50/50'), was also greater in PRS17 than in $PRS17_{V48G}$ during this stage.

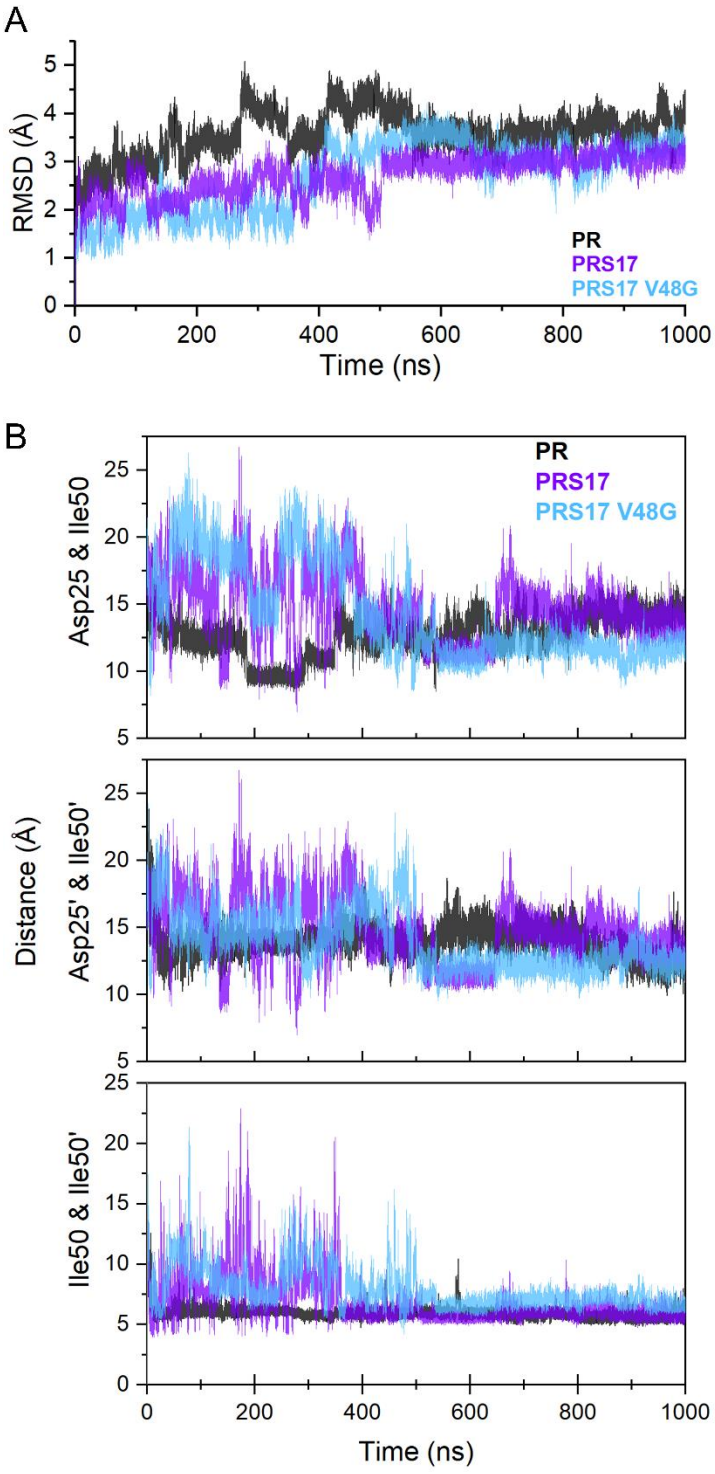

Figure 5: RMSD and flap conformation state analysis from 1 µS MD simulations of ligand-free PRS17, PRS17V48G, and WT PR. *A) RMSD of backbone atoms relative to equilibrated system for each trajectory. B) Intra-subunit distances between Cα atoms of catalytic Asp25/25' and the flap tips at Ile50/50' measure the magnitude of flap openings over time, and distances between Cα atoms of Ile50/50' measures the separation of the two flap tips.*

PRS17 exhibited larger open flap conformations and more frequent and rapid changes in flap behavior than the V48G revertant for the first ~500 ns of the simulation. The nearly simultaneous flap conformation change event from open to tucked closed observed in both mutant proteases is evident by the stable flap distances of ~11-12 Å beginning at ~545 ns. Approximately 100 ns later, the flaps of PRS17 spontaneously reopened and adopted a semi-open conformation with ~15 Å from Asp25/25' to Ile50/50' for the remainder of the trajectory whereas the revertant continued in the closed tucked conformation.

Root mean square fluctuation (RMSF) analysis was used to further describe the flexibility of the flaps for the two distinct time domains observed in the trajectory. To understand how the flap dynamics differ during the two distinct conformational states observed, the 1 μs trajectories were evenly split into two exclusive time domains for RMSF analysis. Time domain 1 represents the shifting wide-open flap states witnessed for both mutants, whereas time domain 2 samples the closed-to-semi-open event for PRS17 and the flap tucked state for $PRS17_{V48G}$. RMSF was calculated by the RMSD from the average position over time for backbone atoms for each subunit during both time domains and plotted by residue (Figure 6A). The flaps from residues 45-58 show the largest RMSF values with Ile50 and Gly51 between ~3-4.5 Å in both subunits of both dimers for time domain 1. PRS17 exhibits overall higher RMSF values than $PRS17_{V48G}$ for both subunits suggesting greater overall flap flexibility. In time domain 2, the flaps of both dimers show much lower RMSF value than in time domain 1. Furthermore, residues for both mutant dimers show similar RMSF profiles in time domain 2 except for the flap tips where PRS17 has RMSF values >2 Å, while

PRS17$_{V48G}$ shows only a modest peak. To visualize the relative flexibility exhibited during the simulation, RMSF was converted to B-factor and mapped to non-physical average coordinate structures (Figure 6B). In time domain 1 where both dimers explore open flap conformations, PRS17 shows greater flap fluctuations throughout the entire length of the flaps compared to the revertant.

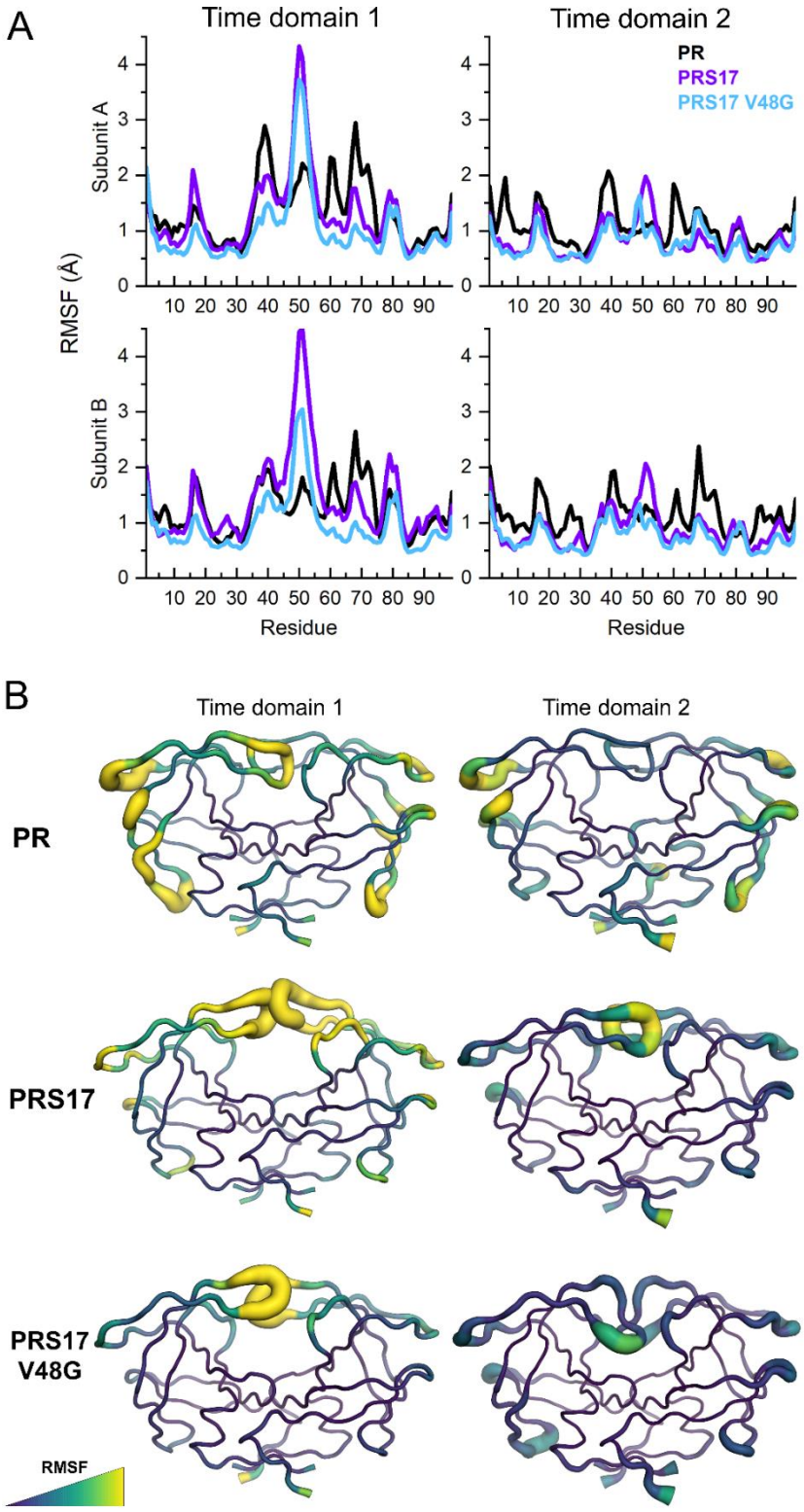

Figure 6: Root mean square fluctuation (RMSF) analysis from a 1 µS molecular dynamics simulation of ligand-free PRS17, PRS17$_{V48G}$, and WT PR. *A) Average RMSF of backbone atoms for subunits A and B plotted by residue for each time domain sampling different flap conformational events exhibited during the simulation. B) RMSF converted to B-*

factors and mapped to non-physical average structures for PRS17 and PRS17$_{V48G}$ for time domains 1 and 2. Structures are shown in the B-factor putty representation colored by backbone RMSF. All four structures are display with the same color scale. Relatively larger RMSF values for backbone atoms are shown with an increasing cartoon radius and viridis color gradient.

In time domain 2, the V48G revertant clearly adopts a tucked conformation for both flaps with relatively little fluctuation. Here, PRS17 exhibits smaller fluctuations than in time domain 1, but the fluctuations are still greater than for PRS17$_{V48G}$. In contrast, wild-type PR does not explore any extreme fluctuations.

MD simulations comparing PRS17, PRS17$_{V48G}$, and wild-type PR suggest the Val48 mutation contributes significantly to the extreme flap dynamics recorded for PRS17 and helps to initiate spontaneous flap opening. Nonetheless, the revertant still displays the extensive flap dynamics that are hallmarks of many drug resistant mutants[35, 36, 65]. Without this G48V mutation, the flaps of the revertant shifted into the inhibitor-binding site and formed a stable closed conformation, a characteristic described previously in MD simulations of WT PR[35, 36]. The long time scale of this simulation allowed observation of the transition of PRS17 from a closed to semi-open flap state that was not observed for the revertant. Therefore, replacing glycine at position 48 with valine appears to play a role in destabilizing the flaps while in the closed conformation and to drive the system towards flap opening.

2.4 Discussion

Mutations in the flaps of HIV-1 protease, such as G48V, play an important role in drug resistance. The G48V mutation of PRS17 was previously suggested to induce a

distinctive curled conformation in the flap tip that propagates through residues 48-52[29]. We have investigated this hypothesis by analyzing the revertant PRS17$_{V48G}$ that replaces Val48 by the glycine of wild type enzyme. Compared to PRS17, this revertant shows 2- to 8-fold increased susceptibility to inhibition by APV and DRV, respectively. The structures of APV complexes with PRS17 and revertant retain the hydrogen bond interactions with inhibitor observed for wild-type PR/APV. Differences are observed in the conformation of Phe53 and hydrophobic interactions of residue 48 and Pro81 with inhibitor. Larger changes are seen for the open conformation structure of inhibitor-free PRS17$_{V48G}$. The curled conformation observed in flap tips of the inhibitor-free PRS17 is abolished, and the flaps of the revertant assume an open conformation more similar to that of the wild type PR. MD simulations illustrate significant differences in flap conformations and dynamics for inhibitor-free dimers of the two mutants and wild-type PR. One microsecond MD simulations suggest the Val48 mutation plays a significant role in the extreme flap dynamics. Notably, the long timescale of the simulation enabled sampling of a spontaneous flap opening event for PRS17 but not for the revertant. The authors found no examples in the literature of comparable time scales for ligand-free forms of such extremely drug resistant mutants. These structural and dynamic changes may account for the improved inhibition by APV and DRV observed for PRS17$_{V48G}$ compared to PRS17.

Molecular dynamics simulations, enzyme kinetics data, and crystallographic studies are consistent with the hypothesis that the G48V mutation in HIV-1 PRS17 has a role in altering the conformational dynamics of the flaps, thereby making it harder for inhibitors to bind and stay in the active site of PR. This insight has implications for the

design of improved inhibitors of HIV-1 PR. Novel inhibitors that incorporate P2 and P2' groups forming better interactions with the protein backbone around residue 48 [25] could result in improved inhibition.

Author Contributions

SHB expressed and purified the protein, performed the enzyme kinetics and crystallographic analysis; DWK and AK ran the molecular dynamics simulations; SHB, YW, DWK and ITW interpreted the structural data; SHB and ITW selected the mutant and planned the experiments; SHB, DWK and ITW wrote the manuscript; all authors edited the manuscript.

Acknowledgements

We are grateful to Johnson Agniswamy for valuable discussions. Clinical inhibitors were obtained from the AIDS Reagent Program, Division of AIDS, NIAID, NIH. We thank the SER-CAT staff at the Advanced Photon Source, Argonne National Laboratory, for assistance during X-ray data collection. 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 No. W-31-109-Eng-38.

Funding sources

The research was supported in part by the National Institute of Health grant AI150461 (I.T.W., R.W.H.) and the Georgia State University Molecular Basis of Disease Fellowship (S.H.B.).

Conflicts of interests

The authors declare no conflicts of interest.

3 NOVEL PROTEASE INHIBITORS GRL-0489 AND GRL-0739 INHIBIT DRUG RESISTANT MUTANT PR20 AS WELL AS CLINICAL INHIBITOR DARUNAVIR

(To be published. Manuscript currently under preparation)

Shelley H. Burnaman[1], Daniel W. Kneller[2], Johnson Agniswamy[1], Yuan-Fang Wang[1], Arun K. Ghosh[3], Irene T. Weber[1]

[1]Department of Biology, Georgia State University, Atlanta GA, USA.

[2]Neutron Scattering Division, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, TN 37831, USA; National Virtual Biotechnology Laboratory, US Department of Energy, Washington, DC, USA.

[3]Departments of Chemistry and Medicinal Chemistry, Purdue University, West Lafayette, Indiana, USA

Corresponding author: Irene T. Weber, Department of Biology, P.O. Box 4010, Georgia State University, Atlanta GA, USA, 404-413-5411, iweber@gsu.edu

Highlights

2 and **3** inhibit drug resistant protease PR20 as well as **1**

PR20/**2** and PR20/**3** are not significantly different than PR20/**1**

2 forms a water-mediated hydrogen bond with Gly48

3 P2 ligand fills the binding pocket more than PR20/**1**

ABSTRACT

HIV-1 infection remains a serious problem worldwide in part due to growing resistance to HIV medications, such as protease inhibitors. HIV-1 protease cleaves Gag and Gag-Pol polyproteins to produce mature infectious virions. Inhibitors of protease prevent viral maturation, but resistance to protease inhibitors is increasingly common. Novel protease inhibitors **2** and **3** were designed based on clinical inhibitor **1** to form more interactions with the active site of protease. **3** has increased central nervous system penetrance, a desirable trait for prevention of HIV-associated neurocognitive disorder and HIV-related dementia. **2** and **3** had no major differences in binding to the HIV-1 protease active site, and they maintained kinetic inhibition similar to darunavir with highly drug resistance PR20. This research offers new data to be used in designing novel inhibitors of HIV-1 protease that have increased penetrance while maintaining effectiveness against highly drug-resistant protease mutants.

Keywords: Drug resistance, HIV protease, novel protease inhibitor, structure-based drug design

3.1 Introduction

Since the beginning of the HIV/AIDS pandemic in the 1980s, 40 million people have become infected with HIV. The Joint United Nations Programme on HIV/AIDS (UNAIDS) has set the “95-95-95 by 2025” goals with the target of having 95% of people with HIV knowing their status, 95% of people that know their status to be on antiretroviral therapy (ART), and 95% of people on ART to have achieved viral suppression[3] .As of 2020, 84% of people living with HIV know their status, 73% of people with HIV are on ART, and 90% of people on ART are virally suppressed[3].

Achieving the 95-95-95 goals requires improving the drugs available for use in ART. Resistance to HIV medications is a growing challenge in the HIV pandemic. HIV drugs work by targeting different parts of the virus life cycle. Protease inhibitors (PIs) target the last step of the life cycle where HIV protease cleaves Gag and Gag-Pol polyproteins to produce a mature infectious virus particle[10]. PIs were first introduced in 1995 (saquinavir; SQV). The latest PI to gain approval was darunavir (DRV; **1**) in 2006. **1** is one of three clinical PIs that are currently in use [11].

Current PIs are effective at treating most HIV infections; however, HIV hides in viral reservoirs in microglial cells, astrocyte, and macrophages in the central nervous system (CNS) where ART drugs like PIs do not easily enter[27]. HIV multiplies and mutates easily in places where ART cannot reach to inhibit the virus life cycle. Viral reservoirs in the CNS lead to HIV-associated neurocognitive disorder (HAND), a syndrome that includes progressive cognitive decline and loss of motor control[26, 27]. This research examines how two novel PR inhibitors, GRL-0489A (**2**) and GRL-0739A

(**3**), interact with highly drug resistant PR mutant PR20. Both **2** and **3** were designed based on **1** to have better active site binding and CNS penetration.

3.1.1 HIV-1 protease mutant PR20

HIV-1 protease (PR) is an aspartic acid protease that self-cleaves from Gag to cleave Gag and Gag-Pol polyproteins in the maturation step of the HIV life cycle[10, 12, 68, 69]. It is a homodimer of 99 amino acid subunits that contain a catalytic aspartic acid (Asp25) in each subunit. PR has flaps above the active site that are in a closed conformation with inhibitor bound and are flexible without inhibitor bound[68]. Mutations that confer drug resistance are typically concentrated near the active site, but many distal mutations also contribute to resistance by altering dimer stability or protein folding [16, 29, 68]. Active site and distal mutations work together to balance the decrease in substrate affinity with prolonging the life of the enzyme or increasing stability to allow PR to stay active and cleave more substrate. PR20 is a clinical mutant selected for its profound resistance to all clinical PIs [16, 17]. It has a much lower affinity for substrate than WT PR (PR20 K_m = 617 µM vs. WT PR K_m = 48 µM), and greater than 3 orders of magnitude poorer affinity for Compound **1** versus WT PR (WT PR K_i = 0.01 nM; PR20 K_i = 40 nM).

3.1.2 Novel protease inhibitors 2 and 3

Protease inhibitors based on the structure of darunavir are designed to combat drug resistance by targeting the active site backbone and filling the expanded binding pocket in highly mutated proteases. PR20 has a larger binding pocket relative to WT

PR, weakening chemical interactions between the inhibitor and protein, decreasing the binding affinity of the inhibitor [16, 17]. The two novel inhibitors in this research were designed to fill the expanded binding pocket of resistant mutants and to interact with the backbone of mutated PR at Gly48. **2** has a *bis*-THF-derived tricyclic ring P2 group [23] and **3** has a cyclopentyl-THF group with an additional hydroxyl attached to the THF ring in place of the *bis*-THF group of **1** [25]. Both contain a methoxy (OMe) P2' group in place of **1**'s NH_2 P2'.

2 is of particular interest due to its impressive antiviral potency (EC_{50} = 0.007 to 0.033 μM for multidrug resistant HIV-1 strains), inhibition of a variety of multidrug-resistant proteases, reduced susceptibility to drug resistance, and ability to penetrate the central nervous system (CNS)[24]. PIs that can reach viral reservoirs in the CNS are useful for treating and preventing HAND and HIV-1 associated dementia[26].

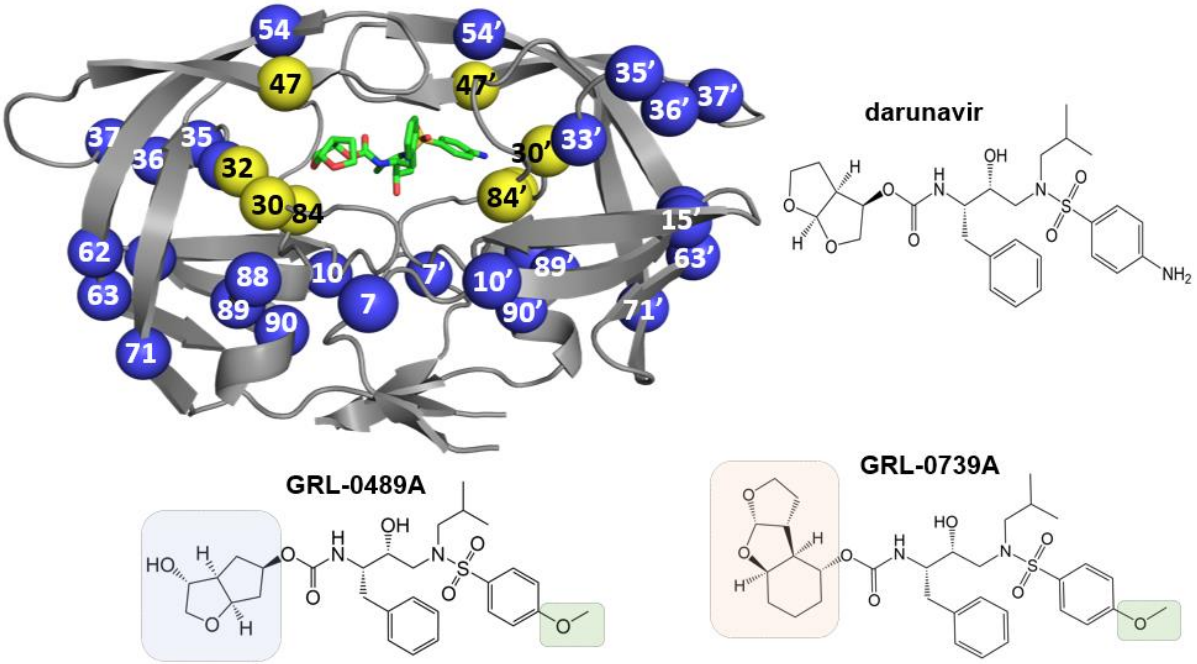

Figure 8: HIV-1 protease multidrug resistant mutant PR20. *Yellow spheres are the locations of the 20 mutations. Compounds **2** (GRL-0739A) and **3** (GRL-0489A) designed based on compound **1** (darunavir; DRV) have differences in the P2 region, highlighted in orange, blue, and green. These inhibitors were designed to make additional interactions with PR in the active site.*

3.2 Methods

3.2.1 Protein Purification

HIV-1 PR20 was purified as previously described [16], with some modification. PR20 in a PET plasmid was transformed into BL21 *E. coli* and grown overnight in Luria-Bertani (LB) media until turbid, then 1 L LB was inoculated with 10 mL overnight culture. Expression of PR20 protein was induced at OD_{600} ~0.4-0.6. The cell pellet was collected via ultracentrifugation and cells were homogenized and sonicated in TE buffer containing lysozyme. Cell homogenate was incubated overnight at 4℃ in TE buffer containing DNase I and Triton X-100. Protein inclusion body was isolated via ultracentrifugation and washing with TE buffers containing urea. The inclusion body was dissolved in 8 M guanidine HCl and purified using size exclusion chromatography (FPLC) and reverse phase chromatography (HPLC). PR20 was dialyzed to refold following the protocol outlined in [70].

3.2.2 Protein X-ray crystallography and structure refinement

PR20 complexed at 1:5 M ratio with **2** was crystallized in 1.3 M NaCl, 0.1 M sodium acetate pH 5.5, and 0.1 M yttrium chloride. PR20 complexed at 1:5 M ratio with **3** was crystallized in 1.1 M NaCl, 0.1 M MES pH 5.5, and 0.1 M yttrium chloride. 1:5 M ratio were crystallized via hanging drop vapor diffusion. Data were collected remotely on Southeastern Regional Collaborative Access Team (SER-CAT) beamline ID-22 of the Advanced Photon Source (APS) at Argonne National Laboratory. Raw data were indexed, integrated, and scaled using HKL-2000[16]. Molecular replacement was competed using PDB structure 3UCB [16] and PHASER MR[46] in CCP4[50]. The initial

structure and electron density map were calculated and further refined using SHELXL [51, 71] and Refmac5 [49].

3.2.3 Kinetic inhibition of PR20 with 2 and 3

Kinetic inhibition of PR20 with **2** and **3** were performed as previously described [22] at 37 °C and pH 5.6 using a FRET-substrate (BACHEM H-2992). Compounds **2** and **3**, synthesized in the lab of Arun Ghosh at Purdue University (>95% purity, HPLC), were dissolved in 100% DMSO.

3.3 Results

3.3.1 Overall Structures

The overall structures of PR20/**2** and PR20/**3** Hydrogen bond interactions between **2** and PR20 were highly conserved compared to between PR20/**1**. PR20/**2** was refined to 1.22 Å in the space group $P6_1$. No structure of PR20/**1** exists in the same space group for comparison, so structure 3UCB in space group $P2_12_12_1$ was used for molecular replacement and superimposition to obtain RMS deviations. PR20/**2** was superimposed onto PR20/**1** with an RMSD of 0.769, and a maximum displacement of 2.438 Å at Arg41. Two hydrogen bonds were lost on the P2 tricyclic group. There were no other major differences in hydrogen bonding between the inhibitor and the ligand binding pocket.

Pr20/**3** was refined to 1.15 Å in the space group $P6_1$. The overall structure compared to PR20/DRV is similar, with an average RMSD of 0.821 Å and a maximum RMSD of 2.487 Å at Arg41. Like PR20 /**2**, any differences on the surface or at crystal

contacts is likely an artifact making comparisons across two different space groups. Hydrogen bonding in the active site was conserved, except for new water-mediated bond between the hydroxyl of the Cp-THF ring and Val47. Like PR20/**2**, PR20/**3** loses one hydrogen bond with the P2' amine and Asn30 carbonyl.

Table 3. Crystallographic and refinement statistics for PR20/2 and PR20/3.

Statistics	PR20/2	PR20/3
Space group	$P6_1$	$P6_1$
Cell dimensions		
a (Å)	60.37	60.45
b (Å)	60.37	60.45
c (Å)	85.58	85.53
γ (°)	120	120
Resolution range (Å)	50-1.22	50-1.15
Unique reflections	52373	62592
Redundancy	5.9 (4.2)	6.4 (2.5)
Completeness	99.2 (94.6)	98.8 (95.0)
$\langle I/\sigma(I)\rangle$	50.4 (3.3)	13.2 (2.0)
R_{sym} (%)	4.8(39.5)	9.1 (52.1)
Refinement resolution range (Å)	33.14-1.22	33.14-1.15
R_{cryst} (%)	15.6	18.0
R_{free} (%)	20.0	21.1
Number of solvent molecules	154	163
Average B-factor ($Å^2$)	27.3	24.6
RMS deviations from ideality		
Bond length (Å)	0.02	0.02
Angles (°)	2.1	2.7

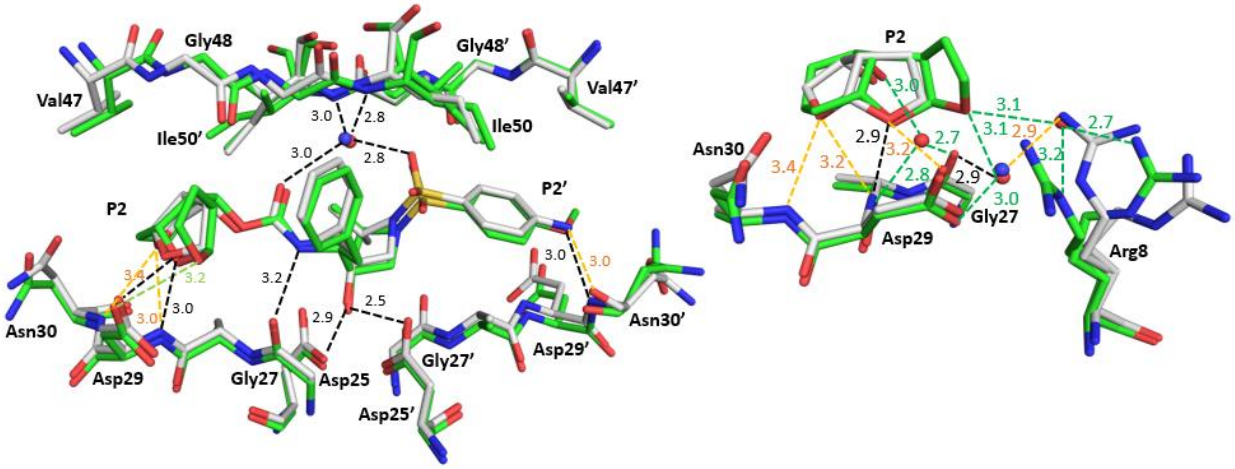

Figure 9: Ligand-binding site of PR20/2. *Overall hydrogen bonds do not differ significantly from PR20/1 (3UCB), but the cyclohexyl-bis-THF ring is a larger group and may fill the expanded binding site better than 1*

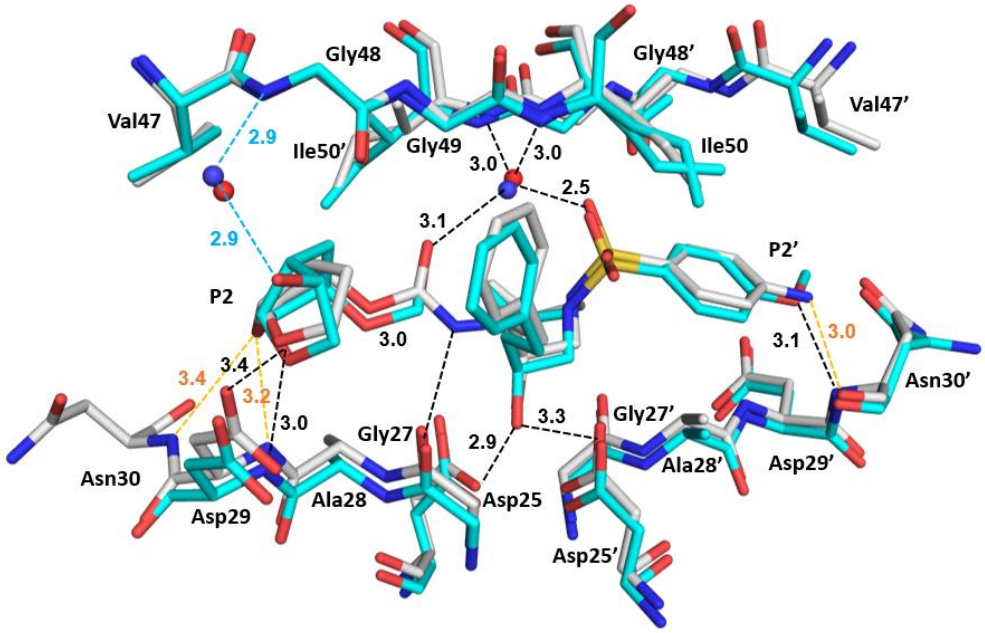

Figure 10: The active site of PR20/3 superimposed with PR20/DRV (PDB 3UCB). *Overall hydrogen bonding is the same as 3UCB. There is one new water-mediated hydrogen bond between 3 and Gly48. Two hydrogen bonds between 1 bis-THF are lost in PR20/3.*

3.3.2 Kinetic inhibition of PR20 with 2 and 3

Kinetic inhibition assays showed that both novel inhibitors in this study are very similar to the inhibitor that their design was based upon. Compared to **1** (K_i PR20 = 41 nM, K_i WT = 0.005 nM), **2** has a K_i of 34 nM for PR20 and 0.1 nM for WT PR [23]. **3** has a K_i of 50 nM for PR20 and a K_i of 0.005 nM for WT PR [25]. Considering that there are few differences in how these two novel inhibitors bind to PR20 compared to **1**, these kinetic inhibition data were as expected.

3.4 Discussion

Overall, the structures and kinetics for P20/**2** and PR20/**3** are not significantly different. For **2**, it is noteworthy that changing one of the THF rings to a cyclopentyl moiety and changing the amine on the P2' side to a slightly larger methoxy group do not have a large impact on the inhibitor's ability to bind to a highly resistant PR mutant.

3 has better blood brain barrier (BBB) penetration than any of the currently approved clinical inhibitors[24]. HIV can enter the CNS through the BBB, creating a reservoir where the virus can continue to replicate outside the pressure of HIV-1 medications. **3** is very much like **1** in terms of enzyme kinetics and active site binding to both WT PR[24] and PR20, but it has the major advantage of being able to penetrate the BBB where it could inhibit HIV-1 replication in the brain.

This research has implications for the design of novel PR inhibitors. Reaching very difficult parts of the body where HIV can easily hide from drug pressure is a highly desirable feature for any new PR inhibitor. It is also useful to know that making changes to the PR inhibitor do not necessarily yield differences in kinetic inhibition or active site

binding. **3** has been studied in complex with a PR inhibitor resistant triple mutant of HIV-1 PR (PR_{TRI})[72], and it was found to have similar inhibition as PR20/**3** and PR20/**1**. Further studies with additional novel inhibitors will add to the repertoire of knowledge on how to design protease inhibitors to combat multidrug resistant HIV-1 PR.

Acknowledgements

Clinical inhibitors were obtained from the AIDS Reagent Program, Division of AIDS, NIAID, NIH. We thank the SER-CAT staff at the Advanced Photon Source, Argonne National Laboratory, for assistance during X-ray data collection. 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 No. W-31-109-Eng-38.

Funding sources

The research was supported in part by the National Institute of Health grant AI150461 (I.T.W., R.W.H.) and the Georgia State University Molecular Basis of Disease Fellowship (S.H.B.).

4 CONCLUSIONS

4.1 PRS17V48G is less resistant to PR inhibitors than PRS17

Enzyme kinetics assays, X-ray crystallography, and molecular dynamics studies indicate that PRS17V48G is less resistant to PR inhibitors than PRS17 because of the change in the conformation of the flap tip. PRS17 has a unique curled flap that propagates from V48 through G52 and changing the valine back to WT glycine reverses the flap curl. This action changes the flap dynamics, allowing the revertant mutant protease to retain the inhibitor in the active site for longer than PRS17. This research supports the hypothesis that the G48V mutation has a role in changing the conformational dynamics of the flaps, making it harder for the inhibitor to bind and stay in the active site.

For future directions, PRS17V48G could be crystallized with CA-p2 and p2-NC substrate analogs for comparison with previously published PRS17 data[22]. PRS17V48G could also be analyzed with urea denaturation to study the dimer stability. V48 in PRS17 is known to destabilize the dimer when bound to CA-p2 substrate analog[22], so it would be rational to study the comparative dimer stability of PRS17 versus PRS17V48G.

The data from these experiments have implications for drug design in highly resistant mutants. Some novel inhibitors that inhibit highly drug resistant mutants like PR20 are designed to form hydrogen bonds or enhanced VDW interactions with G48. This design may not be suitable for mutants that contain G48V mutation, especially if V82S is also present. The combination of these two mutations expands the binding pocket and changes how the substrate analog interacts with the protease[22]. In

addition, the current research demonstrates that the drugs DRV and APV are less effective against PRS17$_{V48G}$. The crystal structure of PRS17$_{V48G}$/APV shows that the flaps are slightly more stable, Drugs that are designed with G48 in mind may need a larger P1 group to fill this expanded binding pocket if G48V or G48V/V82S are present. Future studies of PRS17$_{V48G}$ with novel PIs would be beneficial to further investigate

4.2 PRS17 revertant mutations additional information for future directions

Two revertant mutations of PRS17 were initially considered for study: S82V and V48G. These two mutations work in concert to increase drug resistance in PRS17. S82V and V48G were desirable targets for study because of their location near the active site in the 80's loop and flap, respectively. S82 in PRS17 disrupts the P3 Arg CA-p2 substrate analog residue interaction with Asp29, causing it to instead swing in a different direction a form a water-mediated interaction with Arg8. The Arg8-Asp29' interaction is critical for PR dimer stability, and this interaction is interrupted by the S82 and V48 mutations. V48 breaks the P3 Arg VDW interactions with Phe53 via steric hindrance, and the smaller S82 residue gives P3 Arg room to move to a conformation where it can interact with PR Arg8.

S82V was first mutation to be considered for this research. This proved exceptionally more difficult than anticipated. For revertant mutant S82V, site-directed mutagenesis using an Agilent QuikChange Lightning kit failed. When this method failed, a different site-directed mutagenesis method using ThermoFisher SuperFi polymerase and primers designed with 3' DNA overhangs to facilitate insertion in the correct orientation. After both site-directed mutagenesis methods failed, the gene was ordered from Atum (Newark, CA) in a pre-constructed pJ414 vector. Research scientists at Atum

also had difficulty generating this template, but they eventually had success in stabilizing the gene in a low copy number plasmid, pJ434. This success was not replicated in our lab, as the plasmid we obtained was in too low concentration (less than 15 ng/mg and poor purity as measured by Nano Drop) to be able to be transformed and expressed in our lab. Thus, S82V was abandoned, and I moved on to work with V48G. This construct was also ordered from Atum in a pJ414 vector, and it was easily transformed into BL21(DE3) *E. coli*, overexpressed, and purified according to the methods outlined in section 2 of this dissertation. Though the S82V revertant mutation was never obtained alone from these methods, the ThermoFisher SuperFi polymerase site-directed mutagenesis method produced a double mutation with L46P. This double mutant was not used for this dissertation research, and it is preserved at -80 °C in BL21(DE3) *E. coli*.

4.3 GRL-0739A and GRL-0489A inhibit PR20 as well as darunavir

Darunavir was designed to combat resistance to other clinical protease inhibitors. Darunavir is considered to be third line treatment, only given after treatment with LPV or ATV has failed[9]. Two novel inhibitors synthesized in the lab of Dr. Arun Ghosh (Purdue University) based on the darunavir scaffold were found to be kinetically similar to DRV.

Compared to DRV, GRL-0489A has a cyclopentyl ring in place of one THF ring (Cp-THF) and a methoxy group in place of the P2’ amine. We have learned from enzyme kinetics and protein X-ray crystallography that these two groups can be changed without having a significant effect on kinetic inhibition or hydrogen bonding between inhibitor and protease in the active site.

GRL-0739A has the advantage of being able to enter the CNS by crossing the blood brain barrier (BBB). DRV and other clinical PR inhibitors cannot easily cross the BBB, so viral replication in reservoirs that are known to be present in the CNS is not inhibited by these drugs. GRL-0739A is not significantly different than DRV in terms of kinetic inhibition or binding to the active site of PR, but its ability to cross the BBB is highly desirable for inhibiting HIV-1 replication in viral reservoirs of the CNS.

4.4 General Concluding Remarks

HIV-1 protease is an important target for antiretroviral therapy (ART). After the introduction of the first protease inhibitors in 1995 and the subsequent use of combination ART[9], the number of deaths due to AIDS and the rate of progression to AIDS decreased dramatically[1, 9, 10, 69]. HIV, once considered a death sentence that carried a great stigma, is now considered to be a manageable chronic condition. This stability is threatened by increasing resistance to antiretroviral drugs. HIV-1 is prone to genetic mutations due to the virus' very low fidelity reverse transcriptase. Mistakes in RNA reverse transcription allow the virus to sample many different configurations of protease mutations, as well as other viral proteins, and the combination of mutations that is best suited for producing new mature infectious virus particles in the presence antiretroviral therapy becomes dominant enough to be detected when sequencing the virus. Various drug resistant mutants that have been discovered through viral sequencing are curated in the publicly available Stanford HIVdb (www.hivdb.stanford.edu). Through this database and clinical samples, two mutants of HIV-1 protease were found and used for this research. PR20 was found as a clinical isolate [73] and PRS17 was found using a machine learning algorithm on clinical isolate

genotype-phenotype data from the Stanford HIVdb[29, 30]. Both mutants are highly resistant to clinical protease inhibitors and studying these in complex with substrate analogs and clinical or novel protease inhibitors gives us new insight into how to design new drugs to combat resistance.

Through studies of PR20, PRS17, and a revertant mutant of PRS17, we have learned critical new information. We learned that exchanging one THF ring for a cyclopentyl ring, adding a hydroxyl group to the THF to expand the P2 side of the inhibitor, and exchanging the P2' amine for a larger methoxy group (GRL-0489A) does not result in a change in hydrogen bonding or kinetic inhibition with PR20. Adding a cyclohexane-based ring to the bis-THF of the inhibitor and exchanging the P2' amine for a methoxy group (GRL-0739A) also does not have a significant impact of the number of hydrogen bonds or kinetic inhibition, but these changes to the inhibitor allow it to enter the CNS more easily to inhibit production of infectious virus particles in the brain. Inhibiting HIV-1 in the CNS could prevent or lessen the onset of HAND or HIV-related dementia[26].

Reverting PRS17 V48 back to the WT glycine allowed us to investigate the contribution of this mutation to drug resistance in PRS17. PRS17 contains a unique flap curl conformation that starts at residue 48 and continues through G52. This revertant mutant caused the flap tip to shift back to a more WT-like conformation with no curl in the flap tip. APV and DRV inhibit $PRS17_{V48G}$ better than PRS17, and $PRS17_{V48G}$ has higher affinity for p2-NC substrate analog than PRS17. Taken together, these findings suggest that $PRS17_{V48G}$ is more susceptible to some protease inhibitors, including one designed to work against highly resistant mutants of HIV-1 PR. This mutation

contributes greatly to drug resistance for PRS17. Further studies of substrate analogs, novel and clinical inhibitors, and studies of a double mutant with V48G and S82V would help further elucidate the mechanism of resistance in PRS17.

REFERENCES

1. *HIV/AIDS Fact Sheet*, H.A. Department, Editor. 2020, World Health Organiation: https://www.who.int/news-room/fact-sheets/detail/hiv-aids.
2. UNAIDS, *90-90-90: An ambitious treatment target to help end the AIDS epidemic*. 2014, Joint United Nations Programme of HIV-AIDS (UNAIDS): Geneva, Switzerland.
3. *Political Declaration on HIV and AIDS: Ending Inequalities and Getting on Track to End AIDS by 2030*. 2021, United Nations General Assembly: New York, United States.
4. *HIV Surveillance Supplemental Report*. 2020, Centers for Disease Control and Prevention: Atlanta, Georgia, USA.
5. *HIV Surveillance Report, 2019*. 2021, Centers for Disease Control and Prevention.
6. *Ending the HIV Epidemic: About Ending the HIV Epidemic in the U.S.: Overview*. 2021 June 2, 2021; Available from: hiv.gov/federal-response/ending-the-hiv-epidemic/overview.
7. CDC, *Revised surveillance case definition for HIV infection*. 2014: United States. p. 1-10.
8. *HIV Replication Cycle*. 2018, National Institute of Allergy and Infectious Diseases https://www.niaid.nih.gov/diseases-conditions/hiv-replication-cycle.
9. *Consolidated guidlines on the use of antiretroviral drugs for treating and preventing HIV infection: Recommendations for a public health approach*. 2nd edition ed. 2016, Geneva, Switzerland: World Health Organization.
10. Weber, I. and J. Agniswamy, *HIV-1 Protease: Structural Perspectives on Drug Resistance.* Viruses, 2009. **1**(3): p. 1110-1136.
11. *Update of Recommendations on First- and Second-line Antiretroviral Regimens*. 2019, World Health Organization: Geneva, Switzerland.
12. Weber, I.T., M. Miller, M. Jaskolski, J. Leis, A.M. Skalka, and A. Wlodawer, *Molecular modeling of the HIV-1 protease and its substrate binding site.* Science, 1989. **243**(4893): p. 928-31.
13. Wlodawer, A., M. Miller, M. Jaskolski, B.K. Sathyanarayana, E. Baldwin, I.T. Weber, L.M. Selk, L. Clawson, J. Schneider, and S.B. Kent, *Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease.* Science, 1989. **245**(4918): p. 616-21.
14. Wensing, A.M., V. Calvez, F. Ceccherini-Silberstein, C. Charpentier, H.F. Gunthard, R. Paredes, R.W. Shafer, and D.D. Richman, *2019 update of the drug resistance mutations in HIV-1.* Top Antivir Med, 2019. **27**(3): p. 111-121.
15. Dierynck, I., M. De Wit, E. Gustin, I. Keuleers, J. Vandersmissen, S. Hallenberger, and K. Hertogs, *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.
16. Agniswamy, J., C.-H. Shen, A. Aniana, J.M. Sayer, J.M. Louis, and I.T. Weber, *HIV-1 Protease with 20 Mutations Exhibits Extreme Resistance to Clinical Inhibitors through Coordinated Structural Rearrangements.* Biochemistry, 2012. **51**(13): p. 2819-2828.

17. Agniswamy, J., C.-H. Shen, Y.-F. Wang, A.K. Ghosh, K.V. Rao, C.-X. Xu, J.M. Sayer, J.M. Louis, and I.T. Weber, *Extreme Multidrug Resistant HIV-1 Protease with 20 Mutations Is Resistant to Novel Protease Inhibitors with P1'-Pyrrolidinone or P2-Tris-tetrahydrofuran.* Journal of Medicinal Chemistry, 2013. **56**(10): p. 4017-4027.
18. Agniswamy, J., J.M. Louis, C.-H. Shen, S. Yashchuk, A.K. Ghosh, and I.T. Weber, *Substituted Bis-THF Protease Inhibitors with Improved Potency against Highly Resistant Mature HIV-1 Protease PR20.* Journal of Medicinal Chemistry, 2015. **58**(12): p. 5088-5095.
19. Rhee, S.Y., J. Taylor, W.J. Fessel, D. Kaufman, W. Towner, P. Troia, P. Ruane, J. Hellinger, V. Shirvani, A. Zolopa, and R.W. Shafer, *HIV-1 protease mutations and protease inhibitor cross-resistance.* Antimicrob Agents Chemother, 2010. **54**(10): p. 4253-61.
20. Kantor, R., W.J. Fessel, A.R. Zolopa, D. Israelski, N. Shulman, J.G. Montoya, M. Harbour, J.M. Schapiro, and R.W. Shafer, *Evolution of primary protease inhibitor resistance mutations during protease inhibitor salvage therapy.* Antimicrob Agents Chemother, 2002. **46**(4): p. 1086-92.
21. Clemente, J.C., R.M. Coman, M.M. Thiaville, L.K. Janka, J.A. Jeung, S. Nukoolkarn, L. Govindasamy, M. Agbandje-McKenna, R. McKenna, W. Leelamanit, M.M. Goodenow, and B.M. Dunn, *Analysis of HIV-1 CRF_01 A/E protease inhibitor resistance: structural determinants for maintaining sensitivity and developing resistance to atazanavir.* Biochemistry, 2006. **45**(17): p. 5468-77.
22. Agniswamy, J., D.W. Kneller, R. Brothers, Y.-F. Wang, R.W. Harrison, and I.T. Weber, *Highly Drug-Resistant HIV-1 Protease Mutant PRS17 Shows Enhanced Binding to Substrate Analogues.* ACS Omega, 2019. **4**(5): p. 8707-8719.
23. Ghosh, A.K., G.L. Parham, C.D. Martyr, P.R. Nyalapatla, H.L. Osswald, J. Agniswamy, Y.-F. Wang, M. Amano, I.T. Weber, and H. Mitsuya, *Highly Potent HIV-1 Protease Inhibitors with Novel Tricyclic P2 Ligands: Design, Synthesis, and Protein–Ligand X-ray Studies.* Journal of Medicinal Chemistry, 2013. **56**(17): p. 6792-6802.
24. Amano, M., Y. Tojo, P.M. Salcedo-Gómez, G.L. Parham, P.R. Nyalapatla, D. Das, A.K. Ghosh, and H. Mitsuya, *A Novel Tricyclic Ligand-Containing Nonpeptidic HIV-1 Protease Inhibitor, GRL-0739, Effectively Inhibits the Replication of Multidrug-Resistant HIV-1 Variants and Has a Desirable Central Nervous System Penetration Property In Vitro.* Antimicrobial Agents and Chemotherapy, 2015. **59**(5): p. 2625-2635.
25. Ghosh, A.K., B.D. Chapsal, G.L. Parham, M. Steffey, J. Agniswamy, Y.-F. Wang, M. Amano, I.T. Weber, and H. Mitsuya, *Design of HIV-1 Protease Inhibitors with C3-Substituted Hexahydrocyclopentafuranyl Urethanes as P2-Ligands: Synthesis, Biological Evaluation, and Protein–Ligand X-ray Crystal Structure.* Journal of Medicinal Chemistry, 2011. **54**(16): p. 5890-5901.
26. Ghosh, A.K., A. Sarkar, and H. Mitsuya, *HIV-Associated Neurocognitive Disorder (HAND) and the Prospect of Brain-Penetrating Protease Inhibitors for Antiretroviral Treatment.* Med Res Arch, 2017. **5**(4).
27. Wallet, C., M. De Rovere, J. Van Assche, F. Daouad, S. De Wit, V. Gautier, P.W.G. Mallon, A. Marcello, C. Van Lint, O. Rohr, and C. Schwartz, *Microglial*

Cells: The Main HIV-1 Reservoir in the Brain. Front Cell Infect Microbiol, 2019. **9**: p. 362.
28. Churchill, M.J., S.L. Wesselingh, D. Cowley, C.A. Pardo, J.C. McArthur, B.J. Brew, and P.R. Gorry, *Extensive astrocyte infection is prominent in human immunodeficiency virus-associated dementia.* Ann Neurol, 2009. **66**(2): p. 253-8.
29. Agniswamy, J., J.M. Louis, J. Roche, R.W. Harrison, and I.T. Weber, *Structural Studies of a Rationally Selected Multi-Drug Resistant HIV-1 Protease Reveal Synergistic Effect of Distal Mutations on Flap Dynamics.* PLOS ONE, 2016. **11**(12): p. e0168616.
30. Park, J.H., J.M. Sayer, A. Aniana, X. Yu, I.T. Weber, R.W. Harrison, and J.M. Louis, *Binding of Clinical Inhibitors to a Model Precursor of a Rationally Selected Multidrug Resistant HIV-1 Protease Is Significantly Weaker Than That to the Released Mature Enzyme.* Biochemistry, 2016. **55**(16): p. 2390-2400.
31. *HIV drug resistance report 2019*. 2019, World Health Organization: Geneva, Switzerland.
32. Hogan, A.B., B.L. Jewell, E. Sherrard-Smith, J.F. Vesga, O.J. Watson, C. Whittaker, A. Hamlet, J.A. Smith, P. Winskill, R. Verity, M. Baguelin, J.A. Lees, L.K. Whittles, K.E.C. Ainslie, S. Bhatt, A. Boonyasiri, N.F. Brazeau, L. Cattarino, L.V. Cooper, H. Coupland, G. Cuomo-Dannenburg, A. Dighe, B.A. Djaafara, C.A. Donnelly, J.W. Eaton, S.L. van Elsland, R.G. FitzJohn, H. Fu, K.A.M. Gaythorpe, W. Green, D.J. Haw, S. Hayes, W. Hinsley, N. Imai, D.J. Laydon, T.D. Mangal, T.A. Mellan, S. Mishra, G. Nedjati-Gilani, K.V. Parag, H.A. Thompson, H.J.T. Unwin, M.A.C. Vollmer, C.E. Walters, H. Wang, Y. Wang, X. Xi, N.M. Ferguson, L.C. Okell, T.S. Churcher, N. Arinaminpathy, A.C. Ghani, P.G.T. Walker, and T.B. Hallett, *Potential impact of the COVID-19 pandemic on HIV, tuberculosis, and malaria in low-income and middle-income countries: a modelling study.* Lancet Glob Health, 2020. **8**(9): p. e1132-e1141.
33. Konvalinka, J., H.G. Krausslich, and B. Muller, *Retroviral proteases and their roles in virion maturation.* Virology, 2015. **479-480**: p. 403-17.
34. Weber, I.T., Y.-F. Wang, and R.W. Harrison, *HIV Protease: Historical Perspective and Current Research.* Viruses, 2021. **13**(5): p. 839.
35. Shen, C.H., Y.C. Chang, J. Agniswamy, R.W. Harrison, and I.T. Weber, *Conformational variation of an extreme drug resistant mutant of HIV protease.* J Mol Graph Model, 2015. **62**: p. 87-96.
36. Kneller, D.W., J. Agniswamy, R.W. Harrison, and I.T. Weber, *Highly drug-resistant HIV-1 protease reveals decreased intra-subunit interactions due to clusters of mutations.* FEBS J, 2020. **287**(15): p. 3235-3254.
37. Shafer, R.W., *Rationale and uses of a public HIV drug-resistance database.* J Infect Dis, 2006. **194 Suppl 1**: p. S51-8.
38. Rhee, S.Y., M.J. Gonzales, R. Kantor, B.J. Betts, J. Ravela, and R.W. Shafer, *Human immunodeficiency virus reverse transcriptase and protease sequence database.* Nucleic Acids Res, 2003. **31**(1): p. 298-303.
39. Yu, X., I.T. Weber, and R.W. Harrison, *Identifying representative drug resistant mutants of HIV.* BMC Bioinformatics, 2015. **16 Suppl 17**: p. S1.

40. Yu, X., I.T. Weber, and R.W. Harrison, *Prediction of HIV drug resistance from genotype with encoded three-dimensional protein structure.* BMC Genomics, 2014. **15 Suppl 5**: p. S1.
41. Yu, X., I.T. Weber, and R.W. Harrison, *Sparse Representation for Prediction of HIV-1 Protease Drug Resistance.* Proc SIAM Int Conf Data Min, 2013. **2013**: p. 342-349.
42. 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.
43. Deshmukh, L., V. Tugarinov, J.M. Louis, and G.M. Clore, *Binding kinetics and substrate selectivity in HIV-1 protease-Gag interactions probed at atomic resolution by chemical exchange NMR.* Proc Natl Acad Sci U S A, 2017. **114**(46): p. E9855-E9862.
44. Winters, M.A., J.M. Schapiro, J. Lawrence, and T.C. Merigan, *Human immunodeficiency virus type 1 protease genotypes and in vitro protease inhibitor susceptibilities of isolates from individuals who were switched to other protease inhibitors after long-term saquinavir treatment.* J Virol, 1998. **72**(6): p. 5303-6.
45. Maschera, B., G. Darby, G. Palu, L.L. Wright, M. Tisdale, R. Myers, E.D. Blair, and E.S. Furfine, *Human immunodeficiency virus. Mutations in the viral protease that confer resistance to saquinavir increase the dissociation rate constant of the protease-saquinavir complex.* J Biol Chem, 1996. **271**(52): p. 33231-5.
46. McCoy, A.J., R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, and R.J. Read, *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.
47. Heaslet, H., R. Rosenfeld, M. Giffin, Y.C. Lin, K. Tam, B.E. Torbett, J.H. Elder, D.E. McRee, and C.D. Stout, *Conformational flexibility in the flap domains of ligand-free HIV protease.* Acta Crystallogr D Biol Crystallogr, 2007. **63**(Pt 8): p. 866-75.
48. Emsley, P., B. Lohkamp, W.G. Scott, and K. Cowtan, *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
49. Murshudov, G.N., P. Skubak, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D. Winn, F. Long, and A.A. Vagin, *REFMAC5 for the refinement of macromolecular crystal structures.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 355-67.
50. Winn, M.D., C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, and K.S. Wilson, *Overview of the CCP4 suite and current developments.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 235-42.
51. Sheldrick, G.M., *Crystal structure refinement with SHELXL.* Acta Crystallogr C Struct Chem, 2015. **71**(Pt 1): p. 3-8.
52. Abraham, M.J., Murtola, T., Schulz, R., Pall, S., Smith, J.C., Hess, B. Lindahl, E., *High performance molecular simulations through multi-level parallelism from laptops to supercomputers.* SoftwareX, 2015. **1-2**: p. 19-25.

53. Huang, J., S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B.L. de Groot, H. Grubmuller, and A.D. MacKerell, Jr., *CHARMM36m: an improved force field for folded and intrinsically disordered proteins.* Nat Methods, 2017. **14**(1): p. 71-73.
54. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L., *Comparison of simple potential funtions for simulating liquid water.* Journal of Chemical Physics, 1983. **79**: p. 926-935.
55. Bussi, G., D. Donadio, and M. Parrinello, *Canonical sampling through velocity rescaling.* J Chem Phys, 2007. **126**(1): p. 014101.
56. Berendsen, H.J.C., Postma, J.P.M, Van Gunsteren, W.F., Dinola, A., Haak, J.R., *Molecular dynamics with coupling to an external bath.* Journal of Chemical Physics, 1984. **81**: p. 3684-3690.
57. Hoover, W.G., *Canonical dynamics: Equilibrium phase-space distributions.* Phys Rev A Gen Phys, 1985. **31**(3): p. 1695-1697.
58. Nose, S., *A molecular dynamics method for simulations in the canonical ensemble.* Molecular Physics, 1984. **52**: p. 255-268.
59. Parrinello, W.G., Rahman, A., *Polymorphic transitions in single crystals: A new molecular dynamics method.* Journal of Applied Physics, 1981. **52**: p. 7182-7190.
60. Hess, B., *P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation.* J Chem Theory Comput, 2008. **4**(1): p. 116-22.
61. Hess, B., Bekker, H., Berendsen, H.J.C., Fraaije, J.G.E.M., *LINCS: A linear constraint solver for molecular simulations.* Journal of Computational Chemistry, 1997. **18**: p. 1463-1472.
62. Ohtaka, H., A. Schon, and E. Freire, *Multidrug resistance to HIV-1 protease inhibition requires cooperative coupling between distal mutations.* Biochemistry, 2003. **42**(46): p. 13659-66.
63. King, N.M., M. Prabu-Jeyabalan, R.M. Bandaranayake, M.N. Nalam, E.A. Nalivaika, A. Ozen, T. Haliloglu, N.K. Yilmaz, and C.A. Schiffer, *Extreme entropy-enthalpy compensation in a drug-resistant variant of HIV-1 protease.* ACS Chem Biol, 2012. **7**(9): p. 1536-46.
64. Liu, F., A.Y. Kovalevsky, Y. Tie, A.K. Ghosh, R.W. Harrison, and I.T. Weber, *Effect of Flap Mutations on Structure of HIV-1 Protease and Inhibition by Saquinavir and Darunavir.* Journal of Molecular Biology, 2008. **381**(1): p. 102-115.
65. Chetty, S., S. Bhakat, A.J. Martin, and M.E. Soliman, *Multi-drug resistance profile of PR20 HIV-1 protease is attributed to distorted conformational and drug binding landscape: molecular dynamics insights.* J Biomol Struct Dyn, 2016. **34**(1): p. 135-51.
66. Zhu, Z., D.I. Schuster, and M.E. Tuckerman, *Molecular dynamics study of the connection between flap closing and binding of fullerene-based inhibitors of the HIV-1 protease.* Biochemistry, 2003. **42**(5): p. 1326-33.
67. Hornak, V., A. Okur, R.C. Rizzo, and C. Simmerling, *HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations.* Proc Natl Acad Sci U S A, 2006. **103**(4): p. 915-20.
68. Louis, J.M., R. Ishima, D.A. Torchia, and I.T. Weber, *HIV-1 Protease: Structure, Dynamics, and Inhibition*, in *Advances in Pharmacology*. 2007, Elsevier. p. 261-298.

69. Louis, J.M., I.T. Webert, J. Tözsér, G. Marius Clore, and A.M. Gronenborn, *HIV-1 protease: Maturation, enzyme specificity, and drug resistance*, in *Advances in Pharmacology*. 2000, Elsevier. p. 111-146.
70. Ishima, R., D.A. Torchia, and J.M. Louis, *Mutational and Structural Studies Aimed at Characterizing the Monomer of HIV-1 Protease and Its Precursor.* Journal of Biological Chemistry, 2007. **282**(23): p. 17190-17199.
71. Sheldrick, G.M. and T.R. Schneider, *SHELXL: high-resolution refinement.* Methods Enzymol, 1997. **277**: p. 319-43.
72. Pawar, S., Y.F. Wang, A. Wong-Sam, J. Agniswamy, A.K. Ghosh, R.W. Harrison, and I.T. Weber, *Structural studies of antiviral inhibitor with HIV-1 protease bearing drug resistant substitutions of V32I, I47V and V82I.* Biochem Biophys Res Commun, 2019. **514**(3): p. 974-978.
73. Louis, J.M., A. Aniana, I.T. Weber, and J.M. Sayer, *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.