

https://dspace.org/

Inter-region non-active site mutations in HIV-1 subtype C protease: impact on protein dynamics, stability, and drug-binding

Sherry, Dean

2021
https://hdl.handle.net/10539/33179

Downloaded from a DSpace repository. https://dspace.org/



UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

Inter-region non-active site mutations in HIV-1 subtype C protease: impact on protein dynamics, stability, and drug-binding

Dean Sherry

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Molecular and Cell Biology in the Faculty of Science, University of the Witwatersrand, Johannesburg, South Africa

Supervisor: Professor Yasien Sayed

Protein Structure-Function Research Unit, School of Molecular and Cell Biology, University of the Witwatersrand, Johannesburg, South Africa

Declaration

I, Dean Sherry (722127), am a student registered for the degree of Doctor of Philosophy (PhD) in the academic year 2021.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where explicitly indicated otherwise and acknowledged.
- I have not submitted this work before for any other degree or examination at this, or any other University.
- The information used in the Thesis has not been obtained by me while employed by, or working under the aegis of, any person or organisation other than the University.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.

Signature: 

Date: 09/08/2021

Abstract

Human immunodeficiency virus (HIV) continues to present a significant challenge to the health of millions of individuals globally. The HIV protease enzyme is crucial for viral propagation and, as such, it is a major drug target in anti-retroviral therapy. However, the high replicative rate of the virus and the mutation-prone nature of reverse transcription leads to sequence polymorphisms which compound the difficulty of effective treatment. Indeed, the development of inhibitor resistance remains a significant stumbling block for successful anti-HIV therapy. Drug resistance begins with active site mutations which directly decrease drug binding through altering the geometry of the active site. Non-active site mutations have long been thought to play a compensatory role for the potentially deleterious effects of active site mutations. However, their role in drug resistance cannot be understated as multi-drug resistant variants must accumulate constellations of non-active site mutations in order to establish high levels of resistance. The role that these mutations play in resistance has not been fully explained, however, it has been hypothesised that their presence alters the hydrophobic sliding mechanism which is thought to be responsible for modulating the conformation of the flaps. In this research, a South African subtype C protease variant (named HP3 PR) was investigated. This variant contains the following polymorphisms which have been correlated with decreased drug susceptibility; namely, I13V, I62V, and V77I. A novel crystal structure of the HP3 PR was solved and, through structural, functional, and *in silico* analyses a mechanism of resistance was elucidated. The resistance mechanism appears to work in two parts: eliciting changes in the behaviour of the apo as well as the ligand-bound states of the protein. Thermodynamic and differential scanning calorimetric analyses revealed that the HP3 protease binds to certain clinical inhibitors with decreased binding affinity and that the inhibitor bound protease complex exhibits decreased thermal stability. *In silico* analyses revealed an increased preference for the closed flap conformation when in the apo state alongside increased structural rigidity of distinct regions of the protease. Thus, the apo HP3 protease requires an increased energetic penalty for the conformational changes required for drug binding, which contributes to decreased favourability of drug binding. Secondly, the ligand-bound protease exhibited wider flap conformations (semi-open) with an increased degree of flap flexibility and ligand mobility within the active site. The inhibitor-bound protease complexes therefore exhibit diminished stability though minimising intermolecular interactions and resisting closing of the flaps in the presence of inhibitors. Collectively, these data elucidate the biophysical basis for the selection of non-active site mutations in the protease and how these mutations contribute to inhibitor resistance. By shifting the drug resistance paradigm to incorporate the dynamic nature of the enzyme and enzyme-complexes a better understanding of resistance mechanisms is made salient. [452]

Acknowledgements

I would like to thank my supervisor, Professor Yasien Sayed, for your endless support throughout these years. Your enthusiasm and love for science and for your students has always inspired me. Thank you for having such confidence in me since the very beginning and for giving me the freedom that you did, you have allowed me to grow as a scientist and thinker in a way that I will never truly be able to thank you enough for.

A special thank you to Michelle Bolognesi. You are the love of my life and a constant source of inspiration for me. Thank you for igniting my passion for life and for being so willing to have endless discussions with me. Life would be significantly less appealing without you. I love you.

I would like to thank my family and friends for their love and support, despite my attempts to bore them with the details of my project. Mom, dad, Ryan and Bron; my loving, supportive family.

Thank you to my friends and colleagues at the Protein Structure-Function Research Unit (PSFRU) for the shared laughter, tears, and discussions. A great thank you to Dr Ramesh Pandian for your involvement in solving the crystal structure of the “HP3” protease, without your expertise this would not have been achieved so comprehensively.

I owe an immense amount of gratitude to my friend, co-author, and mentor, Dr Roland Worth. You have been instrumental in my development as a scientist and your friendship has been invaluable.

Thank you to the South African National Research Foundation (NRF) and Wits Postgraduate Research Award (PMA) for financial assistance throughout my degree. This work was also supported by the Global Challenge Research Fund (GCRF) through Science and Technology Facilities Council (STFC), grant number ST/R002754/1: Synchrotron Techniques for African Research and Technology (START).

Research Outputs

1. Sherry D, Worth R, Sayed Y. (2020) Two-Step Preparation of Highly Pure, Soluble HIV Protease from Inclusion Bodies Recombinantly Expressed in *Escherichia coli*. *Curr. Protoc. Protein. Sci*. 100(1):e106. doi: 10.1002/cpps.106. PMID: 32339408.

   Breakdown of contribution: performed all laboratory experiments and data analysis, performed troubleshooting, established the purification protocol, wrote and composed the entire manuscript, edited subsequent drafts of the manuscript, wrote and submitted all of the supporting documentation (e.g., Cover letters) to the journal.

2. Sherry D, Worth R, Sheik Ismail Z, Sayed Y. (2021) Cantilever-centric mechanism of cooperative non-active site mutations in HIV protease: implications for flap dynamics. *J. Mol. Graph. Model.* 106(107931) doi: 10.1016/j.jmgm.2021.107931 ISSN 1093-3263

   Breakdown of contribution: performed docking studies and molecular dynamics simulations of the proteins utilised in the study. Performed data analysis, compiled all figures and tables, wrote the manuscript in its entirety, and edited subsequent drafts of the manuscript.

3. Sherry D, Worth R, Sayed Y. (2021) Plasticity-associated functionality and inhibition of the HIV protease. *Adv Exp Med Biol - Protein Reviews* https://doi.org/10.1007/5584_2021_655

   Breakdown of contribution: performed the literature review and gathered all information in the article, generated and compiled all images in the article, compiled the tables found in the article. Wrote the manuscript in its entirety and edited subsequent drafts before submission to the journal.

4. Sherry D, Pandian R, Sayed Y. (2021) Structural and thermodynamic analysis of non-active site mutations in the cantilever region of the South African HIV subtype C protease. (Manuscript)

   Breakdown of contribution: performed all laboratory experiments including protein expression and purification, isothermal titration calorimetry, differential scanning calorimetry, macromolecular crystallisation, and enzyme assays. Performed data analysis of the experimental outcomes. Generated and compiled all images and tables in the article. Wrote the manuscript in its entirety and edited subsequent drafts.

Conference Outputs

The Molecular Biosciences Research Thrust (MBRT) Research Day, University of the Witwatersrand, (Johannesburg, South Africa) November 2019. *Poster*. “The Structure of a Novel South African Subtype C HIV-1 Protease”. Dean Sherry and Yasien Sayed.

CCP4 Crystallography School in South Africa, Data Collection to Structure Refinement and Beyond Digital Conference, University of Cape Town (South Africa), February 2021. *Poster*. “Overexpression, purification and crystallisation of a unique South African HIV subtype C protease variant”. Sherry D, Pandian R, Achilonu I, and Sayed Y.

CCP4 Crystallography School in South Africa, Data Collection to Structure Refinement and Beyond Digital Conference, University of Cape Town (South Africa), February 2021. *Poster*. “Structural and Functional Studies of Biological Macromolecules Implicated in Disease: Towards Drug Discovery and Development”. Pandian R, Sherry D, Olfsen J, Worth R, Achilonu I, Dirr HW, and Sayed Y

List of Abbreviations

AIDS	Acquired immunodeficiency syndrome
APV	Amprenavir
ART	Anti-retroviral therapy
ARV	Anti-retroviral
ATV	Atazanavir
C_α	Alpha carbon
CM	Carboxy-methyl
CRF	Circulating recombinant forms
C-SA	South African subtype C
DEAE	Diethyl aminoethyl
DMSO	Dimethyl sulphoxide
DRV	Darunavir
DSC	Differential scanning calorimetry
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
FDA	Food and drug administration
FPLC	Fast protein liquid chromatography
FPV	Fosamprenavir
HAART	Highly active anti-retroviral therapy
HDX-MS	Hydrogen-deuterium exchange-mass spectrometry
HIV	Human immunodeficiency virus
HP3	HIV-1 subtype C protease containing I13V, I62V and V77I mutations
HP4	HP3 protease with the additional mutation: L33F
IDV	Indinavir
IFD	Induced fit docking
IN	Integrase
INSTI	Integrase strand transfer inhibitor
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
K_a	Association equilibrium constant

K_d	Dissociation equilibrium constant
k_{off}	Dissociation rate constant
k_{on}	Association rate constant
ln	Natural log
LPV	Lopinavir
MD	Molecular dynamics
MDR	Multi-drug resistant
MWCO	Molecular weight cut-off
NFV	Nelfinavir
NMR	Nuclear Magnetic Resonance
NNRTI	Non-nucleoside reverse transcriptase inhibitor
NOP	Naturally occurring polymorphism
NRTI	Nucleoside reverse transcriptase inhibitor
PDB	Protein data bank
PI	Protease inhibitor
PMSF	Phenylmethylsulfonyl fluoride
PR	Protease
PR·PI	Protease-protease inhibitor complex
ps	Picosecond
R_g	Radius of gyration
RMSD	Root-mean square deviation
RMSF	Root-mean square fluctuation
RT	Reverse transcriptase
RTV	Ritonavir
SASA	Solvent accessible surface area
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SE-HPLC	Size exclusion-high performance liquid chromatography
SQV	Saquinavir
ssRNA	Single-stranded RNA
T	Temperature
T_m	Transition temperature

TPV	Tipranavir
WT	Wild type
Δ*G*	Change in Gibbs free energy
Δ*H*	Change in enthalpy
Δ*S*	Change in entropy

List of Figures

CHAPTER 1 – INTRODUCTION

Figure 1. The lifecycle of the HIV-1. 3

Figure 2. Three-dimensional structure of the HIV-1 protease. 5

Figure 3. Heterogenous flap conformers displayed by the apo HIV-1 PR. 7

Figure 4. Proposed catalytic mechanism of the HIV aspartyl protease. 9

Figure 5. Three-dimensional molecular structures of FDA-approved protease inhibitors for clinical use…. 10

Figure 6. Diagrammatic representation of hydrogen bonding between the 12

Figure 7. Polymorphism locations of mutations in the PR associated with significantly reduced PI binding. 13

Figure 8. The inter-region locations of the I13V/I62V/V77I amino acid polymorphisms. .18

CHAPTER 2 – EXPRESSION AND PURIFICATION OF THE HIV-1 PR

Figure 1. Diagrammatic representation of the Basic Protocol and Support Protocol 1…..24

Figure 2. Expected tricine-SDS-PAGE results following solubilization, purification, and refolding of HIV protease……26

Figure 3. Thermogram following titration of acetyl-pepstatin into a sample of HIV protease purified according to the Basic Protocol……29

CHAPTER 3 – ANALYSES OF BINDING ENERGETICS, DIMER STABILITY AND THREE-DIMENSIONAL STRUCTURE

Figure 1. Fitted enthalpy data of the C-SA PR, WT_{D25A} PR and HP3 PR titrated with the aspartyl PR inhibitor, acetyl-pepstatin……..47

Figure 2. Specific activity determination comparison of enzymatically active PR vs inactive PR containing the D25A active site mutation……..48

Figure 3. Structural alignment of ribbon representations of the C-SA PR, and the HP3 PR ……50

Figure 4. Gibbs free energy of binding for the interactions between WT_{D25A} PR and C-SA PR and four PIs……52

Figure 5. Summary of the thermodynamic parameters obtained for the binding of ATV, DRV, SQV and RTV to WT_{D25A} PR and HP3 PR……53

Figure 6. DSC thermograms of dimeric WT_{D25A} PR and HP3 PR in their apo forms (APO), and in the presence of ATV, DRV, RTV and SQV…….55

Figure 7. The active site of C-SA PR and WT_{D25A} PR……59

Figure 8. Binding geometries of clinical inhibitors to C-SA PR and WT_{D25A} PR……..60

Figure 9. I13V/I62V/V77I mutations occur at the interfaces between internal regions of the PR……..66

CHAPTER 4 – MOLECULAR DYNAMICS AND INDUCED FIT DOCKING STUDIES

Figure 1. The structural architecture of the HIV PR enzyme (PDB ID: 6I45)……..78

Figure 2. Calculated Glide parameters following IFD of PIs to WT PR, HP3 PR and HP4 PR at pH 5……84

Figure 3. RMSD of the PR peptide backbone with respect to time…….……..87

Figure 4. The non-active site mutations change the dynamic profile of the PR enzyme….90

Figure 5. The curling of the backbone of the cantilever tips modulates flap opening……..92

Figure 6. Cross-correlation maps showing the vectors that govern the alpha carbon atoms of the WT, HP3 and HP4 PRs……93

Figure 7. The effect of the A) I13V and B) I62V, V77I and L33F mutations on the Cα atoms of residues of the PR……95

Figure 8. The average Rg of the WT PR, HP3 PR, and HP4 PR when in complex with clinical inhibitors……...98

Figure 9. Interatomic distances measured between the Cα atoms of the flap tips and active site residues and the angles measured between the Ca atoms of the cantilever tip……100

Figure 10. RMSF ratio (mutant/WT) of the atoms of each PI when in complex with the PR variants……104

Figure 11. Protein-ligand interaction diagrams of WT PR and HP3 PR in complex with SQV……108

Figure 12. Protein-ligand interaction diagrams of WT PR and HP4 PR in complex with DRV……110

Figure 13. Molecular surface renderings of the active site cavities of the PR variants in complex with PIs……112

Figure 14. Water molecules associated with the flap tips of the PR variants……114

CHAPTER 5 – THE MECHANISM OF NON-ACTIVE SITE MUTATIONS IN DRUG RESISTANCE

Figure 1. The structure of the WT PR in complex with a natural substrate (CA-SP2) analogue (PDB ID: 6O48)……130

Figure 2. The hydrophobic sliding mechanism for the opening of the PR flaps……132

Figure 3. Two-dimensional structures of the CA-SP2 Gag cleavage site as well as all currently available FDA-approved HIV-1 PIs……133

Figure 4. The dynamic substrate envelope is a probability distribution which is likely to be more accurate to incorporate into structure-based drug design……..136

Figure 5. Structural overlay of WT and mutated PR variants in complex with DRV and the corresponding amino acid sequence alignment of the variants……..141

Figure 6. Diagrammatic representation of the Gag polyprotein……...144

Figure 7. Non-active site mutations alter PR dynamics in favour of resistance……149

Figure 8. Allosteric binding sites identified for the HIV PR……151

Figure 9. The ‘exo site’ of the PR occupied by the small fragment 4D9 (PDB ID: 3KF0). ……..152

Figure 10. The ‘flap top’ pocket located on the surface of the PR flaps with 1F1 bound (PDB ID: 4EJK)……153

Figure 11. PIs bound to the ‘flap top’ binding pocket located on the surface of the flaps…156

Figure 12. The Eye site allosteric binding pocket of the HIV PR (PDB ID: 5VJ3)……157

Figure 13. HV9 bound to a probable allosteric binding site located near the N-terminus of the PR……..159

CHAPTER 6 – GENERAL DISCUSSION AND CONCLUSIONS

Figure 1. The influence of the HP3 mutations on the apo and inhibitor-bound HIV PR….171

List of Tables

CHAPTER 1 – INTRODUCTION

Table 1. Amino acid polymorphism locations in the HIV PR associated with reduced drug susceptibility……11

Table 2. Protease inhibitors associated with the accumulation of the I13V, L33F, I62V and V77I mutations ……...16

CHAPTER 3 – ANALYSES OF BINDING ENERGETICS, DIMER STABILITY AND THREE-DIMENSIONAL STRUCTURE

Table 1. Data collection and refinement statistics for the HP3 PR structure…….49

Table 2. Summarised induced fit docking parameters for C-SA PR and WT_{D25A} PR with clinical inhibitors……...54

Table 3. Comparison of ligand binding parameters between the WT_{D25A} PR and the C-SA PR……..57

Table 4. Comparison of transition temperatures (T_m) data for the apo HP3 PR and WT_{D25A} PR as well as in complex with four clinical protease inhibitors……..…62

Table 5. Comparison of binding affinity (K_d) data for the HP3 PR and WT_{D25A} PR and four clinical protease inhibitors……..63

CHAPTER 4 – MOLECULAR DYNAMICS AND INDUCED FIT DOCKING STUDIES

Table 1. RMSD and RMSF data for WT PR, HP3 PR, and HP4 PR collected over the course of a 50 ns MD simulation……86

Table 2. The increase/decrease (%) in the average Asp25/25'-Ile50/50' distance of the HP3 PR and HP4 PR with respect to the WT PR when in complex with PIs…..102

Table 3. The increase/decrease (%) in the average RMSF of key regions of the HP3 PR and HP4 PR with respect to the WT PR when in complex with PIs……102

Table 4. The average SASA ($Å^2$) of the PIs when in complex with WT PR, HP3 PR and HP4 PR……105

Table 5. The type of major interactions formed between PIs and key residues in the PR active site and flap regions……107

CHAPTER 5 – THE MECHANISM OF NON-ACTIVE SITE MUTATIONS IN DRUG RESISTANCE

Table 1. Resistance mutations of HIV-1 PR in response to PI exposure……..137

Table 2. The locations of PR mutations associated with significantly reduced PI binding……138

Table of Contents

ABSTRACT ..... **II**

ACKNOWLEDGEMENTS ..... **III**

RESEARCH OUTPUTS ..... **IV**

CONFERENCE OUTPUTS ..... **V**

LIST OF ABBREVIATIONS ..... **VI**

LIST OF FIGURES ..... **IX**

LIST OF TABLES ..... **XII**

CHAPTER 1 – INTRODUCTION **1**

OVERVIEW OF HIV/AIDS 1

HIV genetic diversity *1*

The lifecycle of HIV *2*

THE HIV PROTEASE 4

The structure and function of the PR *4*

The catalytic mechanism of the PR *8*

INHIBITION OF THE PROTEASE 8

PROTEASE MUTATIONS AND INHIBITOR RESISTANCE 11

Active site polymorphisms *11*

Non-active site polymorphisms *14*

Non-active site mutations exhibit cooperativity *14*

RESISTANCE ASSOCIATED COOPERATIVE MUTATIONS: I13V/I62V/V77I 16

AIM AND OBJECTIVES 19

Aim *19*

Objectives *19*

CHAPTER 2 – EXPRESSION AND PURIFICATION OF THE HIV-1 PR **20**

TWO-STEP PREPARATION OF HIGHLY PURE, SOLUBLE HIV PROTEASE FROM INCLUSION BODIES EXPRESSED IN ESCHERICHIA COLI 20

Introduction *22*

Basic Protocol *23*

Support Protocol 1 ..... *26*

Support Protocol 2 ..... *28*

Reagents and Solutions *30*

Commentary *32*

Literature Cited *35*

CHAPTER 3 – ANALYSES OF BINDING ENERGETICS, DIMER STABILITY AND THREE-DIMENSIONAL STRUCTURE 37

STRUCTURAL AND THERMODYNAMIC ANALYSIS OF NON-ACTIVE SITE MUTATIONS IN THE CANTILEVER REGION OF THE SOUTH AFRICAN HIV SUBTYPE C PROTEASE 37

Abstract *38*

Introduction *38*

Methodology *41*

Results *46*

Discussion *56*

Supplementary data *67*

References *68*

CHAPTER 4 – MOLECULAR DYNAMICS AND INDUCED FIT DOCKING STUDIES 75

CANTILEVER-CENTRIC MECHANISM OF COOPERATIVE NON-ACTIVE SITE MUTATIONS IN HIV PROTEASE: IMPLICATIONS FOR FLAP DYNAMICS 75

Abstract *76*

Introduction *76*

Methods and Materials *81*

Results and Discussion *83*

Molecular dynamics *85*

Essential dynamics of the apo protease *85*

RMSD and RMSF *86*

The effect of fulcrum and cantilever mutations on PI-bound protease *97*

Concluding remarks *113*

References *121*

CHAPTER 5 – THE MECHANISM OF NON-ACTIVE SITE MUTATIONS IN DRUG RESISTANCE 126

PLASTICITY-ASSOCIATED FUNCTIONALITY AND INHIBITION OF THE HIV PROTEASE 126

ABSTRACT 127

INTRODUCTION 128

The tertiary structure of the HIV protease *129*

The emergence of resistance in HIV protease *131*

Primary resistance mutations: changes in the active site geometry *134*

Non-active site mutations associated with drug resistance *138*

The role of Gag mutations in resistance *142*

Flexibility-assisted functionality of the PR *145*

Allosteric inhibition of the PR *150*

CONCLUSIONS 160

REFERENCES 161

CHAPTER 6 – GENERAL DISCUSSION AND CONCLUSIONS 168

Altered apo PR stability 168

Diminished inhibitor-bound complex stability 172

Conclusion 174

REFERENCES 176

Chapter 1 – Introduction

Overview of HIV/AIDS

The Human Immunodeficiency Virus (HIV), the causative agent of Acquired Immunodeficiency Syndrome (AIDS), continues to impose a significant socioeconomic impact globally, particularly in developing countries. The latest statistics on global infections estimate that ~38 million people were living with HIV as of 2020 (UNAIDS, 2020). Therefore, HIV remains a serious global epidemic, particularly in sub-Saharan Africa where approximately 70% of all infections occur (roughly 26 million individuals). Worryingly, due to the recent global COVID-19 pandemic, the UNAIDS 2020 targets for the AIDS epidemic have been severely affected:

> *"…COVID-19 risks blowing HIV progress way off course. Missed targets have resulted in 3.5 million more HIV infections and 820 000 more AIDS-related deaths since 2015 than if the world was on track to meet the 2020 targets. In addition, the response could be set back further, by 10 years or more, if the COVID-19 pandemic results in severe disruptions to HIV services."*
>
> *UNAIDS, Global AIDS Update 2020*

Furthermore, although antiretroviral therapy has significantly reduced the number of AIDS-related deaths since their peak in 2015, access to antiretroviral therapy is unfortunately not ubiquitous. This fact is reflected in the morbidity and mortality rates occurring in sub-Saharan Africa and South America (UNAIDS, 2020). HIV exhibits a high degree of genetic diversity, which results in varied distributions of HIV subtypes across the globe. The high degree of genetic diversity observed in HIV is attributable to the low fidelity and lack of proof-reading capabilities of the viral reverse transcriptase (RT) enzyme, in combination with the high replication rate of the virus (Roberts et al., 1988). The mutation-prone nature of the RT, therefore, lies at the root of the significant genetic diversity observed between and within subtypes. The RT is responsible for the genetic mutations which arise rapidly in antiretroviral target proteins, which makes developing effective drugs rather challenging (Bebenek et al., 1989).

HIV genetic diversity

Two types of HIV have been identified to date: HIV-1 and HIV-2 (Taylor et al., 2008). The majority of global infections are attributed to HIV-1, while HIV-2 infections are confined to West Africa (Clavel et al., 1986). HIV-1 can be classified into groups (M, N, O and P) which can be further classified into subtypes. Group M is composed of subtypes A, B, C, D, F, G, H, J and K, as wells as circulating recombinant forms (CRFs), of which subtype C is responsible for

approximately 50% of HIV-1 infections worldwide. Subtype C occurs mainly in developing countries in sub-Saharan Africa, India, Brazil and China while the subtype B is responsible for infections that occur in more developed countries such as the United States, Western Europe and Australia (UNAIDS, 2020). Due to the geographical distributions of these subtypes, global research has been largely focused on the subtype B and, as a result, all FDA-approved antiretroviral drugs have been developed using the subtype B. Research into the subtype C is grossly under-represented, for instance, the Protein Data Bank (PDB) is replete with thousands of structures of the HIV protease enzyme, of which a mere five structures are of the subtype C. It is clear, therefore, that global research into HIV-1 subtype C is disproportionately low when considering its prevalence globally (~50%). Furthermore, according to the UNAIDS 2020 report, the prevalence of HIV infections in the adult population of South Africa is approximately 17% which, alarmingly, implicates South Africa as an epicentre of the HIV/AIDS epidemic.

The lifecycle of HIV

HIV belongs to the genus *Lentivirus*, which is a subgroup of the retrovirus family which are enveloped viruses containing single-stranded RNA (ssRNA) genomes. The virus infects T-lymphocytes that express the CD4 glycoprotein on the surface of the plasma membrane. The lifecycle of the virus is briefly summarised in Figure 1. HIV attaches to the surface of the target cell through specific interactions that occur between receptors on the target cell (CD4 and co-receptors CCR5/CXCR4) and viral-envelope glycoproteins (gp120 and gp41)(Chan and Kim, 1998; Wyatt and Sodroski, 1998). Following successful adsorption, the viral envelope fuses with the plasma membrane and the viral ssRNA genome is released from the virion into the cytoplasm. The viral genome is then reverse transcribed into double-stranded DNA (dsDNA) by the aforementioned viral RT enzyme (Jacobo-Molina and Arnold, 1991). The viral dsDNA enters the nucleus where it is then integrated into the host genome through the action of the viral integrase (IN) enzyme. Thereafter, the viral genome is transcribed into messenger RNA (mRNA) as well as copies of the ssRNA viral genome. The mRNA is translated into viral polyprotein precursors; namely, Gag and Gag-Pol. The Gag and Gag-Pol polyproteins are transported to the plasma membrane along with copies of the viral RNA genome where assembly of the viral virion occurs. The polyproteins are anchored to the inner surface of the membrane via a covalently linked N-myristate and a charged region on the surface of the protein (Henderson et al., 1983). The process of recruiting the necessary viral components initiates membrane envelopment, and budding, which results in the production of the immature HIV virion (Briggs and Kräusslich, 2011). The virus requires the viral protease (PR) to produce mature, infectious, viral progeny through the catalytic processing of the Gag and Gag-Pol polyproteins into functional enzymes and structural proteins (Freed, 1998).

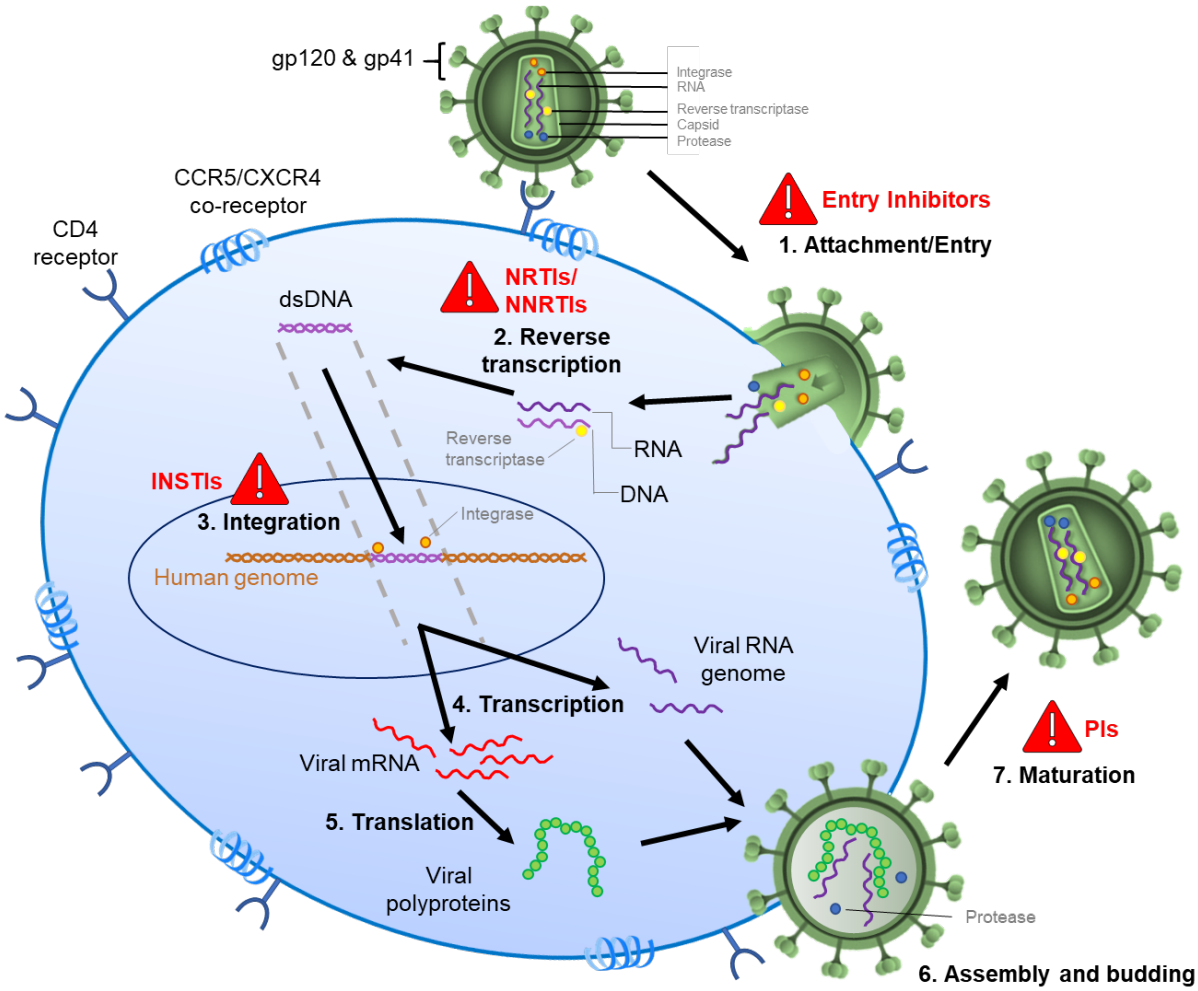


Figure 1. **The lifecycle of the HIV-1.** Stages of the life cycle are numbered. 1: Adsorption of the HIV virion to receptors on the host-cell membrane and subsequent fusion with the host cell. 2: Reverse transcription of viral ssRNA into dsDNA. 3: Entry of viral dsDNA into the nucleus and integration into the host cell's DNA sequence. 4: Transcription of viral ssRNA viral genomic material and mRNA. 5: Translation of viral mRNA into viral Gag/Gag-Pol polyproteins. 6: Assembly and budding of newly translated viral proteins and ssRNA genome. 7: Maturation of the HIV virion into its mature, infectious, form. Red triangles depict the stages of the lifecycle that are major targets for ART. Entry inhibitors inhibit the interactions formed between viral surface glycoproteins and receptors. Nucleoside reverse transcriptase inhibitors (NRTIs) and non-nucleoside reverse transcriptase inhibitors (NNRTIs) target the RT enzyme. Integrase strand transfer inhibitors (INSTIs) target the IN enzyme and pre-integration complex. Protease inhibitors (PIs) target the protease enzyme. (self-generated image using Microsoft PowerPoint)

The identification of the various steps in the lifecycle of the virus has led to the development of a number of antiretroviral drugs that target various stages in the process. Effective treatment with highly active antiretroviral therapy (HAART) can reduce the viral load to near-undetectable levels but the virus cannot be fully eliminated (Tözsér, 2001). HAART makes use of over twenty different drugs to target the various stages of the viral life cycle, including nucleoside reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs), protease inhibitors (PIs), integrase strand transfer inhibitors (INSTIs) and fusion/entry inhibitors (Figure 1). Interestingly, there is evidence which suggests that PIs are the most effective course of therapy in the treatment of HIV infection (Lefebvre and Schiffer, 2008; Thompson et al., 2010; Volberding and Deeks, 2010). The critical role that the PR plays in viral maturation has been illustrated experimentally through its inactivation via site-directed mutagenesis of the active site aspartate residue (Asp25)(Kohl et al., 1988). Effective processing of the Gag/Gag-Pol polyproteins by the PR is crucial for viral proliferation and, as a consequence, the PR has become a major drug targets for anti-HIV therapy (Debouck, 1992).

The HIV protease

The HIV-1 PR (EC 3.4.23.16) is a major target for drug therapy as it is vital for the production of infectious viral progeny through the processing of its substrates; Gag and Gag-Pol. The virus is assembled as an immature, non-infectious, virion containing predominantly uncleaved Gag/Gag-Pol proteins. Maturation is achieved through cleavage of five specific cleavage sites in the Gag by the PR, leading to reorganisation of the internal virion arrangement (Briggs and Kräusslich, 2011; Darke et al., 1988). The inhibition of the action of the PR, therefore, avoids subsequent infections of host cells through rendering the virus non-infectious.

The structure and function of the PR

Experimental techniques such as X-ray crystallography and nuclear magnetic resonance (NMR) have led to the elucidation of a relatively unambiguous picture of the three-dimensional structure of the PR (Figure 2). The PR is a homodimeric enzyme composed of two identical 99-amino acid subunits which associate via the dimer interface to form the active site cavity. Being an aspartyl protease, the PR active site contains a conserved catalytic triad, Asp25-Thr26-Gly27, which is characteristic of this class of enzymes (Mager, 2001; Navia et al., 1989). Substrates entry into the active site cavity is regulated by two “flap regions” of the enzyme, composed of β-strands, which are highly flexible. A dynamic image of the PR has been revealed through NMR and molecular dynamics (MD) simulations which show that the flaps exist in a dynamic equilibrium of heterogenous flap conformations; namely, open, semi-open and closed (Freedberg et al., 2002; Ishima et al., 1999; Nicholson et al., 1995).

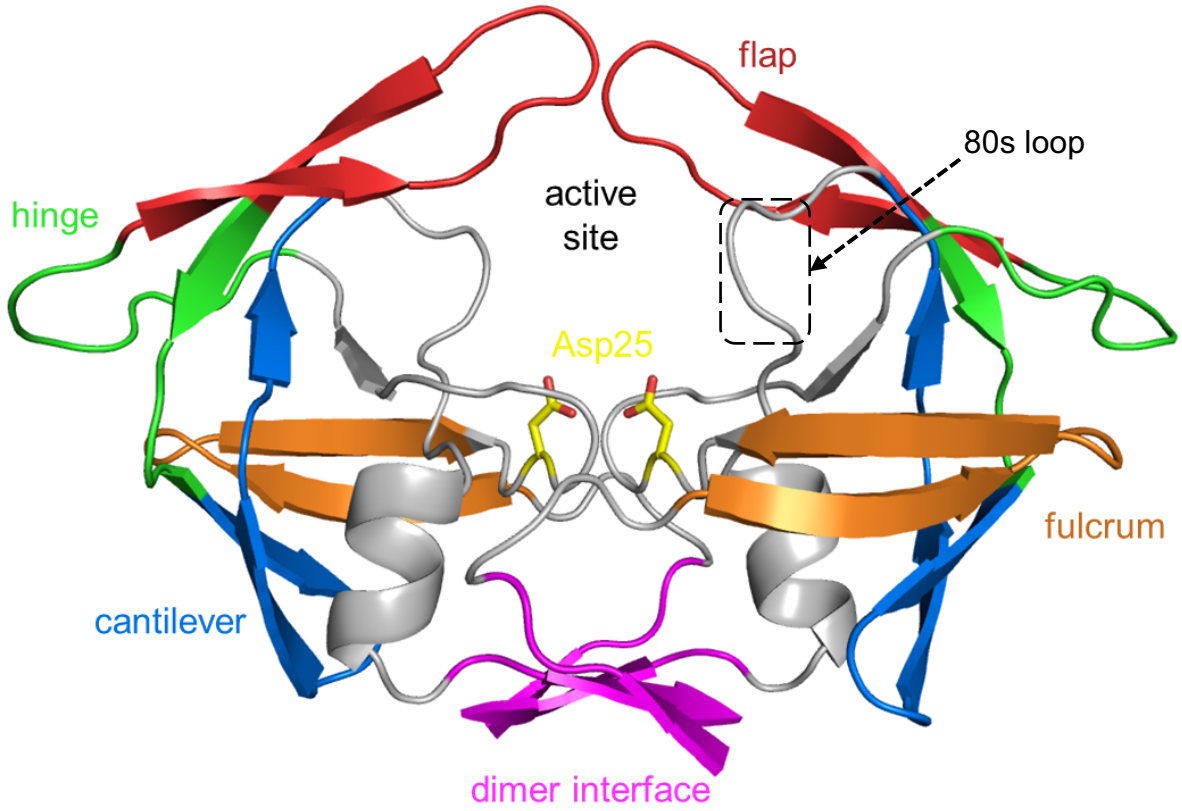


Figure 2. Three-dimensional structure of the HIV-1 protease. The flexible flap regions of the PR (residues 46-54; red) cover the top of the active site cavity. The fulcrum region (residues 10-22; orange), cantilever region (residues 62-78; blue) and hinge regions (residue 35-42 & 57-61; green) of the PR are implicated in maintaining the equilibrium of flap conformations. Each monomer contributes one aspartic acid residue at position 25 (shown in yellow) which forms a part of the conserved catalytic triad (Asp25-Thr26-Gly27). The dimer interface (residues 1-4 & 96-99; purple) is formed by the N- and C-terminus. The 80s loop (residues 79-84) is indicated in the black rectangular box. This figure was generated using PyMOL Molecular Graphics Software. PDB ID: 6I45; Sherry et al., 2020.

Traditionally, the conformations of the flaps have been defined by the distance between the alpha carbon (C_α) of the flap tip residues (Gly51/51'); where interatomic distances less than 7 Å, between 7 and 12 Å, and greater than 12 Å are defined as closed, semi-open and open conformations, respectively. However, definitions using this metric will be affected by both flap asymmetry and flap tip curling. As such, more recently these conformations have been defined by the interatomic distance between the C_α of the catalytic Asp25/25' residues (located at the bottom of the active site) and the flap tip Ile50/50' residues (Perryman et al., 2004). The new definitions of these conformations are defined as: closed, < 17 Å; semi-open, 17-22 Å; open > 22 Å (Figure 3).

The flexibility of the flaps is necessary for the functioning of the PR as the fully open conformation allows entry of substrate into the active site cavity, while the closed conformation promotes the formation of key intermolecular interactions between the PR and its substrate (Gustchina and Weber, 1990). It has been proposed that the PR is composed of five rigid-body “domains” that move synergistically when ligand binding occurs (Rose et al., 1998), a more recent depiction of the enzyme subdivides these “domains” into specific regions. The PR monomer is, therefore, divided into the terminal region, or the dimer interface, (residues 1-9 and 86-99), the flap region (residues 42-56), the hinge region (residues 35-42 and 57-61), the cantilever region (residues 62-78) and fulcrum region (residues 10-22) (Figure 2). A more accurate understanding of the structure of the PR has provided a paradigm for making sense of the movements of these regions during flap opening and closing. The large conformational range available to the flaps can be largely attributed to the hinge regions but the opening motion is also reliant on the synergetic downwards motion of the hinges as well as the fulcrum and cantilever regions (Foulkes-Murzycki et al., 2007; Hornak et al., 2006). This combined movement is referred to as the hydrophobic sliding mechanism, in which the hinge and cantilever regions can slide downward across the surface of the fulcrum region, causing the upward and outward motion of the flaps. This mechanism relies on the van der Waals contacts that stabilise the hydrophobic core of the PR possessing the ability to swap readily between neighbouring internal hydrophobic residues.

The hydrophobic core of the PR has been defined as twenty amino acids with limited solvent accessibility (Foulkes-Murzycki et al., 2007; Naicker et al., 2014); namely, L5, V11, I13, V15, I19, A22, L24, I26, L33, L38, I62, I64, I66, V75, V77, I85, M89, L90, L93, L97. The hydrophobic sliding mechanism involves several of these hydrophobic residues, in particular, I13 and I15 appear to play a central role in the movement of the core regions of the PR (Foulkes-Murzycki et al., 2007; Ishima et al., 1999). Briefly, in the closed conformation M36 and L38 in the hinge region and I62, I64 and V75 in the cantilever region make contacts with I15 of the fulcrum region. Flap opening occurs when the hinge and cantilever regions slide over I15 and

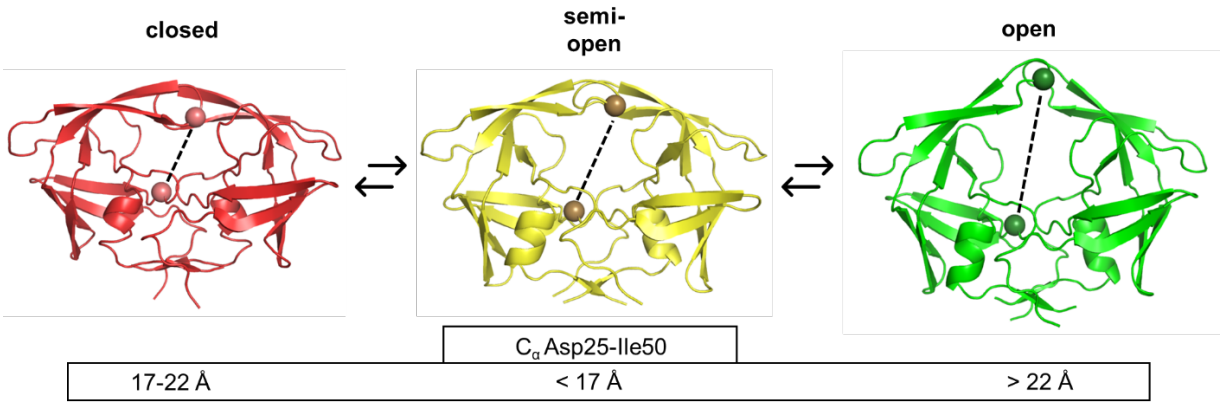


Figure 3. Heterogenous flap conformers displayed by the apo HIV-1 PR. The closed (red), semi-open (yellow) and open (green) conformations of the flaps exist in dynamic equilibrium. Representation of the flaps in the fully open conformer is based on MD simulation models. These conformations are defined by the distance between the C_α of Asp25 and Ile50 as follows; closed < 17 Å; semi-open, 17-22 Å; open > 22 Å. This figure was generated using PyMOL Molecular Graphics Software. PDB ID: 6I45; Sherry et al., 2020.

downwards towards the N- and C-terminus, and the contacts originally made with I15 are transferred to I13 (Foulkes-Murzycki et al., 2007). It would appear, therefore, that the van der Waals contacts in the core serve to stabilise the PR in the closed conformation and the exchange of these contacts allows for the resultant rearrangement of the core and opening of the flaps. A complete three-dimensional understanding of the active site of the PR, including the role of the flaps, has been crucial for the design of inhibitors that are geometrically complementary to the active site. However, when designing transition state analogues, which are a potent class of inhibitors often used to target enzyme activity, an understanding of the catalytic mechanism of the target enzyme is of equal importance.

The catalytic mechanism of the PR

The catalytic mechanism of the HIV PR is believed to be consistent with a general acid-base catalysis mechanism, in which the two Asp25/25' residues play a crucial role through the activation of a coordinated water molecule which acts as a nucleophile. In the active pH range (4 – 4.5), the Asp25 residues of aspartyl proteases exist in opposing states of protonation (Hofmann et al., 2002)(Figure 4). A substrate is initially recognised by the PR and binds to the active site (ES), thereafter the nucleophilic water molecule attacks the carbonyl carbon of the scissile peptide bond and a tetrahedral intermediate complex is formed (ES*). Protonation of the scissile amide nitrogen atom results in the breakdown of the intermediate complex and the release of the two hydrolysis products (P_1 & P_2) whereafter the PR may hydrolyse another substrate (Suguna et al., 1987).

Inhibition of the protease

The knowledge gained about the structure of the PR active site, and its mechanism of action, has allowed for rational drug design of protease inhibitors (PIs) which target the active site and impede the function of the PR. The action of PIs prevents the viral infection from spreading but has no effect on infected cells already carrying the proviral DNA (Robins and Plattner, 1993). There are currently nine PIs (Figure 5) approved for ART by the FDA (excluding the pro-drug fosamprenavir); namely, amprenavir (APV), atazanavir (ATV), darunavir (DRV), indinavir (IDV), lopinavir (LPV), nelfinavir (NFV), ritonavir (RTV), saquinavir (SQV) and tipranavir (TPV).

The difference between inhibitors and the substrates that they mimic lies in the presence of a non-hydrolysable hydroxyethylene core in the inhibitors, which is believed to emulate the tetrahedral transition-state of the scissile bond observed during the proteolytic reaction (Brik and Wong, 2003; Partaledis et al., 1995). The paradigm for drug design generally follows the lock-and-key hypothesis, which is focused on geometric complementarity between the drug and its binding site on the target protein.

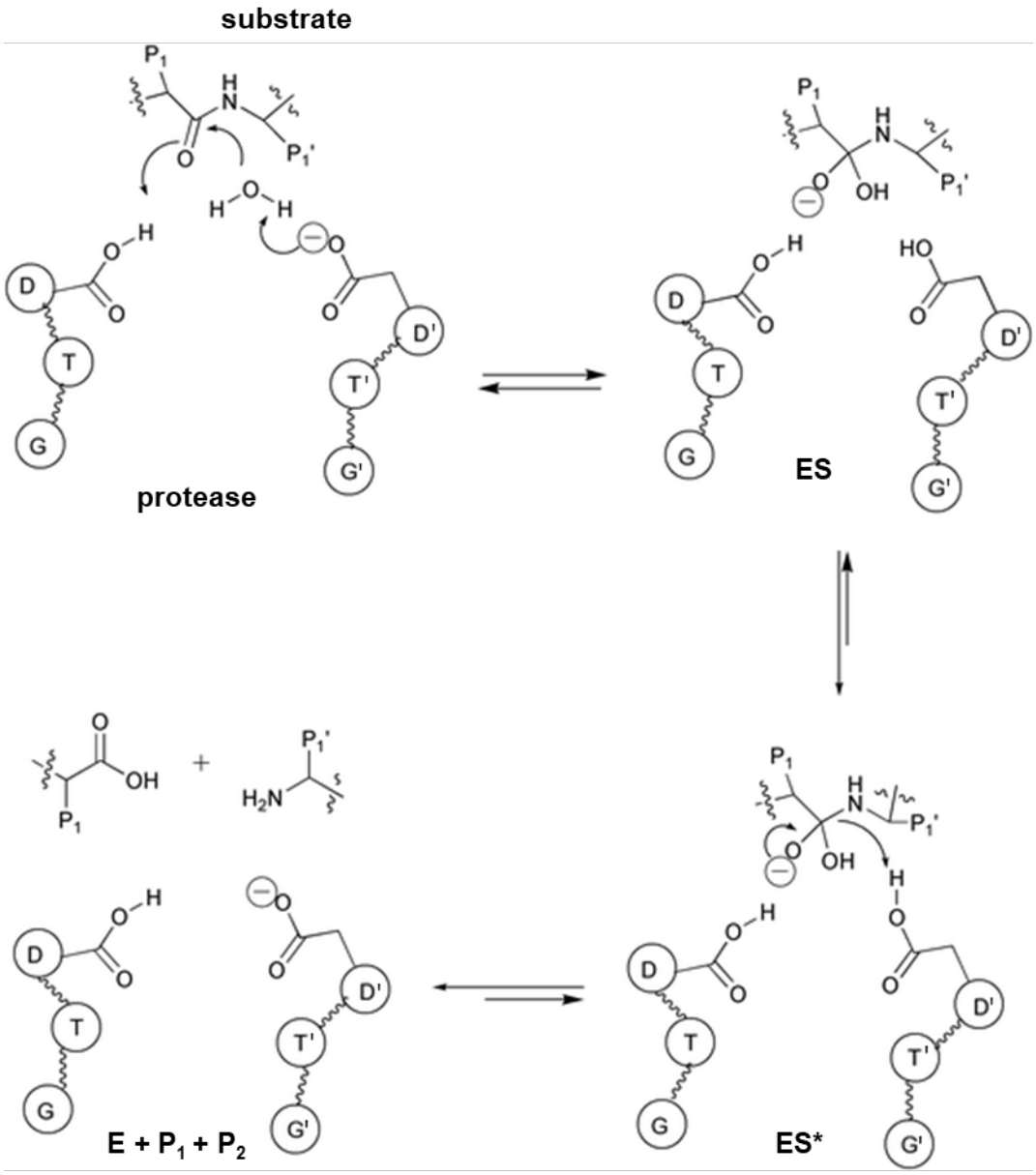


Figure 4. Proposed catalytic mechanism of the HIV aspartyl protease. This reaction mechanism is generalised for all HIV PR substrates. A coordinated water molecule is located at the catalytic aspartate residues (D25/D25') in the unbound PR and acts as a nucleophile when activated by the unprotonated aspartate group. Image adapted from (Suguna et al., 1987).

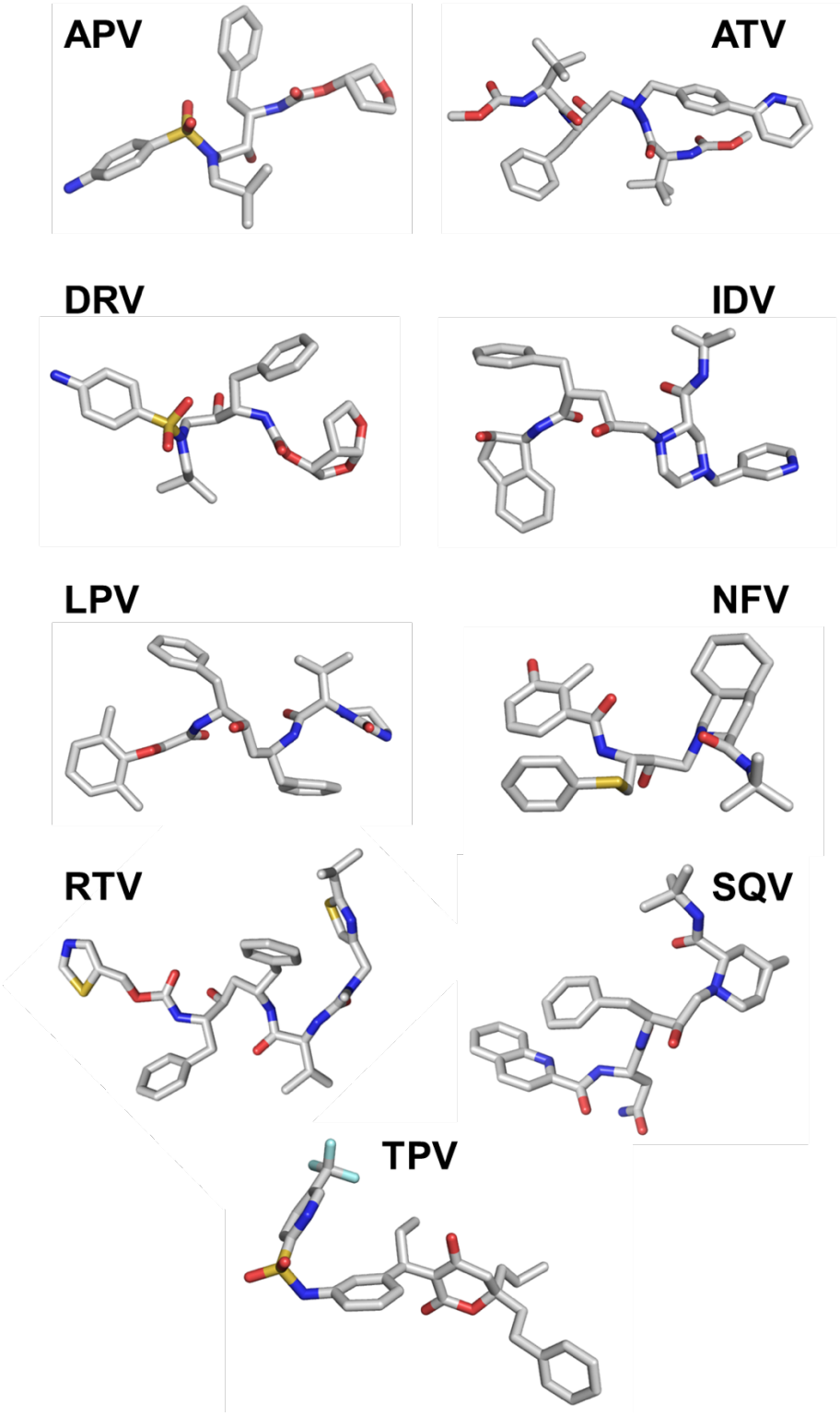


Figure 5. Three-dimensional molecular structures of FDA-approved protease inhibitors for clinical use. The inhibitors are as follows: amprenavir (APV, CID: 65016), atazanavir (ATV, CID: 148192), darunavir (DRV, CID: 213039), indinavir (IDV, CID: 5462355), lopinavir (LPV, CID: 92727), nelfinavir (NFV, CID: 64142), ritonavir (RTV, CID: 392622), saquinavir (SQV, CID: 60787) and tipranavir (TPV, CID: 54682461). This image was generated using PyMOL Molecular Graphics Software.

Due to the highly specific configuration that transition states represent along the reaction coordinate, it is believed that transition state analogues bind to enzymes with higher affinity than their corresponding substrates (Schramm, 2011). As such, PIs are designed to imitate the transition state of the cleavage reaction of the PR which makes these drugs highly specific (Lefebvre and Schiffer, 2008; Thompson et al., 2010; Volberding and Deeks, 2010). The active site cavity of the PR has been divided into eight specific subsites that accommodate the side chains of PIs and the intermolecular contacts formed are relatively uniform for all inhibitor-bound complexes (Figure 6). The standard subsites (S/S') of the PR are Gly27/27', Asp29/29', Gly48/48' and a water molecule coordinated by Ile50/50'. When the substrate is located within the subsites the scissile bond is held between the S1 and S1' subsites (Figure 6) (Schechter and Berger, 1967; Wlodawer and Erickson, 1993).

According to the WHO, first-line ART for adolescents and adults should contain two nucleoside reverse-transcriptase inhibitors (NRTIs), as well as an inhibitor of integrase or a non-nucleoside reverse-transcriptase inhibitor (NNRTI). Upon failure of first-line ART, second-line ART is employed which makes use of NRTIs and PIs. Theoretically, PIs are ideal candidates for successful ART. Unfortunately, however, the selective pressure introduced by ART in conjunction with the aforementioned mutation-prone nature of the viral RT produces numerous mutations in the PR sequence (Roberts et al., 1988). Indeed, mutations observed at over half of the 99 amino acid positions of the PR are implicated in reduced drug susceptibility (Wu et al., 2003)

Protease mutations and inhibitor resistance

Active site polymorphisms

Resistance to PIs generally starts with the onset of primary mutations, which are amino acid polymorphisms that directly impact drug susceptibility and are usually located within the active site cavity (Table 1, Figure 7). Primary mutations, generally, involve substitutions of residues in the active site cavity which affect drug binding through geometric alterations at the protein/inhibitor interface (Chen et al., 1995; Markowitz et al., 1995; Schaffer and Verkhivker, 1998).

Table 1. Amino acid polymorphism locations in the HIV PR associated with reduced drug susceptibility.

Location of mutation	Amino acid position														
Active site	D 30	V 32	M 46	I 47	G 48	I 50	F 53	V 82	N 83	I 84					
Non-active site	L 10	V 11	K 20	L 24	L 33	E 35	K 43	I 54	Q 58	G 73	T 74	L 76	N 88	L 89	L 90

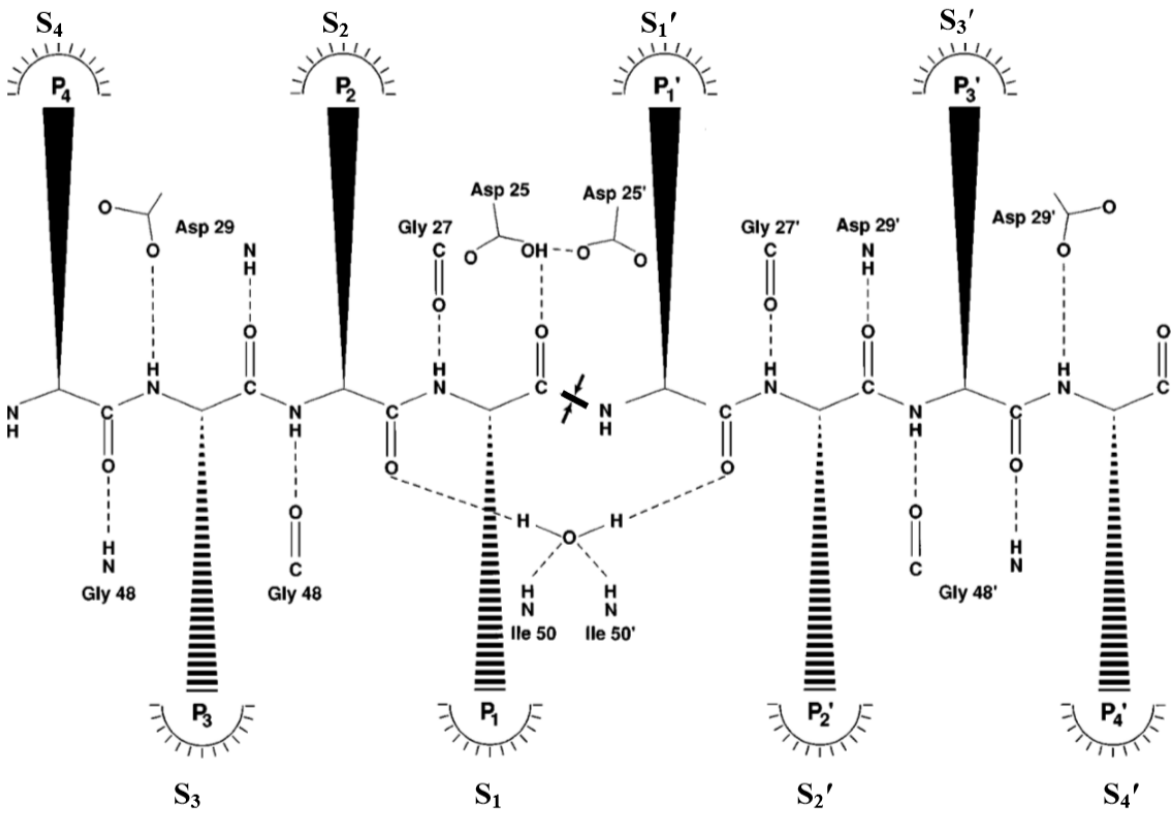


Figure 6. Diagrammatic representation of hydrogen bonding between the active site of the HIV-1 protease and a modelled substrate. The subsites of protease are represented by S4-S1 and S1'-S4' depending on whether the subsite is toward the N- or C-terminus of the scissile bone, respectively. The substrate has equivalent sites denoted as P4-P1 and P1'-P4'. The black arrows indicate the location of the scissile peptide bond (Schechter and Berger, 1967).

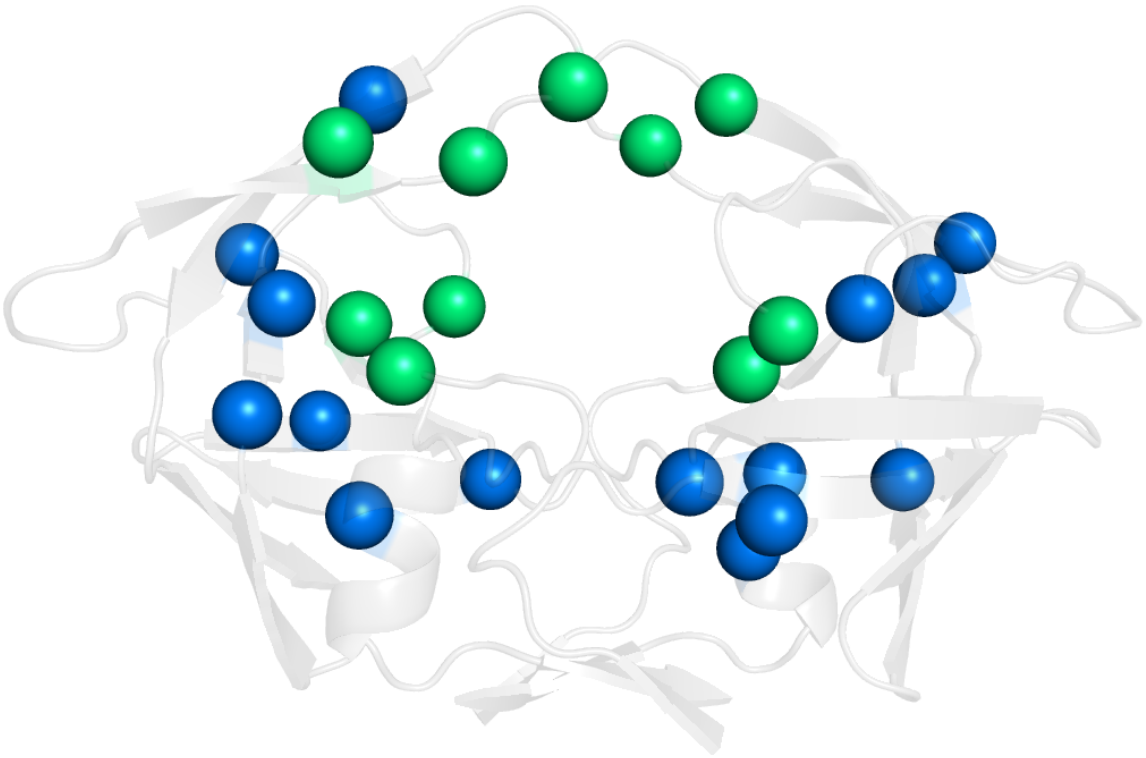


Figure 7. Polymorphism locations of mutations in the PR associated with significantly reduced PI binding. So called primary mutations associated with decreased inhibitor susceptibility (green) occur within the active site while secondary mutations (blue), generally, occur outside of the active site. This figure was generated using PyMOL Molecular Graphics Software. PDB ID: 6I45; Sherry et al., 2020.

The mechanism of action of primary mutations at the active site is, therefore, easily rationalised. PIs are rigid molecules which allows for geometric complementarity and selectivity, but it is this rigidity which limits the ability of PIs to adapt to changes in the active site. As such active site mutations are able to reduce inhibitor binding capacity through directly modulating the geometric complementarity of the active site cavity. Notably, primary mutations possess the potential to diminish binding affinity for substrates in the same manner. As such, subsequent "compensatory" mutations have been seen to accumulate within the hydrophobic core of the PR as well as in the Gag/Gag-Pol cleavage sites (Clavel and Mammano, 2010; Codoñer et al., 2017).

Non-active site polymorphisms

Mutations found outside of the active site, or secondary mutations (Table 1, Figure 7), are thought to compensate for the perturbations generated by primary mutations in order to improve catalytic efficiency and viral fitness (Borman et al., 1996; Nijhuis et al., 1999). An accumulation of mutations within the PR disturbs the balance between substrate turnover and inhibitor binding, in favour of resistance. Primary and secondary mutations result in inhibitor resistance when complex combinations of polymorphisms result in compromised inhibitor binding while selectivity for substrates, and substrate turnover, is maintained (Prabu-Jeyabalan et al., 2002). Secondary mutations have been directly linked with inhibitor resistance (Clemente et al., 2003; Foulkes-Murzycki et al., 2007; Muzammil et al., 2003; Rhee et al., 2010) which suggests that these mutations are not purely compensatory in their nature, however their mechanism of action is not well-understood.

Non-active site mutations exhibit cooperativity

Single amino acid polymorphisms are not generally sufficient to introduce significant inhibitor resistance but, rather, combinations of multiple mutations are necessary to confer resistance to potent inhibitors. For instance MDR strains of the PR have been found to contain anywhere between 10 and 22 polymorphisms (Weber et al., 2015). The role of secondary mutations is likely to be associated with their location in the hydrophobic core of the PR. Such mutations may work synergistically to impart disruptions in drug binding which would otherwise be unfeasible when occurring in isolation. However, the mechanism of action of these mutations is not likely to impart steric irregularities in the active site cavity but, rather, it is likely that their presence results in novel conformational heterogeneity of the PR. Specifically, mutations distal from the active site are thought to alter the conformational equilibrium between open, semi-open and closed conformations of the flaps (Foulkes-Murzycki et al., 2007; Mittal et al., 2012).

Secondary mutations may increase the deleterious effects of a primary mutation through altering the network of intramolecular interactions; a so-called "network hypothesis" has been

proposed as a common mechanism of resistance (Ragland et al., 2014). The network hypothesis describes the complex system of intramolecular interactions through which non-active site mutations may propagate perturbations throughout the PR. However, it is more likely that the changes transmitted by these mutations, ultimately, affect the flexible flap regions rather than the active site residues themselves. The hydrophobic sliding mechanism provides a paradigm for understanding how mutations in the hydrophobic core cause resistance. The dynamics of the PR exist in a finely tuned equilibrium of small-scale (internal) and large-scale (flap and hinge) motions which are largely dependent on the hydrophobic sliding mechanism (Foulkes-Murzycki et al., 2007). Conservative non-active site mutations in the hydrophobic core alter the total internal hydrophobic surface area of the protein, leading to different degrees and strengths of van der Waals contacts, which alters the dynamic range of the flap regions. For instance, mutations that decrease the internal hydrophobic surface area serve to increase the rigidity of the core while concurrently increasing flap/hinge region flexibility (Goldfarb et al., 2015). The more stable core may stabilise the enzyme in a manner that allows a larger array of open conformers, thus, providing a possible explanation for the mechanisms of resistance since increased conformational flexibility is associated with reduced drug susceptibility (Foulkes-Murzycki et al., 2007; Galiano et al., 2009; Goldfarb et al., 2015; Rose et al., 1998).

A key example to consider is a comparison between inhibitor susceptibility between the subtype C PR and the subtype B PR. Thermodynamic analyses of FDA-approved PIs revealed that the subtype C PR exhibits 3- to 24-fold weaker binding for the PIs in comparison to the subtype B (Velázquez-Campoy et al., 2003). These significant differences in binding affinities have been attributed to the eight naturally occurring polymorphisms (NOPs) which occur in the subtype C PR in comparison to the subtype B. The subtype C PR contains NOPs outside of the active site region; namely, T12S, I15V and L19I, which are located in the fulcrum region, M36I and R41K, which are found in the hinge region, and H69K and L89M which are found in the cantilever and 80s loop, respectively. The I93L polymorphism is found in the α-helix present towards the C-terminus. The mutations present in the subtype C have been found to increase the propensity of the flaps to adopt fully open conformers (Ahmed et al., 2013; Kear et al., 2009; Naicker et al., 2013). These findings bolster the hypothesis that conservative mutations in the hydrophobic core of the PR bring about novel conformational heterogeneity in the flaps. Since PIs are designed to bind to the closed conformation, the ability of certain mutations to alter the equilibrium between flap conformations may alter dissociation rates of drugs which would contribute to drug resistance. Non-active site mutations, therefore, change the dynamics of the hydrophobic core which, in turn, alters the dynamic properties of the flap regions. It therefore seems evident that drug resistance arises, at least partially, from

alterations in the dynamics of the flaps which results in an inability of current PIs to induce flap closing (Hornak and Simmerling, 2007).

Resistance associated cooperative mutations: I13V/I62V/V77I

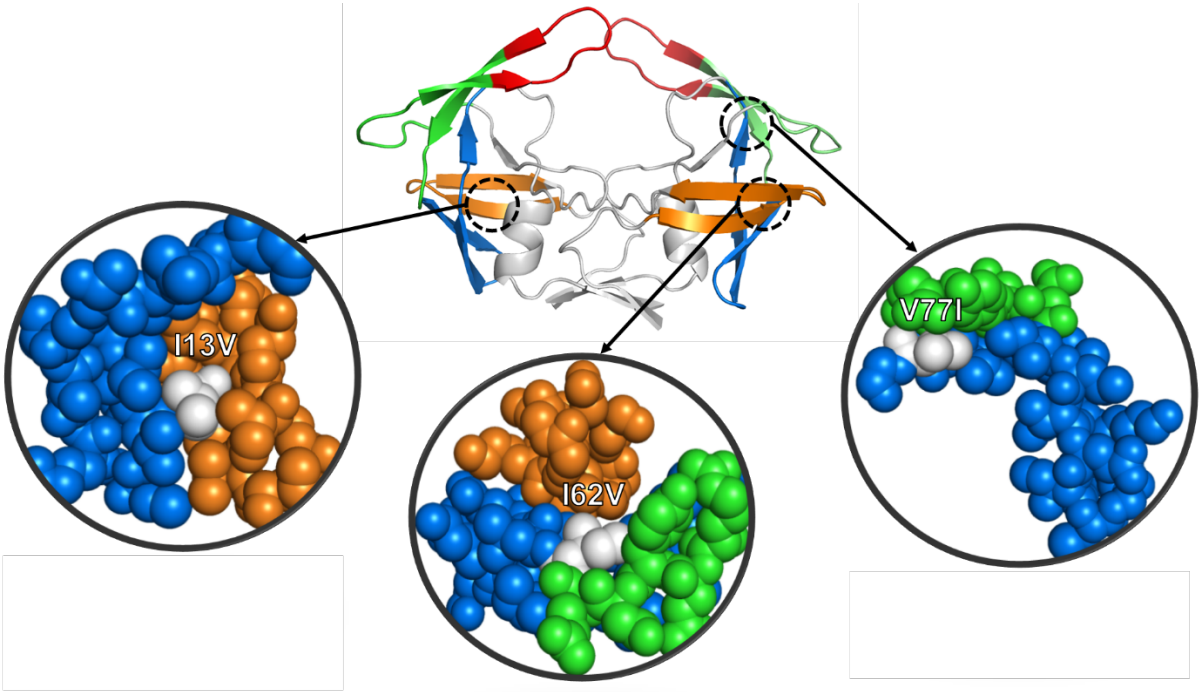
The set of mutations, I13V/I62V/V77I are found in 34%/16%/18% of the general population of individuals infected with HIV, but the propensity of these mutations increase to 51%/48%/35% in the presence of the L33F mutation (Foulkes-Murzycki et al., 2007; Johnson et al., 2003; Rhee et al., 2010; R. W. Shafer et al., 1999). The PIs which bring about the I13V, L33F, I62V and V77I mutations are summarised in Table 2. It is noteworthy that the L33F mutation occurs in response to the most potent generation of PIs, DRV and TPV, as well as LPV and RTV, the latter of which is recommended as a booster PI in the majority of ARV therapy. As such, the propensity of this set of mutations possesses the potential to increase significantly over time due to inhibitor therapy. Although these mutations have been associated with reduced PI susceptibility, their mechanism of action has not yet been fully elucidated. However, it can be hypothesised that these mutations alter the hydrophobic sliding mechanism through changing the network of interactions that constitute the hydrophobic core of the protein. The exact location of the mutations may be essential to their apparent role in altering the conformational dynamics of the PR. It has been hypothesised that mutations at the boundaries of the PR "domains" (or regions) play an important role in resistance (Rose et al., 1998). Furthermore, the presence of mutations at the borders of internal regions appear to increase as a result of exposure to PIs (Rose et al., 1998) and, interestingly, these mutations often result in cross-resistance to PIs (Johnson et al., 2003; Rhee et al., 2010; Shafer et al., 1999). Such observations support the hypothesis that the mechanism of these mutations is not dependent on the specific inhibitor structures but, rather, these mutations impart a non-specific affect by altering the precise dynamic flexibility of the flap regions.

Table 2. Protease inhibitors associated with the accumulation of the I13V, L33F, I62V and V77I mutations.

	APV	ATV	DRV	FPV	IDV	LPV	NFV	RTV	SQV	TPV
I13V							✓			✓
L33F	✓	✓	✓			✓		✓		✓
I62V		✓							✓	
V77I					✓		✓	✓	✓	

(Gathered from: Mata-Munguía et al., 2014; Rhee et al., 2010; Weber et al., 2015)

The I13V/I62V/V77I mutations occur at the internal hydrophobic interface between distinct regions of the core of the PR (Figure 8). The I13V mutation occurs at the interface between the fulcrum and cantilever regions. Similarly, the I62V mutation is seen to occur at the interfaces between the cantilever, fulcrum, and hinge regions. While the V77I mutation occurs at the interface between the cantilever and hinge regions. A mutation from a larger isoleucine residue to a smaller valine would result in a decrease in the total hydrophobic surface area originally occupied inside the core. A smaller hydrophobic surface area may cause the core to compactify and become more stable and, therefore, able to stabilise a wider range of open conformations. Conversely, a mutation of valine to a larger isoleucine may increase the degree of hydrophobic contacts in the hinge and cantilever regions which may provide the framework for larger flap conformations. In summation, mutations resulting in a more stable core may allow the PR to stabilise more open conformers while mutations in the hinge and cantilever regions may augment the effects of core mutations by further increasing the dynamic range of the flaps of the PR (Williams et al., 2019; Zondagh et al., 2019). The selective pressure imparted by PIs introduce this set of cooperative non-active site mutations which occur at the internal interfaces between regions of the PR which may play a crucial role in reduced protease inhibitor susceptibility through modifying the hydrophobic sliding mechanism.




```
WT      PQITLWKRPLVSIKVGGQIKEALLDTGADDTVLEEINLPGKWKPKMIGGIGGFIKVRQYD  60
HP3     PQITLWKRPLVSVKVGGQIKEALLATGADDTVLEEINLPGKWKPKMIGGIGGFIKVRQYD  60

WT      QILIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNF                       99
HP3     QVLIEICGKKAIGTVLIGPTPVNIIGRNMLTQLGCTLNF                       99
```

Figure 8. The inter-region locations of the I13V/I62V/V77I amino acid polymorphisms. Drug resistance mutations (white van der Waals surfaces) are located at hydrophobic interfaces on the internal surfaces of PR regions. These distinct regions of the PR are the cantilever (blue), the fulcrum (orange) and hinge (green). This figure was generated using PyMOL Molecular Graphics Software. PDB ID: 6I45; Sherry et al., 2020. The sequence alignment was generated using the CLUSTAL-Omega tool (EMBL-EBI).

Aim and Objectives

Aim

A particular combination of non-active site mutations; namely, I13V, I62V and V77I, are seen to occur in response to protease inhibitor therapy. These mutations occur at the internal interfaces between distinct regions of the protease. The aim of this study is to assess the impact that this combination of mutations has on the structural stability and dynamics of the South African HIV-1 subtype C protease in order to understand their role in reduced protease inhibitor susceptibility.

Site-directed mutagenesis of the highly conserved active site aspartate residue with an asparagine or alanine has been shown to eliminate proteolytic activity of HIV-1 PR. Thus, to mitigate all auto-proteolytic and catalytic processing of the PR, a D25A active site mutation was incorporated into the PR variants in this study. This is useful when studying the structure of the PR, and not its catalytic function, since auto-catalysis is often a problem for experiments that require high concentrations of the enzyme (such as isothermal titration calorimetry and macromolecular crystallisation).

Objectives

1. Overexpress the inactive wild type (WT) South African subtype C (C-SA) protease (WT_{D25A} PR) as well as the inactive I13V/I62V/V77I protease mutant (referred to as the "HP3 PR") through recombinant protein expression in *Escherichia coli*.
2. Purify the PR variants via ion-exchange chromatography.
3. Elucidate the active site concentration of the purified catalytically inactive PR samples through active site titrations.
4. Evaluate the energetics of binding between the PR variants and PIs via displacement titration calorimetry using selected PIs which cause the mutations at the focus of this study; namely, atazanavir, darunavir, ritonavir, and saquinavir.
5. Elucidate the three-dimensional structure of the apo-HP3 PR via X-ray crystallography. The atomic structure of the C-SA PR has been solved previously (Naicker et al., 2013).
6. Determine the effect that the mutations present in the HP3 PR have on the overall stability of the protease in comparison to the wild type through differential scanning calorimetry.
7. Investigate the effect of these mutations on the dynamic movement of the core regions and flaps of the protease through *in silico* molecular dynamic simulations.
8. Investigate the effect of these non-active site mutations on the binding of protease inhibitors through *in-silico* induced-fit docking studies.

Chapter 2 – Expression and purification of the HIV-1 PR

Two-Step Preparation of Highly Pure, Soluble HIV Protease from Inclusion Bodies Expressed in Escherichia coli

Dean Sherry, Roland Worth and Yasien Sayed.

Current Protocols in Protein Science – Published (2020)

In this article, a protocol for the expression and purification of HIV-1 protease was established. The protocol is composed of a Basic Protocol which describes the purification process involved for recovering the protease from inclusion bodies. The first Support Protocol describes the methods for protein over-expression and extraction of the inclusion bodies containing the PR expressed from *Escherichia coli*. The second Support Protocol describes the procedure for the determination of the active site concentration of the enzyme preparation via isothermal titration calorimetry.

Author contributions: Dean Sherry performed all experimental laboratory work, analysed experimental data, wrote the manuscript and edited subsequent versions of the manuscript for publication. Roland Worth assisted with experimental troubleshooting and revision of manuscript. Yasien Sayed supervised the project and provided revisions for the manuscript.

Two-Step Preparation of Highly Pure, Soluble HIV Protease from Inclusion Bodies Recombinantly Expressed in *Escherichia coli*

Dean Sherry,[1] Roland Worth,[1] and Yasien Sayed[1,2]

[1]Protein Structure-Function Research Unit, School of Molecular and Cell Biology, University of Witwatersrand, Johannesburg, South Africa
[2]Corresponding author: yasien.sayed@wits.ac.za

Heterologous expression of exogenous proteases in *Escherichia coli* often results in the formation of insoluble inclusion bodies. When sequestered into inclusion bodies, the functionality of the proteases is minimized. To be characterized structurally and functionally, however, proteases must be obtained in their native conformation. HIV protease is readily expressed as inclusion bodies, but must be recovered from the inclusion bodies. This protocol describes an efficient method for recovering HIV protease from inclusion bodies, as well as refolding and purifying the protein. HIV protease–containing inclusion bodies are treated with 8 M urea and purified via cation-exchange chromatography. Subsequent refolding by buffer exchange via dialysis and further purification by anion-exchange chromatography produces highly pure HIV protease that is functionally active. © 2020 by John Wiley & Sons, Inc.

Basic Protocol: Recovery, refolding, and purification of HIV protease from inclusion bodies
Support Protocol 1: Expression and extraction of inclusion bodies containing HIV protease expressed in *Escherichia coli*
Support Protocol 2: Determination of the active site concentration of HIV protease via isothermal titration calorimetry

Keywords: HIV protease ● inclusion bodies ● ion-exchange chromatography ● protein purification ● protein refolding

How to cite this article:
Sherry, D., Worth, R., & Sayed, Y. (2020). Two-step preparation of highly pure, soluble HIV protease from inclusion bodies recombinantly expressed in *Escherichia coli*. *Current Protocols in Protein Science*, *100,* e106. doi: 10.1002/cpps.106

Introduction

During the lifecycle of the human immunodeficiency virus (HIV), HIV protease plays a crucial role in virus maturation and has therefore become a major drug target for antiretroviral therapy. HIV protease (EC 3.4.23.16) is an aspartyl endopeptidase capable of cleaving a variety of peptide sequences based on the shape of the substrate envelope (Gottlinger, Sodroski, & Haseltine, 1989; Louis, Ishima, Torchia, & Weber, 2007). Consequently, nonspecific proteolysis as well as autoproteolysis are common in samples containing HIV protease. As a result, HIV protease exhibits cytotoxicity when expressed in a variety of cells, including *Escherichia coli* (Baum, Bebernitz, & Gluzman, 1990), which explains why it is readily sequestered into inclusion bodies, where its functionality is limited (Ido, Han, Kezdy, & Tang, 1991).

The formation of insoluble aggregates of heterologously expressed protein in the form of inclusion bodies is relatively common, especially when employing an *E. coli*–based expression system. Consequently, heterologous expression of viral proteins can present a challenge. In order to maintain the solubility of the target protein, the conditions for protein expression are typically optimized by varying the bacterial strain, expression vector, induction temperature, induction time, and/or concentration of inducing agent. The soluble expression of protein is often preferred because it facilitates downstream processing, such as chromatographic methods. In some instances, however, as in the case of HIV protease, inclusion body expression is still preferred, considering that accumulation of the protein into inclusion bodies largely prevents the protease from adopting its functional form, thus minimizing autoproteolysis. Autoproteolysis is a factor that significantly affects the homogeneity of the sample, which could, for instance, decrease the active site concentration of the final protein yield. Consequently, methodologies for protein solubilization and purification have been developed for proteins that are expressed insolubly. These methods provide a means of obtaining proteins that are difficult to express in a soluble form and generally make use of 8M urea to recover the target protein from inclusion bodies, followed by refolding step where the urea is removed in order to obtain the protein in its native state (Ashnagar et al.,2014; Singh & Panda, 2005; Singh, Upadhyay, Upadhyay, Singh, & Panda, 2015; Wingfield, Palmer, & Liang, 2014). By recovering the protein from insoluble inclusion bodies in this manner, these purification methods provide a means of working with proteins that are otherwise difficult to obtain.

The Basic Protocol outlined below describes an efficient and reproducible method for solubilizing and purifying HIV protease from inclusion bodies. Briefly, the inclusion bodies are

isolated and washed before being denatured using 8 M urea. The denatured protein is then isolated via cation-exchange chromatography and then refolded during buffer exchange via dialysis. The sample of soluble refolded protein is then purified a second time using anion-exchange chromatography, which provides highly pure and soluble HIV protease. In addition, Support Protocol 1 describes the preparation of inclusion bodies heterologously expressed in *E. coli*, and Support Protocol 2 describes the determination of the active site concentration of the purified sample via isothermal titration calorimetry.

Basic Protocol

RECOVERY, REFOLDING, AND PURIFICATION OF HIV PROTEASE FROM ENRICHED INCLUSION BODIES

This protocol describes the unfolding, isolation, refolding, and purification of the recombinant viral protease from inclusion bodies produced in *E. coil*. In particular, the purification protocol involves two steps: isolation of denatured proteins via cation-exchange chromatography, and then refolding and purification of native HIV protease using anion exchange chromatography.

Materials

Washed pellet of HIV protease inclusion bodies (see Support Protocol 1)
Unfolding buffer (see recipe) with and without NaCl
Refolding buffer (see recipe)
Storage buffer (see recipe)
Homogenizer (e.g., Pro Scientific PRO 200, PID: 01-01200)
Rotator
Centrifuge (e.g., Heraues Megafuge 16, PID: 75004271) with 8 × 50-ml fixed-angle rotor and 50-ml centrifuge tubes (e.g., Beckman-Coulter)
FPLC system (e.g., GE AKTA Prime Plus, PID: 11001313) with:
5-ml HiTrap DEAE Fast Flow column (PID: GE17-5154-01)
5-ml HiTrap CM Fast Flow column (PID: GE17-5156-01)
Dialysis tubing, MWCO 3500 (e.g., SnakeSkin)
50-ml stirred-cell concentrator, MWCO 3500 (e.g., Amicon, PID: UFSC05001)
Additional reagents and equipment for standard chromatographic techniques and SDS-PAGE

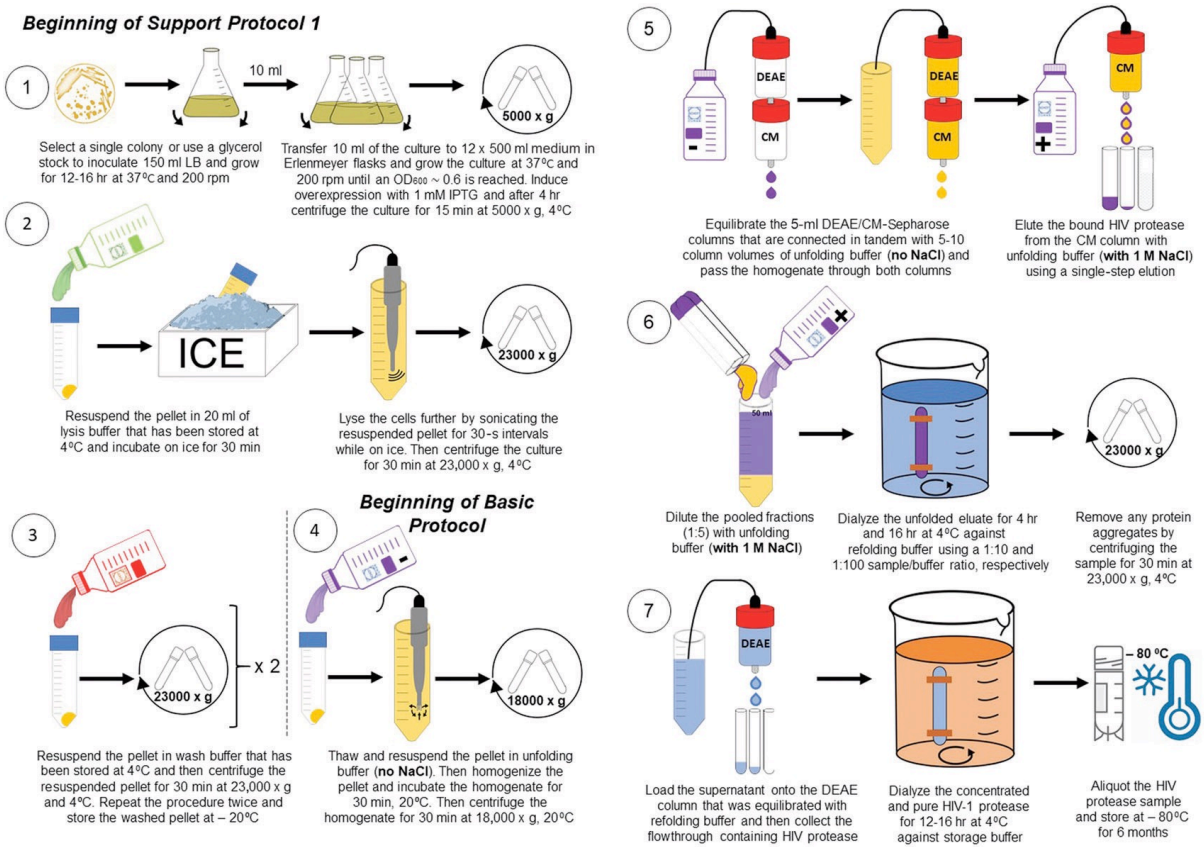


Figure 1 Diagrammatic representation of the Basic Protocol and Support Protocol 1.

Isolate denatured protein via cation exchange

1. Gently thaw an inclusion body pellet (~2 g wet mass) and resuspend in unfolding buffer without NaCl at 100 ml/g pellet mass using a homogenizer for 30 s at 4°C.

   A commonly used denaturant for protein unfolding is guanidine hydrochloride (GdnHCl). Because GdnHCl is a salt, however, it will interfere with protein binding to an ion exchange column, and therefore should not be used. Urea is an ideal denaturant as it is not ionic.

2. Incubate for 30 min at room temperature with gentle inversion on a rotator.
3. Centrifuge suspension for 30 min at 18,000 × *g*, 20°C, and keep the resultant supernatant.
4. Set up a 5-ml HiTrap DEAE column and a 5-ml HiTrap CM Fast Flow column in tandem (Fig. 1) and attach them to the FPLC system.
5. Equilibrate both columns using 5-10 column volumes of unfolding buffer without NaCl at a flow rate of 5 ml/min and 0.3 kPa.
6. Pass the supernatant (step 3) through the columns followed by 5-10 column volumes of unfolding buffer without NaCl to remove any unbound protein from the columns.

   Because the theoretical pI of HIV protease is ~9.1 (Gasteiger et al., 2005) and the pH of the unfolding buffer is 8, the net surface charge of the unfolded protease will be

positive. Thus, the protease binds to the negatively charged resin of the CM-Sepharose column.

7. Remove the DEAE column and elute all bound protein from the CM-Sepharose column using 20 ml unfolding buffer with 1 M NaCl. Collect eluate in 2-ml fractions.

   The DEAE column must be removed before elution (Fig. 1), because unwanted (contaminating) proteins with a net negative surface charge are trapped in the positively charged DEAE column and will also be eluted by the 1 M NaCl.

8. Pool fractions containing protein based on their absorbance at 280 nm.

Refold and purify native HIV protease using anion exchange

9. Dilute pooled fractions 1:5 using unfolding buffer with 1 M NaCl.
10. Dialyze unfolded eluate with gentle stirring against refolding buffer equilibrated at 4°C using the following two-step procedure:

    4 hr at 4°C using a 1:10 sample/buffer ratio

    16 hr at 4°C using a 1:100 sample/buffer ratio

11. Collect dialyzed solution and remove any aggregated protein by centrifuging for 30 min at 23,000 × *g*, 4°C.

    To minimize autoproteolysis, the protein sample should be kept on ice at all times and the remainder of the purification procedure should be performed as soon as possible.

12. Equilibrate a 5-ml HiTrap DEAE Fast Flow column using 5-10 column volumes of refolding buffer and then pass the dialysate through the equilibrated column. Collect flowthrough in 2-ml fractions.

    Because the theoretical pI of HIV protease is ~9.1 (Gasteiger et al., 2005) and the pH of the refolding buffer is 7.5, the net surface charge of the unfolded protease will be positive. Thus, the protease will pass through the positively charged DEAE resin and will be present in the flowthrough.

13. Pool fractions containing protein based on absorbance at 280 nm.
14. Concentrate sample, if necessary, to a concentration of ~0.3-0.5 mg/ml using an Amicon stirred-cell concentrator (3500 MWCO).
15. Evaluate sample purity by tricine-SDS-PAGE.

    A sample analysis is shown in Figure 2.

16. Dialyze concentrated protein sample against storage buffer for 4-8 hr at 4°C.
17. Use an aliquot of purified protein to determine the active site concentration via isothermal titration calorimetry (see Support Protocol 2). Aliquot remaining sample appropriately and store until required.

    The protein can be stored at −20°C and −80°C for 2 and 6 months, respectively.

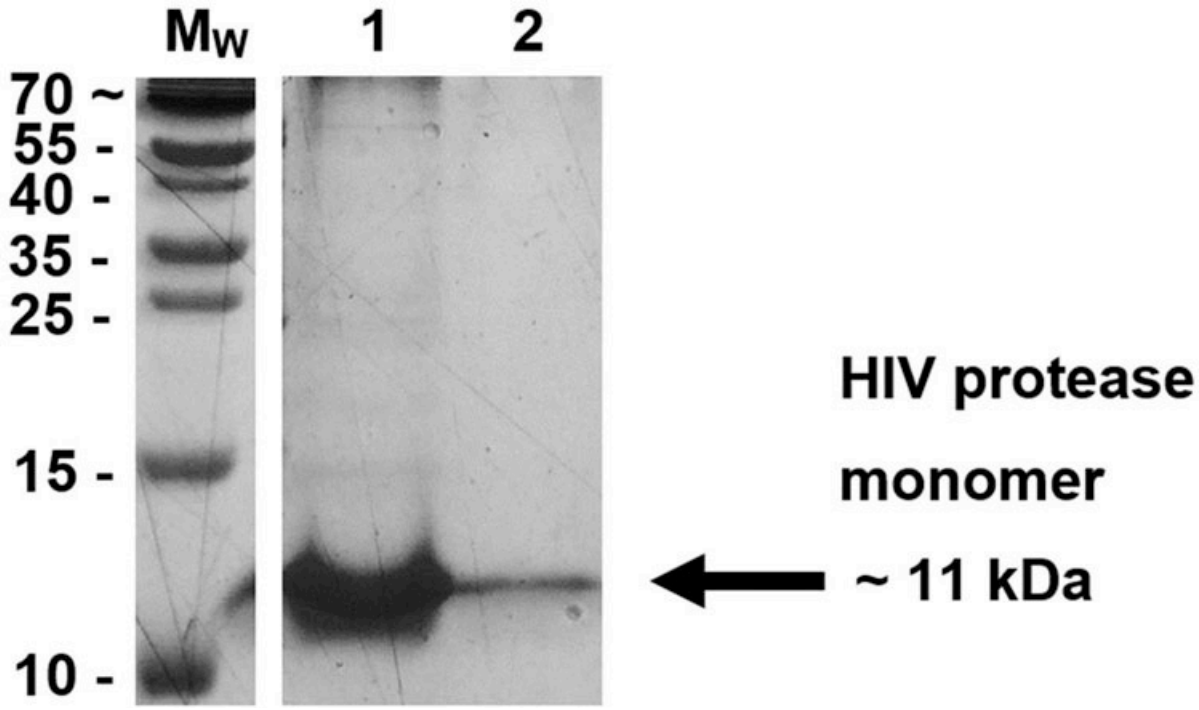


Figure 2 Expected tricine-SDS-PAGE results following solubilization, purification, and refolding of HIV protease. Lane 1, denatured protein eluate; lane 2, pure HIV protease; MW, molecular weight marker. All samples were analysed on the same gel.

Support Protocol 1

EXPRESSION AND EXTRACTION OF INCLUSION BODIES CONTAINING HIV PROTEASE EXPRESSED IN *E. COLI*

The processes involved in successful expression of HIV protease as inclusion bodies and their extraction from *E. coli* are described in three steps: heterologous overexpression of HIV protease in *E. coli*, cell lysis and, recovery of inclusion bodies (Fig. 1).

Materials

Supplemented LB broth (see recipe)

E. coli BL21 (DE3) pLysS strain carrying HIV protease gene cloned into pET-11a expression vector (GenScript Biotech)

1 M IPTG stock solution (see recipe)

Lysis buffer (see recipe), cold

Wash buffer (see recipe), cold

500-ml and 2-L Erlenmeyer flasks (e.g., Duran)

Orbital shaker incubator (e.g., MRC) set to 37°C

UV/visible spectrophotometer (e.g., Ultrospec 1100 pro, Amersham)

750-ml centrifuge bottles (e.g., Beckman)

50-ml centrifuge tubes (e.g., Beckman)
Homogenizer (e.g., Pro Scientific PRO 200, PID: 01-01200)
Sonicator (e.g., QSonica)

Overexpress recombinant HIV protease

1. Inoculate 150 ml supplemented LB broth with a single discrete colony from a selection plate of HIV protease–expressing *E. coli* BL21 (DE3) pLysS *or* a 100-µl aliquot of frozen glycerol stock culture.

   Supplemented LB broth contains 30 µg/ml chloramphenicol and 100 µg/ml ampicillin added fresh before use. These antibiotics are selected in this instance because the pLysS and pET-11a plasmids confer resistance to chloramphenicol and ampicillin, respectively.
2. Incubate 12-16 hr at 37°C with shaking at 200 rpm.
3. Transfer overnight culture in 10-ml aliquots to twelve 2-L Duran Erlenmeyer flasks containing 500 ml supplemented LB broth (1:50 dilution). Grow cultures at 37°C until an OD600 of ~0.6 is obtained.

   The total volume for overexpression should be 6 L in order to obtain a sufficient yield of protein.
4. Induce HIV protease expression by adding 1 mM IPTG to each culture and allow overexpression to occur for 4 hr at 37°C with shaking at 200 rpm.
5. Harvest bacteria by combining in eight 750-ml centrifuge bottles and centrifuging for 15 min at 5000 × *g*, 4°C.

Lyse cells

6. Combine and resuspend the cell pellets in cold lysis buffer (2 ml/g wet cell mass) using a homogenizer for 30 s at 4°C.
7. Incubate suspension for 30 min on ice with continuous stirring using a rotator.
8. Disrupt cells using eight cycles of sonication for 30 s at 70% maximum power followed by equilibration on ice for 30 s. Keep suspension on ice at all times.

Prepare inclusion body pellets

9. Collect inclusion bodies by centrifuging in 50-ml centrifuge tubes for 30 min at 23,000 × *g*, 4°C.
10. Discard supernatant, resuspend pellet in 50ml cold wash buffer, and centrifuge again for 30 min at 23,000 × *g*, 4°C.

    Complete resuspension of the pellets is essential for removing soluble proteins and cellular components, and thus increasing the protease yield and purity at the end of purification.

11. Repeat wash once more.
12. Store washed pellets at −20°C until needed.

Support Protocol 2

DETERMINATION OF THE ACTIVE SITE CONCENTRATION OF HIV PROTEASE VIA ISOTHERMAL TITRATION CALORIMETRY

Successful determination of the concentration of active HIV protease in a purified sample involves two steps: determination of protein concentration via absorbance spectroscopy and determination of active site concentration via isothermal titration calorimetry.

Materials

Purified protein (~1.8 ml, 10-15 μM; see Basic Protocol)
Storage buffer (see recipe)
5.15 mg/ml acetyl-pepstatin stock (see recipe)
Isothermal calorimeter (e.g., TA Instruments Nano ITC)

Determine protein concentration

1. Dilute a sample of protein 1:10 with storage buffer and measure the UV/vis absorbance spectrum from 230 to 400 nm. Repeat for dilutions of 1:12, 1:14, 1:16, and 1:18.

   A simple method for performing this series involves diluting 100 μl protein with 900 μl buffer for the first reading (1:10), and then successively adding 200 μl buffer for each subsequent reading, thereby minimizing sample loss.
2. Plot a standard curve of dilution factor (*x* axis) versus absorbance at 280 nm (*A*280; *y* axis) and determine the protein concentration using the gradient of the linear regression model used to fit the data.

   All A280 values must be corrected for light scattering effects as described by Winder and Gent (1971).

Perform ITC

3. Prepare a working solution of 200 μM acetyl-pepstatin by diluting 12.5 μl of 5.15 mg/ml stock with 487.5 μl storage buffer.

   The ITC technique is highly sensitive to what is referred to as buffer mismatches. In order to prevent a buffer mismatch, the acetyl-pepstatin stock must be diluted with the storage buffer that the purified protein was dialyzed against in the Basic Protocol.
4. Degas the purified protein sample, diluted acetyl-pepstatin, and ~50 ml storage buffer under a vacuum of 0.3 to 0.5 atm for ~15 min.

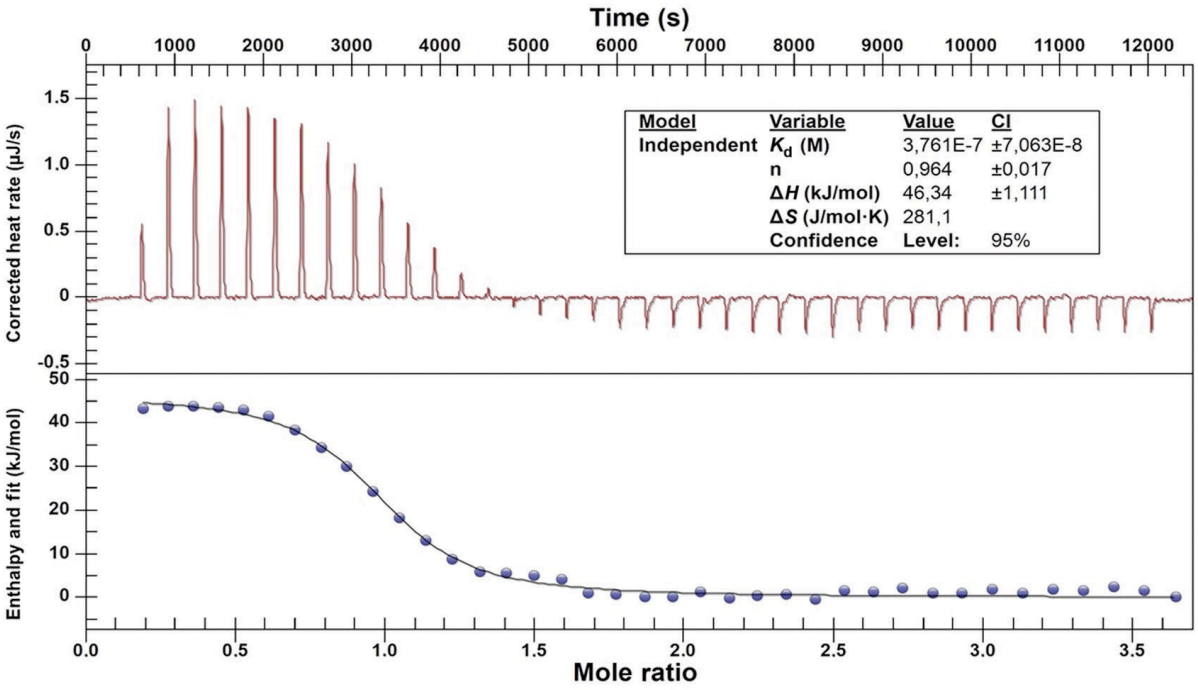


Figure 3 Thermogram following titration of acetyl-pepstatin into a sample of HIV protease purified according to the Basic Protocol. Using a binding stoichiometry of 1:1, it can be seen that the HIV protease active site concentration was equal to ~96% based on the obtained *n* value, where an *n* value equal to 1 represents 100% native (active) protease in the sample.

Any bubbles in the sample cell or the syringe will produce a high heat signal that can mask the heat signal produced by the biomolecular interaction, therefore significantly impacting the accuracy of the data obtained.

5. Fill the ITC sample cell with degassed protein sample and fill the ITC syringe with acetyl-pepstatin.

   For a table outlining the sample cell volume and injection syringe capacity of each particular Nano ITC model, refer to the Nano Isothermal Titration Calorimeter Getting Started Guide for Models 601000, 601001, 601002, page 16.

6. Allow the system to equilibrate to 20°C with stirring at 310 rpm.
7. Perform the experiment with the following parameters:

   300 s initial and final baselines

   300 s injection intervals

   2-µl initial ligand injection volume

   repeated 5-µl injection volumes for a total of 30 injections

8. Analyze data from the isotherm using NanoAnalyze software.

   The software provides integrated heats of the biomolecular interaction (in microjoules), which must be corrected for the heats of dilution before the data are fitted to an

independent binding model. Heats of dilution refer to the signal produced for the titration of ligand into buffer. To correctly account for the heats of dilution, a blank titration must be performed, in which the ligand is titrated into a sample cell containing only the storage buffer.

9. Correct the apparent protein concentration (see step 2) by using the stoichiometric value (*n*) obtained from the fitted data (see step 8). For example:

   Apparent protein stock concentration = 0.32 mg/ml

   Stoichiometric value obtained (*n*) = 0.964 (Fig. 3)

   Corrected protein stock concentration = 0.32 mg/ml × 0.964 = 0.308 mg/ml

10. Determine the active site concentration (c*active*) using the general formula:

$$c\textit{active} = c\textit{observed} \times n$$

where c*observed* is the uncorrected absorbance value obtained at 280 nm and n is the stoichiometry value obtained from the fitted data (see step 8).

A value of 1 indicates a 100% native (active) enzyme preparation. The stoichiometry value n can be used to correct the enzyme concentration values determined by absorbance spectroscopy at 280 nm, since absorbance spectroscopy does not distinguish between native and denatured/degraded proteins in solution.

Reagents and Solutions

Acetyl-pepstatin stock solution, 5.15 mg/ml

5.15 mg/ml lyophilized acetyl-pepstatin powder

10 mM sodium hydroxide

Store aliquots up to 1 year at −20°C

Ampicillin stock solution, 100 mg/ml

Dissolve 1 g ampicillin disodium salt in 10 ml of sterilized water. Filter using a 0.20-µm syringe filter. Store aliquots up to 1 year at −20°C.

Chloramphenicol stock solution, 30 mg/ml

Dissolve 0.3 g chloramphenicol in 10 ml of 70% (v/v) ethanol. Filter using a 0.20-µm syringe filter. Store aliquots up to 1 year at −20°C.

IPTG stock solution, 1 M

Dissolve 0.238 g isopropyl-β-D-thiogalactopyranoside in 10ml sterilized water. Pass through a 0.20-µm filter. Store aliquots up to 2 months at −20°C.

Lysis buffer

10 mM Tris, pH 8.0

1 mg/ml bovine DNase I (see recipe)

10 mM calcium chloride

10 mM magnesium chloride

Store up to 1 year at 4°C

Just before use, add:

2 mM ethylenediaminetetraacetic acid (EDTA)

1 mM phenylmethylsulfonyl fluoride (PMSF)

DNase stock solutions (5 mg/ml) should be prepared in 20 mM Tris·Cl, pH 7.5, with 1 mM MgCl2 and stored at −20°C for up to 18 months. PMSF stocks (1 M) should be prepared in 99% (v/v) ethanol and added immediately before the buffer is used, as its half-life is between 35 and 110 min, depending on the pH of the solution.

Refolding buffer

10 mM Tris, pH 7.5

2 mM dithiothreitol (DTT)

0.02% (w/v) sodium azide

10% (v/v) glycerol

Store up to 1 day at 4°C

Storage buffer

10 mM sodium acetate, pH 5.0

0.02% (w/v) sodium azide

150 mM sodium chloride

Store up to 6 months at 4°C

Supplemented lysogeny broth (LB)

1% (w/v) tryptone powder

0.5% (w/v) yeast extract powder

0.5% (w/v) sodium chloride

Dissolve in 800 ml H2O

Fill to 1 L with H2O

Sterilize by autoclaving at 121°C for 15 min

Store up to 1 month at 4°C

Just before use, add:

30 µg/ml chloramphenicol (see recipe)

100 µg/ml ampicillin (see recipe)

Unfolding buffer with and without NaCl

8 M urea

10 mM Tris, pH 8.0

Store up to 6 months at 20°C

Just before use, add:

2 mM DTT

1 M NaCl, if indicated

Dissolving urea in water is an endothermic reaction, so the solution will become cold to the touch. Heating the solution to 30°C with rapid stirring will help dissolve urea powder more rapidly. Due to the short half-life of DTT, it should only be added immediately before use.

Wash buffer

10 mM Tris, pH 8.0

2 mM EDTA

2% (v/v) Triton X-100

Store up to 1 year at 4°C

Commentary

Background Information

Not surprisingly, heterologous protein expression in *E. coli* disrupts the naturally occurring bacterial cytoplasmic environment. As a result, the intracellular concentration levels of proteins, such as chaperones and proteases, are regulated through aggregation and disaggregation in response to altered protein homeostasis. HIV protease has a history of being sequestered into inclusion bodies when heterologously expressed in *E. coli* (Ido et al., 1991), creating a need for purification methods suitable for proteins that are difficult to express as a soluble fraction, such as membrane proteins and highly hydrophobic proteins (see *Current Protocols* article Wingfield et al., 2014). A protocol for expression and purification of HIV protease from inclusion bodies was previously developed in our laboratory (Naicker et al., 2013). Although it successfully recovers the protease from inclusion bodies and provides pure protein samples, in our experience, it unfortunately results in a relatively low protein yield of ~0.2 mg/liter of culture (Naicker et al., 2013). In an attempt to increase the yield, we developed a purification protocol using a thioredoxin and His-tag protein fusion system. When compared to the standard overexpression and purification protocol, the fusion-derived wildtype protease yield was increased three fold, from 0.83 to 2.5 mg/liter of culture (Zondagh et al., 2019). However, this protocol introduces more steps to the purification process, because the HIV protease must be separated from the fusion constructs using thrombin, which must then be removed to obtain pure protease. The protocol described herein may prove to be useful because it provides high yields (as much as 2.5 mg/liter of culture) of pure HIV protease, offering quantities that are sufficient for structural

and functional characterization, without the need for a fusion tag.

Critical Parameters and Troubleshooting

Cell lysis and isolation of inclusion bodies

Effective sonication is a crucial step in obtaining as much of the target protein as possible. Insufficient cell lysis will significantly diminish the efficacy of the purification process. After sonication and centrifugation, the resulting pellet should be uniform in consistency and color. Inadequate cell lysis will result in the formation of a second uniform layer on top of the densely packed insoluble pellet formed during centrifugation. If this occurs, a second sonication step (or several) is recommended until no visible cell mass is present after centrifugation. A typical sign that unlysed cells are still present in the suspension can be seen after resuspension in unfolding buffer—the suspension becomes thick and grumous. In this event, DNase may be added at a final concentration of 1 mg/ml followed by homogenization and then incubation at room temperature until the viscosity of the suspension decreases significantly. During the wash steps, almost all soluble protein and unwanted cell matter should be removed from the pellets containing the inclusion bodies. If the wash steps are not performed effectively, the yield and purity of the sample may be compromised downstream. Therefore, care must be taken to thoroughly resuspend the pellet in the wash buffer, as the pellet will become more resistant to resuspension after each subsequent centrifugation cycle.

Protein refolding via dialysis

A critical step in the purification process is the dialysis-based buffer exchange, which removes urea and allows the proteins to refold. One-step dialysis is a simple method in which the denaturant (urea) concentration decreases with time, allowing the rate of refolding to increase with time. However, rapid removal of the denaturant may also increase the rate of protein misfolding and aggregation, possibly due to increased exposure of the hydrophobic surface area of the proteins coming into contact with one another (Clark, 2001). If a large amount of the target protein is observed in the aggregate from the post-dialysis centrifugation step, the dialysis may require optimization. Practical investigation has revealed that diluting the denatured protein sample improves the efficacy of refolding, most likely because the lower protein concentration limits nonspecific interprotein hydrophobic interactions during refolding. Therefore, stepwise dialysis can be used to minimize unwanted aggregation. In stepwise dialysis, the denatured proteins are brought to equilibrium first with a small buffer volume (e.g., 10× the sample volume) and then with a larger buffer volume (e.g., 100× the sample volume). The gradual removal of denaturant prevents the resolubilized

proteins from aggregating and coming out of solution, thus increasing the refolding efficiency (Tsumoto, Ejima, Kumagai, & Arakawa, 2003). However, the step wise method is a time-consuming (multi-day) procedure, which must be taken into consideration because HIV protease is likely to undergo autoproteolysis over extended periods of time. Therefore, we do not recommend a multi-day dialysis method, but rather the two-step method described here, which can be completed in a single day while still effectively removing the denaturant. A second aspect to consider when refolding denatured proteins is the choice of dialysis (refolding) buffer. After extensive practical investigation, we have found that the ideal buffer for HIV protease must contain a stabilizing agent (10% [v/v] glycerol) and a reducing agent (such as DTT) capable of preventing formation of nonspecific disulfide bridges. However, buffer choice will vary depending on the intrinsic characteristics of the target protein, and thus varying the ionic strength or pH, and/or supplementing the solution with other additives, may be useful in some cases (Basu, Li, & Leong, 2011).

Understanding the Results

The purity of the target protein will rely heavily on successful sonication of all of the cells in the sample before the wash steps are performed. The purity of the sample can reach >99%, as reflected in the SDS-PAGE analysis depicted in Figure 2, lane 2. The yield of HIV protease obtained will depend largely on successful execution of the refolding step. As is evident in Figure 3, an active site concentration of >96% can be achieved. Because HIV protease has a propensity for autoproteolysis, it is important to determine the true active site concentration of the purified sample. This is achieved by performing an active site titration and then utilizing the binding stoichiometry between the protein sample and the ligand to correct for misfolded/degraded proteins in solution. As A_{280} values do not distinguish between native and misfolded/degraded protein, it is imperative that Support Protocol 2 be performed following successful purification. This protocol has the potential to recover a large amount of target protein (as seen in Fig. 2, lane 1); however, the final yield obtained is dependent on the refolding efficacy. We have observed yields as high as 2.5 mg/liter of culture. Furthermore, the stability of the purified protease will also depend largely on the individual characteristics of the enzyme variant under investigation. For example, inactive HIV protease containing an active site knockout mutation (D25A) will not undergo autoproteolysis. However, the inactive protease is more readily expressed than the active protease, and thus aggregation propensity will become a consideration. In general, HIV protease should be handled on ice or at 4°C during all purification stages. The protocol described here attempts to minimize autoproteolysis by utilizing the biophysical characteristics of

the denatured protease during purification while the enzyme is in its inactive state. Once purified and concentrated, the protein should be aliquoted into small volumes (<1 ml) and stored at −80°C for ∼6 months. To prevent degradation, samples should be thawed gently on ice and should not be refrozen for future use. Thawed samples of HIV protease can be screened using SE-HPLC in order to determine whether samples stored long-term are still useable. A chromatogram produced by a sample of intact HIV protease will produce a single sharp peak, whereas HIV protease that has undergone autoproteolysis will produce a shoulder and/or smaller peaks at higher retention times.

Time Considerations

In total, the time needed to complete all three protocols is ∼5 days. An overnight starter culture is required for efficient overexpression of protease on the following day. The time taken to lyse the *E. coli* cells and prepare pellets containing inclusion bodies of HIV protease (see Support Protocol 1) is about half a day. Once pellets are prepared, the time required to unfold, purify, and refold the protease (see Basic Protocol) is ∼1 day. The ITC procedure (see Support Protocol 2) can be performed in half a day. A step in the Basic Protocol that may require further optimization between protease variants is the refolding procedure. Buffer optimization is a crucial step in successful refolding of the protease, because this step will determine how much denatured protein is recovered from the inclusion bodies. If a multistep dialysis method is required for refolding, the protocol may require an extra day or two.

Acknowledgments

This work was supported by the South African Medical Research Council (SAMRC) with funding to Yasien Sayed under a Self-Initiated Research Grant as well as the Competitive Programme for Rated Researchers (to Y.S., no. CSRP170428229183).

Conflicts of Interest

The authors declare no conflict of interest.

Literature Cited

Ashnagar, F., Khodabandeh, M., Arpanaei, A., Sadigh, Z. A., Rahimi, F., & Shariati, P. (2014). Optimizing primary recovery and refolding of human interferon-b from *Escherichia coli* inclusion bodies. *Iranian Journal of Biotechnology*, *12*(4), 26–34. doi: 10.15171/ijb.1157.

Basu, A., Li, X., & Leong, S. S. J. (2011). Refolding of proteins from inclusion bodies: Rational design and recipes. *Applied Microbiology and Biotechnology*, *92*(2), 241–251. doi: 10.1007/s00253-011-3513-y.

Baum, E. Z., Bebernitz, G. A., & Gluzman, Y. (1990). Isolation of mutants of human immunodeficiency virus protease based on the toxicity of the enzyme in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America, 87*(14), 5573–5577. doi: 10.1073/pnas.87.14.5573.

Clark, E. D. (2001). Protein refolding for industrial processes. *Current Opinion in Biotechnology*, *12*(2), 202–207. doi: 10.1016/s0958-1669(00) 00200-7.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In J. M. Walker (Ed.), *The proteomics protocols handbook* (pp. 571–607). Totowa, NJ: Humana Press. doi: 10.1385/ 1-59259-890-0:571.

Gottlinger, H. G., Sodroski, J. G., & Haseltine, W. A. (1989). Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(15), 5781–5785. doi: 10.1073/pnas.86.15.57 81.

Ido, E., Han, H. P., Kezdy, F. J., & Tang, J. (1991). Kinetic studies of human immunodeficiency virus type 1 protease and its active-site hydrogen bond mutantA28S. *Journal of Biological Chemistry*, *266*(36), 24359–24366.

Louis, J. M., Ishima, R., Torchia, D. A., & Weber, I. T. (2007). HIV-1 protease: Structure, dynamics, and inhibition. *Advances in Pharmacology*, *55*, 261–298. doi: 10.1016/S1054-3589(07) 55008-8.

Naicker, P., Achilonu, I., Fanucchi, S., Fernandes, M., Ibrahim, M. A. A., Dirr, H. W., Sayed, Y. (2013). Structural insights into the South African HIV-1 subtype C protease: Impact of hinge region dynamics and flap flexibility in drug resistance. *Journal of Biomolecular Structure and Dynamics*, *31*(12), 1370–1380. doi: 10.1080/07391102. 2012.736774.

Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, S. M., & Panda, A. K. (2015). Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microbial Cell Factories*, *14*, 41. doi: 10.1186/s12934-015-0222-8.

Singh, S. M., & Panda, A. K. (2005). Solubilization and refolding of bacterial inclusion body proteins. *Journal of Bioscience and Bioengineering*, *99*(4), 303–310. doi: 10.1263/jbb.99.303.

Tsumoto, K., Ejima, D., Kumagai, I., & Arakawa, T. (2003). Practical considerations in refolding proteins from inclusion bodies. *Protein Expression and Purification*, *28*(1), 1–8. doi: 10.1016/s1046-5928(02)00641-1.

Winder, A. F., & Gent, W. L. (1971). Correction of light-scattering errors in spectrophotometric protein determinations. *Biopolymers*, *10*(7), 1243–1251. doi: 10.1002/bip.360100713.

Wingfield, P. T., Palmer, I., & Liang, S.-M. (2014). Folding and purification of insoluble (inclusion body) proteins from *Escherichia coli*. *Current Protocols in Protein Science*, *78*, 6.5.1–6.5.30. doi: 10.1002/047114 0864.ps0605s78.

Zondagh, J., Basson, A. E., Achilonu, I.,Morris, L., Dirr, H. W., & Sayed, Y. (2019). Drug susceptibility and replication capacity of a rare HIV-1 subtype C protease hinge region variant. *Antiviral Therapy*, *24*(5), 333–342. doi: 10.3851/ IMP3308.

Key Reference

Naicker et al. (2013). See above.

Original publication on which this protocol is based.

Chapter 3 – Analyses of binding energetics, dimer stability and three-dimensional structure

Structural and thermodynamic analysis of non-active site mutations in the cantilever region of the South African HIV subtype C protease

Dean Sherry, Ramesh Pandian, Yasien Sayed

Manuscript in preparation (2021)

In this article, the effect of the non-active site mutations at the focus of this study are analysed using conventional *in vitro* biochemical techniques in order to provide a greater understanding of their mechanism of action in protease inhibitor resistance. These techniques include isothermal titration calorimetry, differential scanning calorimetry, enzyme assays and macromolecular crystallisation.

Author contributions: Dean Sherry performed the laboratory experiments and the subsequent data analysis of the results from these experiments and wrote the manuscript. Ramesh Pandian provided critical expertise in the solving of the three-dimensional structure of the HP3 PR (PDB ID: 6I45). Yasien Sayed supervised the project and provided proofreading and revisions for the manuscript.

Structural and thermodynamic analysis of non-active site mutations in the cantilever region of the South African HIV subtype C protease

Dean Sherry, Ramesh Pandian, Yasien Sayed*

Protein Structure-Function Research Unit, School of Molecular and Cell Biology, University of the Witwatersrand, Johannesburg, 2050, South Africa

*Corresponding author: Yasien Sayed; e-mail address: yasien.sayed@wits.ac.za.

Abstract

Human immunodeficiency virus (HIV) infections in sub-Saharan Africa represent more than half of global infections. The protease of the virus is a major drug target in antiviral therapy, and its rapid rate of mutation contributes significantly to clinical failure. The set of cooperative mutations, I13V/I62V/V77I, have been associated with reduced inhibitor susceptibility and are the focus of the current study. The mutant protease decreased the binding affinities of ATV and DRV by 64- and 12-fold, respectively, and decreased the overall favourable Gibbs free energy for ATV, DRV, RTV and SQV. Moreover, these mutations decreased the thermal stability of the protease when in complex with ATV and DRV by approximately 6.4 and 4.2 °C, respectively. The crystal structure of the mutant protease revealed that the location of these mutations and their effect on the hydrophobic sliding mechanism may be crucial in their role in resistance. Additionally, the importance of the active site aspartyl residue 25 for the binding of the four clinical protease inhibitors was assessed by ITC, DSC and docking studies. The T_m values for the WT_{D25A} PR were similar to the C-SA PR while the Gibbs free energy of binding decreased for all the inhibitors in the presence of the D25A mutation. This change was mirrored in the calculated docking scores for the inhibitors showing decreased affinity. The D25A knock-out mutation increases protein yield and eliminates autocatalysis of purified protein samples; however, its removal corresponds to a decrease in inhibitor binding affinity.

Key words: HIV protease subtype C, expression and purification, crystallisation, displacement titration calorimetry, differential scanning calorimetry.

Introduction

The HIV-1 protease (PR) is one of the most important drug targets for anti-HIV therapy due to the critical role that the enzyme plays in viral maturation. The PR cleaves viral Gag and Gag-

Pol polyproteins into functional enzymes and structural proteins which cause a significant rearrangement of the interior of the virion, resulting in mature and infectious viral progeny (Göttlinger et al., 1989; Louis et al., 2007). Anti-HIV therapy has been bolstered by the advent of protease inhibitors (PIs) which target the active site cavity of the PR homodimer; formed at the interface between two subunits each comprised of 99 amino acid residues. Access to the active site it regulated by two dynamic flaps, composed of β-hairpins, which cover the active site cavity (Wlodawer et al., 1989). The dynamic nature of the PR has been revealed through NMR experiments (Freedberg et al., 2002; Nicholson et al., 1995) as well as molecular dynamics (MD) simulations (Ishima et al., 1999). The flaps exist in a dynamic equilibrium between open, semi-open and closed conformations. The open conformation allows entry of substrates/inhibitors into the active site whereafter the closed conformation permits binding through the formation of several intermolecular interactions (Gustchina and Weber, 1990). The ability of the flaps to exist in an equilibrium of conformations has been largely attributed to the flexible hinge regions (residues 35-42 & 56-61) which undergo large conformational changes during flap movement (Gustchina and Weber, 1990). However, more recently, the overall dynamics of distinct regions of the PR have been implicated in the motions of the flaps. Specifically, flap opening has been attributed to the concerted downward movement of the fulcrum (residues 10-22), cantilever (residues 62-78) and hinge regions as well as a slight rotation about the dimer interface (residues 1-4 & 95-99) in what is referred to as the hydrophobic sliding mechanism (Foulkes-Murzycki et al., 2007; Hornak et al., 2006).

PIs are active site inhibitors designed to mimic the tetrahedral transition state of the proteolytic cleavage of the natural substrates of the PR; Gag and Gag-Pol (Brik and Wong, 2003). A major limiting factor in the treatment of HIV-1 infection is a result of the emergence of viral strains that exhibit resistance to PIs (Baxter et al., 2016; Boden and Markowitz, 1998; Johnson et al., 2003). Inhibitor resistance arises from the large genetic variation within HIV strains; a consequence of the high viral replication rate and the low fidelity and lack of proof-reading of viral reverse transcriptase (Roberts et al., 1988). Under the selective pressure of PIs, the virus selects for PR variants containing specific amino acid mutations that lower the affinity for inhibitors while maintaining sufficient affinity for the substrate. Mutations within the active site cavity have the potential to diminish binding affinity for Gag and Gag-Pol and, as such, secondary mutations outside of the active site accumulate in the PR sequence. These secondary mutations are linked to inhibitor resistance (Clemente et al., 2004; Muzammil et al., 2003; Rhee et al., 2010); however, their role has long been thought to be largely compensatory in nature and, thus, their mechanism of action in drug resistance has remained largely undefined.

It has been hypothesised that non-active site mutations may change the dynamic equilibrium of the flaps, which is to say, these mutations may result in novel conformational heterogeneity that selects for conformers that preferentially bind to the substrate and simultaneously diminish inhibitor binding. Cross-resistant PR variants are known to evolve complex combinations of non-active site mutations alongside relatively few active site mutations (Baxter et al., 2016; Clemente et al., 2004; Kozal, 2004; Ohtaka et al., 2002; Olsen et al., 1999) which alludes to a common cooperative mechanism between numerous mutations outside the active site. A particular set of non-active site mutations which display cooperativity are the mutations: I13V, L33F, I62V and V77I. The I13V, I62V and V77I mutations occur in 34%, 16% and 18% of the general population, respectively (Johnson et al., 2003; Rhee et al., 2010; Robert W. Shafer et al., 1999). Interestingly, in the presence of the L33F mutation these incidences increase to 51%, 48% and 35%, respectively, relative to the general population (Foulkes-Murzycki et al., 2007). This set of mutations has been correlated with decreased inhibitor binding affinities, but their mechanism of action has yet to be elucidated. However, their proximity within the PR and their apparent cooperative movements during MD simulations (Foulkes-Murzycki et al., 2007) has led to the hypothesis that these mutations contribute to resistance through modulating the flexibility of the hydrophobic core of the PR. Mutations of this nature possess the potential to alter the mobility of the hydrophobic core through changing the total hydrophobic surface area as well as the nature of hydrophobic contacts present within the core of the PR. These differences, in turn, may alter the functional properties of the active site through modulating the equilibrium of the flaps, shifting the comparative binding affinities of inhibitors and substrates. Since PIs are designed to bind the closed conformation of the PR and are structurally rigid, mutations that shift the aforementioned equilibrium of flap conformations to favour open or semi-open conformations, may alter the dissociation rates of inhibitors, resulting in decreased PI susceptibility (Clemente et al., 2003; Perryman et al., 2004; Rose et al., 1998).

The global AIDS epidemic currently involves approximately 38 million infected individuals with an estimated 70% of infections occurring in sub-Saharan Africa (UNAIDS, 2020). Shockingly, close to 20% of the South African population is infected with HIV, with the subtype C being the most prevalent in the region. Globally, about half of HIV infections occur because of the subtype C; however, all currently available ARVs are designed against the subtype B due to its geographical distribution. Subtype B is the most prominent subtype in North America and Europe and, as a result, global research into the subtype C is wholly underrepresented. The South African subtype C PR (C-SA PR) is, in essence, the consensus subtype C PR (Los Alamos HIV sequence database, http://www.hiv.lanl.gov/). The C-SA PR contains eight naturally occurring polymorphisms (NOPs) with respect to the amino acid sequence of the

subtype B PR. These NOPs are found outside of the active site and, therefore, are not directly associated with inhibitor binding. However thermodynamic analyses of the efficacy of FDA-approved PIs against subtype C PR indicate that, in general, subtype C PR exhibits 3- to 24-fold weaker binding affinity for PIs when compared to subtype B PR (Velázquez-Campoy et al., 2003). These observed differences in binding affinity between PR subtypes strengthens the hypothesis that mutations in the hydrophobic core play a role in resistance in a manner that is likely due to novel conformational heterogeneity which is accessible to the mutated PR.

In the present study the mechanism of action of the I13V/I62V/V77I mutations in altered drug binding characteristics of the C-SA PR was elucidated. The nature of the effect of these mutations on the thermodynamic profile of several PIs (atazanavir, darunavir, saquinavir, ritonavir) was investigated via calorimetric analyses. Furthermore, the effect of these mutations on the structure and stability of the PR was elucidated via macromolecular crystallography and differential scanning calorimetry, respectively. A secondary aim of this study was to elucidate the effect of the D25A active site knock-out mutation on the stability and binding characteristics of the HIV PR.

Methodology

Inhibitors

Clinical inhibitors of HIV-1 protease, ATV, DRV, RTV and SQV were obtained from the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH. Stock solutions of the inhibitors (10 mM) were prepared in 100% (v/v) dimethyl-sulphoxide (DMSO). Secondary stock solutions (typically 200 µM) were prepared by dilution of the DMSO stock solution into 10 mM sodium acetate buffer (pH 5) and diluted as required with storage buffer. Due to limited solubility, the secondary stock solution of SQV was prepared to 100 µM.

Expression and Purification

Two pET-11a plasmids encoding for the wild-type (WT) PR and I13V/I62V/V77I PR (hereafter referred to as HP3 PR, meaning “the HIV PR variant with these 3 mutations”) genes, respectively, were purchased from GenScript (Hong Kong). A D25A point mutation was introduced into both sequences to remove the active site aspartic acid residue which eliminated the enzymatic activity of the PR and prevents autocatalysis (Darke et al., 1989). Elimination of the catalytic function of the PR is useful when studying the structure, thermodynamics and dynamics of the PR, without its catalytic function, since auto-catalysis is often a problem for experiments that require high concentrations of the enzyme (such as isothermal titration calorimetry and macromolecular crystallisation). The C-SA PR sequence

data were obtained from Prof Lynn Morris (AIDS Virus Research Unit, NICD, Johannesburg, South Africa). The vector encoding C-SA PR was available in our laboratory.

Escherichia coli BL21 (DE3) pLysS cells were transformed with the vectors encoding the PR variants using the heat-shock method. Thereafter, vector samples were obtained via Miniprep and DNA sequencing verified the coding sequences of the C-SA PR as well as the catalytically inactive WT PR (hereafter referred to as WT_{D25A} PR) and HP3 PR. Overexpression and purification were performed using a protocol previously established in our laboratory (Sherry et al., 2020). Overexpression of HIV PR using isopropyl-β-D-thiogalactopyranoside results in the formation of inclusion bodies (Ido et al., 1991). The cells containing the heterologously expressed PR were disrupted via sonication (30 sec pulse/30 sec interval, 10 times) using a Q-Sonica 125 sonicator. The PR was recovered from inclusion bodies by resuspending the pellet in 8 M urea, 10 mM Tris-HCl, 2 mM dithiothreitol (DTT) at pH 8. The soluble, unfolded PR was subsequently purified via ion-exchange chromatography using DEAE-CM columns coupled in tandem as described previously (Sherry et al., 2020). Since the theoretical *p*I of HIV PR is ~9.1 (Gasteiger et al., 2005) and the pH of the urea buffer is 8, the net surface charge of the unfolded PR will be positive and, therefore, binds to the negatively charged resin of the CM-Sepharose column. Thereafter, the unfolded PR was recovered from the CM column in a one-step elution using urea buffer containing 1 M NaCl. Refolding of the denatured PR was achieved by dialysing against 10 mM formic acid, 10% (v/v) glycerol (pH 3) followed by dialysis against 10 mM sodium acetate, 2 mM DTT (pH 5). The PR was further purified using CM-Sepharose ion-exchange chromatography, with elution via a 0-1 M NaCl gradient. The eluent was then dialysed into storage buffer comprised of 10 mM sodium acetate, 0.01% (w/v) sodium azide (pH 5). Purity was assessed via electrophoresis using 16% tricine-SDS-PAGE (Schägger and von Jagow, 1987). The concentration of the PR was determined using a molar extinction coefficient (ε) of 25 480 $M^{-1}.cm^{-1}$ from absorbance spectra obtained on a Jasco V-630 spectrophotometer using the Beer-Lambert law:

$$A = \varepsilon c l \qquad \text{[equation 1]}$$

Where A is the absorbance value at 280 nm, ε is the molar extinction coefficient ($M^{-1}.cm^{-1}$), c is the concentration (M) of the protein solution and l is the pathlength of the light passing through the sample (cm).

Active site titration

To assess the active site concentration of the purified PR samples, isothermal titration calorimetry (ITC) was performed using acetyl pepstatin, an aspartyl protease inhibitor, which binds relatively weakly to HIV PR. Since the WT_{D25A} and HP3 PR variants at the focus of this

study were catalytically inactive, this procedure was performed to determine to success of the refolding procedure as opposed to the degree of autocatalysis in the sample. The ITC experiments were performed at 20 °C using a Nano-ITC instrument (TA Instruments, Delaware, USA). Acetyl-pepstatin (200 μM) was titrated against 15 μM PR. The heat due to dilution of acetyl-pepstatin was subtracted from the data set and the baseline was adjusted using NanoAnalyze™. The change in heats were integrated and fitted using NanoAnalyze™. The active site concentration was determined from the calculated stoichiometry value (n) which is determined by the point of inflection of the resultant thermogram. Here, an 'n' value of 1 indicates a 100% active enzyme preparation. The n-value is, therefore, used as a correction factor for the concentration of active (or correctly folded) PR in a purified sample.

Enzyme activity determination

The specific activities of the purified PR variants were determined following the hydrolysis of a fluorogenic substrate (Abz-Arg-Val-Nle-Phe(NO_2)-Glu-Ala-Nle-NH_2) which mimics the conserved ^{360}ARVL/AEA^{366} sequence which comprises the capsid/SP1 cleavage site in the HIV-1 Gag polyprotein. The quenching amino group, Phe(NO_2), is held in close proximity to the fluorogenic 2-aminobenzoyl group (Abz), which reduces its fluorescence emission. Cleavage of the Nle-Phe(NO_2) peptide bond eliminates the quenching effect of the Phe(NO_2) group and allows the fluorophore to fluoresce (Carmel and Yaron, 1978; Szeltner and Polgár, 1996). An increase in fluorescence emission was detected at an excitation wavelength of 337 nm (2.5 nm bandwidth) and an emission wavelength of 425 nm (5 nm bandwidth). The experiments were performed using a substrate concentration of 50 μM and enzyme concentrations of 10-50 nM in a reaction buffer comprised of 1 M NaCl, 50 mM sodium acetate (pH 5.0). Measurements were taken for 1 min during the steady-state and the reaction velocity (μmol.min^{-1}) was plotted as a function of enzyme (mg). The intensity of the signal associated with complete cleavage of 1 nmol of substrate was used to convert intensity to activity.

Crystallisation and data collection

Crystallisation conditions resulting in optimal crystal growth were screened using the Hampton Research Index HR2-144 (Hampton Research). Crystals of the HP3 PR were grown at 20 °C using the sitting-drop vapour diffusion method on a 96-well microplate. Diffraction quality crystals were obtained using a reservoir buffer comprising 10 mM HEPES (pH 7.5) and 300 mM NaCl in 20% polyethylene glycol 6 000. The stock protein concentration used for crystallisation was 1.6 mg.mL^{-1} in 10 mM sodium acetate (pH 5.0) with 0.01% (w/v) sodium azide. Each hanging drop (2 μL or 4 μL) comprised equal volumes of reservoir buffer and stock protein. Single crystals were mounted on a cryoloop, and briefly soaked in cryoprotectant solution containing mother liquor with 50% (v/v) of polyethylene glycol 3 350, and frozen in

liquid nitrogen. X-ray diffraction data were collected on the I03 light source at the Diamond Light Source synchrotron (Oxfordshire, UK) using a Pixel Dectris Pilatus 36M detector. During data collection, the crystal was maintained at 100 K in a stream of nitrogen gas. Images were collected covering an oscillation angle of 0.2° per image at a wavelength of 0.9763 Å. The data set was processed using Xia2 software (Winter, 2010).

Model building and structure refinement

The phases of the PR structure were solved by the molecular replacement method using PHASER (McCoy et al., 2007). The C-SA PR was used as a search probe for molecular replacement (PDB ID: 3U71; Naicker et al., 2013). Model building was performed using Coot (Emsley and Cowtan, 2004); thereafter, cycles of global reciprocal space refinement using PHENIX (Liebschner et al., 2019) and local real space refinement using Coot (Emsley and Cowtan, 2004) were performed. Stereochemical validation of the HP3 PR target model was performed using MolProbity (Chen et al., 2010) and PROCHECK (Laskowski et al., 1993). The PyMOL Molecular Graphics System (Schrödinger LLC., Portland, USA) was used to generate images of the structure.

Differential scanning calorimetry

The heat capacity function of the WT_{D25A} and HP3 PRs was measured as a function of temperature using a highly precise Nano-DSC microcalorimeter (TA Instruments, Delaware, USA). The protein samples and reference solutions (buffer) were degassed to avoid bubble formation in the machine during sample loading and subsequent experimentation. The thermal denaturation profiles were performed on the apo form of the PRs (10-15 µM dimer) as well as in the presence of ATV and DRV at concentrations of 100 µM. The PR/PR-PI complexes were heated linearly from 20-100 °C at a rate of 1 °C.min^{-1} using storage buffer (10 mM sodium acetate, 0.01% (w/v) sodium azide, pH 5.0) in the reference cell. Data were analysed by non-linear least squares fitting models, as described previously (Sayer and Louis, 2009), as part of the NanoAnalyze™ software package.

Displacement titration calorimetry

Determining the energetics of ligands with extremely high binding affinities can be difficult when using the ITC technique as is the case with PIs which bind in the nM-pM range (Velázquez-Campoy et al., 2001). High affinity ligands are assessed via displacement titration in which a ligand that binds with lower affinity to the PR is displaced with the higher affinity ligand such as the PIs (Velázquez-Campoy et al., 2001). The most commonly utilised lower affinity binder for displacement titration experiments is the aspartyl protease inhibitor, acetyl pepstatin. As such, the PR samples (~15 µM) were incubated in 200 µM acetyl pepstatin for

30 minutes at 20 °C in order to equilibrate the mixture and to allow for the PR active sites to become saturated with acetyl pepstatin. Thereafter, displacement titration calorimetry was performed at 20 °C using PI concentrations of 80-200 µM. Displacement titrations were performed in 10 mM sodium acetate, pH 5, and 2% (v/v) DMSO using a Nano-ITC microcalorimeter (TA Instruments, Delaware, USA). After baseline correction for the dilution enthalpies of the PIs, the displacement binding energetics were determined using NanoAnalyse™ software. The data were fit using an independent binding model for a 1:1 ratio of active site to ligand.

Induced-fit ligand docking

Induced-fit docking (IFD) studies were performed on the South African subtype C PR (C-SA PR) and WT_{D25A} PR using Glide software (Schrödinger LLC., Portland, USA) implemented on Windows architecture. The effect of the mutations in the HP3 PR have been reviewed extensively (manuscript under review) and, thus, were not included in the docking procedure analysis. The WT_{D25A} PR model was generated with the mutagenesis plug-in present in the PyMOL Molecular Graphics System (Schrödinger LLC., Portland, USA), using the C-SA PR (PDB ID: 3U71) as a starting model. The WT_{D25A} PR model was validated using PROCHECK (Laskowski et al., 1993). The following Schrödinger modules were used for protein preparation: Glide, Prime and Protein Preparation Wizard. The bond orders of the PR models were corrected and hydrogen atoms were added to the models whereafter the ionisable residues were assigned charges in accordance with an experimental pH of 5.0 using the Epik prediction tool (Greenwood et al., 2010; Shelley et al., 2007). Thereafter, the models were subjected to energy minimisation until the average RMSD reached 0.3 Å.

The two-dimensional structure data files for the protease inhibitors to be used for IFD were obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov/). The PIs used in this docking study were: atazanavir (ATV, CID: 148192), darunavir (DRV, CID: 213039), ritonavir (RTV, CID: 392622) and saquinavir (SQV, CID: 60787). In order to prepare the two-dimensional structure data files of the PIs for the docking studies, the LigPrep tool was used (Schrödinger LLC, Portland, USA). The preparation involved the addition of hydrogens to the structures as well as the assignment of the correct bond lengths and angles before assigning the correct chirality. The Epik tool (Greenwood et al., 2010; Shelley et al., 2007) was used to select the tautomers and ring structures with the lowest energy. Thereafter, the PIs were subjected to energy minimisation.

A 20 Å cubic box surrounding the active site D25/A25 residues of the energy-minimised PR models was defined as the region where IFD would occur. Residues of the protein that were within 5 Å of the ligand poses were refined and subjected to side chain optimisation. A scaling

factor of 0.5 was used for the van der Waals radii of the protein and ligand during the docking procedure since Glide does not allow for receptor flexibility. Instead, Glide allows for the magnitude of van der Waals radii of non-polar residues to be scaled in a manner that decreases the penalty scores for close contacts. In this manner, the docking procedure affords some minor flexibility within the protein and ligand. The OPLS3 force field (Harder et al., 2016) was used when performing the IFD simulations which generated 20 conformational poses for each protein-inhibitor complex. The poses with the top IFD scores were selected for further analyses.

Results

Protease preparation

Recombinant expression of the C-SA PR, WT_{D25A} PR and HP3 PR yielded approximately 0.2, 2.36 and 1.32 mg of PR per litre of culture, respectively. The purity of the PR variants was estimated to be > 99% using tricine-SDS-PAGE. An active site titration of purified PR samples allows for the calculation of the percentage of active PR in solution. The stoichiometry values obtained from the thermograms for each titration indicated that 63% of the C-SA PR sample was active, while 58% and 95% of the WT_{D25A} PR and HP3 PR samples were correctly folded (Table S1). The distinction between “active preparation” and “correctly folded” lies in the fact that purified inactive PR cannot be defined by activity but the ability to bind acetyl pepstatin is still indicative of correctly folded active sites. The free energy associated with the binding of acetyl pepstatin to the PR variants was found to be the largest for the C-SA PR and lower for the PR variants containing the D25A active site mutation. Specifically, the free energies for these interactions were found to be -46.26, -33.64 and -27.02 $kJ.mol^{-1}$ for the C-SA, WT_{D25A} and HP3 PRs, respectively (Figure 1, Table S1).

The C-SA PR was shown to be enzymatically active by following hydrolysis of the fluorogenic substrate (Abz-Arg-Val-Nle-Phe(NO_2)-Glu-Ala-Nle-NH_2) while, in contrast, the presence of the D25A active site mutation was shown to eliminate enzymatic activity (Figure 2). The specific activities of the C-SA PR and WT_{D25A} PR were derived from the slope of the plots in Figure 2 and found to be 21.05 $\mu mol.min^{-1}.mg^{-1}$ and 0.03 $\mu mol.min^{-1}.mg^{-1}$ (where this value represents zero activity), respectively. These data illustrate that the purified PR samples were correctly refolded since all the variants possess the ability to bind the active site inhibitor acetyl pepstatin, while it was also evident that the D25A mutation eliminated catalytic activity of the PR.

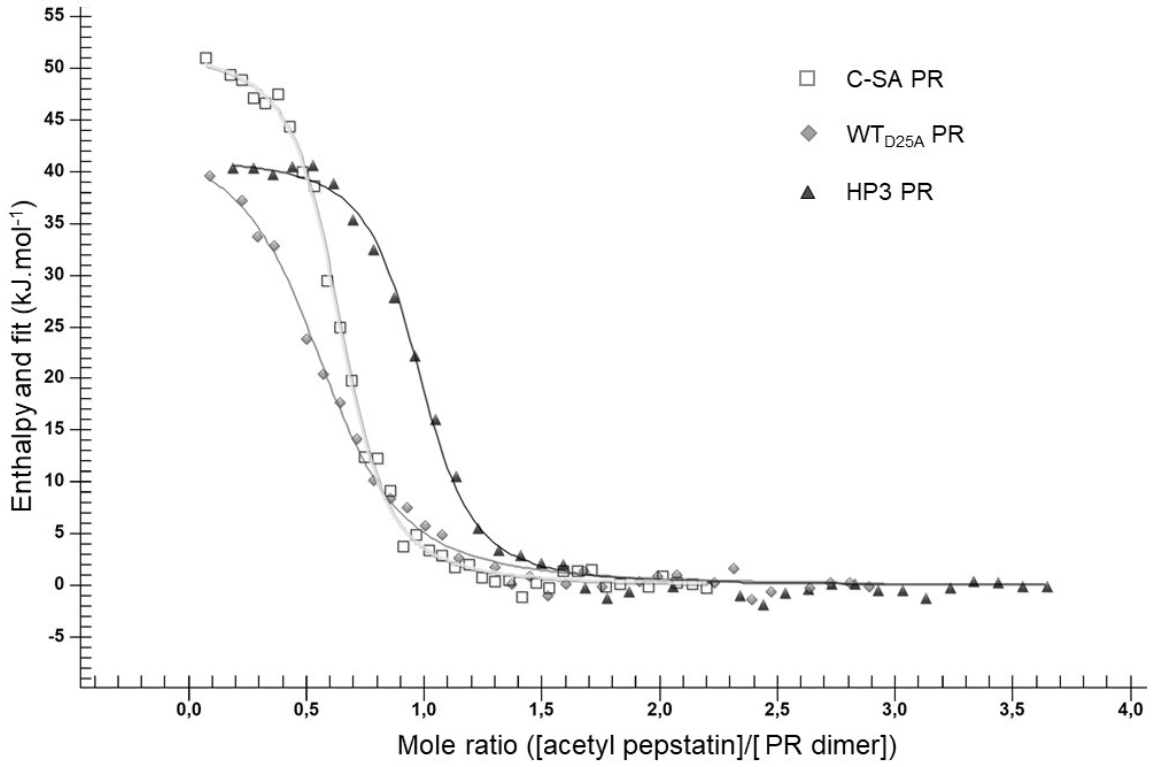


Figure 1. Fitted enthalpy data of the C-SA PR, WT_{D25A} PR and HP3 PR titrated with the aspartyl PR inhibitor, acetyl-pepstatin. The integrated heats for the heat evolved for each titration are plotted against the molar ratio of acetyl pepstatin to dimeric PR. The data were corrected for the heat of dilution of titrating acetyl pepstatin into buffer.

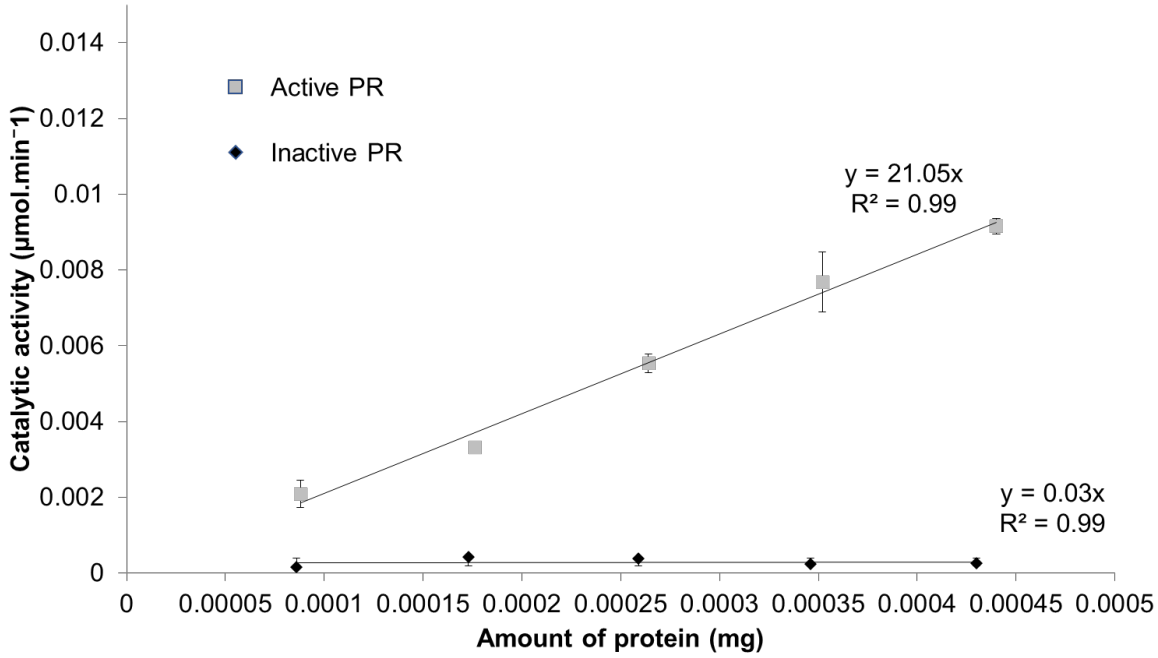


Figure 2. Specific activity determination comparison of enzymatically active PR vs inactive PR containing the D25A active site mutation. The specific activity is determined following the hydrolysis of the fluorogenic HIV-1 PR substrate (Abz-Arg-Val-Nle-Phe(NO_2)-Glu-Ala-Nle-NH_2) in 50 mM sodium acetate and 1 M NaCl (pH 5.0) at 20 °C.

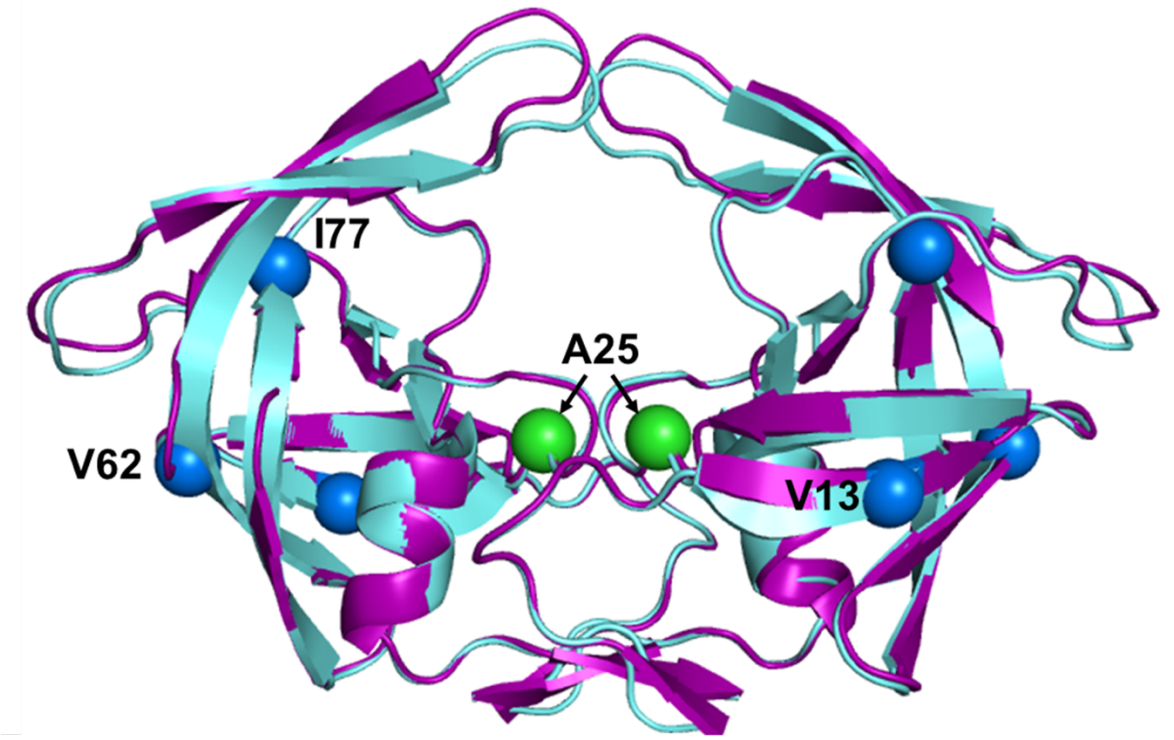
Crystal structure of the HP3 protease

The subtype C of HIV is the most prevalent subtype worldwide which emphasises the importance of increasing our understanding of the structure and function of this subtype. Analysis of the structural homology of the HP3 PR and the C-SA PR (PDB ID: 3U71, Naicker et al., 2013) revealed that the overall quaternary structure of the PR was not significantly affected by the presence of the non-active site mutations: I13V, I62V and V77I (Figure 3). The coordinates for the HP3 PR X-ray crystal structure resolved at a resolution of 1.8 Å were deposited to the Protein Data Bank (PDB ID: 6I45). The structure was shown to have no stereochemical errors such as disallowed bond angles or rotamers (Table 1). Furthermore, this space group and unit cell parameters of the structure were the same as those observed previously in the C-SA PR and unbound HIV PR crystals (Heaslet et al., 2007; Naicker et al., 2013; Yedidi et al., 2011). The solvent content and Matthews coefficient (V_M) of the HP3 PR crystal are highly similar to that of the C-SA PR which indicates that the asymmetric unit is the PR monomer (Naicker et al., 2013).

Table 1. Data collection and refinement statistics for the HP3 PR structure.

Data collection and processing statistics	
X-ray source	Synchrotron (I03, DLS Ltd., UK)
Wavelength (Å)	0.9763
Temperature (K)	100
Oscillation angel (◦)	0.2
Crystal to detector distance (mm)	313.25
Space group	P 4_1 2_1 2
Unit cell parameters (Å)	a = b = 45.58 & c = 105.01
Resolution range (Å)	35.00-1.80 (1.83-1.80)*
No. of unique reflections	10931 (547)*
R_{merge}	0.147 (2.512)*
Completeness (%)	100 (100)*
I/σ(I)	9.52 (0.94)*
Matthews coefficient (V_M) (Å^3/Dalton)	2.55
Solvent content (%)	51.70
Asymmetric unit content	Monomer
Structure Refinement	
Resolution limit (Å)	1.80-34.42
No of reflections	10877
Completeness (%)	99.84
R-factor / R_{free} (%)	20.85 / 25.71
No. of protein / ligand atoms	750 / 0
RMSD from standard geometry	
Bond-length (Å)	0.008
Bond-angles (˚)	0.889
Chiral volume (Å^3)	0.064
Average *B* value (Å^2)	34.25
Ramachandran statistics	
Residues in favourable regions (%)	98
Residues in allowed regions (%)	2
Residues in outliers regions (%)	0
PDB ID	6I45

*Values in parenthesis show the statistics of the highest resolution shell.




```
C-SA PR    PQITLWKRPLVSIKVGGQIKEALLDTGADDTVLEEINLPGKWKPKMIGGIGGFIKVRQYD  60
HP3  PR    PQITLWKRPLVSVKVGGQIKEALLATGADDTVLEEINLPGKWKPKMIGGIGGFIKVRQYD  60

C-SA PR    QILIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNF      99
HP3  PR    QVLIEICGKKAIGTVLIGPTPVNIIGRNMLTQLGCTLNF      99
```

Figure 3. Structural alignment of ribbon representations of the C-SA PR (cyan, PDB ID: 3U71), and the HP3 PR (purple, PDB ID: 6I45). The position of the three non-active site mutations that HP3 PR carries with respect to C-SA PR are indicated as blue spheres and labeled. Similarly, the mutations of the catalytic aspartic acid residues to alanine residues, which are present in the HP3 PR, are indicated by green spheres. Sequence alignment of the C-SA PR and HP3 PR with amino acid variations highlighted for clarity.

Isothermal titration calorimetry

The effect of the D25A active site mutation on the energetics of binding was examined by comparing the change in Gibbs free energy (Δ*G*) of the interactions between the selected PIs and the WT_{D25A} PR and C-SA PR. It is evident that the change in free energy is negative in all of the binding interactions studied (Figure 4, 5) which indicates that these interactions are energetically favourable. However, it is noteworthy that Δ*G* decreases in magnitude for all of the PIs examined when binding to the WT_{D25A} PR in comparison to the C-SA PR (Figure 4). Specifically, the ΔΔ*G* was approximately 12, 17, 7 and 10 $kJ.mol^{-1}$ for ATV, DRV, SQV and RTV, respectively.

The binding energies of the WT_{D25A} and HP3 PR variants to the selected PIs indicate that the enthalpy change (Δ*H*) upon binding was similar in magnitude for ATV, DRV and RTV (Figure 5). However, the Δ*H* was significantly smaller for both PR variants when binding to SQV. It is noteworthy that the ΔΔ*H* showed a decrease in magnitude for SQV and RTV when binding to HP3 PR in comparison to WT_{D25A} PR of 4.38 and 6.46 $kJ.mol^{-1}$, respectively (Figure 5). In contrast, both ATV and DRV exhibited an increase in the magnitude of the change in enthalpy when binding to HP3 PR in comparison to WT_{D25A} PR. The effect that the mutations in HP3 PR have amounted to a ΔΔ*H* gain of 4.69 and 24.51 $kJ.mol^{-1}$ for ATV and DRV, respectively and ΔΔ*H* loss of 6.46 and 4.38 $kJ.mol^{-1}$ for RTV and SQV, respectively. The change in entropy of binding between SQV and RTV exhibited minor decreases in magnitude when binding to the HP3 PR in comparison to the WT_{D25A} PR (Figure 5). Interestingly, however, the entropy changes upon binding of ATV and DRV showed significant differences between the WT_{D25A} PR and the HP3 PR. Specifically, the effect of the mutations on ΔΔ*S* amounted to a decrease in entropy of 14.82, 31.32, 3.37 $kJ.mol^{-1}$ for ATV, DRV, RTV, respectively and an increase of 0.38 $kJ.mol^{-1}$ for SQV. Since the Δ*G* of the systems must be negative for the protein-ligand interaction to be favourable it is notable that the Δ*G* decreases in magnitude for the inhibitors when binding to the HP3 PR in comparison to the WT_{D25A} PR. In the case of ATV and DRV the observed changes between the HP3 PR and WT_{D25A} PR systems with regard to ΔΔ*S* are significant. These changes indicate that the binding of ATV and DRV to the HP3 PR are less entropically favourable.

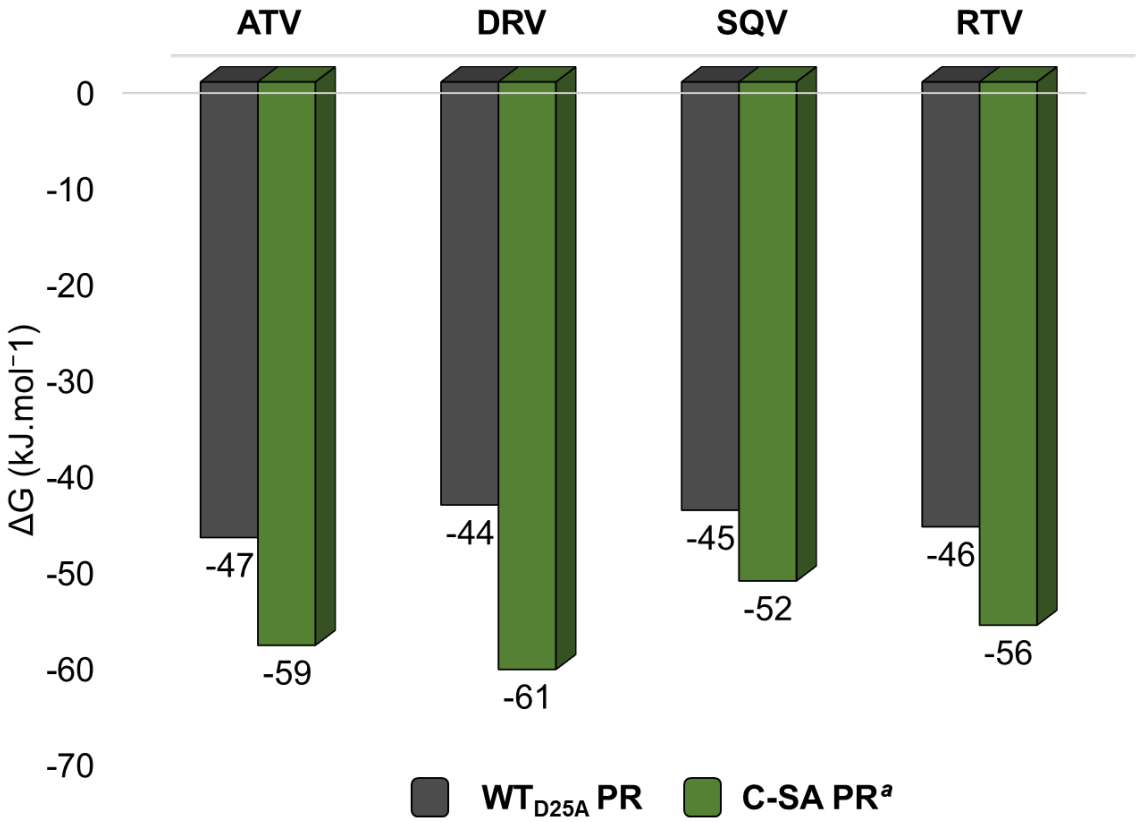


Figure 4. Gibbs free energy of binding for the interactions between WT_{D25A} PR and C-SA PR and four PIs. [a] *The interaction between ATV, DRV, SQV and RTV and C-SA PR as reported by (Mosebi et al., 2008; Naicker et al., 2014).*

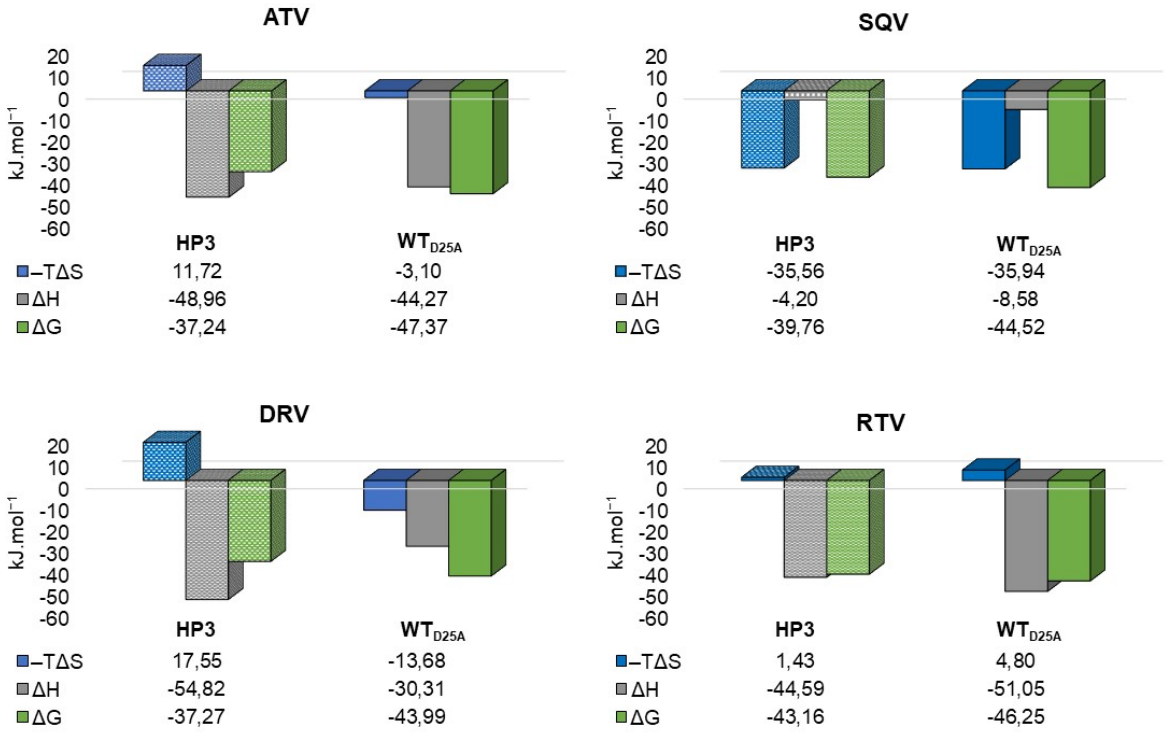


Figure 5. Summary of the thermodynamic parameters obtained for the binding of ATV, DRV, SQV and RTV to WT_{D25A} PR and HP3 PR.

Differential scanning calorimetry (DSC)

DSC was used to measure the heat capacity of the PR solutions as a function of temperature. In this manner, the structural stabilities of the WT_{D25A} PR and HP3 PR samples were compared in the absence of ligand as well as in the presence of ATV, DRV, RTV and SQV. As is evident from Figure 6, the stabilities of the two native PR variants were comparable, with T_m values of 66.8 °C and 67.5 °C for the HP3 PR and WT_{D25A} PR, respectively. For WT_{D25A} PR in the presence of DRV a biphasic transition was observed (Figure 6) which resulted in two distinct transition peaks for this complex. The melting curve for WT_{D25A} PR·DRV produced two peaks with an area ratio of approximately 60/40 and transition temperatures of 76.7 °C and 84.9 °C, respectively. The melting curves for WT_{D25A} in complex with ATV, RTV and SQV exhibited single transition peaks corresponding to T_m 73.6 °C, 80.4 °C and 65.3 °C, respectively. The melting curves for HP3 PR exhibited singular transition temperatures with peaks corresponding to 66.9 °C, 80.7 °C, 74.2 °C and 66.3 °C in the presence of ATV, DRV, RTV and SQV, respectively (Figure 6).

Ligand docking

The induced fit ligand docking studies revealed changes in the calculated binding parameters for the C-SA PR and WT_{D25A} PR, summarised in Table 2. Notably, the G_{score} values decreased in magnitude by 1.16, 1.22, 4.37 for ATV, RTV and SQV, respectively, for the WT_{D25A} PR in comparison to the C-SA PR. However, in contrast, the G_{score} increased in magnitude by 1.6 when DRV was docked to the WT_{D25A} PR. These changes in the G_{score} correlate with similar changes in the $Glide_{hbond}$. The $Glide_{hbond}$ parameter provides an estimation of the contribution of hydrogen bonds to the interaction between the protein and ligand where a more negative value is indicative of a larger degree of hydrogen bonding. Furthermore, the $Glide_{ecoul}$ parameter decreased in magnitude for ATV, DRV, SQV, and increased for RTV when docked to the WT_{D25A} PR.

Table 2. Summarised induced fit docking parameters for C-SA PR and WT_{D25A} PR with clinical inhibitors.

PR variant	PI	Glide parameter					
		G_{score}	ligand efficiency SA	lipo	hbond	evdw	ecoul
C-SA	ATV	-9.87	-0.70	-4.01	-0.71	-58.38	-15.78
	DRV	-6.35	-0.56	-2.29	-0.18	-40.80	-12.99
	RTV	-10.20	-0.75	-5.31	-0.46	-65.17	-4.99
	SQV	-14.15	-1.06	-5.86	-1.52	-74.09	-17.88
WT_{D25A}	ATV	-8.71	-0.59	-3.97	-0.37	-60.75	-8.42
	DRV	-7.95	-0.70	-3.87	-0.46	-47.16	-8.03
	RTV	-8.98	-0.47	-3.98	-0.38	-61.84	-9.35
	SQV	-9.78	-0.73	-5.54	-0.64	-56.69	-4.47

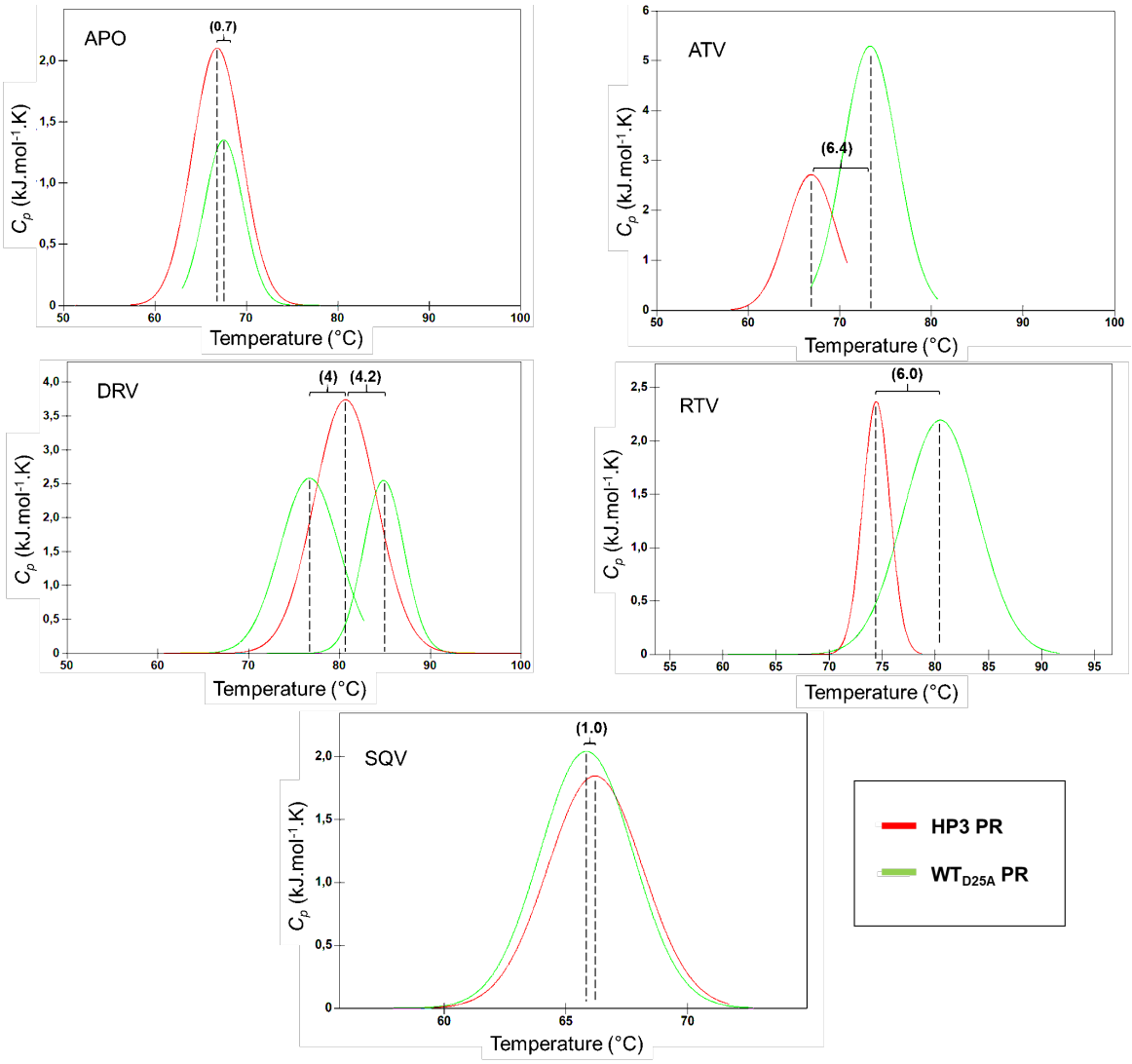


Figure 6. DSC thermograms of dimeric WT_{D25A} PR (green) and HP3 PR (dark red) in their apo forms (APO), and in the presence of ATV, DRV, RTV and SQV. These data have been baseline corrected and normalised to molar heat capacity using the PR dimer concentrations. The change in transition temperature (ΔT_m) of each protein/protein-inhibitor complex is indicated in brackets.

Discussion

The effect of the D25A active site mutation on HIV PR

In order to understand the effect that the D25A active site mutation has on the structural and functional characteristics of the HIV PR, the WT_{D25A} PR was compared with the C-SA PR. First, the lack of catalytic activity of the WT_{D25A} PR was determined via a specific activity assay; following the cleavage of a fluorogenic substrate. From Figure 2 it is evident that the D25A mutation successfully eliminated the catalytic activity of WT_{D25A} PR. The C-SA PR shows an increase in the fluorogenic signal produced as a function of the amount of protein (Figure 2) while, in comparison, the WT_{D25A} PR produces no fluorescent signal at any protein concentration which is illustrated by the unchanging gradient of the graph (m = 0.03). This is because the quenching Phe(NO_2) group is held in close proximity to the fluorogenic 2 aminobenzoyl group (Abz), resulting in fluorescence quenching. The cleavage of the Nle-Phe(NO_2) scissile peptide bond eliminates its quenching effect and allows fluorescence to be detected at 425 nm (Carmel and Yaron, 1978). In the absence of the Asp25 residue, the WT_{D25A} PR is unable to cleave the substrate and produce a fluorescent signal. These data therefore indicate a decrease of ~99.9% of the specific activity of the HIV PR due to the D25A mutation.

Overexpression of the WT_{D25A} PR and C-SA PR revealed that the yield of the inactive PR was approximately 10-fold higher per litre of expression culture while the active site titrations revealed that the WT_{D25A} PR and C-SA PR had approximately equal stoichiometry values (~60%) (Table S1). Together these data indicate that the catalytically inactive PR expresses readily in comparison to the C-SA PR. The HIV PR is notoriously difficult to obtain in sufficient quantities for *in vitro* analyses due to the cytotoxic effect that it exhibits as a result of recombinant expression (Herber et al., 1991). Therefore, the benefit of the D25A mutation when working with HIV PR *in vitro* lies in the nullification of its cytotoxic effects which provides greater potential for protein yield per volume of expression culture. Furthermore, purified catalytically active PR is liable to undergo degradation due to the propensity for the PR to perform autocatalysis. This is especially problematic when performing experiments that require high concentrations of the protease for analyses such as ITC, HDX-MS and macromolecular crystallisation. Therefore, the inactive PR would likely prove useful for such experimental procedures without the concern of autocatalysis resulting in deterioration of the sample quality.

While the benefits of the D25A mutation are evident in terms of protein yield and experimental concerns while working with the PR *in vitro*, alterations in the active site cavity are likely to change the binding affinity of PIs. Therefore, the effect of this mutation needs to be understood

before drawing meaningful conclusions based on experimental data using inactive PR variants. The effect of the D25A mutation on the stability of the homodimeric PR was analysed by scanning calorimetry (Figure 6). The T_m of the WT_{D25A} PR was marginally higher than that of the C-SA PR (~1.5 °C) which exhibit maximal transition peaks at 67.5 °C and 66 °C (Velázquez-Campoy et al., 2001), respectively. As such it was concluded that the D25A mutation did not have a significant effect on the thermal stability of the PR dimer. However, the D25A mutation did appear to impact the efficacy of ligand binding as is summarised in Table 3. Here, a consistent decrease in the magnitude of Δ*G* was observed in the WT_{D25A} PR when compared to the C-SA PR. This decrease in the change in free energy suggests that the favourability of ligand binding is affected by the absence of the Asp25 residues. These findings are echoed by the increase in the dissociation constant (K_d) of the inhibitors when binding to the WT_{D25A} PR (Table 3). The K_d values give an indication of the binding affinity of the ligands to the PRs and it is therefore noteworthy that the K_d values of the PIs are 10-1000 larger for the WT_{D25A} complexes. Larger K_d values indicate less tight binding of the ligand to the protein and, thus, these data indicate that the D25A mutation significantly decreases the favourability of PI binding. Similarly, with the exception of DRV, the induced-fit docking data indicate that the PIs bind to the WT_{D25A} PR less favourably as indicated by less negative G_{score} values in comparison to C-SA PR.

Table 3. Comparison of ligand binding parameters between the WT_{D25A} PR and the C-SA PR.

		Ligand				
Parameter	PR	Ac-PEP	ATV	DRV	SQV	RTV
Δ*G* (kJ.mol^{-1})						
	C-SA[a]	-46.26	-58.58	-61.09	-51.88	-56.48
	WT_{D25A}	-33.64	-47.37	-43.99	-44.52	-46.25
K_d (nM)						
	C-SA	430	0.036	0.012	0.84	0.083
	WT_{D25A}	1010	3.6	14	11	5.7
G_{score}						
	C-SA	-	-9.87	-6.35	-14.15	-10.2
	WT_{D25A}	-	-8.71	-7.95	-9.78	-8.98

[a] *(Mosebi et al., 2008; Naicker et al., 2014).*

The importance of the Asp25 residue in the functionality of the HIV PR has been described previously (Ido et al., 1991). Briefly, it is suggested that the carboxyl groups of two Asp25 residues in the active site exist in opposing states of protonation (COO^-/COOH). The opposing states of protonation of the Asp25 residues allow for the coordination of an essential water molecule into the active site which is then able to perform a nucleophilic attack on the scissile peptide bond of the PR substrates (Ido et al., 1991). Similarly, the Asp25 residues in the active

site of the PR form hydrogen bonds with PIs. The side chain of alanine is significantly smaller than that of the aspartic acid which brings alanine further away from substrates and inhibitors during binding (Figure 7). In addition, the side chain of the alanine side chain does not possess the ability to participate in hydrogen bonding due to the lack of an electronegative atom. Therefore, the active site mutation (D25A) changes the shape (Figure 7) and charge state of the active site cavity of the PR which is likely to play a significant role in altered binding affinity of substrates and inhibitors.

The reasons for the change observed in the binding affinity for the WT_{D25A} PR may be further elucidated by the induced fit docking study data. The three-dimensional docked structures indicate that the PIs occupy altered orientations within the active site cavity in the absence of the Asp25 residue (Figure 8). This is likely due to the change in the charge, size and shape of the active site cavity which causes the PR·PI complexes to sample different conformations and binding orientations in order to find their energy minima. For instance, the strength of hydrogen bonding ($Glide_{hbond}$) decreases in the absence of the Asp25 residue, with the exception of the DRV-bound WT_{D25A} PR complex which increased from -0.18 to -0.46 (Table 2). We hypothesised that the increased hydrogen bonding calculated for this complex is likely due to the drug forming novel contacts with the flaps of the PR when occupying an altered binding orientation. Similarly, significant changes occur in the calculated electrostatic forces during binding ($Glide_{ecoul}$). The removal of the electronegative oxygen atoms which form part of the carboxyl side chain of Asp25 result in a decrease in the electrostatic forces present in the active site cavity of the PR. The contribution of electrostatic interactions decreased by 7.36, 4.96 and 13.41 for the ATV, DRV and SQV complexes, respectively, in the presence of the mutations (Table 2). Notably, the electrostatic contribution increased by 4.36 in the RTV-bound HP3 PR complex in comparison to the WT_{D25A} PR complex (Table 2).

As such, the benefit of eliminating the catalytic activity of the PR during *in vitro* studies are apparent with respect to protein yield and handling. However, as is evident in the current study, the D25A mutation affects the efficacy with which ligands bind to the active site of the PR relative to the WT. Therefore, in order to study PR variants that have been catalytically nullified via the D25A mutation the experimental data must be compared to the corresponding inactive WT PR, as is the case in the current study. Data such as these may prove useful in the analyses of mutations in the PR sequence in isolation; without the associated cleavage reaction or autocatalysis and may provide insight into the mechanism of action of such mutations. However, since the D25A mutation is not naturally occurring the effect of this mutation on ligand binding could be a limiting factor in extrapolation of data into understanding clinical isolates.

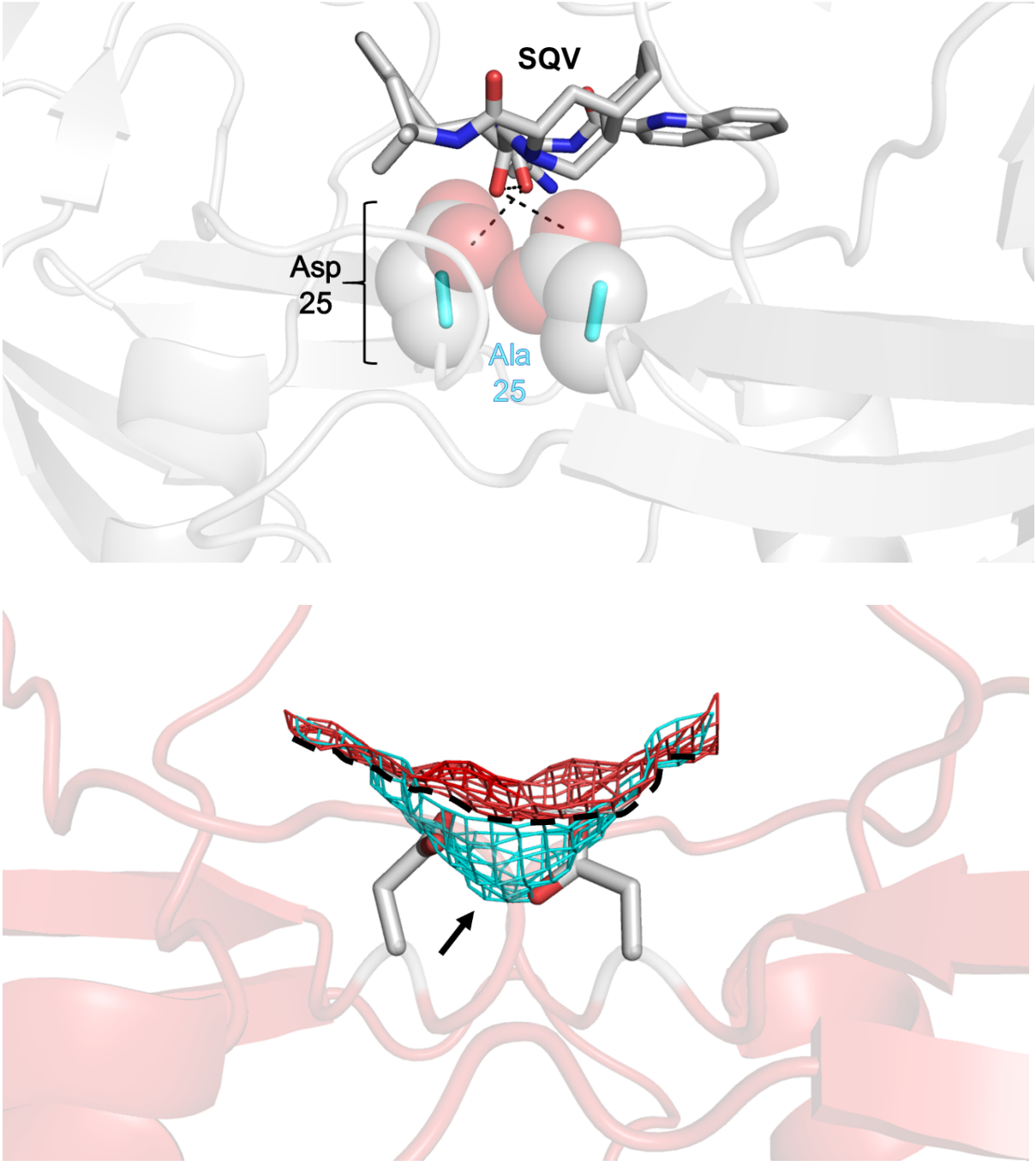


Figure 7. The active site of C-SA PR and WT_{D25A} PR. The oxygen atoms of the carboxyl group of the Asp side chain assist with coordinating SQV into the active site through the formation of hydrogen bonds (upper panel, black dotted lines). The D25A mutation (cyan) illustrates the difference in size of the side chains and the lack of hydrogen bond formation when Ala25 is present. The change in the shape of the active site due to the D25A mutation is illustrated by the different contours of the surface represented as a mesh (lower panel, black arrow).

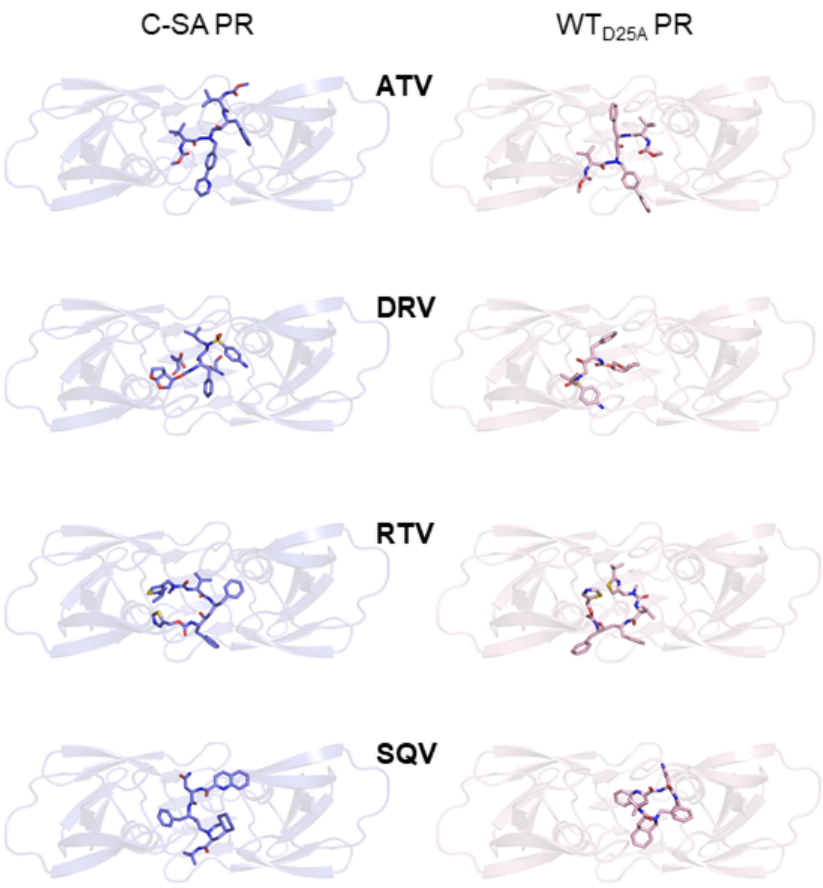


Figure 8. Binding geometries of clinical inhibitors to C-SA PR and WT_{D25A} PR. The PIs exhibit altered binding orientations within the active site of the PR when the Asp25 residue is replaced with an alanine.

The effect of the I13V/I62V/V77I mutations on the HIV-1 PR

Alterations in the dimer stability

The DSC technique measures the magnitude of the forces that stabilise the native structure of proteins (Todd et al., 1998). Additionally, the relative stabilities of native proteins and their ligand-bound counterparts change as a result of differential interactions between polar and non-polar atoms of the proteins and ligands, in addition to the interactions of these atoms with the surrounding solvent (Connelly, 1994). If a ligand binds to a protein with high affinity, the heat capacity of the protein-ligand complex would be different from that of the ligand or protein alone. As a consequence of this, the DSC profiles of the apo protein as well as the protein in complex with different ligands will have distinct shapes depending on the effect the ligand poses on the structural stability of the complex. Indeed, dramatic deformations of the thermograms observed from ligated proteins have been reported (Lin et al., 1994; Shrake and Ross, 1990; Takahashi and Fukada, 1985). As such the impact of the mutations present in the HP3 PR on the stability of the apo and ligand-complexed state were analysed. In the apo state, the melting temperature of the HP3 PR was very similar to that of the WT_{D25A} PR with the HP3 PR showing a slight decrease in stability of 0.7 °C.

The mutations in the HP3 PR occur in the hydrophobic core and result in changes in the total hydrophobic surface area and side chain packing of the PR which has the potential to alter the number and strength of interactions that stabilise the native conformation of the PR. In the case of I13V and I62V, the mutation from a larger isoleucine side chain to a smaller valine would decrease the hydrophobic surface area whereas the V77I mutation would increases the internal hydrophobic surface area. The melting curve data for the apo PRs illustrate that these mutations present are not sufficient to elicit a significant change in the overall stability of the apo PR. Binding of PIs to the PR brings the complex into a more energetically stable state which is exhibited by the increase in the T_m values observed between the apo and inhibitor-complexed PRs. The WT_{D25A} PR exhibited a biphasic transition when melting in the presence of DRV (Figure 6). This biphasic transition has been observed in similar calorimetry experiments and has been hypothesised to occur due to the different orientations that these drugs can occupy in the active site cavity (Sayer et al., 2008). It is noteworthy that in the case of ATV, DRV and RTV the thermal stability of the HP3 PR complexes is lower than that of the major transition peak of the corresponding WT_{D25A} PR complexes (Table 4). Interestingly, however, the HP3 PR·SQV complex exhibited a transition peak that was ~1 °C more stable than the corresponding WT_{D25A} PR complex (Table 4).

Table 4. Comparison of transition temperatures (T_m) data for the apo HP3 PR and WT_{D25A} PR as well as in complex with four clinical protease inhibitors.

	T_m		
	HP3 PR	WT_{D25A} PR	Δ
Apo	66.8	67.5	+0.7
ATV	67.2	73.6	+6.4
DRV	80.7	76.7/84.9	-4/+4.2
RTV	74.2	80.4	+6
SQV	66.3	65.3	-1

The lower T_m values for these complexes indicate that the mutations present in the HP3 PR decrease the stability of the drug-bound complex. The implications for the decrease in the stability of these complexes and their mechanism of action with respect to inhibitor resistance may lie in the ease with which the PIs dissociate from the mutated PR. Drug binding is a dynamic process which can be defined by the following dynamic equilibrium equation:

$$k_{on}[PR][PI] = k_{off}[PR \cdot PI] \quad \text{[equation 2]}$$

And the dissociation constant is given by:

$$K_d = \frac{[PR][PI]}{[PR \cdot PI]} = \frac{k_{off}}{k_{on}} \quad \text{[equation 3]}$$

As such, the role that these mutations play in drug binding may lie in this dynamic equilibrium. If the stability of the inhibitor-bound complex is lowered in the presence of the non-active site mutations, the energetic penalty associated with the structural rearrangements necessary for displacing the PI from the active site would, in turn, be lowered. This could result in a greater off-rate of the inhibitor leaving the active site of the PR. An increase in k_{off} would result in an increase in the dissociation constant thereby shifting the equilibrium in favour of the unliganded protein and, therefore, in favour of resistance.

Variations in the thermodynamics of PI binding

Analysing the thermodynamic parameters can provide insight into the changes in binding characteristics as a result of the mutations present in HP3 PR. The favourability of the thermodynamics of ligand binding relies upon two essential components: namely, the change in enthalpy (Δ*H*) and entropy (Δ*S*) of the system. Binding enthalpy represents the breaking and/or formation of bonds, such as hydrogen bonds, van der Waals interactions and various electrostatic interactions. A negative ΔH corresponds to the formation of these types of bonds and which contribute to the favourability of binding and can be an indication of protein-ligand specific interactions (Chen et al., 2016). On the other hand, the change in entropy of the

system represents the change in the dynamics or disorder within the system (Perozzo et al., 2004). Ligand binding results in an unfavourable loss of conformational entropy in the PR·PI system due to the fact that the binding partners can no longer sample different rotameric states. This loss of conformational entropy is compensated for by the increase in the entropy of the surrounding solvent which occurs due to desolvation of the hydrophobic surface of the PIs upon binding to the active site. Indeed, a major driving force for the favourability of ligand binding is due to the increase in solvation entropy (S_{solv}) as a result of the hydrophobic effect (Perozzo et al., 2004). Since the change in free energy of the system needs to be negative for a protein-ligand interaction to be favourable, it is often observed that changes in enthalpy and entropy are complementary in nature, a phenomenon termed enthalpy-entropy compensation (Chodera and Mobley, 2013).

The entropy change was positive, and therefore favourable, when the PIs bound to WT_{D25A} PR, with the exception of RTV which was slightly negative. Therefore, the positive ΔS_{bind} upon ligand binding contributed favourably to the free energy of the interactions involving ATV, DRV and SQV. A negative ΔH was observed for the binding of all four of the inhibitors to the WT_{D25A} PR and formed the majority of the binding energy for ATV, DRV and RTV while the binding of SQV was entropically driven. The effect of the mutations present in the HP3 PR resulted in a decrease in the ΔG for all of the inhibitors tested when compared to the WT_{D25A} PR which was indicative of decreased binding strength. Moreover, the binding affinity for the HP3 PR for ATV and DRV was reduced approximately 64- and 16-fold, respectively (Table 5), in comparison to the WT_{D25A} PR. The binding affinity for the HP3 PR for SQV and RTV bound to the mutant PR with approximately 1.5- and 3-fold higher binding affinity, respectively (Table 5).

Table 5. Comparison of binding affinity (K_d) data for the HP3 PR and WT_{D25A} PR and four clinical protease inhibitors.

	K_d (nM)		
	HP3 PR	WT_{D25A} PR	Fold change
ATV	230	3.6	63.88
DRV	230.00	14.50	15.86
SQV	7.78	11.7	0.66
RTV	2	5.7	0.35

The binding of ATV and DRV were enthalpically driven for both PR variants; however, binding to the HP3 PR displayed an unfavourable entropy change ($-T\Delta S > 0$). These data suggest that the mutations alter the PR in a manner that decreases the entropic favourability of the binding mechanism of ATV, DRV and RTV to the HP3 PR. On the other hand, the magnitude of ΔS remained unchanged when SQV bound to the HP3 PR when compared to the WT_{D25A}

PR, while there was a marginal decrease in the Δ*H* which contributed to the binding affinity loss. RTV appeared to be the least affected by the mutations, with minimal changes observed in Δ*G*, which correlated with a slight decrease in the Δ*H* term. It has been suggested that distortions induced by mutations that are distal from the active site may decrease the binding efficacy of clinical inhibitors (Hong et al., 2000; Mosebi et al., 2008).

The resistance mechanism of cooperative secondary mutations

Mutations outside of the active site of the HIV PR are thought to be largely compensatory in their nature; working in a manner that offsets the potential deleterious effects of active site mutations. These secondary mutations contribute to an increase in viral fitness in the presence of active site mutations which have been shown to affect processing of the PR substrates. However, some secondary mutations have been shown to correlate with resistance to PIs, but their mechanism of action has yet to be fully clarified. Multi-drug resistant PR variants are known to rely on the accumulation of numerous non-active site mutations with only one or two active site mutations in order to achieve suitable levels of resistance to the most potent PIs (Johnson et al., 2003; Weber et al., 2015). Furthermore, there are non-active site mutations which appear to occur in distinct combinations which suggest cooperativity in their nature. For example, the L76V mutation substantially increases lopinavir resistance but only when found alongside “partner substitutions”; namely, M46I, I54V, V82A and L90M (Young et al., 2010). Similarly, the mutations present in the HP3 PR (I13V/I62V/V77I) have been correlated with decreased drug susceptibility and have been shown to exhibit cooperative movements during molecular dynamics simulations (Foulkes-Murzycki et al., 2007).

The network of backbone hydrogen bonds is largely conserved during large-scale motions of the PR while, in contrast, van der Waals contacts readily swap between hydrophobic residues in the core of the PR (Foulkes-Murzycki et al., 2007; Ragland et al., 2014). This transferal of van der Waals contacts allows for hydrophobic regions inside the PR to slide across one another in a manner which facilitates the conformational changes associated with flap opening. Therefore, the location of the mutations present in the HP3 PR may provide insight into the nature of their function in decreasing inhibitor susceptibility. Conservative hydrophobic mutations in the core of the PR could alter the conformational range of the PR depending on the nature of the mutations. Sets of mutations, such as I13V/I62V/V77I, would alter the total hydrophobic surface area inside the protein, which may lead to a defective hydrophobic sliding mechanism and/or a more compact and stable core. A more compact core may complement the flexibility of the flap regions in a manner that allows the PR to sustain wider flap conformations. This mechanism has been further elucidated via molecular dynamics simulations (Sherry et al., 2021), where the HP3 PR sustains wider flap conformations when in complex with ATV, DRV and SQV. The wider flap conformations result in an expanded

active site cavity which brings the inhibitors “up and away” from the Asp25 residues, preventing the formation of stabilising intermolecular hydrogen bonds. The location of these mutations seems to be of importance to their mechanism of action. Each of the mutations present in the HP3 PR occur at the interface between defined regions of the hydrophobic core (Figure 9). The I13V mutation occurs on the internal hydrophobic surface of the fulcrum region and makes van der Waals contacts with several residues of the cantilever (Figure 9A) while the I62V mutation occurs at the interface formed between the fulcrum, cantilever and hinge regions (Figure 9B). Similarly, the V77I mutation is found at the interface formed between the cantilever and the hinge region and interacts with several hinge residues (Figure 9C).

The flexibility of the flap regions allows the PR to establish a dynamic equilibrium between the open, semi-open and closed flap conformations. Non-active site mutations may accumulate in a manner that shifts this equilibrium that diminishes drug binding through sustaining wider flap conformations, for example. The rigid nature of PIs renders their efficacy highly susceptible to alterations in the geometry of the active site while the PR substrates, on the other hand, are more flexible in nature and can accommodate such changes more readily. Conservative mutations in the hydrophobic core which change the size or shape of the amino acid side chain may affect the atomic packing of residues in the hydrophobic core of the native structure of the PR. Since the hydrophobic sliding mechanism plays a key role during flap movement and relies on internal hydrophobic residues readily sliding over one another (Foulkes-Murzycki et al., 2007), this dynamic process is precisely controlled and, therefore, depends largely on the internal hydrophobic surface area of the PR as well as the quantity of van der Waals contacts involved. As such, mutations in the hydrophobic core that change these qualities are likely to affect the movement of the flaps. The role that non-active site mutations have on drug susceptibility could, therefore, be linked to the dynamic binding equilibrium of the PR/PR·PI complexes rather than through directly imposing steric or geometric irregularities as is the case with active site mutations. Therefore, combinations of non-active site mutations could be additive in nature, affecting the ease by which inhibitors may bind to the active site of protease through shifting the dynamic equilibrium of PR flap conformations. It can be hypothesised that such mutations contribute to drug resistance when the equilibrium between substrate binding and catalysis, a dynamic process, and inhibition (a restriction of dynamics) is altered in a manner that opposes drug binding through modulating the dynamic ensemble sampled by the PR.

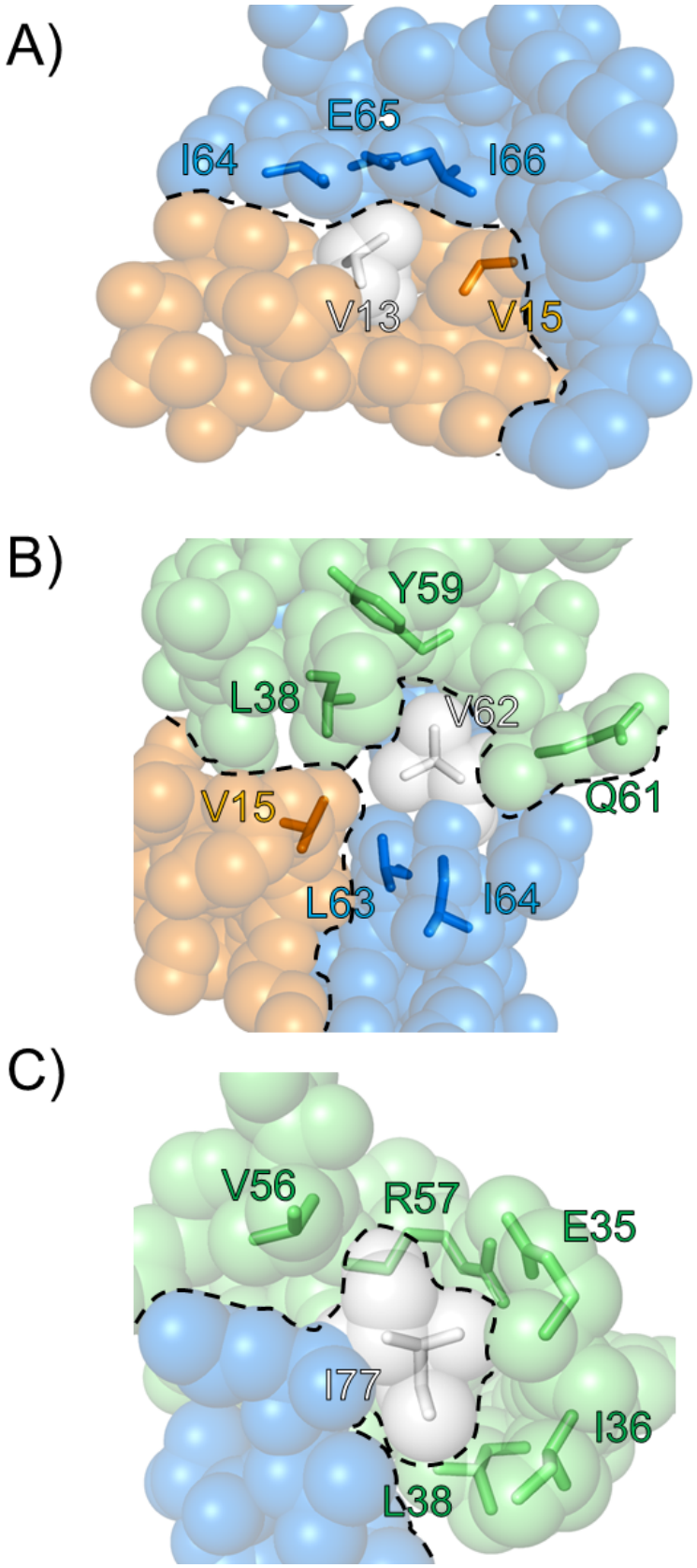


Figure 9. I13V/I62V/V77I mutations occur at the interfaces between internal regions of the PR. The mutated residues (white) are in contact with several residues that form part of the fulcrum (orange), cantilever (blue) and hinge (green) regions.

Supplementary data

Supplementary table 1: Thermodynamic parameters for the binding of acetyl-pepstatin to the C-SA PR, WT_{D25A} PR and HP3 PR variants.

PR	ΔG (kJ/mol)	ΔH (kJ/mol)	–TΔS (kJ/mol)	Stoichiometry (n)	K_d (μM)
C-SA	-46.26	41.54	-87.80	0.63	0.43
WT_{D25A}	-33.64	43.90	-77.54	0.58	1.01
HP3	-27.02	52.04	-79.10	0.95	0.21

References

Baxter, J.D., Chasanov, W.M., Adams, J., 2016. An Update on HIV-1 Protease Inhibitor Resistance. J AIDS Clin Res 7. https://doi.org/10.4172/2155-6113.1000581

Boden, D., Markowitz, M., 1998. Resistance to Human Immunodeficiency Virus Type 1 Protease Inhibitors. Antimicrob Agents Chemother 42, 2775–2783.

Brik, A., Wong, C.-H., 2003. HIV-1 protease: mechanism and drug discovery. Org. Biomol. Chem. 1, 5–14.

Carmel, A., Yaron, A., 1978. An intramolecularly quenched fluorescent tripeptide as a fluorogenic substrate of angiotensin-I-converting enzyme and of bacterial dipeptidyl carboxypeptidase. Eur J Biochem 87, 265–273. https://doi.org/10.1111/j.1432-1033.1978.tb12375.x

Chen, D., Oezguen, N., Urvil, P., Ferguson, C., Dann, S.M., Savidge, T.C., 2016. Regulation of protein-ligand binding affinity by hydrogen bond pairing. Science Advances 2, e1501240. https://doi.org/10.1126/sciadv.1501240

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66, 12–21. https://doi.org/10.1107/S0907444909042073

Chodera, J.D., Mobley, D.L., 2013. Entropy-Enthalpy Compensation: Role and Ramifications in Biomolecular Ligand Recognition and Design. Annual Review of Biophysics 42, 121–142. https://doi.org/10.1146/annurev-biophys-083012-130318

Clemente, J.C., Hemrajani, R., Blum, L.E., Goodenow, M.M., Dunn, B.M., 2003. Secondary Mutations M36I and A71V in the Human Immunodeficiency Virus Type 1 Protease Can Provide an Advantage for the Emergence of the Primary Mutation D30N. Biochemistry 42, 15029–15035. https://doi.org/10.1021/bi035701y

Clemente, J.C., Moose, R.E., Hemrajani, R., Whitford, L.R.S., Govindasamy, L., Reutzel, R., McKenna, R., Agbandje-McKenna, M., Goodenow, M.M., Dunn, B.M., 2004. Comparing the accumulation of active- and nonactive-site mutations in the HIV-1 protease. Biochemistry 43, 12141–12151. https://doi.org/10.1021/bi049459m

Connelly, P.R., 1994. Acquisition and use of calorimetric data for prediction of the thermodynamics of ligand-binding and folding reactions of proteins. Current Opinion in Biotechnology 5, 381–388. https://doi.org/10.1016/0958-1669(94)90046-9

Darke, P.L., Leu, C.T., Davis, L.J., Heimbach, J.C., Diehl, R.E., Hill, W.S., Dixon, R.A., Sigal, I.S., 1989. Human immunodeficiency virus protease. Bacterial expression and characterization of the purified aspartic protease. J. Biol. Chem. 264, 2307–2312.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60, 2126–2132. https://doi.org/10.1107/S0907444904019158

Foulkes-Murzycki, J.E., Scott, W.R.P., Schiffer, C.A., 2007. Hydrophobic Sliding: A Possible Mechanism for Drug Resistance in Human Immunodeficiency Virus Type 1 Protease. Structure 15, 225–233. https://doi.org/10.1016/j.str.2007.01.006

Freedberg, D.I., Ishima, R., Jacob, J., Wang, Y.-X., Kustanovich, I., Louis, J.M., Torchia, D.A., 2002. Rapid structural fluctuations of the free HIV protease flaps in solution: Relationship to crystal structures and comparison with predictions of dynamics calculations. Protein Sci 11, 221–232.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A., 2005. Protein Identification and Analysis Tools on the ExPASy Server, in: Walker, J.M. (Ed.), The Proteomics Protocols Handbook. Humana Press, Totowa, NJ, pp. 571–607. https://doi.org/10.1385/1-59259-890-0:571

Göttlinger, H.G., Sodroski, J.G., Haseltine, W.A., 1989. Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. Proc Natl Acad Sci U S A 86, 5781–5785.

Greenwood, J.R., Calkins, D., Sullivan, A.P., Shelley, J.C., 2010. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. J. Comput. Aided Mol. Des. 24, 591–604. https://doi.org/10.1007/s10822-010-9349-1

Gustchina, A., Weber, I.T., 1990. Comparison of inhibitor binding in HIV-1 protease and in non-viral aspartic proteases: the role of the flap. FEBS Lett 269, 269–272. https://doi.org/10.1016/0014-5793(90)81171-J

Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J.Y., Wang, L., Lupyan, D., Dahlgren, M.K., Knight, J.L., Kaus, J.W., Cerutti, D.S., Krilov, G., Jorgensen, W.L., Abel, R., Friesner, R.A., 2016. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. J. Chem. Theory Comput. 12, 281–296. https://doi.org/10.1021/acs.jctc.5b00864

Heaslet, H., Rosenfeld, R., Giffin, M., Lin, Y.C., Tam, K., Torbett, B.E., Elder, J.H., McRee, D.E., Stout, C.D., 2007. Conformational flexibility in the flap domains of ligand-free HIV

protease. Acta Crystallogr D Biol Crystallogr 63, 866–875. https://doi.org/10.1107/S0907444907029125

Herber, W.K., Bailey, F.J., Carty, C.E., Heimbach, J.C., Maigetter, R.Z., 1991. Production of cytotoxic proteins in Escherichia coli: a fermentation process for producing enzymatically active HIV-1 protease. Appl Microbiol Biotechnol 36, 149–152. https://doi.org/10.1007/BF00164410

Hong, L., Zhang, X.C., Hartsuck, J.A., Tang, J., 2000. Crystal structure of an in vivo HIV-1 protease mutant in complex with saquinavir: insights into the mechanisms of drug resistance. Protein Sci 9, 1898–1904. https://doi.org/10.1110/ps.9.10.1898

Hornak, V., Okur, A., Rizzo, R.C., Simmerling, C., 2006. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. PNAS 103, 915–920. https://doi.org/10.1073/pnas.0508452103

Ido, E., Han, H.P., Kezdy, F.J., Tang, J., 1991. Kinetic studies of human immunodeficiency virus type 1 protease and its active-site hydrogen bond mutant A28S. J. Biol. Chem. 266, 24359–24366.

Ishima, R., Freedberg, D.I., Wang, Y.X., Louis, J.M., Torchia, D.A., 1999. Flap opening and dimer-interface flexibility in the free and inhibitor-bound HIV protease, and their implications for function. Structure 7, 1047–1055.

Johnson, V.A., Brun-Vézinet, F., Clotet, B., Conway, B., D'Aquila, R.T., Demeter, L.M., Kuritzkes, D.R., Pillay, D., Schapiro, J.M., Telenti, A., Richman, D.D., International AIDS Society-USA Drug Resistance Mutations Group, 2003. Drug resistance mutations in HIV-1. Top HIV Med 11, 215–221.

Kozal, M., 2004. Review: Cross-Resistance Patterns Among HIV Protease Inhibitors. AIDS Patient Care and STDs 18, 199–208. https://doi.org/10.1089/108729104323038874

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography. https://doi.org/10.1107/S0021889892009944

Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., Moriarty, N.W., Oeffner, R.D., Poon, B.K., Prisant, M.G., Read, R.J., Richardson, J.S., Richardson, D.C., Sammito, M.D., Sobolev, O.V., Stockwell, D.H., Terwilliger, T.C., Urzhumtsev, A.G., Videau, L.L., Williams, C.J., Adams, P.D., 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent

developments in Phenix. Acta Crystallogr D Struct Biol 75, 861–877. https://doi.org/10.1107/S2059798319011471

Lin, L.N., Mason, A.B., Woodworth, R.C., Brandts, J.F., 1994. Calorimetric Studies of Serum Transferrin and Ovotransferrin. Estimates of Domain Interactions, and Study of the Kinetic Complexities of Ferric Ion Binding. Biochemistry 33, 1881–1888. https://doi.org/10.1021/bi00173a035

Louis, J.M., Ishima, R., Torchia, D.A., Weber, I.T., 2007. HIV-1 protease: structure, dynamics, and inhibition. Adv. Pharmacol. 55, 261–298. https://doi.org/10.1016/S1054-3589(07)55008-8

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. J Appl Cryst 40, 658–674. https://doi.org/10.1107/S0021889807021206

Mosebi, S., Morris, L., Dirr, H.W., Sayed, Y., 2008. Active-site mutations in the South african human immunodeficiency virus type 1 subtype C protease have a significant impact on clinical inhibitor binding: kinetic and thermodynamic study. J Virol 82, 11476–11479. https://doi.org/10.1128/JVI.00726-08

Muzammil, S., Ross, P., Freire, E., 2003. A major role for a set of non-active site mutations in the development of HIV-1 protease drug resistance. Biochemistry 42, 631–638. https://doi.org/10.1021/bi027019u

Naicker, P., Achilonu, I., Fanucchi, S., Fernandes, M., Ibrahim, M.A.A., Dirr, H.W., Soliman, M.E.S., Sayed, Y., 2013. Structural insights into the South African HIV-1 subtype C protease: impact of hinge region dynamics and flap flexibility in drug resistance. J. Biomol. Struct. Dyn. 31, 1370–1380. https://doi.org/10.1080/07391102.2012.736774

Naicker, P., Stoychev, S., Dirr, H.W., Sayed, Y., 2014. Amide hydrogen exchange in HIV-1 subtype B and C proteases--insights into reduced drug susceptibility and dimer stability. FEBS J. 281, 5395–5410. https://doi.org/10.1111/febs.13084

Nicholson, L.K., Yamazaki, T., Torchia, D.A., Grzesiek, S., Bax, A., Stahl, S.J., Kaufman, J.D., Wingfield, P.T., Lam, P.Y., Jadhav, P.K., 1995. Flexibility and function in HIV-1 protease. Nat. Struct. Biol. 2, 274–280. https://doi.org/10.1038/nsb0495-274

Ohtaka, H., Velázquez-Campoy, A., Xie, D., Freire, E., 2002. Overcoming drug resistance in HIV-1 chemotherapy: the binding thermodynamics of Amprenavir and TMC-126 to wild-type and drug-resistant mutants of the HIV-1 protease. Protein Sci. 11, 1908–1916. https://doi.org/10.1110/ps.0206402

Olsen, D.B., Stahlhut, M.W., Rutkowski, C.A., Schock, H.B., vanOlden, A.L., Kuo, L.C., 1999. Non-active Site Changes Elicit Broad-based Cross-resistance of the HIV-1 Protease to Inhibitors. Journal of Biological Chemistry 274, 23699–23701. https://doi.org/10.1074/jbc.274.34.23699

Perozzo, R., Folkers, G., Scapozza, L., 2004. Thermodynamics of protein-ligand interactions: history, presence, and future aspects. J Recept Signal Transduct Res 24, 1–52. https://doi.org/10.1081/rrs-120037896

Perryman, A.L., Lin, J.-H., McCammon, J.A., 2004. HIV-1 protease molecular dynamics of a wild-type and of the V82F/I84V mutant: possible contributions to drug resistance and a potential new target site for drugs. Protein Sci. 13, 1108–1123. https://doi.org/10.1110/ps.03468904

Ragland, D.A., Nalivaika, E.A., Nalam, M.N.L., Prachanronarong, K.L., Cao, H., Bandaranayake, R.M., Cai, Y., Kurt-Yilmaz, N., Schiffer, C.A., 2014. Drug Resistance Conferred by Mutations Outside the Active Site through Alterations in the Dynamic and Structural Ensemble of HIV-1 Protease. Journal of the American Chemical Society 136, 11956–11963. https://doi.org/10.1021/ja504096m

Rhee, S.-Y., Taylor, J., Fessel, W.J., Kaufman, D., Towner, W., Troia, P., Ruane, P., Hellinger, J., Shirvani, V., Zolopa, A., Shafer, R.W., 2010. HIV-1 protease mutations and protease inhibitor cross-resistance. Antimicrob. Agents Chemother. 54, 4253–4261. https://doi.org/10.1128/AAC.00574-10

Roberts, J.D., Bebenek, K., Kunkel, T.A., 1988. The accuracy of reverse transcriptase from HIV-1. Science 242, 1171–1173.

Rose, R.B., Craik, C.S., Stroud, R.M., 1998. Domain Flexibility in Retroviral Proteases: Structural Implications for Drug Resistant Mutations † , ‡. Biochemistry 37, 2607–2621. https://doi.org/10.1021/bi9716074

Sayer, J.M., Liu, F., Ishima, R., Weber, I.T., Louis, J.M., 2008. Effect of the Active Site D25N Mutation on the Structure, Stability, and Ligand Binding of the Mature HIV-1 Protease. J Biol Chem 283, 13459–13470. https://doi.org/10.1074/jbc.M708506200

Sayer, J.M., Louis, J.M., 2009. INTERACTIONS OF DIFFERENT INHIBITORS WITH ACTIVE-SITE ASPARTYL RESIDUES OF HIV-1 PROTEASE AND POSSIBLE RELEVANCE TO PEPSIN. Proteins 75, 556–568. https://doi.org/10.1002/prot.22271

Schägger, H., von Jagow, G., 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal Biochem 166, 368–379. https://doi.org/10.1016/0003-2697(87)90587-2

Shafer, R.W., Hsu, P., Patick, A.K., Craig, C., Brendel, V., 1999. Identification of Biased Amino Acid Substitution Patterns in Human Immunodeficiency Virus Type 1 Isolates from Patients Treated with Protease Inhibitors. J Virol 73, 6197–6202.

Shelley, J.C., Cholleti, A., Frye, L.L., Greenwood, J.R., Timlin, M.R., Uchimaya, M., 2007. Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. J. Comput. Aided Mol. Des. 21, 681–691. https://doi.org/10.1007/s10822-007-9133-z

Sherry, D., Worth, R., Ismail, Z.S., Sayed, Y., 2021. Cantilever-centric mechanism of cooperative non-active site mutations in HIV protease: Implications for flap dynamics. Journal of Molecular Graphics and Modelling 106, 107931. https://doi.org/10.1016/j.jmgm.2021.107931

Sherry, D., Worth, R., Sayed, Y., 2020. Two-Step Preparation of Highly Pure, Soluble HIV Protease from Inclusion Bodies Recombinantly Expressed in Escherichia coli. Current Protocols in Protein Science 100, e106. https://doi.org/10.1002/cpps.106

Shrake, A., Ross, P.D., 1990. Ligand-induced biphasic protein denaturation. J. Biol. Chem. 265, 5055–5059.

Szeltner, Z., Polgár, L., 1996. Rate-determining steps in HIV-1 protease catalysis. The hydrolysis of the most specific substrate. J Biol Chem 271, 32180–32184. https://doi.org/10.1074/jbc.271.50.32180

Takahashi, K., Fukada, H., 1985. Calorimetric studies of the binding of Streptomyces subtilisin inhibitor to subtilisin of Bacillus subtilis strain N'. Biochemistry 24, 297–300.

Todd, M.J., Semo, N., Freire, E., 1998. The structural stability of the HIV-1 protease. J. Mol. Biol. 283, 475–488. https://doi.org/10.1006/jmbi.1998.2090

Velazquez-Campoy, A., Kiso, Y., Freire, E., 2001. The binding energetics of first- and second-generation HIV-1 protease inhibitors: implications for drug design. Arch. Biochem. Biophys. 390, 169–175. https://doi.org/10.1006/abbi.2001.2333

Velázquez-Campoy, A., Todd, M.J., Vega, S., Freire, E., 2001. Catalytic efficiency and vitality of HIV-1 proteases from African viral subtypes. Proceedings of the National Academy of Sciences 98, 6062–6067. https://doi.org/10.1073/pnas.111152698

Velázquez-Campoy, A., Vega, S., Fleming, E., Bacha, U., Sayed, Y., Dirr, H.W., Freire, E., 2003. Protease inhibition in African subtypes of HIV-1. AIDS Rev 5, 165–171.

Weber, I.T., Kneller, D.W., Wong-Sam, A., 2015. Highly resistant HIV-1 proteases and strategies for their inhibition. Future Med Chem 7, 1023–1038. https://doi.org/10.4155/fmc.15.44

Winter, G., 2010. xia2: an expert system for macromolecular crystallography data reduction. J Appl Cryst 43, 186–190. https://doi.org/10.1107/S0021889809045701

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

Yedidi, R.S., Proteasa, G., Martinez, J.L., Vickrey, J.F., Martin, P.D., Wawrzak, Z., Liu, Z., Kovari, I.A., Kovari, L.C., 2011. Contribution of the 80s loop of HIV-1 protease to the multidrug-resistance mechanism: crystallographic study of MDR769 HIV-1 protease variants. Acta Crystallogr D Biol Crystallogr 67, 524–532. https://doi.org/10.1107/S0907444911011541

Young, T.P., Parkin, N.T., Stawiski, E., Pilot-Matias, T., Trinh, R., Kempf, D.J., Norton, M., 2010. Prevalence, mutation patterns, and effects on protease inhibitor susceptibility of the L76V mutation in HIV-1 protease. Antimicrob. Agents Chemother. 54, 4903–4906. https://doi.org/10.1128/AAC.00906-10

Chapter 4 – Molecular dynamics and induced fit docking studies

Cantilever-centric mechanism of cooperative non-active site mutations in HIV protease: implications for flap dynamics

Dean Sherry, Roland Worth, Zaahida Sheik Ismail, and Yasien Sayed.

The Journal of Molecular Graphics and Modelling – Published (2021)

In this article, the effect of the non-active site mutations is investigated using computational biological techniques. Induced fit molecular docking studies and molecular dynamics simulations were performed in order to provide a better understanding of the mechanisms at play in the mutated protease systems.

Author contributions: Dean Sherry performed *in silico* experiments represented in the article, generated images and tables, wrote the manuscript in its entirety and edited subsequent drafts of the manuscript. Roland Worth provided guidance for the *in silico* protocols used, assisted with understanding some of the data obtained, as well as with proofreading of the manuscript. Zaahida Sheik Ismail assisted with editing and proofreading of the manuscript. Yasien Sayed supervised the project and provided proofreading and revisions for the manuscript.

Note on the D25A mutation: The molecular dynamics studies herein were performed on the catalytically active protease containing the naturally-occurring aspartate residues (D25/D25'). The elimination of the catalytic activity of the enzyme is helpful for in vitro analyses but it is not a concern during in silico analyses.

Cantilever-centric mechanism of cooperative non-active site mutations in HIV protease: implications for flap dynamics

Dean Sherry[1], Roland Worth[1], Zaahida Sheik Ismail[1], Yasien Sayed[1]*

1 Protein Structure-Function Research Unit, School of Molecular and Cell Biology, University of the Witwatersrand, Johannesburg, 2050, South Africa

*Corresponding author: Yasien Sayed; e-mail address: Yasien.Sayed@wits.ac.za.

Abstract

The HIV-1 protease is an important drug target in antiretroviral therapy due to the crucial role it plays in viral maturation. A greater understanding of the dynamics of the protease as a result of drug-induced mutations has been successfully elucidated using computational models in the past. We performed induced-fit docking studies and molecular dynamics simulations on the wild-type South African HIV-1 subtype C protease and two non-active site mutation-containing protease variants; HP3 PR and HP4 PR. The HP3 PR contained the I13V, I62V, and V77I mutations while HP4 PR contained the same mutations with the addition of the L33F mutation. The simulations were initiated in a cubic cell universe containing explicit solvent, with the protease variants beginning in the fully closed conformation. The trajectory for each simulation totalled 50 nanoseconds. The results indicate that the mutations increase the dynamics of the flap, hinge, fulcrum and cantilever regions when compared to the wild-type protease while in complex with protease inhibitors. Specifically, these mutations result in the protease favouring the semi-open conformation when in complex with inhibitors. Moreover, the HP4 PR adopted curled flap tip conformers which coordinated several water molecules into the active site in a manner that may reduce inhibitor binding affinity. The mutations affected the thermodynamic landscape of inhibitor binding as there were fewer observable chemical contacts between the mutated variants and saquinavir, atazanavir and darunavir. These data help to elucidate the biophysical basis for the selection of cooperative non-active site mutations by the HI virus.

Introduction

The HIV/AIDS epidemic continues to be a significant socioeconomic burden, particularly in developing countries in sub-Saharan Africa. According to the UNAIDS 2019 global HIV/AIDS statistics, approximately 38 million people are estimated to be living with HIV globally – with an astonishing 7.7 million of those individuals living in South Africa (UNAIDS, 2020). Moreover,

the subtype C virus is the most commonly occurring subtype globally and accounts for an estimated 50% of all infections, the majority of which is concentrated in sub-Saharan African countries, including South Africa (UNAIDS, 2020).

The virus encodes Gag and Gag-Pol polyprotein precursors which contain several structural proteins and enzymes. Specifically, the viral protease (PR) is responsible for cleavage and subsequent activation of the viral Gag and Gag-Pol polyprotein precursors in order to produce mature, infectious, viral progeny (Göttlinger et al., 1989; Louis et al., 2007). Thus, HIV PR remains an important target for rational drug design (Thompson et al., 2010; Volberding and Deeks, 2010). Knowledge of the structure of the HIV PR has led to the design of nine FDA-approved protease inhibitors (PIs) targeted to the active site of the PR with high affinity. The PIs were developed to target the subtype B PR and due to the existence of naturally occurring polymorphisms between subtypes, the PIs designed to target the subtype B PR have been shown to have lower efficacy against some subtype C variants (Velázquez-Campoy et al., 2001c). In order to understand the inhibition of the PR, the biochemical basis for PI susceptibility or, conversely, PI resistance mechanisms must be elucidated.

A more detailed understanding of the PR structure will provide a model that elucidates the movements associated between distinct regions of the PR during drug binding. For instance, entry of the substrate/inhibitor into the active site of PR is regulated by two β-hairpin loops, known as the flaps (Figure 1A), which regulate entry of substrates/inhibitors into the active site cavity (Wlodawer et al., 1989). In turn, the stability and movement of the flaps is closely associated with the hinge region of the PR (residues 35-42 and 56-61), as the conformational flexibility of these regions allow for substantial movement of the flaps (Gustchina and Weber, 1990). However, the overall movement of the flaps into the open conformation is achieved through a concerted synergistic movement of the hinge in conjunction with the fulcrum (residues 10-22) and the cantilever (residues 62-78) (Foulkes-Murzycki et al., 2007; Hornak et al., 2006). It is important to note that although the PR is homodimeric in nature, asymmetric movement of the flap regions has been observed in numerous molecular dynamics (MD) simulations (Hornak et al., 2006; Perryman et al., 2004; W. R. P. Scott and Schiffer, 2000)

Twenty hydrophobic residues with limited solvent accessibility have been identified as forming the hydrophobic core of the PR (Figure S1) (Foulkes-Murzycki et al., 2007; Naicker et al., 2014). It has been suggested that the near-isoenergetic rearrangement of van der Waals contacts between hydrophobic residues in the core of PR, or “hydrophobic sliding”, allows for the different regions within PR to slide across one another – thereby facilitating the

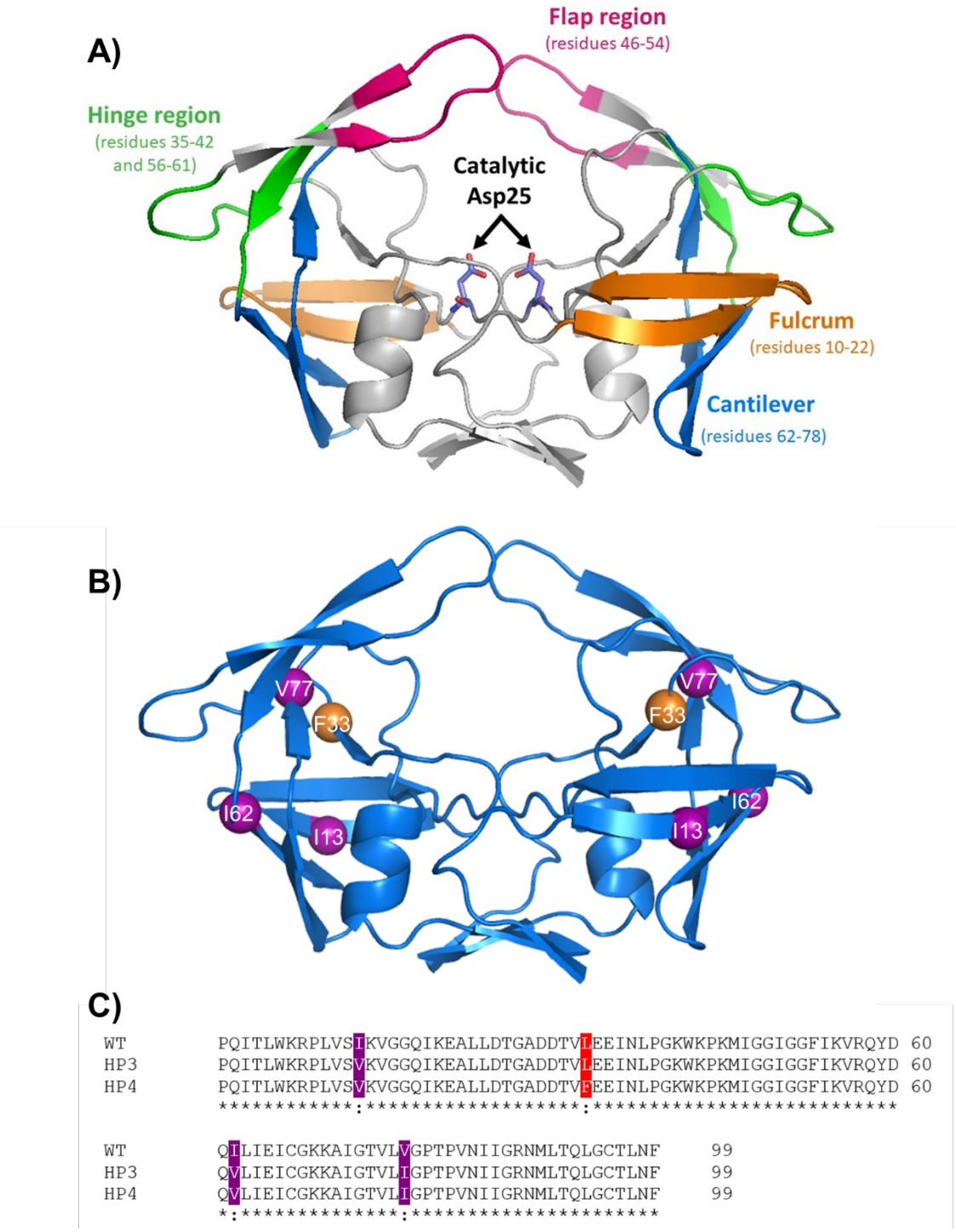


Figure 1. A) The structural architecture of the HIV PR enzyme (PDB ID: 6I45). The movement of the flap regions (pink) which cover the active site Asp25 residues (stick representation) relies on the synergistic movement of the hinge (green), fulcrum (orange) and cantilever (blue) regions. B) The PR with the relative position of mutations indicated by spheres. The purple spheres indicate the locations of the three mutations present in HP3 PR while the orange sphere represents the additional mutation present in HP4 PR. C) Amino acid sequence alignment of the wild-type (WT) South African subtype C PR, HP3 PR and HP4 PR.

conformational changes associated with flap opening (Foulkes-Murzycki et al., 2007; Ragland et al., 2014). These hydrophobic residues may play a crucial role in the conformational adaptability exhibited by the PR. In particular, two isoleucine residues located in the fulcrum region of the PR, Ile13 and Ile15, have been implicated in the rearrangement of hydrophobic contacts in the core of the PR in a manner that facilitates flap movement. Briefly, Ile15 forms hydrophobic interactions with two residues located in the hinge region, Met35 and Leu38, as well as three residues that form part of the cantilever, Ile62, Ile64 and Val75. During flap opening, these hydrophobic contacts involving Ile15 are lost which allows the hinge and cantilever regions to slide over Ile15 and downwards towards the dimer interface. To stabilise the open conformation, the hydrophobic contacts that are lost with Ile15 are reformed with Ile13 (Foulkes-Murzycki et al., 2007). The internal rearrangement of hydrophobic interactions in the core of the PR, therefore, appears to be vital in the functionality of the PR through modulation of flap conformations. Hence, non-active site mutations, mutations located outside of the active site region within the hydrophobic core of the PR, may alter the network of interactions and modify the stability of the core in a manner that allows the enzyme to sample an ensemble of altered conformations. A possible explanation for the mechanisms of resistance of secondary mutations may be, in part, due to an increased conformational flexibility of the flaps in a manner that favours more open conformations, which has been associated with reduced drug susceptibility (Foulkes-Murzycki et al., 2007; Galiano et al., 2009).

Resistance to PIs starts with primary mutations which involve amino acid substitutions within the active site cavity. Primary resistance mutations possess the potential to diminish the binding affinity of the naturally occurring Gag and Gag-Pol substrates leading to the hypothesis that non-active site polymorphisms compensate for alterations generated by primary mutations in ways that improve catalytic efficiency and viral replication (Borman et al., 1996; Nijhuis et al., 1999). For instance, mutations at amino acid positions 82 and 84 (non-active site mutations), which are located outside of the active site, are associated with drug resistance and, in fact, the I84V mutation is associated with resistance to all currently available PIs (Wensing et al., 2017). However, in order to achieve high levels of resistance while maintaining suitable viral fitness levels, active site mutations are accompanied by a constellation of non-active site mutations (Baxter et al., 2016; Clemente et al., 2004; Kozal, 2004; Ohtaka et al., 2002; Olsen et al., 1999). It would appear that single non-active site amino acid substitutions do not introduce significant reductions in inhibitor potency but, rather, specific combinations of multiple mutations are required to confer high levels of resistance (Ishima et al., 2001; Johnson et al., 2003).

Mutations in the hydrophobic core of the PR are likely to bring about alterations in the network of hydrophobic interactions formed within the core due to changes in the size and flexibility of their associated side chains. Many non-active site mutations are conservative substitutions of hydrophobic amino acids such as leucine, isoleucine, valine and methionine, all of which have been found to contribute significantly to drug resistance (Clemente et al., 2004; Foulkes-Murzycki et al., 2007; Muzammil et al., 2003; Rhee et al., 2010; Weber et al., 2015). The mechanism of action of non-active site mutations is not well understood, but it is hypothesised that mutations of this nature could impact the hydrophobic sliding mechanism and, in turn, lead to altered flap dynamics in a manner that diminishes drug binding.

This investigation focuses on four non-active site amino acid substitutions which arise from inhibitor exposure; namely, I13V, L33F, I62V and V77I, which occur in the core of the PR and have been hypothesised to play a role in reduced inhibitor susceptibility (Foulkes-Murzycki et al., 2007). The I13V, I62V and V77I mutations occur in 34%, 16% and 18% of the general population of individuals infected with HIV, respectively (Johnson et al., 2003; Rhee et al., 2005; Robert W. Shafer et al., 1999). Interestingly, the L33F mutation significantly increases the occurrence of I13V from 34% to 51%, I62V from 16% to 48%, and V77I from 18% to 35% relative to the general population (Foulkes-Murzycki et al., 2007). This set of “partner substitutions” are of particular interest since the L33F mutation is a minor resistance mutation for atazanavir, lopinavir and darunavir, and is a major resistance mutation for tipranavir (Baxter et al., 2016; Weber and Agniswamy, 2009). Thus, in future, there is an increased potential for these mutations to occur in patients who are undergoing continued treatment using the newest generation of PIs. While these mutations are correlated with reduced drug susceptibility their mechanism of action has not yet been elucidated. The close proximity of these mutations to one another within the hydrophobic core of the PR (Figure 1B, S1) could explain the possibility of their substitutions being co-dependent. This theory is bolstered by the observed cooperativity of the movements of residues I13, I62 and V77 during MD simulations (Foulkes-Murzycki et al., 2007). Moreover, these mutations have been shown to decrease the favourable free energy change (ΔG) of binding for ATV, DRV, RTV and SQV which correlated with a 64- and 16-fold increase in the dissociation constants of ATV and DRV, respectively. While the dissociation constants for SQV and RTV were relatively unaffected (manuscript in preparation).

To the best of our knowledge, these mutations have not been studied *in silico* in isolation without the presence of active site mutations or additional non-active site mutations. Here, we propose a mechanism of resistance that is explained by the aforementioned set of four non-active site mutations. The mechanism was elucidated through comparative induced-fit docking (IFD) studies and MD simulations using the wild-type (WT) South African subtype C (C-SA)

PR, the HP3 PR and the HP4 PR. In addition, we sought to determine the effect of lowering the environmental pH to 5 as opposed to physiological pH, since a pH of 5 is the experimentally determined working pH.

Methods and Materials

Induced-fit ligand docking

The induced-fit docking (IFD) studies were performed using Glide software (Schrödinger LLC 2009, USA) implemented on Windows architecture. The PR models were generated using PyMOL software (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC). The models of the WT C-SA PR and the HP4 PR were constructed from the coordinates of the HP3 PR, resolved at a resolution of 1.8 Å (PDB ID: 6I45). The models were generated using the built-in PyMOL mutagenesis plug-in and all of the models were validated using PROCHECK (Laskowski et al., 1993). All of the protein structures utilised contained the Q7K mutation which is incorporated in experimental PR constructs because it decreases auto-proteolysis without affecting the kinetic properties of the PR (Louis et al., 1991). The Schrödinger modules Glide, Prime and Protein Preparation Wizard were used for protein preparation. The bond orders of the models were corrected and hydrogen atoms were added to the protein models, while the ionisable residues were assigned a charge corresponding to an experimental pH of 5.0 using the Epik prediction software tool (Greenwood et al., 2010; Shelley et al., 2007). We used an average RMSD threshold of 0.3 Å when performing energy minimisation on the protein models as this is the default RMSD threshold in the Protein Preparation Wizard module. This default RMSD threshold ensures that the receptor is optimally constrained for induced fit molecular docking by relaxing strained bonds, angles and clashes (Schrödinger LLC 2009, USA).

The two dimensional structure data files of the ligands used for IFD were obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov/) which included nine FDA-approved PIs: atazanavir (ATV, CID: 148192), darunavir (DRV, CID: 213039), fosamprenavir (FPV, CID: 131536), indinavir (IDV, CID: 5462355), lopinavir (LPV, CID: 92727), nelfinavir (NFV, CID: 64142), ritonavir (RTV, CID: 392622), saquinavir (SQV, CID: 60787) and tipranavir (TPV, CID: 54682461). The LigPrep tool (Schrödinger LLC 2009, USA), was used to add hydrogens, correct bond angles and lengths and to assign the correct chirality before performing energy minimisation on the ligands. The Epik tool was used to select the tautomers and ring structures with the lowest energy. The energy-minimised protein models were loaded into the workspace, and a 20 Å cubic box surrounding the active site D25 residues was defined at the region where IFD would occur. During IFD, protein residues within 5 Å of the ligand poses were refined and subjected to side chain optimisation. Glide docking does not allow for receptor flexibility but,

rather, allows for the scale of van der Waals radii of non-polar residues to be scaled in order to decrease the penalty scores for close contacts in order to afford some minor flexibility within the receptor and ligand. As such, a scaling factor of 0.50 was selected for the van der Waals radii of the ligand and protein during the Glide docking procedure. The IFD simulations were then performed using the OPLS3 force field (Harder et al., 2016). Twenty conformational poses (system default setting) were generated for each inhibitor during Glide docking and based on their respective top IFD scores the best poses were selected for further investigation using MD simulations. These docked top poses were aligned with corresponding existing crystallographic structures as a method of verifying the validity of the poses selected.

Computational data acquisition

The following PIs were selected for further investigation via MD simulations: ATV, LPV, DRV, SQV, and TPV. ATV, DRV and LPV were selected as these are the three most commonly used PIs for ARV therapy in South Africa. SQV was selected due to its propensity to induce both the I62V and V77I mutations (Baxter et al., 2016). Finally, TPV was selected because exposure to this drug increases the propensity of the I13V, I62V, and V77I mutations in the general population (Foulkes-Murzycki et al., 2007; Johnson et al., 2003; Rhee et al., 2005; Schaffer and Verkhivker, 1998). MD simulation studies were performed on Desmond software version 11.2 (Schrödinger LLC 2009, USA), running on two 3.68 GHz Intel core i7 5960x computers implemented on LINUX architecture. The OPLS3 force field was used for all of the systems (Harder et al., 2016). The simulations were conducted with explicit solvent in a cubic box universe that was generated by using the TIP3P solvation model within the System Builder module (Jorgensen et al., 1983; Mark and Nilsson, 2001). Different box shapes have shown to impart small, but statistically significant, differences during MD simulations (Wassenaar and Mark, 2006). Therefore, in the present study the cubic box universe was selected as has been used previously (Naicker et al., 2013; Zondagh et al., 2018). As such, the protein was centred 10 Å from the edges of a cubic box. The apo, DRV-, LPV- and TPV-bound systems were neutralised by the addition of 11 (charge -11) chloride ions while the ATV- and SQV-bound systems utilised 12 (charge -12) chloride ions. After building the solvated systems, the protein/protein-inhibitor complexes were subjected to minimisation and relaxation under the NPT ensemble (300 K, 1 atm) using the default protocol of Desmond, as described elsewhere (Reddy et al., 2015). This process performs a series of minimisations and short MD simulations which equilibrate the system before the experimental simulation is performed. Briefly, a 12 ps MD simulation is performed under the NVT ensemble (constant number of particles (N), constant volume (V) and constant temperature (T)) at a temperature of 10 K. Thereafter, a 12 ps simulation is performed under the NPT ensemble (constant number of particles (N), constant pressure (P) and constant temperature (T)) at a temperature of 10 K

and a pressure of 1 atm. This is followed by a further two 24 ps simulations under the NPT ensemble at a temperature of 300 K and a pressure of 1 atm. The first simulation is performed with non-hydrogen solute atoms restrained while the second does not have these restrictions. Following equilibration, the MD simulations were performed for 50 ns under a constant temperature of 300 K with a Berendsen thermostat and an average pressure of 1 atm maintained by means of a Berendsen barostat. These data were collected over 1000 frames with data being collected every 0.05 ns.

Data analysis

Initially, the quality of the trajectories was assessed by examining the volume, temperature and pressure parameters of the simulations, which were analysed in the Simulation Quality Analysis Module in Maestro (Maestro, Schrödinger LLC 2020, USA). The RMSF, RMSD, and ligand interaction-diagram analyses were performed using the Simulation Interaction Diagram Module. The radius of gyration (R_g) and solvent accessible surface area (SASA) analyses were performed using the Simulation Event Analysis Module and Simulation Interaction Diagram Module, respectively. The resultant trajectories from the MD simulations were analysed using the molecular visualisation software UCSF Chimera version 1.11.2 (Pettersen et al., 2004) and the PR models were viewed with the molecular visualisation software PyMOL (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.). Sequence alignment was performed using the Clustal Omega tool (EMBL-EBI) (Sievers et al., 2011). Principal component analyses were performed using the Bio3D package using the software environment R (Grant et al., 2006; Skjærven et al., 2014).

Results and Discussion

Induced-fit docking studies

The data obtained from induced-fit docking (IFD) was utilised to select and compare the top poses for each of the different protein/ligand combinations based on the IFD score, that is, the pose which had the lowest IFD score was selected to represent the most favourable complex. Thereafter, the scoring functions, such as the Glide score (G_{score}), where examined for comparative purposes.

Altered protease inhibitor binding due to non-active site mutations: I13V, L33F, I62V and V77I

The calculated Gscore values produced from IFD provide an approximation of the ligand binding free energy and, thus, an approximate ligand binding affinity (Friesner et al., 2004). As such, in order to identify possible changes in the calculated binding scores of the PIs that occur due to the presence of the non-active site mutations, we investigated the G_{score} values for the WT PR, HP3 PR, and HP4 PR (Figure 2). The HP3 PR and HP4 PR did not exhibit

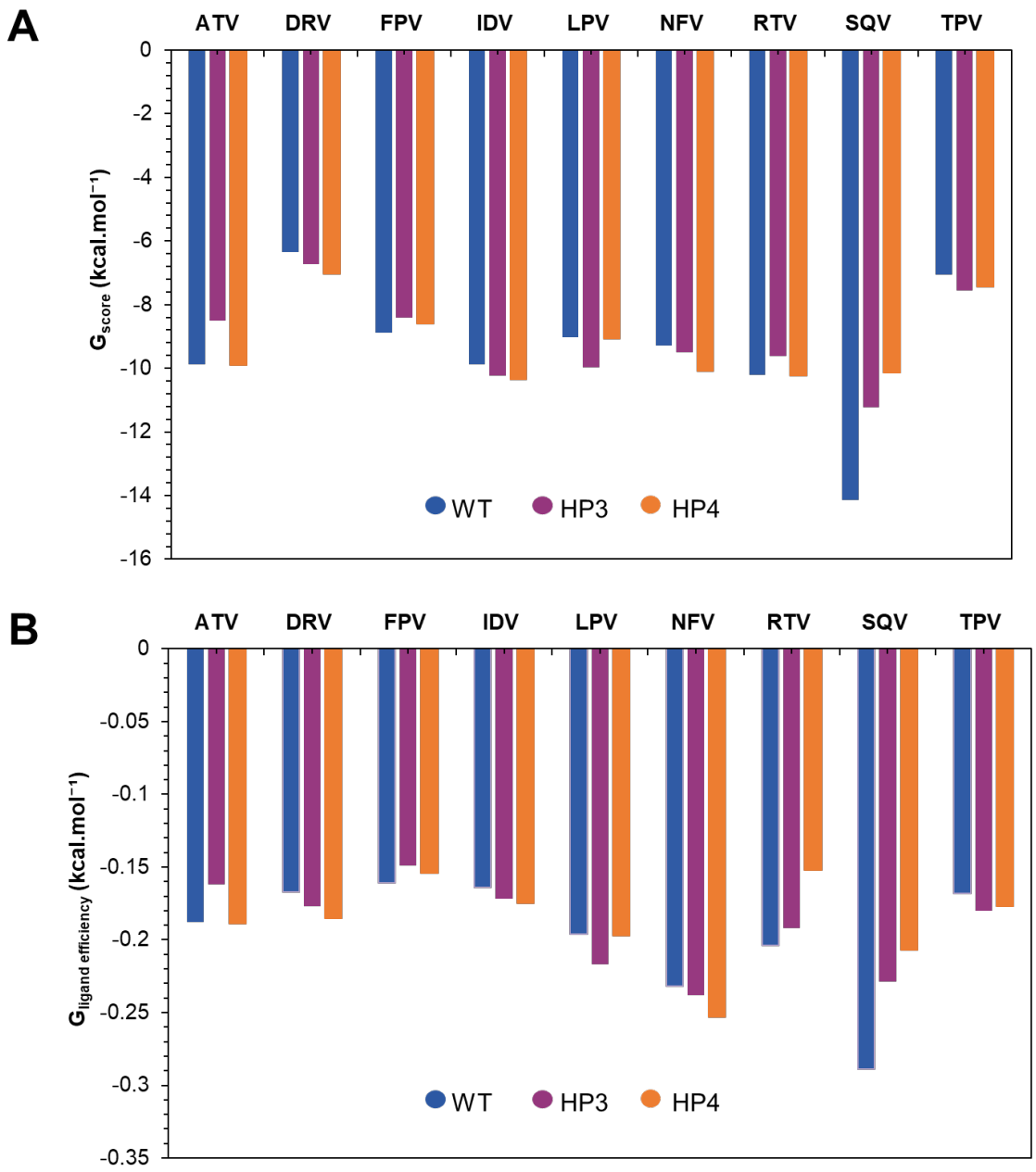


Figure 2. Calculated Glide parameters following IFD of PIs to WT PR, HP3 PR and HP4 PR at pH 5. A) G_{score} values of the PIs when docked to the three PR structures B) $G_{ligand\ efficiency}$ parameter, representative of the component between the protein and inhibitor.

significant changes in the G_{score} for the majority of the PIs when compared to the WT PR. However, the SQV-bound complexes displayed the largest decrease in the G_{score} for both the mutated PR variants. This difference in the G_{score} reflected the decrease in the various components that are used to determine this value, including the G_{hbond}, G_{ecoul} and G_{esite}; the latter describing the polar interactions that occur within the active site. Furthermore, the decrease in the G_{score} is reflected in an increase in the $G_{ligand\ efficiency}$ component which provides a metric for the calculated buried surface area of the ligand. Therefore, the $G_{ligand\ efficiency}$ change illustrates a reduction in the buried surface area of SQV. Since desolvation of the hydrophobic surface area of the SQV contributes significantly to its binding favourability, this change may have significant implications for drug susceptibility. Given that IFD data provides a “snapshot” of the inhibitor-bound PR complexes in an energy minimised and less dynamic conformation, there is a limit to the inferences that can be drawn from the data. Since proteins are dynamic in nature, MD simulations provide a deeper understanding of the effect of non-active site mutations on these dynamic complexes. Therefore, by using MD simulations, we investigated the dynamics of all three PR structures in complex with selected PIs; namely, ATV, DRV, LPV, SQV, and TPV.

Molecular dynamics

Essential dynamics of the apo protease

The essential dynamics refers to the correlated and anti-correlated motions between different regions within each protein over time. These motions of the three PR systems investigated; namely, WT, HP3 and HP4, were studied in order to understand which of these protein motions relate to or govern specific patterns of movement in the PR. The analyses of the essential dynamics performed in this study were restricted to the backbone C_α atoms of each protein.

MD simulations were performed on solvated apo protein systems for 50 ns and all simulations were initiated with each protein in the closed and energy-minimised conformation. The starting structures were aligned with the reference crystallographic structure (PDB ID:6I45). The root-mean-square standard deviation (RMSD) values for the alignment of the WT and HP4 PR models were 0.09 Å and 0.04 Å, respectively. The PR is known to sample a number of heterogeneous flap conformations which exist in an equilibrium; namely, closed, semi-open and open conformations (Wlodawer et al., 1989). Therefore, the simulation length was selected to provide sufficient time for these different flap conformations to be observed. Previously, the various conformations were defined by interatomic distance between Ile50/50' located in the flap tips of each monomer. By means of this method, the closed, semi-open, and open conformations were defined by Ile50/50' distances of <7 Å, 7-12 Å and >12 Å,

respectively. However, since this distance is likely to be affected by both flap asymmetry and flap tip curling, we thought that it would be more reasonable to measure the distance between Asp25/25' and Ile50/50'. The semi-open conformation, in particular, can be rather subjective in its definition and, as such, we attempted to define these distances in a manner that has been used previously (Perryman 2004). We, therefore, defined the closed conformation as Asp25-Ile50 distances of < 17 Å, semi-open conformations as 17-22 Å and the open conformation being anything > 22 Å.

RMSD and RMSF

Root-mean-square deviation (RMSD) calculations reveal the dynamic motion of all C_α atoms in the protein backbone and provide insight into the overall structural stability of a protein during the simulation. The RMSD data indicated that there were variations in the dynamic motion of the mutated proteases (HP3 and HP4) with respect to the WT PR. The RMSD values for the HP3 PR were notably greater than those of the WT PR and HP4 PR for most of the simulation (Figure 3A). The HP3 variant exhibited the highest RMSD with an average RMSD of 2.53 Å (Table 1). In contrast, the WT PR and HP4 PR exhibited average RMSD values of 2.22 Å and 2.17 Å, respectively (Table 1). These data indicate that the I13V, I62V and V77I mutations may decrease the overall structural rigidity of the HP3 PR resulting in greater dynamic motion associated with the backbone C_α atoms. However, inclusion of the L33F mutation appears to restore some stability resulting in reduced overall dynamics of the HP4 PR in comparison to both the WT and HP3 PRs.

Table 1. RMSD and RMSF data for WT PR, HP3 PR, and HP4 PR collected over the course of a 50 ns MD simulation.

Average RMSD	WT	2.22	± 0.36
	HP3	2.53	± 0.32
	HP4	2.17	± 0.22

		Hinge		Flap	
		Chain A	Chain B	Chain A	Chain B
Change in average RMSF wrt WT (%)	HP3	-10.88	-9.11	-30.55	-36.85
	HP4	-23.36	-14.60	-40.88	-48.55

		Cantilever		Fulcrum	
		Chain A	Chain B	Chain A	Chain B
Change in average RMSF wrt WT (%)	HP3	-16.89	-17.82	-16.27	-14.28
	HP4	-22.94	-37.43	-17.88	-27.54

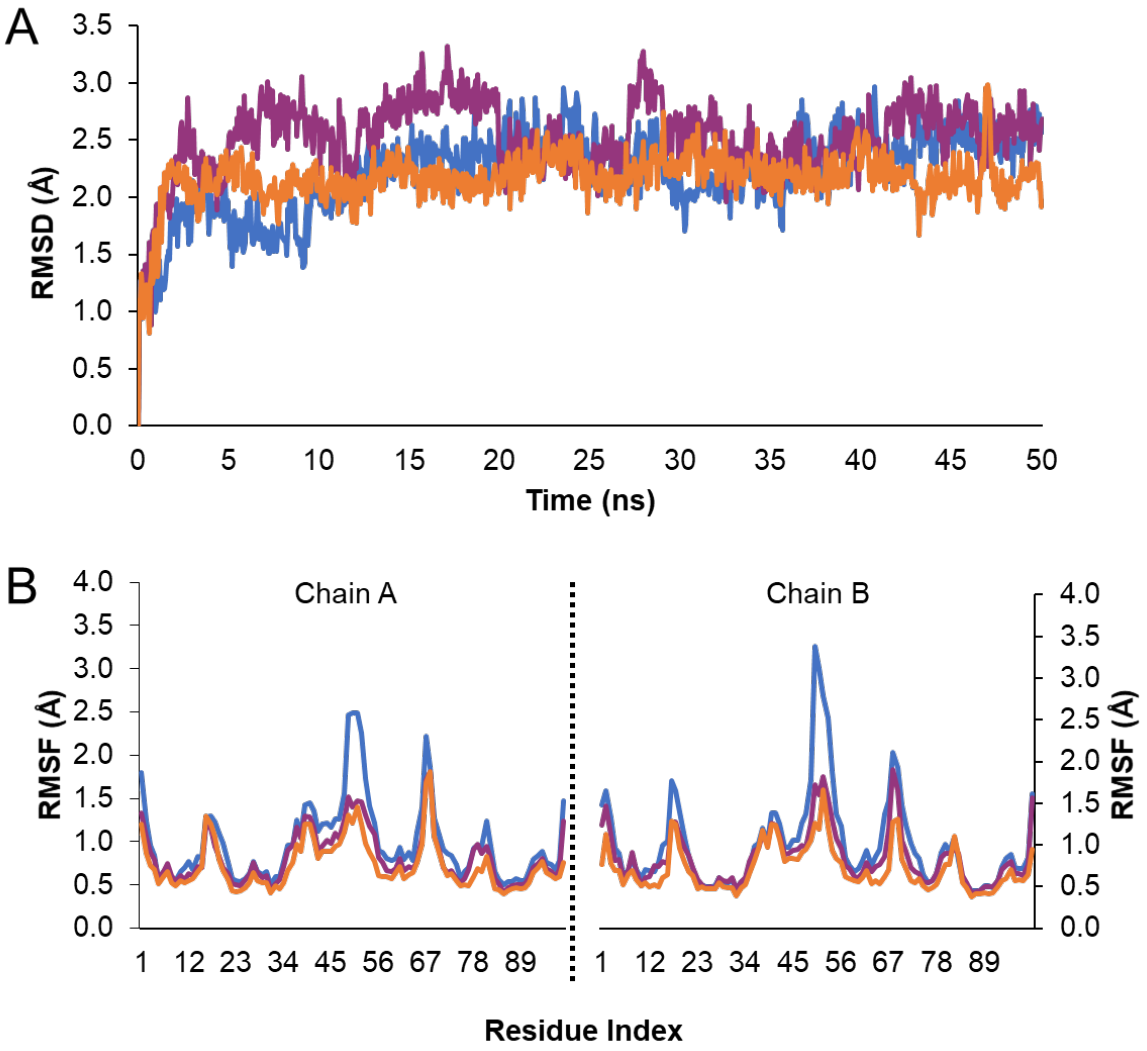


Figure 3 A) RMSD of the PR peptide backbone with respect to time. The RMSD of the backbone C_α atoms were calculated over the course of a 50 ns simulation for the WT (blue), HP3 (purple) and HP4 (orange) PR systems. The simulations were initiated with each PR variant in the closed conformation. B) RMSF of the PR structures with respect to amino acid number.

Furthermore, the root-mean-square fluctuation (RMSF) data from the simulations described the average deviation of the C_α of the amino acid residues from a reference point over time. The reference points for these calculations were the starting structures for each simulation in the closed conformation. Therefore, the RMSF provided insight into the degree to which regions of the PR variants fluctuated during the simulations. The regions of the WT PR that displayed the greatest RMSF values included amino acid residues 46-54 and 62-78 (Figure 3B), which encompass the flap region and cantilever region, respectively. A separate, but smaller peak in RMSF occurred at residues 35-42 which form the hinge region of the PR. The most striking difference in the average RMSF of HP3 PR and HP4 PR occurs in the flap regions where the average RMSF of the flaps decreased by 30-50% in the mutated PR variants (Table 1). Additionally, the average RMSF of the hinge and cantilever regions decreased for both mutated PR variants. Interestingly, the decrease in the average RMSF observed in the HP3 PR appears to be consistently augmented in the HP4 PR in these key regions (Table 1). These data indicate that in the apo form, HP3 PR and HP4 PR display higher degrees of structural rigidity in the hinge, flap, and cantilever regions in comparison to the WT PR.

The plotted RMSD curves (Figure 3A) indicate that the WT PR and HP4 PR exhibited lower deviations in their respective RMSD which could be indicative of greater resistance to structural alterations induced by the MD force field. Moreover, lower RMSD values may be attributed to more stable apo conformations of the PR. Although the RMSF data suggest that the amino acid residues within WT PR exhibited the greatest fluctuation, the average deviation of the backbone and, therefore, the overall dynamics was greatest for the HP3 PR. As noted previously, the flexibility of the flap regions of the mutants decreased up to 30-50% relative to WT PR (Table 1). Non-active site mutations that reduce flap flexibility can attenuate the affinity of the enzyme for natural substrates as well as PIs (Collins et al., 1995; Mahalingam et al., 2002, 2001). In-depth thermodynamic analyses of WT and PI resistant PRs suggest that drug resistance mutations can affect inhibitor binding indirectly by altering the relative stability of the apo (semi-open) and inhibitor-bound (closed) PR conformations upon inhibitor binding (Todd et al., 2000). Specifically, these mutations stabilise the semi-open conformation which lowers the binding affinity due to an increased energetic penalty associated with closing the flaps. Conversely, if these mutations destabilise the inhibitor-bound closed conformation, a smaller energetic penalty would be associated with the dissociation of the inhibitor. As such, it has been suggested that the free energy change associated with the conformational change of the PR must be included when accurately determining binding affinity (Todd et al., 2000).

Flap opening and flap tip curling

Upon substrate recognition and binding, the PR flap regions cover the substrate and then open again when the products are released. Additionally, the flap residues form interactions with the substrate and, therefore, play a key role in substrate (or inhibitor) recognition and binding (Gustchina and Weber, 1991). Given this scenario, distance measurements between the flap tips and Asp25 provides a metric of the exposure of substrates/inhibitors to the active site cavity. The distances between the flap tips and catalytic residues of each monomer were calculated for the WT PR and two PR variants over the course of the 50 ns MD simulations (Figure 4). The trajectory data show the altered MD between the WT, HP3 and HP4 proteases as a result of the non-active site mutations. The MD simulations revealed that the distance between the active site (Asp25) and the flap tips (Ile50) was significantly greater in the case of the HP4 PR when compared to the WT PR and HP3 PR. Interestingly, the flaps exhibited a semi open conformation (~20 Å) for HP4 in both monomers for the entirety of the simulation. However, both the WT PR and HP3 PR maintained a closed conformation (~15 Å) in both monomers for the majority of the simulations, with the exception of WT chain B which sampled the open conformation at the beginning and end of the simulation. The ability of the HP4 PR to stabilise and maintain a more open conformation would cause an increase in the volume of the active site cavity which could have a significant impact on drug binding.

Curling of the flap tips acts as a trigger for conformational changes associated with the PR shifting from both closed to semi-open and/or semi-open to open conformations (Freedberg et al., 2002; W. R. P. Scott and Schiffer, 2000). The importance of this event lies in the burial of the conserved hydrophobic Ile50/50' which maximises thermodynamic favourability when the flap regions open. In order to monitor flap tip curling during the simulations we analysed the angle produced between the C_α of three adjacent residues located in the flap tips; namely, Gly48/48', Gly49/49', and Ile50/50'. Interestingly, the flap tip angle data directly corresponded with the fluctuations observed with the opening of the flap regions. Which is to say, that the flap tips of the HP4 PR remained in the curled-in (~100°) conformation in both monomers for the entirety of the simulation. However, both the WT and HP3 PR maintained a curled-out flap tip conformation (~145°) for the majority of the simulation except in chain B of the WT PR which sampled curled-in conformations which were associated with the opening of the flaps. It appears that the mutations present in the HP4 PR allow this variant to sustain an active site cavity with a larger volume which is facilitated by an increase in the propensity of the flap tips to maintain the curled-in conformation. The curled nature of the flap tips is emphasised in Figure 4 (centre panel) where the average structure of the flap tips of both monomers of the PRs illustrated that the HP4 variant adopted an open conformation with a larger degree of flap tip curling. The significance of sustaining a wider flap conformation may lie in the energetic

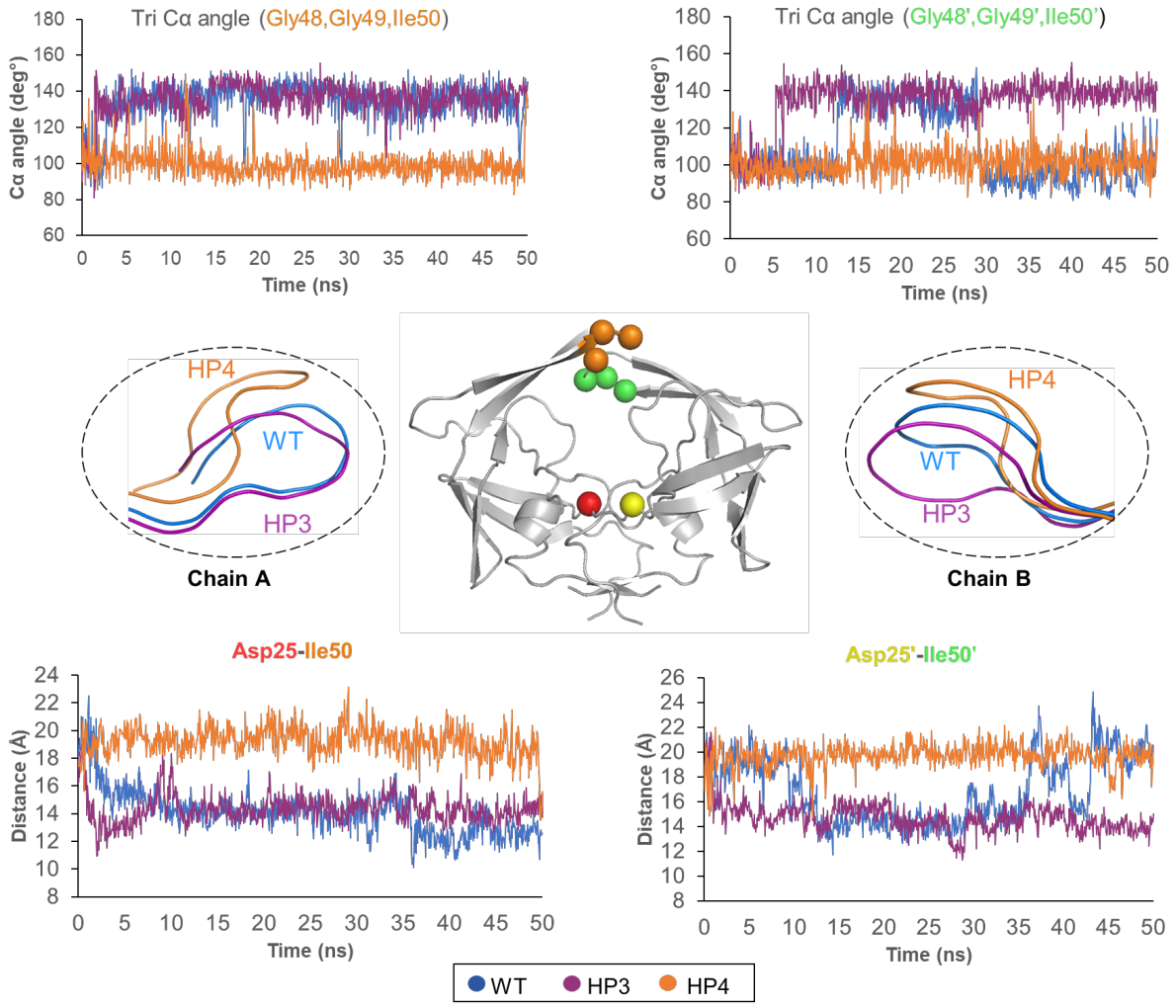


Figure 4. The non-active site mutations change the dynamic profile of the PR enzyme. The MD data are mapped around the HP4 PR which is depicted in the open conformation (centre). The average structures of the flap tip regions of the PR variants are enlarged and depicted aside the PR structure for clarity (presented in the dotted circles). The C_α atoms of key residues that were investigated are represented as spheres. For simplicity, the spheres were colour labelled according to the figure headings. Residues of chain B are represented by a prime (') symbol.

penalty associated with bringing the PR into the closed conformation for ligand binding. Since the HP4 apo PR sustained wider flap conformations, the energy associated with the backbone rearrangements needed to bring the flaps into a closed conformation would be greater than that of the WT. Therefore, such unfavourable changes may decrease the energetic favourability of inhibitor binding and decrease the binding association constant (K_a) of PIs by shifting the equilibrium in favour of open conformers that bind less favourably to PIs.

Sustained curling of the cantilever tips promotes opening of the flaps

As previously mentioned, the curling of the flap tips is an event necessary for the PR to undergo conformational changes from closed to semi-open and/or semi-open to open conformations (Freedberg et al., 2002; W. R. P. Scott and Schiffer, 2000). However, our current simulation data provide insight into a separate event that appears to be associated with the opening of the flap regions. That is, we analysed the angle produced between the C_α of a second set of three adjacent residues located in the tips of the cantilever region: namely, Ile66/66', Cys67/67', and Gly68/89'. Figure 6A illustrates the inverse correlation between the degree of curling associated with the tips of the cantilever and the distance between the flap tips and the active site catalytic residues. When the cantilever tip C_α angle is curled-in (≤100°) the flaps adopt a semi-open conformation while, in contrast, when this angle is curled-out (>100°), the flaps adopt a closed conformation. This concerted movement is evident in the WT PR during the first 30 ns of the simulation; from 0-10 ns the cantilever tips remain curled-in and the flaps sustain the open conformation. However, as the cantilever tips curl-out from 10-33 ns and 40-50 ns, the flaps assume a closed conformation (Figure 5A). Interestingly, HP4 PR maintained a curled-in conformation of the cantilever tip with an average C_α angle of 100° for the simulation which corresponded with the flaps sustaining a semi-open conformation. Although it is evident that flap tip curling precedes movement of the flaps, our findings suggest that the curling of the cantilever tips may play a similar role in the overall mobility of the flaps. These data suggest that cantilever tip angles between Ile66/66', Cys67/67', and Gly68/68' that are greater than 100° promote flap closure. The structural differences in the conformation of the cantilever tips and flap regions are highlighted in Figure 6B, where the semi-open conformation is associated with curling of the cantilever tips and a smaller angle between the C_α atoms of residues 66-68.

Protease region covariance analyses

The comparison of the correlated and anti-correlated fluctuations of the distinct regions of the apo PR provide insight into the interdependence of these regions during dynamic motions of the protein. The residue cross-correlation analysis of the WT PR illustrates that the hinge and flap regions of the PR exhibit anti-correlated movements (Figure 6). These findings support our understanding that the motion of the hinges regulate the stability and movement of the

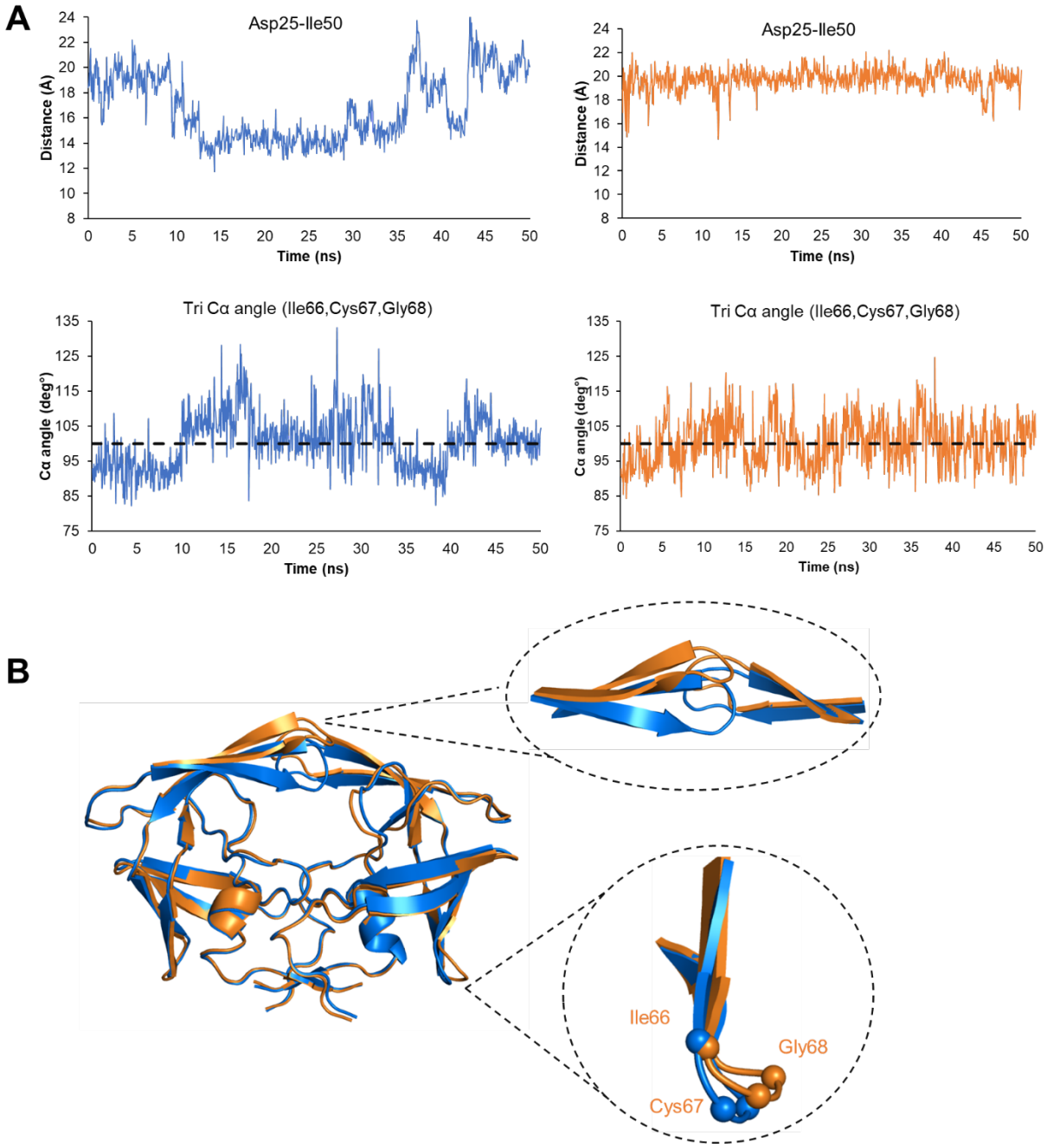


Figure 5. The curling of the backbone of the cantilever tips modulates flap opening. A) The angles formed between the C_α of Ile66, Cys67, and Gly68 residues were measured over the trajectory of both WT PR (blue) and HP4 PR (orange). The dotted black line indicates 100°. B) A structural alignment of WT PR (blue) and HP4 PR (orange) highlights key conformational differences in the degree of opening of the flaps and illustrates cantilever tip curling.

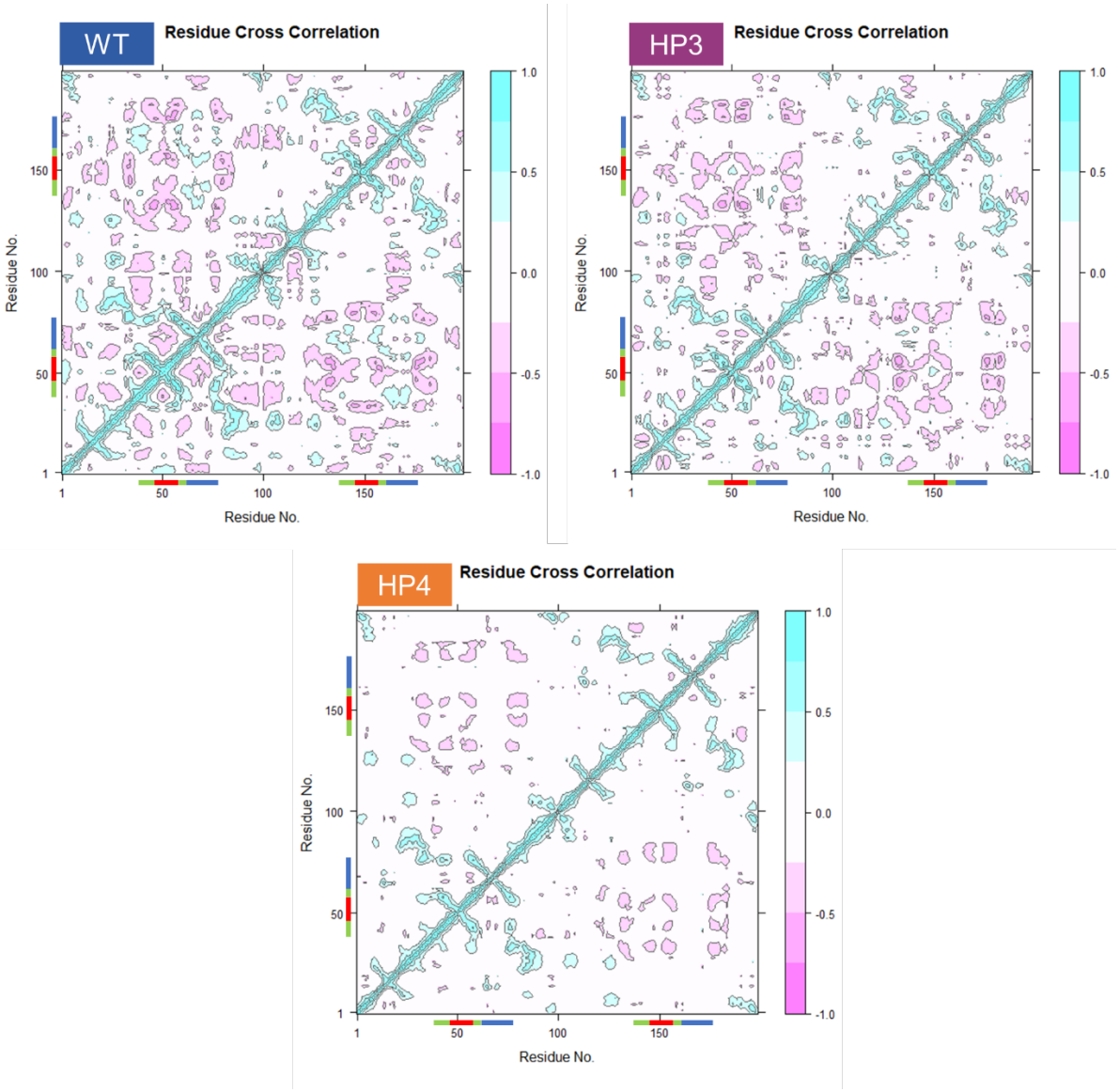


Figure 6. Cross-correlation maps showing the vectors that govern the alpha carbon atoms of the WT, HP3 and HP4 PRs. Amino acids that exhibit correlated (positive) movements are coloured blue while amino acids with anti-correlated (negative) movements are coloured pink. For easy comparison of the motions of key regions of the PR, the x- and y-axis possess colour coordinated bars which span the regions of the axis corresponding to the hinge (green), flaps (red) and cantilever (blue) regions. Image generated using the Bio3D package using the software environment R (Grant et al., 2006; Skjærven et al., 2014).

flaps (Gustchina and Weber, 1990). The upward movement of the flaps during opening relies on the concerted downward motion of residues that comprise the hinges, or conversely, during flap closing the hinges move upward to accommodate the motion of the flaps. Interestingly, a similar anti-correlated relationship is exhibited by the fluctuations of the flap region of the WT PR and the cantilever region (Figure 6). These data indicate that as the cantilever slides downwards towards the N- and C-terminus, the flap regions move upwards. This suggests that the downward motion of the cantilever regions is important in the mobility of the flaps as postulated in the hydrophobic sliding mechanism (Foulkes-Murzycki et al., 2007). The HP3 and HP4 PR exhibit similar patterns of correlation and anti-correlation to those of the WT PR (Figure 6). Most notably, however, the number of residues which exhibit atomic fluctuations becomes decreased in the HP3 PR and HP4 PR, as indicated by the decreased area of both the correlated (blue) and anti-correlated (pink) fluctuations in Figure 6. These findings suggest that the presence of the non-active site mutations in HP3 PR may serve to increase the rigidity of the apo PR, and that the L33F mutation may further augment this action in the HP4 PR. These findings agree with the decreased RMSF values represented in Table 1 which illustrate the increased rigidity of regions of the mutated PRs in comparison to the WT PR. An increase in the rigidity of the apo PR may lead to a larger energetic penalty associated with the opening of the flap regions.

Altered amino acid residue interactions

In order to better understand how the I13V, I62V, V77I and L33F mutations impart their effect on the PR structure and dynamics we studied the changes in the average distance between Cα atoms of the amino acid residues of the PR in both the WT and mutated PR systems. Ile13 is located in the fulcrum region on the PR and is theorised to play an important role in the hydrophobic sliding mechanism (Foulkes-Murzycki et al., 2007). The effect of the I13V substitution was revealed by the average distances observed between Val13 and neighbouring residues of the fulcrum “teens” (residues 13-19) β-strand: Val11, Ser12 and Val15 (Figure 7A). The substitution of isoleucine for a valine residue appears to compactify the fulcrum region by bringing neighbouring residues closer together to maintain the integrity of the hydrogen bond network involving these residues. As a result of the shift in the location of Ser12, the average distances between Ser12, Lys20 and Glu21 are also decreased which brings the second β-strand of the fulcrum region closer towards the “teens” β-strand. Thus, the propensity of the residues between the neighbouring β-strands of the fulcrum to form hydrogen bonds was increased as a result of the presence of Val13. In combination, the minimisation of distances within and between β-strands of the fulcrum region results in an increase in the stability of the fulcrum which can be attributed to the formation of hydrogen

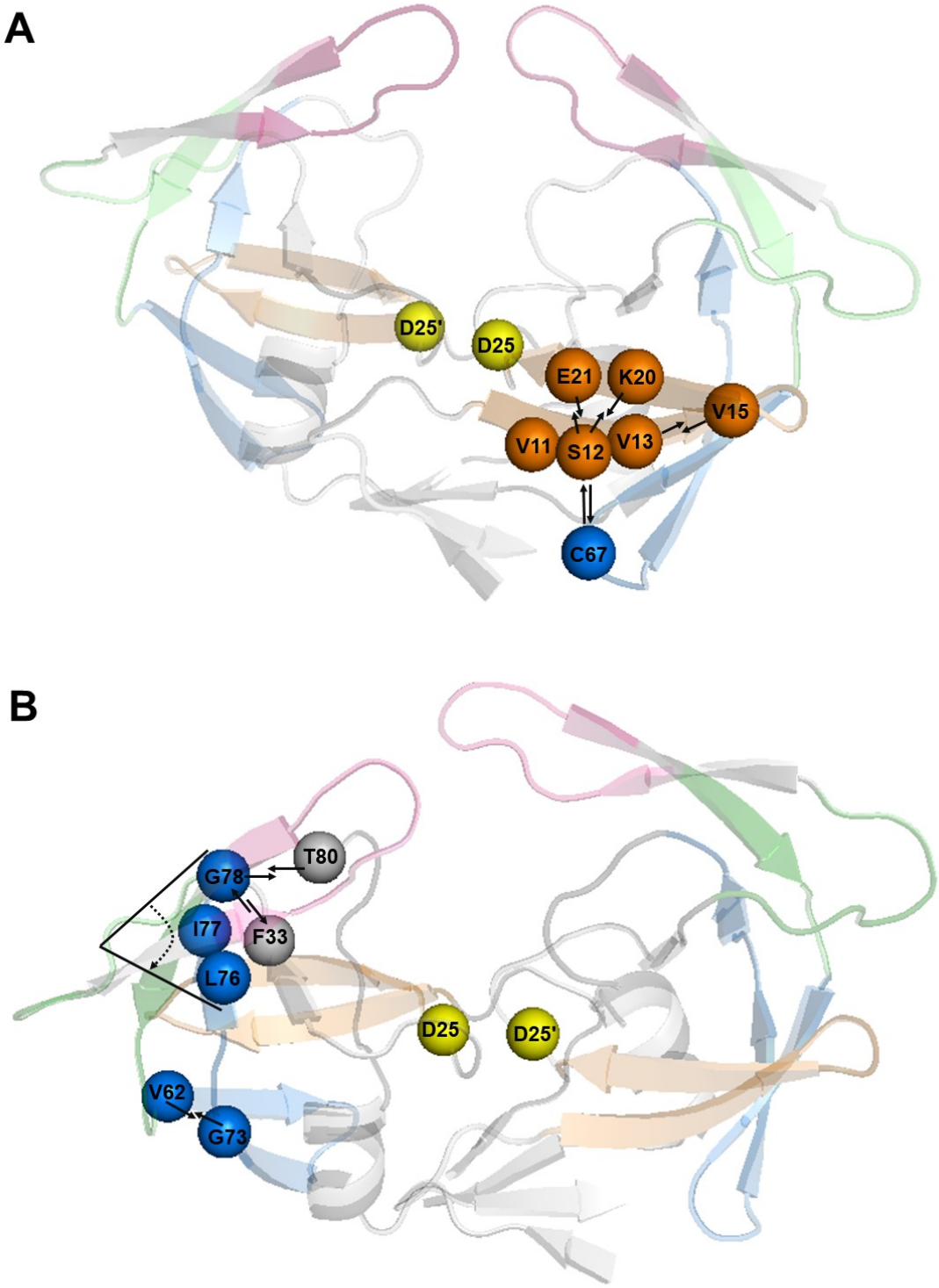


Figure 7. The effect of the A) I13V and B) I62V, V77I and L33F mutations on the Cα atoms of residues of the PR. The effect of the primary interactions of the mutated residues with their neighbouring residues, as well as notable secondary interactions that are altered are depicted. The black arrows indicate whether the residues are brought together (→←), move apart (←→) or exist in equilibrium (□). The decreased backbone Cα angle between residues 76-78 is highlighted.

bonds in the presence of the I13V mutation which is reflected in a decrease in the average RMSF of this region (Table 1). In addition, the hydrophobic sliding mechanism suggests that Val15 makes hydrophobic contacts with residues Leu38, Ile62, Ile64 and Val75 when the WT PR is in the closed conformation (Foulkes-Murzycki et al., 2007). When the flaps open, these contacts are transferred from Val15 to Ile13, which maintains the stability of the core while allowing for larger structural rearrangements associated with opening of the flaps. The mutation of Ile13 to the smaller, less flexible Val13 residue decreases the total hydrophobic surface area inside the core of the protein which could lead to a defective hydrophobic sliding mechanism. This, in conjunction with the more compact nature of the fulcrum, could result in a more stable core of the mutated PRs. There is evidence to suggest an increase in the stability of the core when examining the average RMSF of the hinge, fulcrum, and cantilever (Figure 3, Table 1). Furthermore, we noted that the average distances between Ser12 and Cys67, located in the cantilever, were unaltered in the mutated proteases which indicates that this interaction remains in equilibrium in all three structures. However, the distance between Ser12, Lys20 and Glu21 decreases in the mutated proteases, shifting Ser12 further away from Cys67. In order to maintain the equilibrium between Ser12 and Cys67, the cantilever tips adopt an altered backbone conformation which shifts Cys67 upwards, resulting in the curled cantilever tip conformation observed. It is unclear whether cantilever tip curling causes flap opening or vice versa; however, there is evidence that these motions are correlated (or anti-correlated).

The I62V and V77I mutations located in the cantilever have profound effects on the backbone of this region. For instance, the distances between Val62 and Gly73 were decreased in the presence of the I62V substitution (Figure 7B) which in turn brings the alternate β-strands of the cantilever closer together in order to stabilise this region of the mutated PRs. Moreover, the average C_α angle between Leu76, Ile77, and Gly78 is smaller in the mutated proteases (Figure 7B) which resulted in a reduced average distance between Gly78 and its neighbouring residues Glu34 and Thr80. The newly adopted conformation of the backbone region between residues 76-78 may be stabilised by the higher tendency for Gly78 to form interactions with Glu34 (hinge region) and Thr80 (loop region). Therefore, the I62V and V77I cantilever mutations may serve to stabilise the flap region of the HP3 and HP4 PR through a similar compactification of this region as is observed in the fulcrum region in the presence of the I13V mutation. This is, once again, reflected in a decrease in the average RMSF of this region (Table 1). However, these mutations alone do not sustain a wider flap conformation for the HP3 PR as is evidenced by the distances observed during the simulations (Figure 4). Interestingly, the average distance between Leu33 and Gly78 was increased for HP3 PR, thereby diminishing the propensity for these residues to form significant contacts, but with the

addition of the L33F substitution, the HP4 PR sustained a wider flap conformation. The increased bulk of the phenylalanine side chain residue could likely restore contact that is lost between Leu33 and Gly78 in the HP3 PR, and act as a scaffold which supports the flaps in a wider conformation. In summation, the I13V, I62V and V77I mutations serve to compactify the fulcrum and cantilever regions of the HP3 and HP4 PRs which may serve to stabilise these regions of the PR. As a result, the flap regions exhibit an increased stability which is exhibited in their reduced RMSF fluctuation (Figure 3). However, in the absence of the L33F mutation, the apo HP3 PR appears to be unable to maintain wide flap conformations. We postulated that the increased bulk of the phenylalanine residue at position 33 in the HP4 PR restores some stability to the cantilever region through the restoration of interactions with this region. The cantilever is closely associated with the hinge regions and the L33F mutation may provide the stability necessary to permit the PR to sustain a wider and open flap conformation.

The effect of fulcrum and cantilever mutations on PI-bound protease

Sustained wider flap conformations

The highly flexible flap regions of the PR adopt heterogeneous conformations; namely, closed, semi-open, and open conformations which exist in an equilibrium when the PR is in the apo state (Wlodawer et al., 1989). However, the inhibitor binding event is considered to be a "locking down" of the PR into a particular conformation, or ensemble of closed conformations, which exhibit minimal conformational entropy. The flap region, in particular, is stabilised as a result of the additional binding free energy from associations made with the inhibitor (Todd and Freire, 1999). As such, any mutations that decrease the favourable change in binding free energy of the inhibitor-bound complex may increase the dissociation constant (K_d) of inhibitors and, thus, decrease their efficacy (Maschera et al., 1996; Shuman et al., 2003). Furthermore, mutations that shift the equilibrium towards favouring more semi-open or open conformations may decrease drug binding efficacy by increasing the energetic penalty associated with bringing the flaps into the closed conformation. The radius of gyration (R_g) of the PR variants provides a measurement of the compactness of the proteins during the MD simulations. Should the PR become more dynamic in the presence of the non-active site mutations, the average R_g value of the HP3 PR and HP4 PR would be elevated. The average R_g values of the HP3 PR in complex with the PIs was found to be higher than those of the corresponding values observed in the WT PR complexes (Figure 8). Similarly, with the exception of the LPV-bound complex, the HP4 PR complexes displayed increased average R_g values in comparison the corresponding WT PR complexes (Figure 8). These findings suggest that the compactness of the mutated PR variants exhibit greater dynamics when in complex with PIs in comparison to the WT. With respect to the flaps of the PR, a wider flap conformation for a greater period

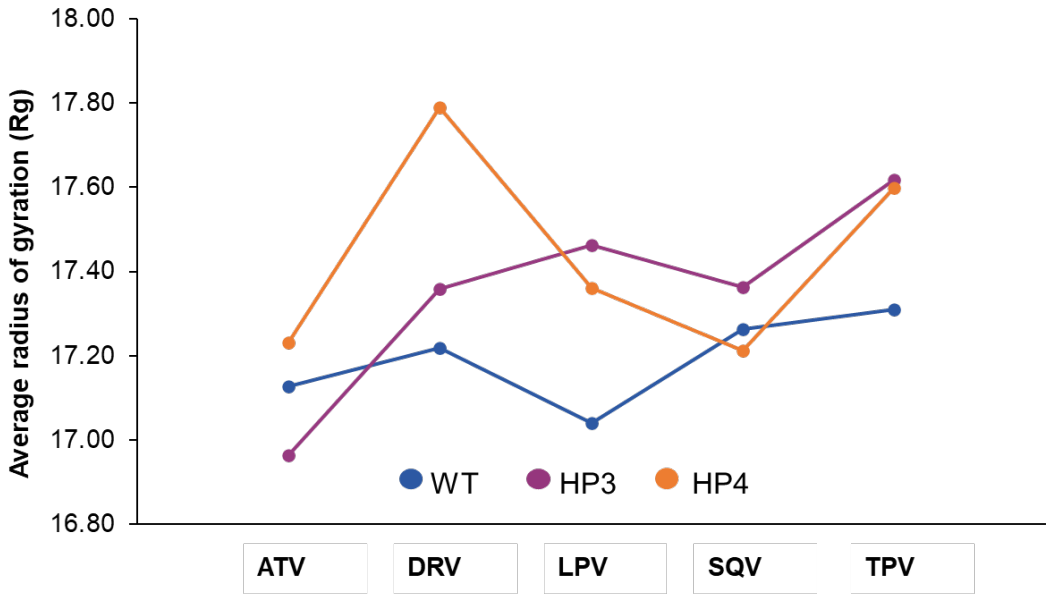


Figure 8. The average R_g of the WT PR, HP3 PR, and HP4 PR when in complex with clinical inhibitors. The average R_g was calculated for the PR variants in complex with SQV, ATV, DRV, LPV and TPV. These data were measured over the course of 50 ns MD simulations.

of time during simulations is associated with greater average R_g values (Chetty et al., 2016). The radius of gyration of the mutated inhibitor-bound PR complexes is exaggerated in comparison to the WT PR which could be indicative of greater dynamic motions of the mutant complexes. Therefore, it is noteworthy that when in the presence of the PIs, the PR variants containing the non-active site mutations displayed the tendency to sustain wider flap conformations in comparison to the WT PR (Figure 9). It is important to note that we define the flap conformations using the distance between the Cα atoms of Asp25/25' and Ile50/50' as has been performed previously (Perryman et al., 2004). The closed, semi-open and open conformations are defined by distances of < 17 Å, 17-22 Å and > 22 Å, respectively. In accordance with this, the WT PR remained in the closed conformation with an Asp25-Ile50 distance <17 Å when in complex with SQV, ATV and DRV. However, when in complex with LPV and TPV, the WT PR sampled both closed and semi-open (17-22 Å) conformations. PIs are designed as substrate transition-state analogues and, thus, they bind to the active site catalytic residues and bring the flaps into the closed conformation (Tóth and Borics, 2006). Indeed, this phenomenon is observed in the majority of three-dimensional PR structures deposited to the Protein Data Bank (PDB), which exhibit closed flap conformations when in complex with substrates or inhibitors.

We postulated that mutations that widen the conformation of the PR flaps when bound to inhibitors could result in reduced inhibitor susceptibility through reducing the formation of intermolecular interactions such as hydrogen bonds and/or van der Waals contacts. Therefore, it is significant to observe that the HP3 PR and HP4 PR possess the ability to sustain greater flap distances in comparison to the WT PR when in complex with PIs. The WT complexes favourably sample closed conformations for the majority of the 50 ns simulations with flap distances < 17 Å (Figure 9). The mutated proteases, however, favour semi-open (17-22 Å) and, in some cases, open (>22 Å) flap conformations (Figure 9). Due to the rigid nature of the PIs, any increase in the interatomic distance between the inhibitor and the flaps or the inhibitor and the catalytic residues would not be easily compensated for. Therefore, the greater propensity that the mutated proteases exhibit for wider flap conformations could have a significant impact on drug susceptibility and the manner in which the inhibitors interact with the active site and flap residues. Interestingly, we noted that curling-in of the cantilever tips of the WT PR occurred when the flaps assumed the closed conformation when in complex with PIs (Figure 8). While, in contrast, the semi-open flap conformations of the HP3 PR and HP4 PR were associated with the cantilever tips curling-out when in complex with PIs; particularly, SQV, ATV, and DRV (Figure 9). The findings are of particular interest given that in the apo form curling-in of the cantilever tips was associated with sustaining the semi-open flap conformation (Figure 4). Again, our findings suggest that the mobility of the cantilever tips may

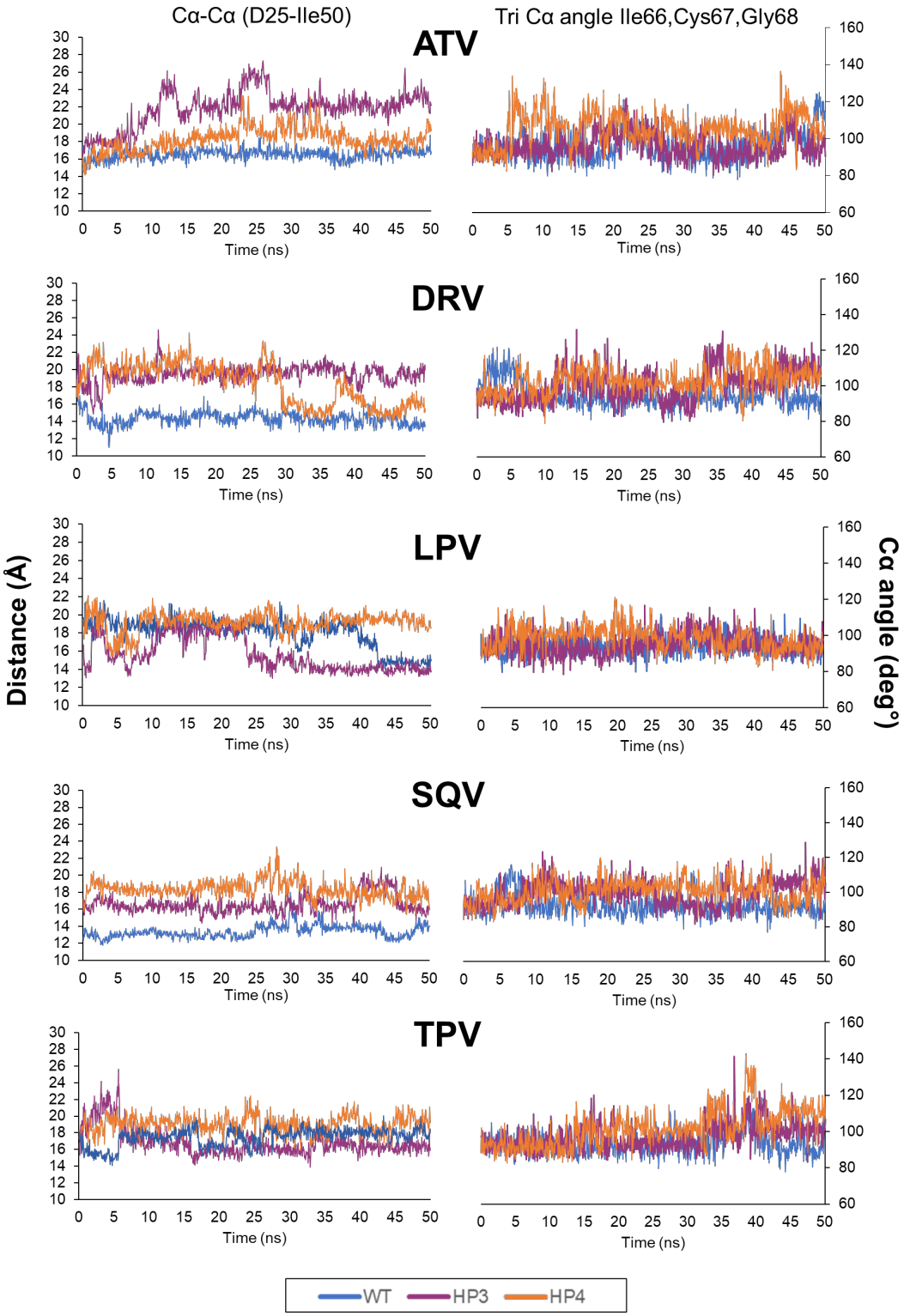


Figure 9. Interatomic distances measured between the C_α atoms of the flap tips (Ile50/50') and active site residues (Asp25/25') (left) and the angles measured between the C_a atoms of Ile66/66', Cys67/67', and Gly68/68' of the cantilever tip (right). These measurements were performed for the WT, HP3 and HP4 PR variants in complex with SQV, ATV, DRV, LPV and TPV. These data were measured over the course of 50 ns MD simulations, beginning with the PR variants in the closed conformation.

play a significant role in the dynamic motions of the PR, and that the curling-in or -out of the cantilever tips is an important event that relates to the flap conformation.

As such, developing allosteric inhibitors that can modulate or immobilise the cantilever tips of the PR may prove to be a promising avenue for future PIs that could be co-administered with active site inhibitors, to further increase the efficacy of drug regimens.

Increased protein and ligand fluctuation

The stability and motion of the flap regions is known to be closely associated with, and largely dependent on, the substantial flexibility observed in the hinge regions (Gustchina and Weber, 1990). Moreover, the concerted movement of the hinges, fulcrum, and cantilever regions is proposed to precede shifting the flap conformation from the closed to the open conformation (Foulkes-Murzycki et al., 2007; Hornak et al., 2006). Mutations that modify the flexibility of the hinges and these regions of the core of the PR are, therefore, likely to translate to changes in the mobility of the flaps. Moreover, our data suggest that curling-in of the cantilever tips allows the apo PR to sustain semi-open flap conformations whereas curling-out of the cantilever tips allows the flaps to relax into a closed conformation. Consequently, we hypothesised that changes in the flap dynamics rely not only on the mobility of the hinges but, additionally, their movements can be directly affected by changes in the mobility of the cantilever.

The RMSF data of the inhibitor-bound PR variants illustrates any changes in the degree of fluctuation of key regions of the PR complexes which result from the presence of the non-active site mutations. The RMSF data of the WT PR in the absence of an inhibitor illustrated that the flap regions of both peptide chains are the most mobile regions of the PR, exhibiting average fluctuations of ~2.5 Å (Figure 3). When in complex with DRV and ATV the flap regions of both chains fluctuate to a lesser extent with average fluctuations of ~1 Å (Table S1). Moreover, while bound to SQV, LPV, and TPV the average fluctuations of the flaps were ~1.5 Å, ~1.4 Å and ~1.6 Å, respectively (Table S1). The decrease in the fluctuation of the flaps in comparison to the apo PR is observed since inhibitor binding is promoted by the formation of key interactions with the residues of the flaps (46-MIGGIGGFI-54) when the flaps are in the closed conformation (Gustchina and Weber, 1991). The dissimilar average fluctuations of the flap regions of the PR (Table S1) are likely a result of the asymmetric nature of PIs, which would result in the formation of dissimilar interactions with each monomer depending on the orientation of the drug molecule in the active site. Relative to the WT PR, both the HP3 PR and HP4 PR displayed an increased change in their average flap distances when in complex with ATV, DRV and SQV (Table 2). Similarly, the flap regions of HP3 PR and HP4 PR fluctuate more readily than those of the WT when in complex with these three PIs, with the exception of the HP3 PR in complex with SQV (Table 3).

Table 2. The increase/decrease (%) in the average Asp25/25'-Ile50/50' distance of the HP3 PR and HP4 PR with respect to the WT PR when in complex with PIs

	PR chain	ATV	DRV	SQV	TPV	LPV
HP3	A	32	5	24	-5	-13
	B	24	36	6	-3	3
HP4	A	11	16	37	-4	6
	B	23	28	19	9	2

Table 3. The increase/decrease (%) in the average RMSF of key regions of the HP3 PR and HP4 PR with respect to the WT PR when in complex with PIs.

		Hinge region		Flap region		Cantilever region	
	Protease inhibitor	A	B	A	B	A	B
HP3	**SQV**	11	-14	-39	-20	12	-4
	ATV	59	39	154	69	23	0
	DRV	24	9	56	29	-5	3
	LPV	-8	0	-4	0	-2	-7
	TPV	-23	-7	-48	23	-13	-16
HP4	**SQV**	31	31	20	30	29	44
	ATV	14	20	81	18	21	12
	DRV	24	33	53	116	25	20
	LPV	9	0	-12	52	6	-7
	TPV	-8	-7	-32	23	2	-16

The HP3 PR displayed increased mobility in the hinge and flap regions when in complex with ATV and DRV in comparison to the WT PR (Table 2, Figure 3). The flap regions of the HP3 PR in particular exhibited an increase in the average RMSF as large as 154% and 56% when in complex with ATV and DRV, respectively (Table 3). The HP4 PR displayed a global increase in the average RMSF of the hinge, flap, and cantilever regions when in complex with SQV, ATV and DRV. That is, the flaps of the HP4 PR demonstrated an increase of as much as 30%, 81%, and 116% in their average fluctuations when in complex with SQV, ATV, and DRV, respectively (Table 3). The hinge regions of the mutated PR variants displayed a tendency towards greater mobility than those of the WT PR, with the exception of the LPV- and TPV-bound mutant complexes (Table 3).

Interestingly, when in complex with the PIs, the cantilever region of the HP4 PR showed a tendency to be more mobile than that of the WT PR, but this trend was not observed to the same degree in the HP3 PR complexes. These average RMSF data suggest that the mobility of the flaps may not only rely on the flexibility of the hinges, but the cantilever regions as well. Our evidence suggests that these mutations in the fulcrum and cantilever regions increase the mobility of these regions when in complex with PIs (Table 3). The increased flexibility of these regions correlated with an increase in the mobility of the flap regions of the PR. As mentioned previously, the mobility of the flaps is closely related to that of the hinges; however, flexibility of the cantilever may be equally important in the dynamics of the PR. Drug resistance may arise, at least partly, from these changes in the flap conformation seemingly due to the inability of current PIs to induce sufficient flap closing (Hornak and Simmerling, 2007).

In order to determine how the changes in flap flexibility affected the motion of the PIs in the active site we examined the ratios of the RMSF data between the HP3 PR and HP4 PR with respect to the WT PR (Figure 10). Here, $\Delta RMSF_{PI}$ values that were >1 indicate higher mobility of the PI while values <1 indicate reduced mobility of the PI in comparison to the WT PR complexes. The $\Delta RMSF_{PI}$, data for SQV, LPV and TPV for the HP3 complex were <1 for 66%-100% of the atoms of the PIs and which is indicative that these drugs were largely restricted in the active site of the HP3 PR. These findings correlate with the mobility of the flap regions of HP3 PR which were less mobile than those of the WT PR when in complex with these inhibitors (Table 3). Conversely, the $\Delta RMSF_{PI}$ data for ATV, DRV were >1 for and 67% and 100% of the atoms of the PIs which correlated with exaggerated flap movement and, therefore, these PIs are more mobile in the active site of HP3 PR. The DRV-bound HP3 complex in particular exhibited the greatest increase in mobility out of all of the inhibitors examined. With the exception of TPV, the PIs displayed a larger degree of motion within the active site of the HP4 PR in comparison to the WT PR with increases in fluctuation of between 76%-100% of

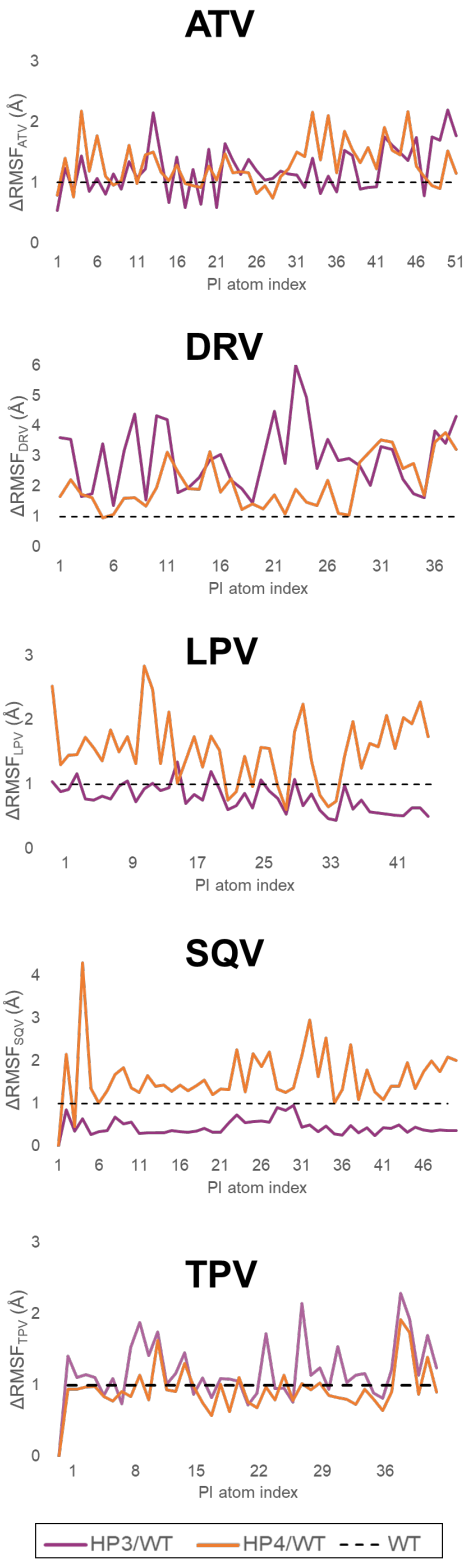


Figure 10. RMSF ratio (mutant/WT) of the atoms of each PI when in complex with the PR variants. The WT RMSF data for each PI serves as a reference point to which the RMSF data for the HP3 and HP4 complexes are compared. The RMSF of the PIs in complex with WT is standardised to a value of 1 (dotted line). An increase/decrease in RMSF when in complex with the mutant PRs is represented by values >1 and <1, respectively.

the PI atoms. These findings corresponded with the higher mobility of the HP4 flap regions which was observed in these complexes (Table 3).

Furthermore, changes in the solvent accessible surface area (SASA) of the PIs when bound to the PR variants correlated with the increased mobility of the inhibitors in the active site. The SASA of the HP3 PR complexes was greater than that of the corresponding WT PR complexes (Table 4), with the exception of LPV. This increase in SASA was further increased for the HP4 PR when in complex with DRV, SQV, and TPV. The burial of hydrophobic moieties of PIs within the active site of the PR excludes solvent interactions with the PIs which, in doing so, contributes to the entropic favourability of inhibitor binding (Kožíšek et al., 2014).

Table 4. The average SASA (Å2) of the PIs when in complex with WT PR, HP3 PR and HP4 PR.

	Average PI SASA (Å^2)				
	ATV	**DRV**	**LPV**	**SQV**	**TPV**
WT	162	34	127	109	95
HP3	219	126	98	160	194
HP4	216	151	113	263	214

Thus, mutations which increase the SASA of the PIs when in complex with the mutant PRs are likely to diminish binding efficacy by decreasing the entropic favourability of the interaction. Collectively, these findings suggest that when the PR exhibits increased flap flexibility, the PI bound to the active site experiences a greater degree of mobility within the active site (Figure 10). Therefore, exaggerated flap and cantilever mobility can be implicated in an increase in the mobility of the inhibitors when bound to the active site of the PR.

The PIs currently employed in ARV therapy suppress the flexibility of the flaps (Ishima et al., 1999; Nicholson et al., 1995). Under the selective pressure of these inhibitors, HIV selects for PR mutations that reduce the ability of the inhibitors to suppress movement of the flaps. The viral fitness of the mutant PR is further maintained through changes in the flexibility and dynamics of the natural Gag and Gag-Pol substrates which coevolve corresponding mutations at the sites where PR performs cleavage (Doyon et al., 1996; Mammano et al., 1998). Since the competitive advantage of inhibitors likely arises from their structural rigidity, it is difficult to design flexible inhibitors that can bind with a higher affinity than the natural substrates of PR, while still retaining the ability to adapt to variations in the active site geometry. Here, it is evident that these non-active site mutations increase the flexibility of the flap, hinge, and cantilever regions (Table 2.2) which results in an increased mobility of the drug within the

active site (Figure 10). Such changes in the dynamics of the inhibitor-complexed PR systems could result in diminished, or even eliminated, intermolecular interactions. Since the strong binding (in the nM range) of PIs is mainly driven by a favourable enthalpy change (Velazquez-Campoy et al., 2001a; Velazquez-Campoy and Freire, 2001b), mutations that minimise inhibitor-protein contacts would minimise binding enthalpy and result in decreased inhibitor efficacy. Moreover, the increased mobility observed in the inhibitor-bound complexes results in an increase in the SASA of the PIs (Table 3) in the active site. This could further decrease the favourability of binding by reducing the hydrophobic surface area buried upon PI binding, leading to reduced binding favourability and, therefore, reduced PI susceptibility.

Altered drug-binding interactions

Our data established that the PIs are more mobile within the active site of the mutated PR variants and that the flap regions of these PR variants are more flexible in comparison to the WT when bound to certain inhibitors. In addition, the flaps of both PR variants adopt wider conformations than the WT PR when bound to the PIs. These changes will have significant effects on the number and nature of the intermolecular interactions that the active site and flap residues form with the inhibitors during binding. PIs are designed to target the active site in the closed conformation and to suppress flexibility of the flaps, as mentioned previously. As such, we investigated how the mutations changed the location of the drug in the active site as well as the difference in the interactions formed between the inhibitors and the PR variants. In all inhibitor-complexed forms, the WT PR forms interactions with the inhibitors via one or more active site residues (Asp25, Gly27, Asp29, and/or Asp30) as well as with residues of the flaps (Gly48, Gly49, and/or Ile50). In the WT PR, the substrate or inhibitor forms interactions with these key residues in both monomers (Table 5). The utilisation of these residues is consistent with our understanding of the substrate-envelope hypothesis (See Ch. 5, pg. 134, Primary active site mutations) and the residues that are known to be associated with substrate and inhibitor binding (Nalam et al., 2010; Prabu-Jeyabalan et al., 2002; Wlodawer and Erickson, 1993). Briefly, the substrate binding specificity of the PR appears to be determined by an asymmetrical shape of the substrate rather than a particular amino acid sequence.

The PR can bind several substrates that possess different amino acid sequences. These various substrates occupy a consensus volume within the active site and form a number of conserved hydrogen bonds with the PR; namely, with residues Asp25/25', Gly27/27', Asp29/29', Gly48/48' and a partial hydrogen bond conservation involving Asp30/30'. Additionally, a conserved water molecule mediates interactions with Ile50/50' and the PIs (Prabu-Jeyabalan et al., 2002). The effect of the non-active site mutations which are present in the HP3 PR result in similar changes in the nature of interactions formed between HP3 and the PIs investigated, with the exception of TPV. A common mechanism of binding that occurs

with PIs involves the formation of hydrogen bonds between the hydroxyl group of the inhibitor and the carboxyl group of the active site Asp25/25′ residues (Lv et al., 2015).

Table 5. The type of major interactions formed between PIs and key residues in the PR active site and flap regions.

	Active site								Flaps					
WT	Asp25	Asp25'	Gly27	Gly27'	Asp29	Asp29'	Asp30	Asp30'	Gly48	Gly48'	Gly49	Gly49'	Ile50	Ile50'
ATV	H bond	H bond	Water	H bond	Water	H bond	Water	H bond	H bond	Water	-	-	HI	HI
DRV	H bond	H bond	-	-	Water	H bond	Water	H bond	Water	-	-	H bond	-	-
LPV	H bond	-	Water	-	H bond	-	H bond	-	-	-	-	-	-	Water
SQV	H bond	-	Water	-	Water	-	-	-	-	-	-	-	H bond	-
TPV	-	-	-	-	-	Water	-	Water	-	H bond	-	-	-	Water
	Active site								Flaps					
HP3	Asp25	Asp25'	Gly27	Gly27'	Asp29	Asp29'	Asp30	Asp30'	Gly48	Gly48'	Gly49	Gly49'	Ile50	Ile50'
ATV	Water	Water	-	-	H bond	-	H bond	-	Water	-	-	-	-	-
DRV	Water	Water	-	-	-	Water	-	Water	H bond	-	-	-	H bond	-
LPV	H bond	Water	H bond	-	H bond	-	Water	-	-	-	-	-	-	HI
SQV	Water	-	Water	H bond	Water	H bond	-	-	-	-	-	-	HI	HI
TPV	-	-	-	-	-	H bond	-	H bond	-	Water	-	-	-	-
	Active site								Flaps					
HP4	Asp25	Asp25'	Gly27	Gly27'	Asp29	Asp29'	Asp30	Asp30'	Gly48	Gly48'	Gly49	Gly49'	Ile50	Ile50'
ATV	Water	Water	-	Water	-	-	-	H bond	Water	Water	-	-	HI	HI
DRV	-	-	-	-	-	-	H bond	-	H bond	Water	Water	-	Water	Water
LPV	H bond	H bond	-	Water	H Bond	-	H bond	-	-	-	-	-	-	HI
SQV	-	-	-	-	-	H bond	-	-	Water	-	Water	-	Water	Water
TPV	-	-	-	-	-	H bond	-	-	-	Water	-	-	-	-

Hydrogen bonds Hydrophobic Interactions (HI) Water bridges

These conserved hydrogen bonds are observed during the WT PR simulations, but, are notably absent in the HP3 PR simulations (Table 5). Rather, the HP3 PR requires the incorporation of water molecules to form water-bridges between the active site residues and the inhibitor which spontaneously enter the complex during the simulations. The effect of these mutations on inhibitor binding is most notably exemplified when examining the active site of the WT PR and HP3 PR when in complex with SQV (Figure 11), but can be seen in the ATV- and DRV-bound complexes as well (Figure S3). Since the flaps adopt a wider conformation in the HP3 PR, the inhibitor is located higher up in the active site. In this position, the inhibitor is beyond hydrogen bonding distance with the Asp25/25' residues which necessitates water-bridges to simultaneously maintain contacts with these residues while still making significant contacts with the flap residues (Table 5, Figure 11). We postulate that these water molecules rotate and slide across each other in the active site, such that these water-bridges are more transient than those that would form directly between the inhibitor and the protein. In effect, the water-bridges would diminish the enthalpic favourability of inhibitor binding by diminishing the propensity for which hydrogen bonds are formed and maintained between the active site and inhibitor. As such, the wider flap conformations

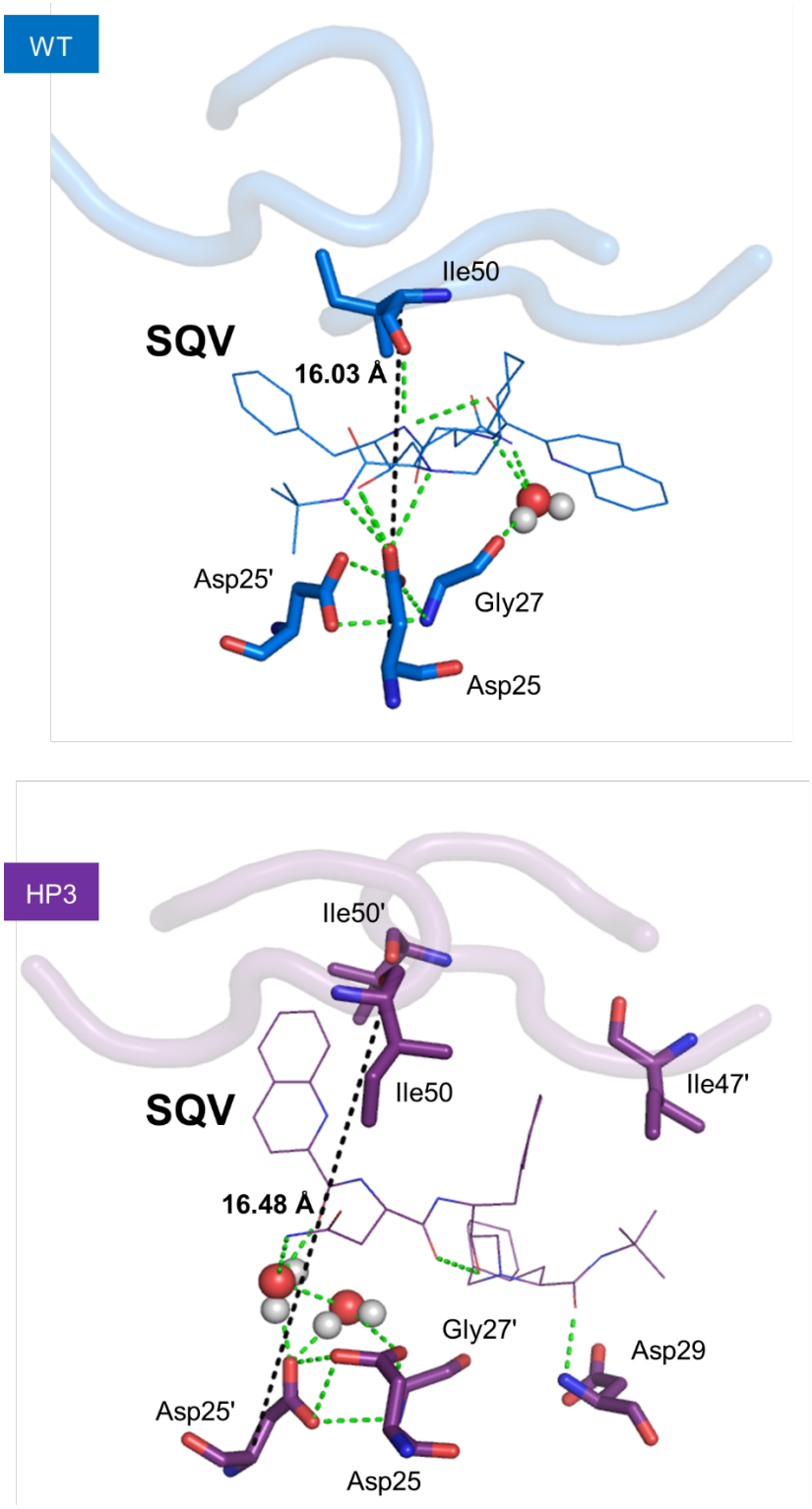


Figure 11. Protein-ligand interaction diagrams of WT PR and HP3 PR in complex with SQV. Specific and nonspecific protein-ligand interactions were monitored throughout 50 ns simulations to identify the most prominent interactions formed. Hydrogen bonds are denoted by dotted green lines. Water molecules associated with the active site are represented as spheres. The interatomic C_α distance between Asp25 and Ile50 is denoted by the dotted black line, and the distance is labelled.

adopted by the HP3 PR may decrease inhibitor susceptibility by diminishing interactions between the active site Asp25/25' residues and the inhibitor.

In fact, replacing conserved water-bridges that form between currently available inhibitors and the enzyme is a proposed method of making more effective and specific PIs (Swain et al., 1990). The observed displacement of the inhibitors within the active site of the PR is likely to contribute to inhibitor resistance by hindering the mechanism of action of the PIs, which is, to associate with the active site residues and, simultaneously, bring the flaps into the closed conformation. The HP4 PR contains the same non-active site mutations present in the HP3 PR, with an additional L33F mutation. In a similar manner to the HP3 PR, the changes observed in the residues associated with PI binding follow a similar trend for all PIs with the exception of LPV. The HP4 PR exhibits a significantly higher propensity towards the formation of water-bridges between the inhibitors and the flap residues. Again, these water molecules associated with the flap residues spontaneously enter the active site during the course of the simulations. Moreover, markedly few of the conserved hydrogen bonds are formed with the residues of the active site, particularly Asp25/25' and Gly27/27' (Table 5). Only LPV was able to form hydrogen bonds with the catalytic residues of HP4 PR, while ATV sustains contact with the catalytic residues via water-bridges. In contrast DRV, TPV and SQV formed no contacts with the catalytic residues during the simulations (Table 5, Figure S2). However, TPV forms no hydrogen bonds with the Asp25/25' residues in any of the PR variants studied which indicates that this observation is a result of the mechanism of binding of TPV rather than the impact of the mutations.

The effect of the addition of the L33F mutation is demonstrated by the difference in the active site of the WT PR and HP4 PR when in complex with DRV (Figure 12). The HP4 PR complexes were more reliant on hydrogen bonds between the inhibitors and either Asp29/29' or Asp30/30' residues, typically from a single monomer (Table 5). These hydrogen bonds formed tentative anchors for the inhibitors to remain in contact with the active site; however, the majority of the stability of the complexes relied on hydrogen bonds, water-bridges, and hydrophobic interactions with the flaps (Table 5).

These observations mirror the impact that the I13V, I62V and V77I mutations impart on the HP3 PR and correspond with the increased flap distance sustained by the HP4 PR during the simulations (Table 2). However, a significant difference between the HP3 PR and HP4 PR complexes is evident in the flap region. Here, the HP4 PR exhibited the propensity to coordinate more than a single water molecule that is capable of forming water-bridges between the flaps and the inhibitors. These observations are in contrast to the fact that a single conserved water molecule is necessary for PI interactions with the flap tip residues within WT

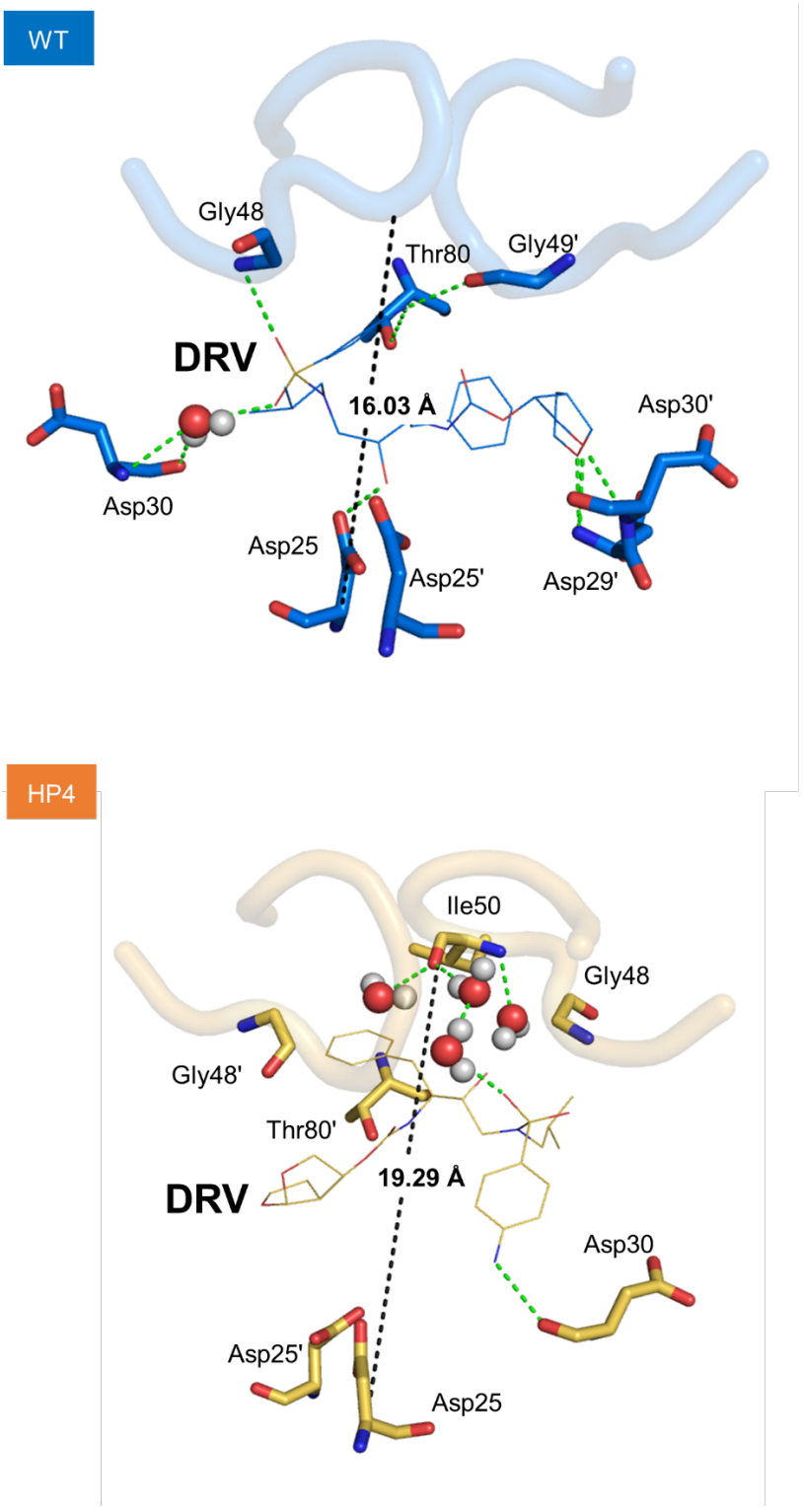


Figure 12. Protein-ligand interaction diagrams of WT PR and HP4 PR in complex with DRV. Specific and nonspecific protein-ligand interactions were monitored throughout 50 ns simulations to identify the most prominent interactions formed. Hydrogen bonds are denoted by dotted green lines. Water molecules associated with the active site are represented as spheres. The interatomic C_α distance between Asp25 and Ile50 is denoted by the dotted black line, and the distance is labelled.

PR (Wlodawer and Erickson, 1993). The multiple water molecules associated with the flaps are likely to have a similar effect on binding, to those which are associated with the active site catalytic residues.

Particularly when multiple water molecules are present in the flap region, hydrogen bonds are readily swapped between these water molecules which could increase the overall dynamics of the complex. Furthermore, the flap regions are more flexible than the active site and, thus, the effect of these water molecules is likely to be further augmented by the inherent flexibility of the flaps. Therefore, the effect of the addition of the L33F mutation appears to be two-fold. Firstly, the L33F mutation stabilises the flaps in an open conformation with curled flap tips which reduces contact with the active site residues and can eliminate contact with the Asp25/25' residues all together. Secondly, this mutation serves to stabilise the curled flap tip conformation in a manner that coordinates multiple water molecules into the flap region which associate with the inhibitor and, most likely, increase the dynamics of this region of the complexes.

Changes in the active site geometry

There is evidence from these simulations to show that the active site of the PR may be expanded by the presence of the I13V, I62V, and V77I non-active site mutations. The wider flap conformation adopted by the flaps would undoubtedly increase the volume of the active site. As a result, the PIs are located higher up within the active site of the mutated proteases and some associations with the active site residues are lost. Moreover, the mutated proteins, in particular HP4, displayed an increased propensity towards adopting curled-in flap tips. These observations, therefore, led us to investigate how the combination of these two effects altered the geometry of the active site.

Through investigating the molecular surface of the active site of the proteases, it was evident that the HP3 PR and HP4 PR were able to sustain larger active sites than the WT PR when in complex with ATV, DRV, and SQV (Figure 13). When in complex with LPV and TPV, the mutated proteases have similarly sized active sites in comparison to the WT PR, but the flaps of the mutants produce pockets in the active site above the inhibitors (Figure S4). Clearly, the inhibitor would be unable to bind to the active site residues and the flap residues simultaneously without the incorporation of water molecules to fill the increased volume. Therefore, the increased volume of the active site of the mutants clarifies the need for water-bridge formation. In order to better understand the effect of the flap tips curling-in, the presence of water molecules in the flap tip region of the PR variants was further examined. The WT PR contained a water molecule between the flap tips which is conserved amongst inhibitor-bound complexes and is believed to be crucial in inhibitor binding (Wlodawer and Erickson, 1993).

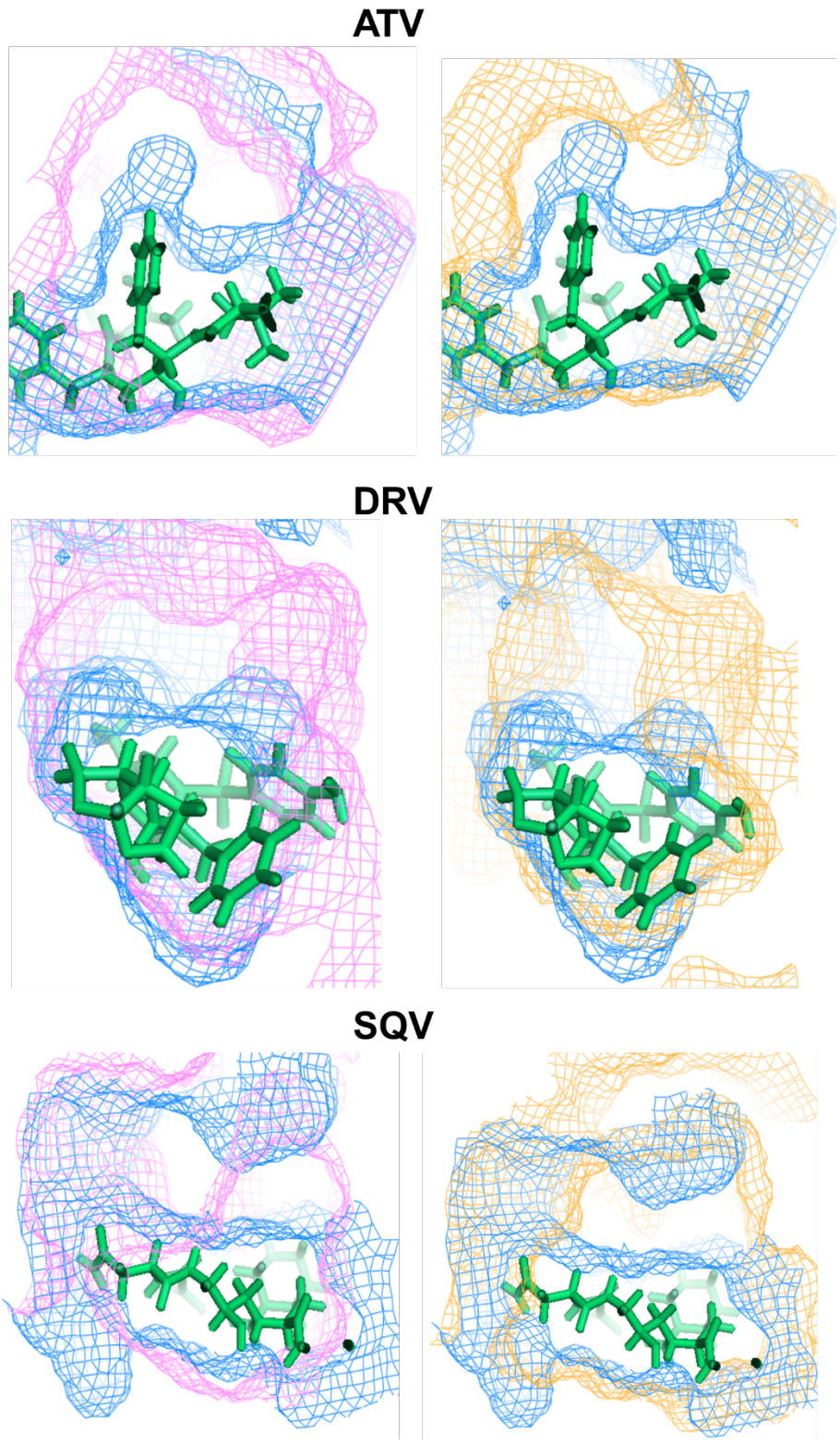


Figure 13. Molecular surface renderings of the active site cavities of the PR variants in complex with PIs. The molecular surface of the PR is represented as a mesh and colour coded (blue = WT, purple = HP3, orange = HP4). The PI (green) bound to the active site in the WT complex is represented as sticks.

The flap regions of the HP3 PR do not contain a water molecule between the flap tips but, rather, the flaps take on an expanded conformation which produces a large cavity at the top of the active site.

The cavity produced by this change in conformation is spontaneously filled by water molecules (Figure 14). Similarly, the HP4 PR adopts a conformation that produces an expansion of the active site cavity, but the flaps of the HP4 PR create a well-defined pocket in a manner that HP3 PR is unable to do (Figure 14). This pocket likely comes about as a result of the increased propensity for flap tip curling observed in the HP4 PR.

The I13V, I62V and V77I mutations appear to stabilise wider flap conformations in the inhibitor-bound state which may result in a decrease in PI susceptibility through minimising significant contacts between the inhibitor and the active site residues. The flap tips of HP3 PR do not sustain the curled conformation as readily as those of the HP4 PR which may diminish their ability to form coordination complexes with water molecules, hence the absence of water-bridges in the flap regions of the inhibitor-bound HP3 complexes. The addition of the L33F mutation allows the PR to sustain a curled flap tip conformation, which is evident in the apo HP4 PR as well as when in complex with PIs (Figure 4 and 14, respectively). When in complex with inhibitors the curled nature of the flap tips of HP4 PR coordinates several water molecules into this region. The ability that water molecules possess to rotate, slide, and rapidly swap hydrogen bonds may influence ligand binding when multiple water molecules are present. The dynamic nature of these interactions may decrease the stability of inhibitor-bound complexes through the minimisation of sustained intermolecular interactions and, therefore, decrease drug susceptibility through increasing the global dynamics of the complexes.

Concluding remarks

The exact impact that non-active site mutations have on the PR is not wholly understood. However, their importance in drug resistance mechanisms and the maintenance of viral fitness cannot be ignored. The prevalence of certain non-active site mutations, such as I13V, I62V and V77I and L33F, rises with the increased usage of second-generation inhibitors in anti-retroviral therapy. Therefore, their importance in the development of successful drug regimens relies heavily on understanding the mechanisms behind reduced drug susceptibility which are imparted by non-active site mutations such as these. In this study, we report how these mutations impacted the conformation and dynamics of the flap regions of the PR in a manner that is likely to result in reduced drug susceptibility. Our data indicate that these mutations may increase the global rigidity of the apo PR which would likely result in an increased energetic penalty associated with conformational changes associated with drug binding. Thus, a more rigid apo PR may decrease the energetic favourability of drug binding. Furthermore,

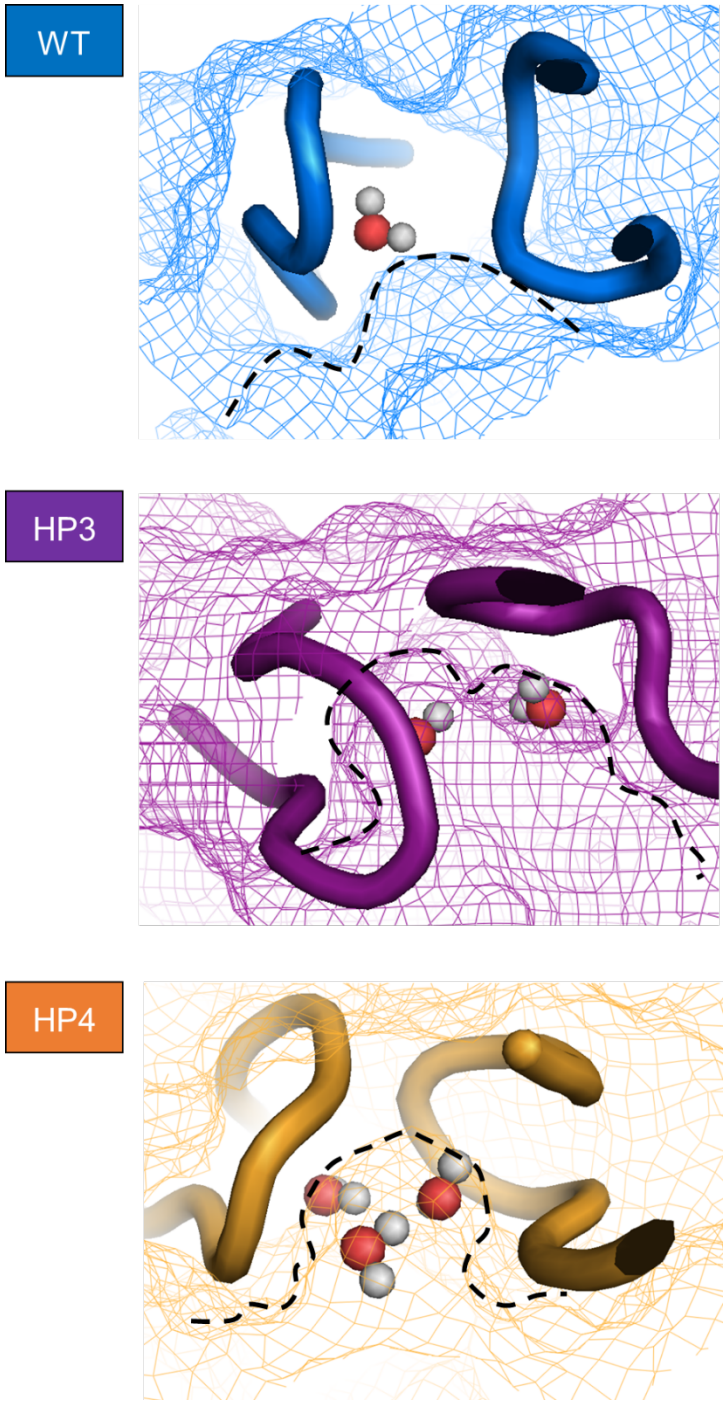
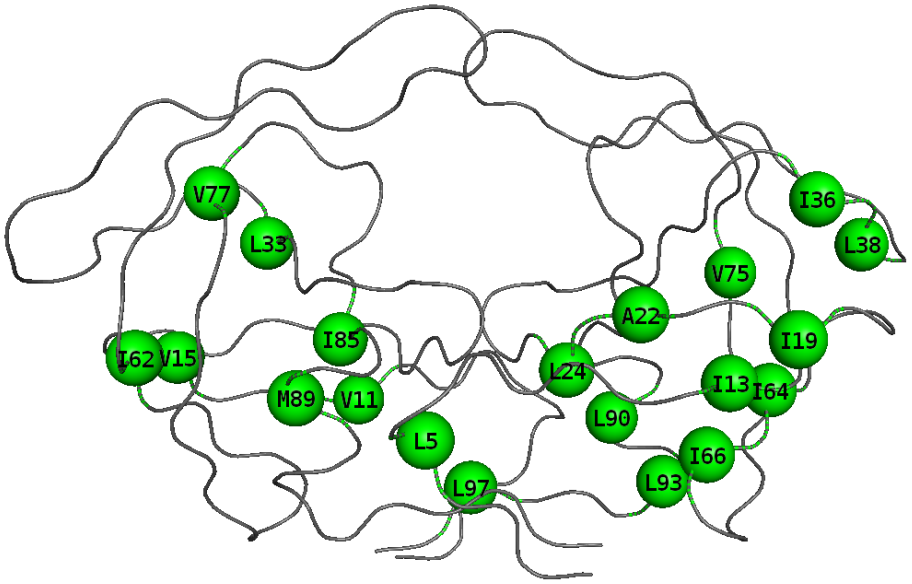


Figure 14. Water molecules associated with the flap tips of the PR variants. The surface of the active site cavity is clarified through the use of dotted black lines.

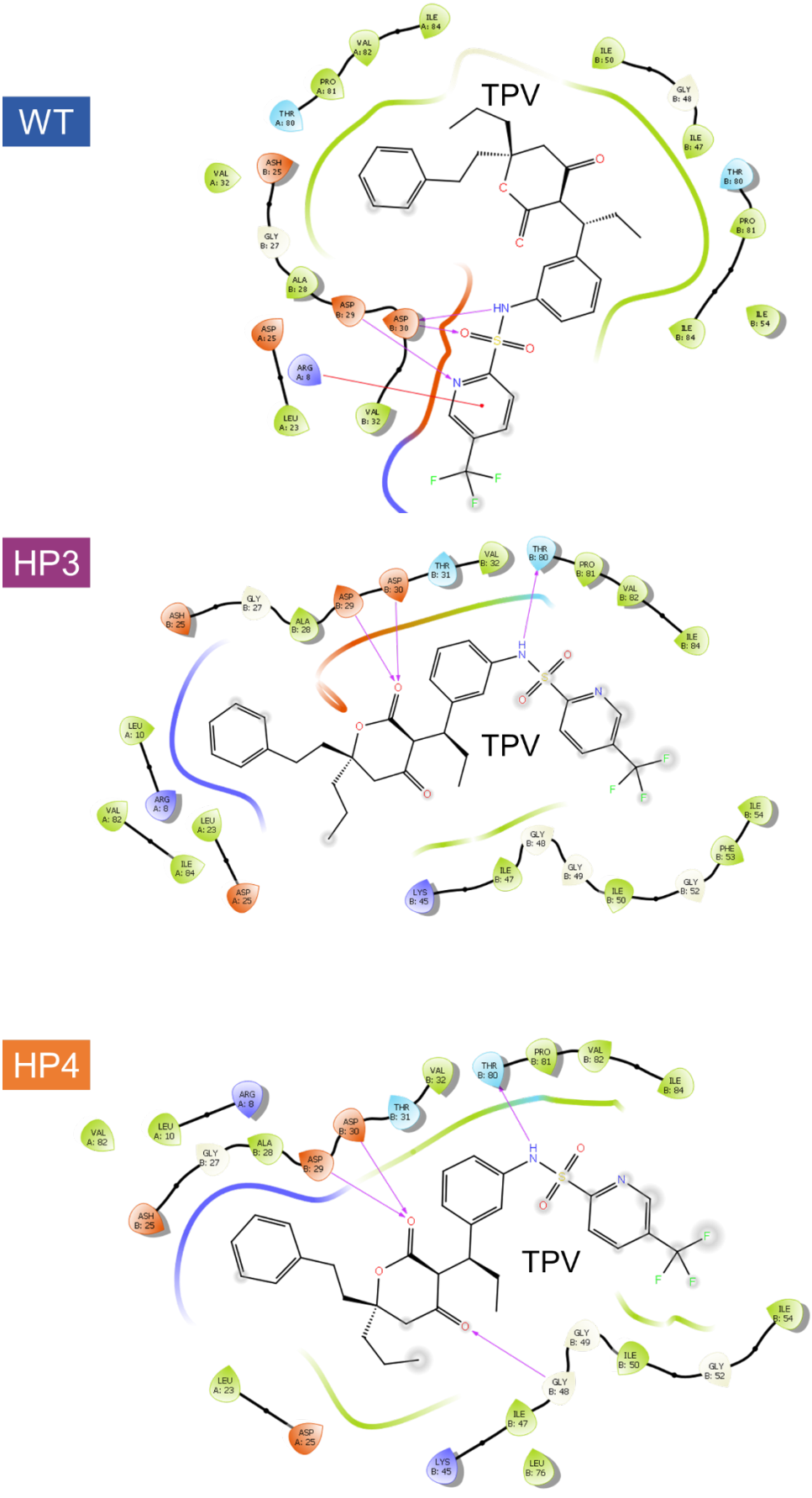
these mutations may allow the PR to stabilise wider flap conformations due to increased flexibility of other key regions of the PR, such as the hinge, fulcrum, and cantilever when complexed with clinical inhibitors.

These wider conformers result in the loss of certain key contacts formed between the PR and the inhibitor, which results in greater mobility of the PIs within the active site. Additionally, in the presence of all four mutations the PR adopts curled flap tips which spontaneously coordinate several water molecules between the flaps and the inhibitors which results in greater dynamics within these complexes. We noted that the conformational flexibility of the cantilever tip region may contribute to the mobility of the flap regions. In particular, curling-in and curling-out of the cantilever tips may be correlated with movement of the flaps. Therefore, allosteric inhibition of flexible regions of the PR that lie outside of the active site, but play a role in the dynamics of the PR, may be a promising avenue for the development of drug regimens which are effective against drug resistant PR. The dynamic scheme of the PR which is presented here provides an explanation for an experimentally determined decrease in drug susceptibility as a result of the mutations we have investigated. However, the limited duration of the simulations in the present study suggests that the conclusions drawn from these trajectories are not fully quantitative since large scale motions of the flaps occur within similar timescales. As such, further exploration of these mutations should be performed *in silico* over longer timescales as well as *in vitro* using conventional biochemical techniques.

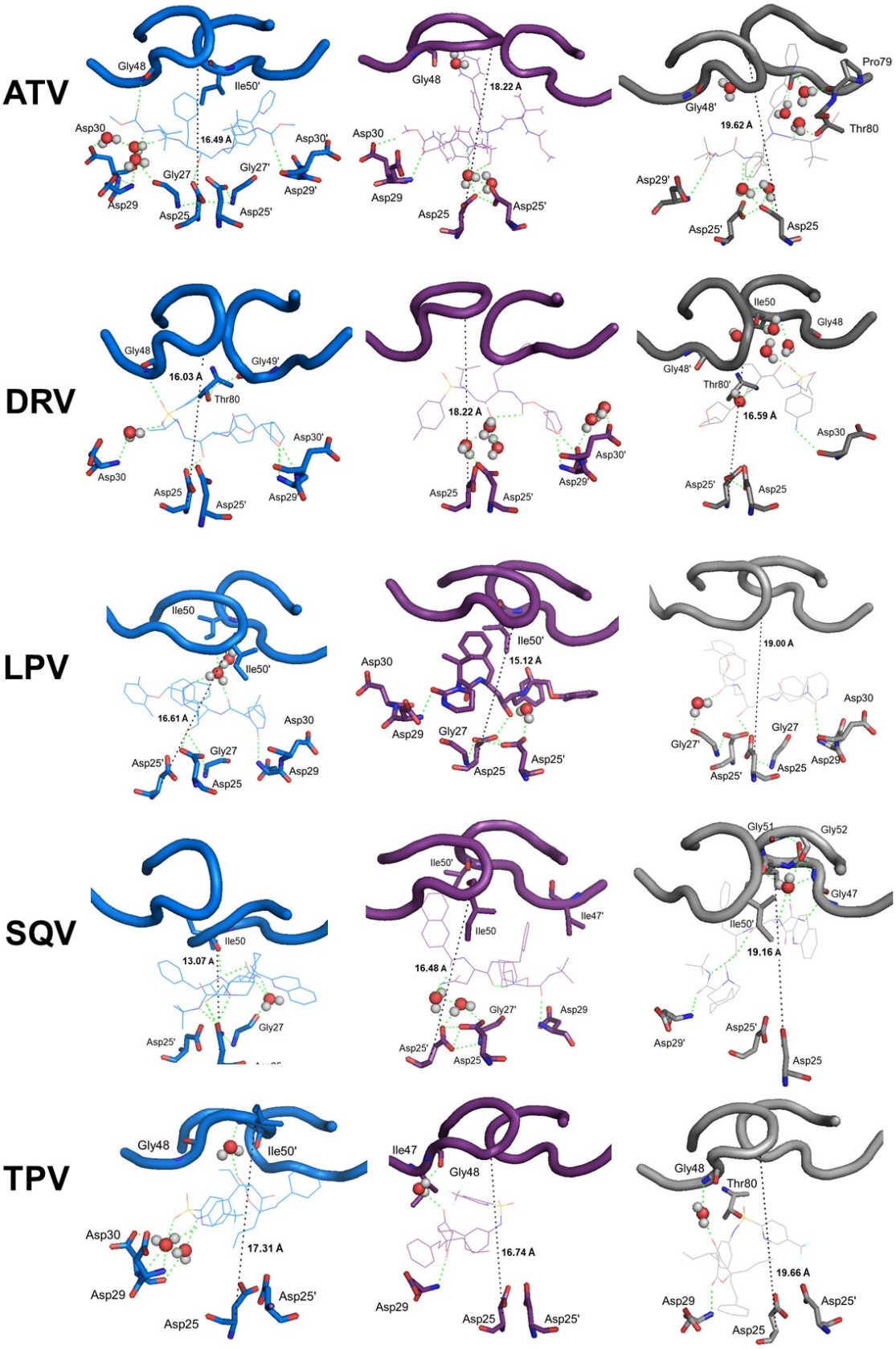
Supplementary Figures



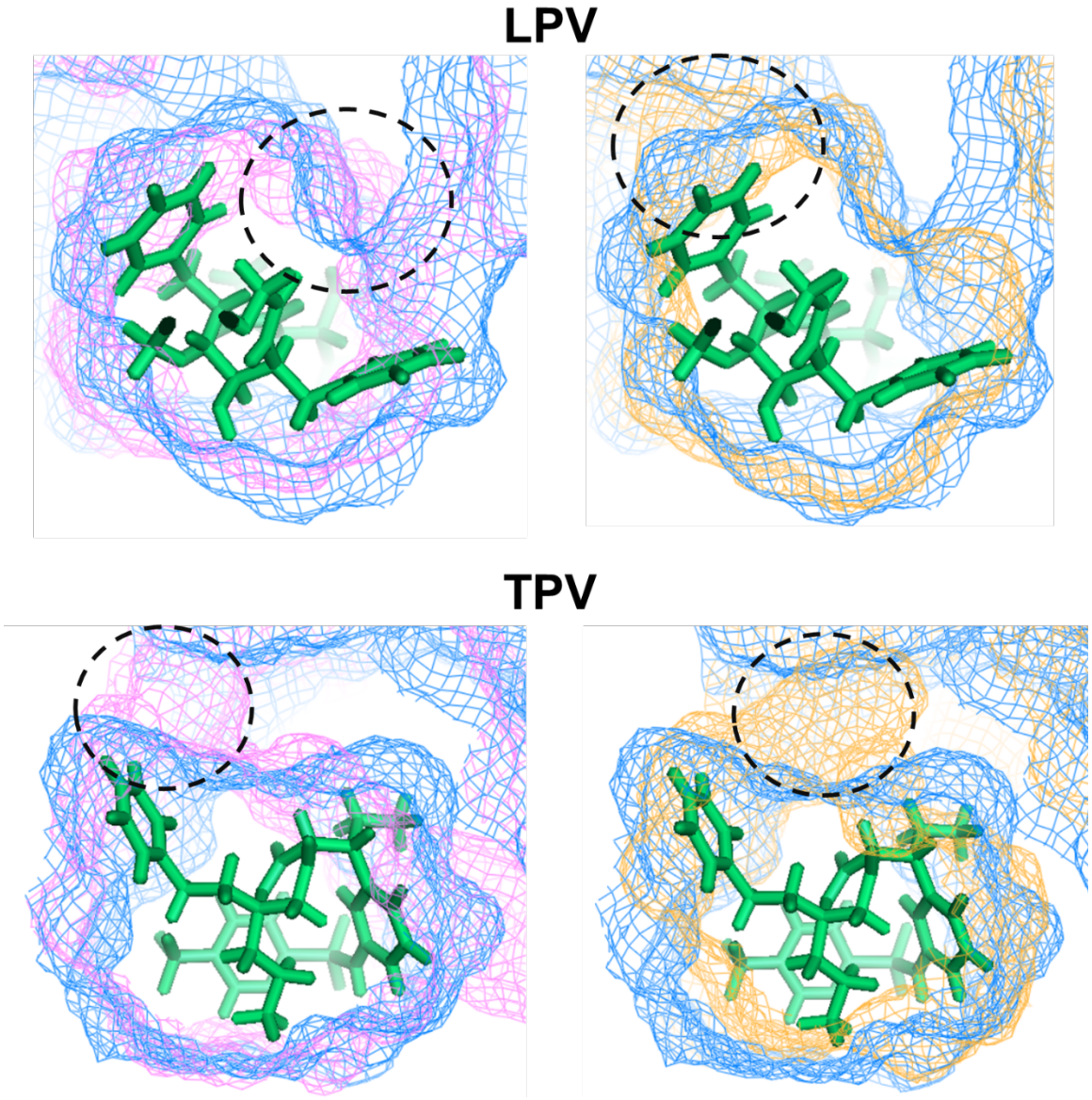
Supplementary Figure 1: The hydrophobic core of the PR. The twenty hydrophobic amino acids that comprise the hydrophobic core of the PR are represented as spheres and labelled accordingly. These amino acids include: L5, V11, I13, V15, I19, A22, L24, L33, I26, L38, I62, I64, I66, V75, V77, I85, M89, L90, L93 and L97.



Supplementary Figure 2. Ligand interaction diagrams of the PR variants in complex with TPV.



Supplementary Figure 3. Protein-ligand interaction diagrams for the WT (blue), HP3 (purple) and HP4 (grey) PR variants in complex with PIs. The residues of the flaps (46-MIGGIGGFI-54) are shown in the cartoon representation.



Supplementary Figure 4. Molecular surface renderings of the active site cavities of the PR variants in complex with LPV and TPV. The molecular surface of the WT (blue), HP3 (purple) and HP4 (orange) proteases. The PIs (green) when in complex with the WT PR are represented as sticks. The pockets formed at the top of the active site of the HP3 PR and HP4 PR are indicated by the dotted circles.

Supplementary Tables

Supplementary Table 1. Average RMSF of key regions of the WT PR, HP3 PR, and HP4 PR when in complex with PIs

	PI	Hinge	Hinge'	Flap	Flap'	Cantilever	Cantilever'
WT	**SQV**	0.77	0.83	1.77	1.34	0.72	0.78
	ATV	0.77	0.81	0.93	1.15	0.77	0.85
	DRV	0.78	0.77	0.88	1.06	0.76	0.78
	LPV	0.85	0.84	1.61	1.13	0.80	0.83
	TPV	1.04	0.94	2.04	1.26	0.90	0.94
HP3	**SQV**	0.85	0.72	1.08	1.08	0.80	0.75
	ATV	1.23	1.13	2.35	1.94	0.95	0.85
	DRV	0.97	0.84	1.37	1.37	0.73	0.81
	LPV	0.79	0.84	1.54	1.12	0.78	0.77
	TPV	0.80	0.87	1.05	1.55	0.78	0.79
HP4	**SQV**	1.01	1.09	2.13	1.75	0.92	1.13
	ATV	0.88	0.98	1.68	1.36	0.94	0.95
	DRV	0.97	1.03	1.35	2.30	0.96	0.94
	LPV	0.93	0.84	1.42	1.71	0.85	0.77
	TPV	0.96	0.87	1.40	1.55	0.92	0.79

Supplementary Table 2. Average distance between Asp25/25'-Ile50/50' distance of the WT PR, HP3 PR, and HP4 PR when in complex with PIs

	PR chain	**ATV**	**DRV**	**SQV**	**TPV**	**LPV**
WT	A	16.5	16.2	13.4	19.0	18.0
	B	15.8	14.3	16.8	17.4	16.3
HP3	A	21.8	17.1	16.6	18.2	15.7
	B	19.5	19.5	17.8	16.9	16.8
HP4	A	18.3	18.8	18.4	18.3	19.2
	B	19.4	18.4	20.0	19.0	16.7

Acknowledgement of funding

The research reported in this publication was supported by the South African Medical Research Council under a Self-Initiated Research Grant to Yasien Sayed. The views and opinions expressed are those of the authors and do not necessarily represent the official views of the SA MRC. YS would like to thank the South African National Research Foundation (NRF) for financial support via the Competitive Support for Rated Researchers grant (CPRR210118582137).

References

Baxter, J.D., Chasanov, W.M., Adams, J., 2016. An Update on HIV-1 Protease Inhibitor Resistance. J AIDS Clin Res 7. https://doi.org/10.4172/2155-6113.1000581

Borman, A.M., Paulous, S., Clavel, F., 1996. Resistance of human immunodeficiency virus type 1 to protease inhibitors: selection of resistance mutations in the presence and absence of the drug. J. Gen. Virol. 77 (Pt 3), 419–426. https://doi.org/10.1099/0022-1317-77-3-419

Chetty, S., Bhakat, S., Martin, A.J.M., Soliman, M.E.S., 2016. 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. 34, 135–151. https://doi.org/10.1080/07391102.2015.1018326

Clemente, J.C., Moose, R.E., Hemrajani, R., Whitford, L.R.S., Govindasamy, L., Reutzel, R., McKenna, R., Agbandje-McKenna, M., Goodenow, M.M., Dunn, B.M., 2004. Comparing the accumulation of active- and nonactive-site mutations in the HIV-1 protease. Biochemistry 43, 12141–12151. https://doi.org/10.1021/bi049459m

Collins, J.R., Burt, S.K., Erickson, J.W., 1995. Flap opening in HIV-1 protease simulated by “activated” molecular dynamics. Nat. Struct. Biol. 2, 334–338. https://doi.org/10.1038/nsb0495-334

Doyon, L., Croteau, G., Thibeault, D., Poulin, F., Pilote, L., Lamarre, D., 1996. Second locus involved in human immunodeficiency virus type 1 resistance to protease inhibitors. J Virol 70, 3763–3769.

Foulkes-Murzycki, J.E., Scott, W.R.P., Schiffer, C.A., 2007. Hydrophobic Sliding: A Possible Mechanism for Drug Resistance in Human Immunodeficiency Virus Type 1 Protease. Structure 15, 225–233. https://doi.org/10.1016/j.str.2007.01.006

Freedberg, D.I., Ishima, R., Jacob, J., Wang, Y.-X., Kustanovich, I., Louis, J.M., Torchia, D.A., 2002. Rapid structural fluctuations of the free HIV protease flaps in solution: Relationship to crystal structures and comparison with predictions of dynamics calculations. Protein Sci 11, 221–232.

Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., Repasky, M.P., Knoll, E.H., Shelley, M., Perry, J.K., Shaw, D.E., Francis, P., Shenkin, P.S., 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J. Med. Chem. 47, 1739–1749. https://doi.org/10.1021/jm0306430

Galiano, L., Ding, F., Veloro, A.M., Blackburn, M.E., Simmerling, C., Fanucci, G.E., 2009. Drug pressure selected mutations in HIV-1 protease alter flap conformations. J. Am. Chem. Soc. 131, 430–431. https://doi.org/10.1021/ja807531v

Göttlinger, H.G., Sodroski, J.G., Haseltine, W.A., 1989. Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. Proc Natl Acad Sci U S A 86, 5781–5785.

Grant, B.J., Rodrigues, A.P.C., ElSawy, K.M., McCammon, J.A., Caves, L.S.D., 2006. Bio3d: an R package for the comparative analysis of protein structures. Bioinformatics 22, 2695–2696. https://doi.org/10.1093/bioinformatics/btl461

Greenwood, J.R., Calkins, D., Sullivan, A.P., Shelley, J.C., 2010. Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. J. Comput. Aided Mol. Des. 24, 591–604. https://doi.org/10.1007/s10822-010-9349-1

Gustchina, A., Weber, I.T., 1991. Comparative analysis of the sequences and structures of HIV-1 and HIV-2 proteases. Proteins 10, 325–339. https://doi.org/10.1002/prot.340100406

Gustchina, A., Weber, I.T., 1990. Comparison of inhibitor binding in HIV-1 protease and in non-viral aspartic proteases: the role of the flap. FEBS Lett 269, 269–272. https://doi.org/10.1016/0014-5793(90)81171-J

Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J.Y., Wang, L., Lupyan, D., Dahlgren, M.K., Knight, J.L., Kaus, J.W., Cerutti, D.S., Krilov, G., Jorgensen, W.L., Abel, R., Friesner, R.A., 2016. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. J. Chem. Theory Comput. 12, 281–296. https://doi.org/10.1021/acs.jctc.5b00864

Hornak, V., Okur, A., Rizzo, R.C., Simmerling, C., 2006. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. PNAS 103, 915–920. https://doi.org/10.1073/pnas.0508452103

Hornak, V., Simmerling, C., 2007. Targeting structural flexibility in HIV-1 protease inhibitor binding. Drug Discov Today 12, 132–138. https://doi.org/10.1016/j.drudis.2006.12.011

Ishima, R., Freedberg, D.I., Wang, Y.X., Louis, J.M., Torchia, D.A., 1999. Flap opening and dimer-interface flexibility in the free and inhibitor-bound HIV protease, and their implications for function. Structure 7, 1047–1055.

Ishima, R., Louis, J.M., Torchia, D.A., 2001. Characterization of two hydrophobic methyl clusters in HIV-1 protease by NMR spin relaxation in solution. J Mol Biol 305, 515–521. https://doi.org/10.1006/jmbi.2000.4321

Johnson, V.A., Brun-Vézinet, F., Clotet, B., Conway, B., D'Aquila, R.T., Demeter, L.M., Kuritzkes, D.R., Pillay, D., Schapiro, J.M., Telenti, A., Richman, D.D., International AIDS Society-USA Drug Resistance Mutations Group, 2003. Drug resistance mutations in HIV-1. Top HIV Med 11, 215–221.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L., 1983. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79, 926–935. https://doi.org/10.1063/1.445869

Kozal, M., 2004. Review: Cross-Resistance Patterns Among HIV Protease Inhibitors. AIDS Patient Care and STDs 18, 199–208. https://doi.org/10.1089/108729104323038874

Kožíšek, M., Lepšík, M., Grantz Šašková, K., Brynda, J., Konvalinka, J., Rezáčová, P., 2014. Thermodynamic and structural analysis of HIV protease resistance to darunavir - analysis of heavily mutated patient-derived HIV-1 proteases. FEBS J. 281, 1834–1847. https://doi.org/10.1111/febs.12743

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography. https://doi.org/10.1107/S0021889892009944

Louis, J.M., Ishima, R., Torchia, D.A., Weber, I.T., 2007. HIV-1 protease: structure, dynamics, and inhibition. Adv. Pharmacol. 55, 261–298. https://doi.org/10.1016/S1054-3589(07)55008-8

Louis, J.M., McDonald, R.A., Nashed, N.T., Wondrak, E.M., Jerina, D.M., Oroszlan, S., Mora, P.T., 1991. Autoprocessing of the HIV-1 protease using purified wild-type and mutated fusion proteins expressed at high levels in Escherichia coli. Eur. J. Biochem. 199, 361–369. https://doi.org/10.1111/j.1432-1033.1991.tb16132.x

Lv, Z., Chu, Y., Wang, Y., 2015. HIV protease inhibitors: a review of molecular selectivity and toxicity. HIV AIDS (Auckl) 7, 95–104. https://doi.org/10.2147/HIV.S79956

Mahalingam, B., Boross, P., Wang, Y.F., Louis, J.M., Fischer, C.C., Tozser, J., Harrison, R.W., Weber, I.T., 2002. Combining mutations in HIV-1 protease to understand mechanisms of resistance. Proteins 48, 107–116. https://doi.org/10.1002/prot.10140

Mahalingam, B., Louis, J.M., Hung, J., Harrison, R.W., Weber, I.T., 2001. Structural implications of drug-resistant mutants of HIV-1 protease: high-resolution crystal structures of the mutant protease/substrate analogue complexes. Proteins 43, 455–464. https://doi.org/10.1002/prot.1057

Mammano, F., Petit, C., Clavel, F., 1998. Resistance-associated loss of viral fitness in human immunodeficiency virus type 1: phenotypic analysis of protease and gag coevolution in protease inhibitor-treated patients. J. Virol. 72, 7632–7637.

Mark, P., Nilsson, L., 2001. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. J. Phys. Chem. A 105, 9954–9960. https://doi.org/10.1021/jp003020w

Maschera, B., Darby, G., Palú, G., Wright, L.L., Tisdale, M., Myers, R., Blair, E.D., Furfine, E.S., 1996. 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. 271, 33231–33235. https://doi.org/10.1074/jbc.271.52.33231

Muzammil, S., Ross, P., Freire, E., 2003. A major role for a set of non-active site mutations in the development of HIV-1 protease drug resistance. Biochemistry 42, 631–638. https://doi.org/10.1021/bi027019u

Naicker, P., Achilonu, I., Fanucchi, S., Fernandes, M., Ibrahim, M.A.A., Dirr, H.W., Soliman, M.E.S., Sayed, Y., 2013. Structural insights into the South African HIV-1 subtype C protease: impact of hinge region dynamics and flap flexibility in drug resistance. J. Biomol. Struct. Dyn. 31, 1370–1380. https://doi.org/10.1080/07391102.2012.736774

Naicker, P., Stoychev, S., Dirr, H.W., Sayed, Y., 2014. Amide hydrogen exchange in HIV-1 subtype B and C proteases--insights into reduced drug susceptibility and dimer stability. FEBS J. 281, 5395–5410. https://doi.org/10.1111/febs.13084

Nalam, M.N.L., Ali, A., Altman, M.D., Reddy, G.S.K.K., Chellappan, S., Kairys, V., Ozen, A., Cao, H., Gilson, M.K., Tidor, B., Rana, T.M., Schiffer, C.A., 2010. Evaluating the substrate-envelope hypothesis: structural analysis of novel HIV-1 protease inhibitors designed to be robust against drug resistance. J. Virol. 84, 5368–5378. https://doi.org/10.1128/JVI.02531-09

Nicholson, L.K., Yamazaki, T., Torchia, D.A., Grzesiek, S., Bax, A., Stahl, S.J., Kaufman, J.D., Wingfield, P.T., Lam, P.Y., Jadhav, P.K., 1995. Flexibility and function in HIV-1 protease. Nat. Struct. Biol. 2, 274–280. https://doi.org/10.1038/nsb0495-274

Nijhuis, M., Schuurman, R., de Jong, D., Erickson, J., Gustchina, E., Albert, J., Schipper, P., Gulnik, S., Boucher, C.A., 1999. Increased fitness of drug resistant HIV-1 protease as a result of acquisition of compensatory mutations during suboptimal therapy. AIDS 13, 2349–2359.

Ohtaka, H., Velázquez-Campoy, A., Xie, D., Freire, E., 2002. Overcoming drug resistance in HIV-1 chemotherapy: the binding thermodynamics of Amprenavir and TMC-126 to wild-type and drug-resistant mutants of the HIV-1 protease. Protein Sci. 11, 1908–1916. https://doi.org/10.1110/ps.0206402

Olsen, D.B., Stahlhut, M.W., Rutkowski, C.A., Schock, H.B., vanOlden, A.L., Kuo, L.C., 1999. Non-active Site Changes Elicit Broad-based Cross-resistance of the HIV-1 Protease to Inhibitors. Journal of Biological Chemistry 274, 23699–23701. https://doi.org/10.1074/jbc.274.34.23699

Perryman, A.L., Lin, J.-H., McCammon, J.A., 2004. HIV-1 protease molecular dynamics of a wild-type and of the V82F/I84V mutant: possible contributions to drug resistance and a potential new target site for drugs. Protein Sci. 13, 1108–1123. https://doi.org/10.1110/ps.03468904

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25, 1605–1612. https://doi.org/10.1002/jcc.20084

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

Ragland, D.A., Nalivaika, E.A., Nalam, M.N.L., Prachanronarong, K.L., Cao, H., Bandaranayake, R.M., Cai, Y., Kurt-Yilmaz, N., Schiffer, C.A., 2014. Drug Resistance Conferred by Mutations Outside the Active Site through Alterations in the Dynamic and Structural Ensemble of HIV-1 Protease. Journal of the American Chemical Society 136, 11956–11963. https://doi.org/10.1021/ja504096m

Reddy, S.V.G., Reddy, K.T., Kumari, V.V., Basha, S.H., 2015. Molecular docking and dynamic simulation studies evidenced plausible immunotherapeutic anticancer property by Withaferin A targeting indoleamine 2,3-dioxygenase. J. Biomol. Struct. Dyn. 33, 2695–2709. https://doi.org/10.1080/07391102.2015.1004834

Rhee, S.-Y., Fessel, W.J., Zolopa, A.R., Hurley, L., Liu, T., Taylor, J., Nguyen, D.P., Slome, S., Klein, D., Horberg, M., Flamm, J., Follansbee, S., Schapiro, J.M., Shafer, R.W., 2005. HIV-1 Protease and reverse-transcriptase mutations: correlations with antiretroviral therapy in subtype B isolates and implications for drug-resistance surveillance. J. Infect. Dis. 192, 456–465. https://doi.org/10.1086/431601

Rhee, S.-Y., Taylor, J., Fessel, W.J., Kaufman, D., Towner, W., Troia, P., Ruane, P., Hellinger, J., Shirvani, V., Zolopa, A., Shafer, R.W., 2010. HIV-1 protease mutations and protease inhibitor cross-resistance. Antimicrob. Agents Chemother. 54, 4253–4261. https://doi.org/10.1128/AAC.00574-10

Schaffer, L., Verkhivker, G.M., 1998. Predicting structural effects in HIV-1 protease mutant complexes with flexible ligand docking and protein side-chain optimization. Proteins 33, 295–310.

Scott, W.R.P., Schiffer, C.A., 2000. Curling of Flap Tips in HIV-1 Protease as a Mechanism for Substrate Entry and Tolerance of Drug Resistance. Structure 8, 1259–1265. https://doi.org/10.1016/S0969-2126(00)00537-2

Shafer, R.W., Hsu, P., Patick, A.K., Craig, C., Brendel, V., 1999. Identification of Biased Amino Acid Substitution Patterns in Human Immunodeficiency Virus Type 1 Isolates from Patients Treated with Protease Inhibitors. J Virol 73, 6197–6202.

Shelley, J.C., Cholleti, A., Frye, L.L., Greenwood, J.R., Timlin, M.R., Uchimaya, M., 2007. Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. J. Comput. Aided Mol. Des. 21, 681–691. https://doi.org/10.1007/s10822-007-9133-z

Shuman, C.F., Markgren, P.O., Hämäläinen, M., Danielson, U.H., 2003. Elucidation of HIV-1 protease resistance by characterization of interaction kinetics between inhibitors and enzyme variants. Antiviral Res 58, 235–242. https://doi.org/10.1016/S0166-3542(03)00002-0

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539. https://doi.org/10.1038/msb.2011.75

Skjærven, L., Yao, X.-Q., Scarabelli, G., Grant, B.J., 2014. Integrating protein structural dynamics and evolutionary analysis with Bio3D. BMC Bioinformatics 15, 399. https://doi.org/10.1186/s12859-014-0399-6

Swain, A.L., Miller, M.M., Green, J., Rich, D.H., Schneider, J., Kent, S.B., Wlodawer, A., 1990. X-ray crystallographic structure of a complex between a synthetic protease of human immunodeficiency virus 1 and a substrate-based hydroxyethylamine inhibitor. Proc. Natl. Acad. Sci. U.S.A. 87, 8805–8809. https://doi.org/10.1073/pnas.87.22.8805

Thompson, M.A., Aberg, J.A., Cahn, P., Montaner, J.S.G., Rizzardini, G., Telenti, A., Gatell, J.M., Günthard, H.F., Hammer, S.M., Hirsch, M.S., Jacobsen, D.M., Reiss, P., Richman, D.D., Volberding, P.A., Yeni, P., Schooley, R.T., International AIDS Society-USA, 2010. Antiretroviral treatment of adult HIV infection: 2010 recommendations of the International AIDS Society-USA panel. JAMA 304, 321–333. https://doi.org/10.1001/jama.2010.1004

Todd, M.J., Freire, E., 1999. The effect of inhibitor binding on the structural stability and cooperativity of the HIV-1 protease. Proteins: Structure, Function, and Bioinformatics 36, 147–156. https://doi.org/10.1002/(SICI)1097-0134(19990801)36:2<147::AID-PROT2>3.0.CO;2-3

Todd, M.J., Luque, I., Velázquez-Campoy, A., Freire, E., 2000. Thermodynamic basis of resistance to HIV-1 protease inhibition: calorimetric analysis of the V82F/I84V active site resistant mutant. Biochemistry 39, 11876–11883. https://doi.org/10.1021/bi001013s

Tóth, G., Borics, A., 2006. Closing of the Flaps of HIV-1 Protease Induced by Substrate Binding: A Model of a Flap Closing Mechanism in Retroviral Aspartic Proteases. Biochemistry 45, 6606–6614. https://doi.org/10.1021/bi060188k

Velazquez-Campoy, A., Freire, E., 2001. Incorporating target heterogeneity in drug design. Journal of Cellular Biochemistry 84, 82–88. https://doi.org/10.1002/jcb.10068

Velazquez-Campoy, A., Kiso, Y., Freire, E., 2001. The binding energetics of first- and second-generation HIV-1 protease inhibitors: implications for drug design. Arch. Biochem. Biophys. 390, 169–175. https://doi.org/10.1006/abbi.2001.2333

Velázquez-Campoy, A., Todd, M.J., Vega, S., Freire, E., 2001. Catalytic efficiency and vitality of HIV-1 proteases from African viral subtypes. Proceedings of the National Academy of Sciences 98, 6062–6067. https://doi.org/10.1073/pnas.111152698

Volberding, P.A., Deeks, S.G., 2010. Antiretroviral therapy and management of HIV infection. Lancet 376, 49–62. https://doi.org/10.1016/S0140-6736(10)60676-9

Wassenaar, T.A., Mark, A.E., 2006. The effect of box shape on the dynamic properties of proteins simulated under periodic boundary conditions. J Comput Chem 27, 316–325. https://doi.org/10.1002/jcc.20341

Weber, I.T., Agniswamy, J., 2009. HIV-1 Protease: Structural Perspectives on Drug Resistance. Viruses 1, 1110–1136. https://doi.org/10.3390/v1031110

Weber, I.T., Kneller, D.W., Wong-Sam, A., 2015. Highly resistant HIV-1 proteases and strategies for their inhibition. Future Med Chem 7, 1023–1038. https://doi.org/10.4155/fmc.15.44

Wensing, A.M., Calvez, V., Günthard, H.F., Johnson, V.A., Paredes, R., Pillay, D., Shafer, R.W., Richman, D.D., 2017. 2017 Update of the Drug Resistance Mutations in HIV-1. Top Antivir Med 24, 132–133.

Wlodawer, A., Erickson, J.W., 1993. Structure-based inhibitors of HIV-1 protease. Annu. Rev. Biochem. 62, 543–585. https://doi.org/10.1146/annurev.bi.62.070193.002551

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

Zondagh, J., Balakrishnan, V., Achilonu, I., Dirr, H.W., Sayed, Y., 2018. Molecular dynamics and ligand docking of a hinge region variant of South African HIV-1 subtype C protease. J. Mol. Graph. Model. 82, 1–11. https://doi.org/10.1016/j.jmgm.2018.03.006

Chapter 5 – The mechanism of non-active site mutations in drug resistance

Plasticity-associated functionality and inhibition of the HIV protease

Dean Sherry, Roland Worth and Yasien Sayed.

Adv Exp Med Biol: Protein Reviews – Published (2021)

In this review article, a current understanding of the protease enzyme as a dynamic ensemble of conformers was highlighted. In this context, the role of dynamic equilibria between substrate and inhibitor binding to the protease are explored. Through a review of primary and secondary mutations, the salience of the mechanism of non-active site mutations in inhibitor resistance is illuminated.

Author contributions: Dean Sherry performed the literature review, wrote the manuscript, generated the images in the manuscript and performed subsequent editing of subsequent manuscript revisions. Roland Worth assisted with proofreading and revision of the manuscript. Yasien Sayed supervised the project and provided proofreading and revisions for the manuscript.

Elasticity-associated functionality and inhibition of the HIV protease

Dean Sherry[1], Roland Worth[1] and Yasien Sayed[1]*

[1] Protein Structure-Function Research Unit, School of Molecular and Cell Biology, University of the Witwatersrand, Johannesburg, 2050, South Africa

*Corresponding author: Yasien Sayed; e-mail address: Yasien.Sayed@wits.ac.za.

Abstract

HIV protease plays a critical role in the life cycle of the virus through the generation of mature and infectious virions. Detailed knowledge of the structure of the enzyme and its substrate has led to the development of protease inhibitors. However, the development of resistance to all currently available protease inhibitors has contributed greatly to the decreased success of antiretroviral therapy. When therapy failure occurs, multiple mutations are found within the protease sequence starting with primary mutations, which directly impact inhibitor binding, which can also negatively impact viral fitness and replicative capacity by decreasing the binding affinity of the natural substrates to the protease. As such, secondary mutations which are located outside of the active site region accumulate to compensate for the recurrently deleterious effects of primary mutations. However, the resistance mechanism of these secondary mutations is not well understood, but what is known is that these secondary mutations contribute to resistance in one of two ways, either through increasing the energetic penalty associated with bringing the protease into the closed conformation, or, through decreasing the stability of the protein/drug complex in a manner that increases the dissociation rate of the drug, leading to diminished inhibition. As such, the elasticity of the enzyme-substrate complex has been implicated in the successful recognition and catalysis of the substrates which may be inferred to suggest that the plasticity of the enzyme/drug complex plays a role in resistance.

Introduction

According to the UNAIDS 2020 global HIV & AIDS statistics, approximately 38 million people are estimated to be living with HIV globally (UNAIDS, 2020). There is currently no cure or vaccine for HIV/AIDS; however, through the development of several therapeutic agents which target various steps in the life cycle of the HI virus, infection by the virus has been transformed from an inevitably fatal disease into a manageable chronic illness. Since the advent of antiretroviral therapy, it has become evident that therapy involving a combination of drugs is more effective for the treatment of HIV infection. The introduction of effective treatment, such as highly active anti-retroviral therapy (HAART), can reduce viral load to almost undetectable levels in infected individuals (Tözsér, 2001). The identification of critical events in the HIV replication cycle has led to the discovery of several important biochemical targets; primarily, reverse transcriptase (RT), integrase (IN), and protease (PR) and as a result, HAART makes use of over twenty different drugs that target different stages of the viral life cycle (Menéndez-Arias, 2010).

The drugs targeting the HIV PR enzyme may represent the most effective course of therapy in the treatment of HIV infection (Lefebvre and Schiffer, 2008; Thompson et al., 2010; Volberding and Deeks, 2010); consequently, the PR is a major drug target for anti-retroviral therapy (ART). This stems from the fact that PR cleaves virally-encoded group specific antigen (Gag) and Gag-Pol polyprotein precursors into structural proteins and functional enzymes, which results in the production of mature and infectious viral progeny (Göttlinger et al., 1989; Louis et al., 2007). However, several factors have contributed to the emergence of drug resistance in HIV-infected patients including: low drug adherence and the low fidelity of the HIV RT, which in turn introduce various mutations into the PR primary sequence (Menéndez-Arias, 2010; Wensing et al., 2010). The development of a vaccine would, therefore, circumvent some of the stumbling blocks associated with drug therapy. However, successful development and implementation has proven to be challenging (Hsu and O'Connell, 2017). Instead, the natural substrate cleavage sites of the PR currently represent feasible vaccine targets (Li et al., 2017). However, there are currently no vaccines which specifically target the PR enzyme itself. It is worth noting that, since the maturation process takes place within the micro-environment of the enclosed immature virion, large biologics, such as antibodies, are less likely to be successful in comparison to small molecule inhibitors.

The tertiary structure of the HIV protease

The HIV PR is a homodimeric aspartyl PR containing a characteristic active site which is composed of the conserved Asp25-Thr26-Gly27 sequence common to all aspartyl PRs. Dimerisation is essential for the function of the enzyme as each monomer contributes to the formation of the central active site (Mager, 2001; Navia et al., 1989). Importantly though, entry of the substrate into the active site (residues: 25-32, 47-53 and 80-84) of the PR is regulated by two β-hairpin loops, known as the flaps, which cover the active site cavity (Figure 1A) (Wlodawer et al., 1989).

A dynamic picture of the apo enzyme has been obtained through nuclear magnetic resonance (NMR) experiments (Freedberg et al., 2002; Ishima et al., 1999) and molecular dynamic (MD) simulations (W. R. Scott and Schiffer, 2000). Collectively, these studies have revealed that in the apo-conformation, substantial conformational alterations occur in the flaps (residues 43-56) of the PR and this large degree of conformational flexibility of the flaps is attributed to the conserved glycine-rich sequence, 46-MIGGIGGFI-54, at the flap tips. In particular, the PR samples a number of heterogeneous flap conformations which exist in equilibrium; namely, closed, semi-, and fully- open conformations (Wlodawer et al., 1989). Previously, these conformations were defined by the distance between the alpha carbon (C_α) atom of the two flap tip residues (Gly51 and Gly51'). However, due to the asymmetry of the flaps and the curling motion of the flap tips during MD simulations, it is believed that measuring the distance between the C_α atoms from the flap tips (Ile50 and Ile50') and the catalytic Asp residues (Asp25 and Asp25') provides a more reasonable metric for flap opening (Perryman et al., 2004). The free energy difference between the fully open, semi-open, and closed conformations is minimal (Rick et al., 1998; Rose et al., 1998) and, as such, a wide variety of heterogenous conformations are accessible to the apo enzyme. Although several studies have shown that the semi-open conformation is the most commonly adopted state for the apo PR (Freedberg et al., 2002; Ishima et al., 1999; Nicholson et al., 1995), the fully open conformation allows the substrate to enter and leave the active site of the PR while the closed conformation promotes the formation of key interactions important for substrate recognition and binding (Gustchina and Weber, 1990). The highly conserved nature of this sequence amongst isolates of HIV-1 and HIV-2 highlights its importance in substrate recognition and binding (Gustchina and Weber, 1991). It is noteworthy that the stability and movement of the flaps is closely related to the hinge region of PR (residues 35-42 and 57-61; Figure 1), as these regions allow for substantial conformational flexibility of the flaps (Gustchina and Weber, 1990). It is, therefore, important to note that the opening of the flaps does not occur independently, instead it is achieved through the synergistic movement of distinct regions within the enzyme.

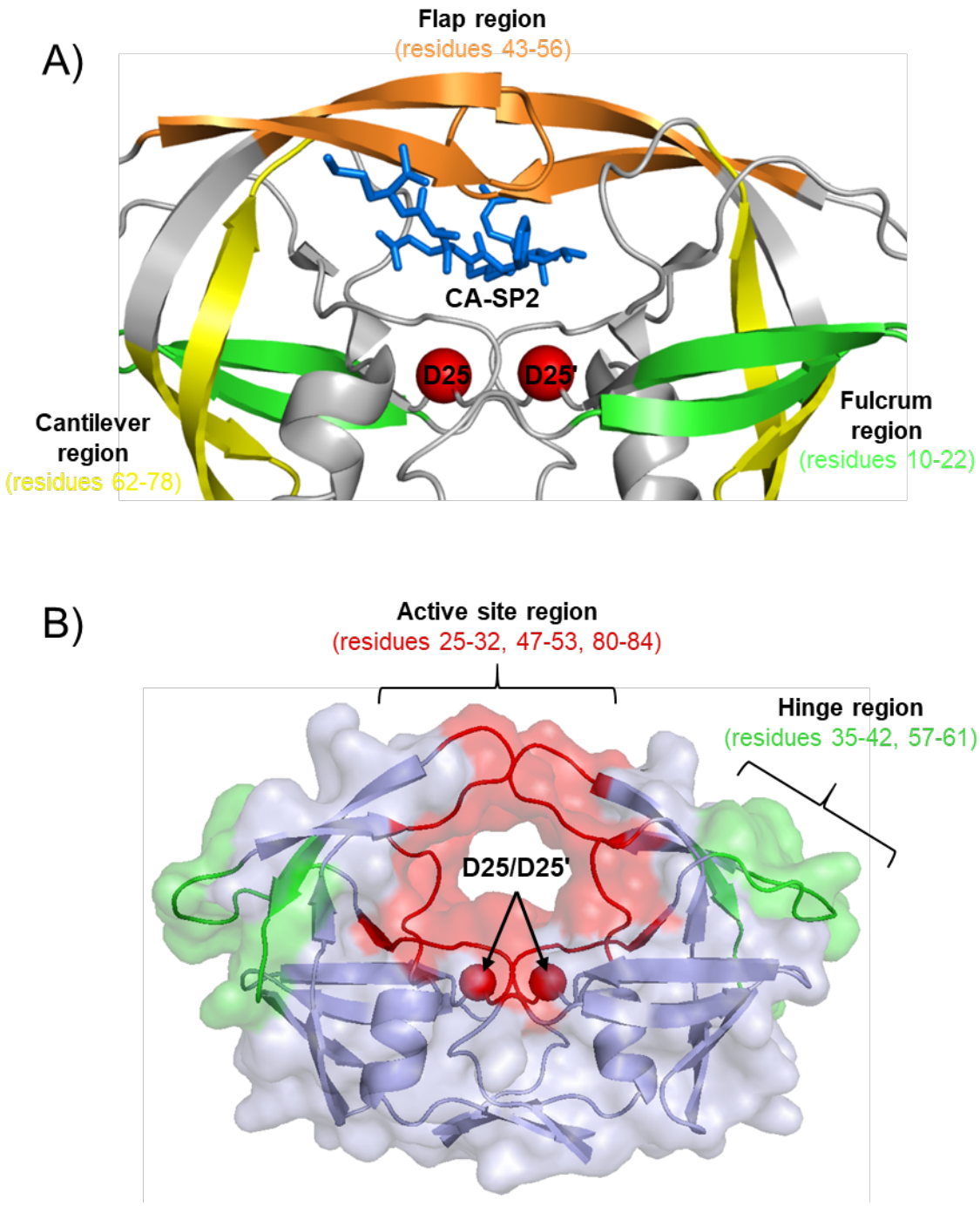


Figure 1. A) The structure of the WT PR in complex with a natural substrate (CA-SP2) analogue (PDB ID: 6O48). The flaps regulate substrate entry into the active site region of the PR. The active site Asp25 residues are represented as red spheres and the CA-SP2 analogue is represented as blue sticks. **B) Surface representation of the WT PR in the apo state.** The dynamics of the flaps is largely modulated by the flexible hinge regions of the PR. Figure generated using PyMOL Molecular Graphics Software (Schrödinger, LLC).

That is, the concerted downward movements of the fulcrum (residues 10-22), cantilever (residues 62-78) and hinge (residues 35-42, 57-61) region (Figure 1), results in the outward motion of the flaps as well as a slight rotation about the dimer interface (Foulkes-Murzycki et al., 2007; Hornak et al., 2006). Twenty hydrophobic residues with limited solvent accessibility form the hydrophobic core of the PR; namely, L5, V11, I13, V15, I19, A22, L24, I26, L33, L38, I62, I64, I66, V75, V77, I85, M89, L90, L93, L97 (Foulkes-Murzycki et al., 2007; Naicker et al., 2014). It is suggested that the rearrangement of hydrophobic residues within the core, or "hydrophobic sliding", modulates changes in the conformation of HIV PR in order to enable the opening and closing of the flaps (Foulkes-Murzycki et al., 2007). During hydrophobic sliding the near-isoenergetic transfer of van der Waals contacts between hydrophobic residues of the core allow the regions of the PR to slide across one another and, consequently, this facilitates conformational changes that result in the outward movement of the flaps (Foulkes-Murzycki et al., 2007; Ragland et al., 2014). In particular, two isoleucine residues, I13 and I15, appear to play a crucial role in the movement of the flap regions of HIV PR (Foulkes-Murzycki et al., 2007; Ishima et al., 2001). Briefly, I15 makes hydrophobic contacts with I36 and L38 in the hinge region and I62, I64 and V75 in the cantilever region (Figure 2). During flap movement, some hydrophobic contacts involving I15 are lost, allowing the hinge and cantilever regions to slide over I15 and downwards towards the N- and C-terminus (Figure 2). However, these lost contacts are instead reformed with I13, therefore ensuring that the fully open conformation is in fact stabilised and that when the core regions of the PR slide over one another it permits the opening of the flaps (Foulkes-Murzycki et al., 2007).

The emergence of resistance in HIV protease

Determining the structure and function of the PR, in particular the active site, has enabled the rational drug design of PR inhibitors (PIs). In general, drug design paradigms follow the lock-and-key hypothesis which focuses on geometric complementarity between the drug and its target protein. PIs are competitive inhibitors that bind at the active site of the PR when the flaps are open and, subsequently, bring the flaps into the closed conformation, folded over the active site. The dynamics of the flap region is important for the activity of the enzyme (Piana et al., 2002); thus, the inhibitor-bound active site with the flaps in a closed conformation will keep the enzyme in a locked-down state and prevent the processing of natural substrates. There are currently nine FDA-approved PIs which are used to target the PR during ART (Figure 3). PIs are designed to mimic the natural substrates cleaved by PR; however, PIs differ in that they are non-hydrolysable due to the presence of an hydroxyethylene core (Partaledis et al., 1995). The core is believed to mimic the tetrahedral transition state of the proteolytic reaction because it is an effective isostere replacement of the scissile bond targeted by PR (Figure 3) (Brik and Wong, 2003).

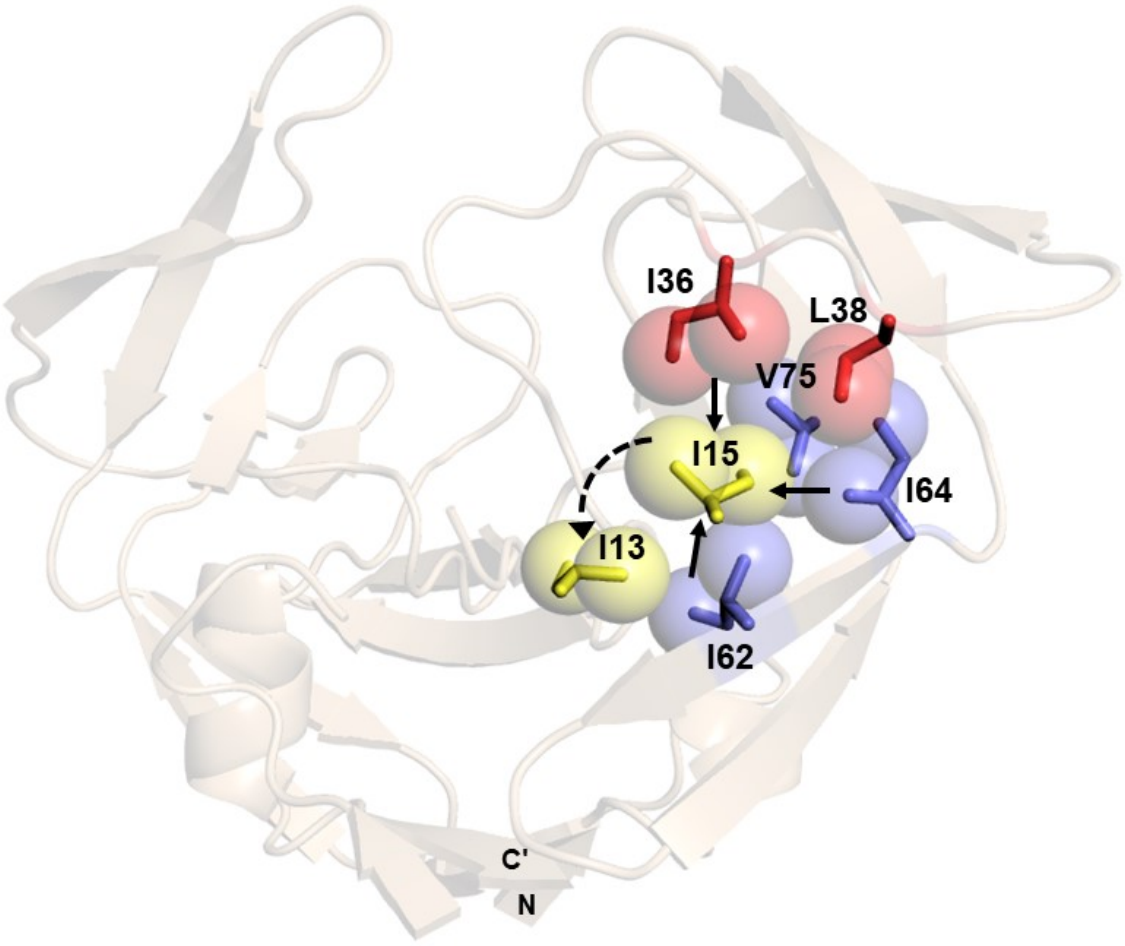


Figure 2. The hydrophobic sliding mechanism for the opening of the PR flaps. In the closed conformation, residues Ile36, Leu38, Ile62, Ile64 and Val75 make contact with Ile15 (solid arrows). Upon flap opening these residues slide over Ile15 and the contacts are instead transferred to Ile13 (dashed arrow). The N- and C-terminus of one of the PR monomers are labelled. This image was generated using the PyMOL Molecular Graphics Software (Schrödinger, LLC), PDB ID: 6I45.

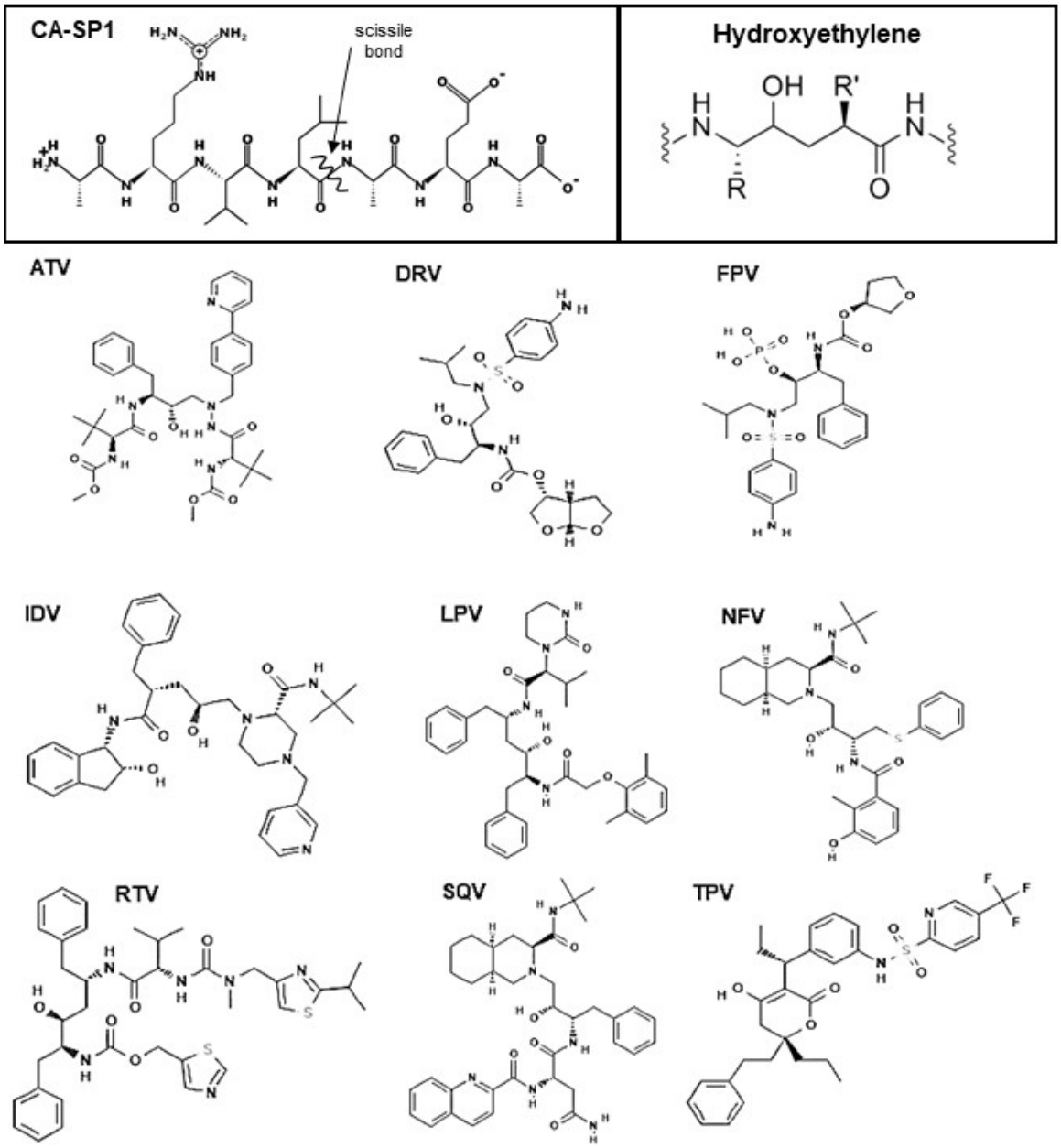


Figure 3. Two-dimensional structures of the CA-SP2 Gag cleavage site as well as all currently available FDA-approved HIV-1 PIs. The location of the scissile bond in the CA-SP2 sequence is indicated. The structure of the non-hydrolysable hydroxyethylene core designed to replace the scissile bond is also depicted. The PIs structures obtained from PubChem (https://pubchem.ncbi.nlm.nih.gov/) are as follows: atazanavir (ATV, CID: 148192), darunavir (DRV, CID: 213039), fosamprenavir (FPV, CID: 131536), indinavir (IDV, CID: 5462355), lopinavir (LPV, CID: 92727), nelfinavir (NFV, CID: 64142), ritonavir (RTV, CID: 392622), saquinavir (SQV, CID: 60787) and tipranavir (TPV, CID: 54682461).

The development of PI resistance has arisen since the advent of PIs during ARV therapy. The viral RT enzyme is extremely error prone due to its lack of proofreading mechanisms and the rapid replication rate of the virus. As a result, the viral genome develops many polymorphic mutations which, in turn, increases the number of variants within the population. This phenomenon is exacerbated by the evolutionary pressures imparted by PIs and, as a consequence, the likelihood for the development of drug resistance is increased (Roberts et al., 1990). Initially, viral variants are preferentially selected with mutations located at the active site commonly referred to as primary mutations. These primary mutations directly compromise inhibitor binding while maintaining substrate specificity and substrate turnover (Prabu-Jeyabalan et al., 2002). PIs are designed to be rigid to provide a large degree of specificity for the PR active site due to their geometric complementarity. This rigidity, however, also limits the ability of PIs to adapt to structural changes within the target as is the case with amino acid polymorphisms in the active site. This lack of adaptability has led to decreased efficacy of first-generation PIs such as saquinavir (SQV), ritonavir (RTV), nelfinavir (NFV) and indinavir (IDV) (Ghosh et al., 2016). Thus, second-generation inhibitors (lopinavir (LPV), fosamprenavir (FPV), atazanavir (ATV), tipranavir (TPV) and darunavir (DRV)) were developed to be more structurally adaptable which is an important feature as these drugs must maintain selectivity for a heterogenous population of PR variants which arise in a given viral population (Ohtaka and Freire, 2005).

Primary resistance mutations: changes in the active site geometry

Primary resistance mutations begin to emerge rapidly at the onset of PI exposure and reduce the affinity for drugs, usually through disrupting the drug/target interface interactions (Chen et al., 1995; Markowitz et al., 1995; Schaffer and Verkhivker, 1998). However, primary mutations can be defined as mutations, either within the active site or outside the active site, which are directly associated with decreased inhibitor binding affinity. In particular, the substrate envelope hypothesis helps explain the mechanism by which primary mutations affect molecular recognition and how the PR evolves to diminish inhibitor binding while still recognising its natural substrate. The natural substrates occupy a consensus volume, within the active site of the PR which is defined as the substrate envelope. As a result, the conserved shape of the substrates is hypothesised to determine the specificity of the PR for its natural substrates (Prabu-Jeyabalan et al., 2002). In comparison, some structures of PIs in complex with the wild-type (WT) PR have atoms of the inhibitors which protrude beyond the substrate envelope (King et al., 2004). Interestingly enough, the atoms of the PIs that protrude from the substrate envelope often make contact with PR residues that commonly mutate to develop drug resistance in response to PI therapy. It would appear, therefore, that certain residues are important for inhibitor binding but not for the binding of the natural substrates. In this manner

(PR) mutations in these locations would specifically affect inhibitor binding while the substrate envelope remains relatively unaffected. However, if the substrate were to protrude out of the envelope and contact the residues associated with drug resistance, this would lead to diminished substrate recognition and processing, which would lead to the co-evolution of the substrate in order to compensate for this decrease in efficiency (King et al., 2004). This information led to the hypothesis that PIs which fit within the substrate envelope are less likely to be influenced by active site mutations (King et al., 2004; Prabu-Jeyabalan et al., 2002). More recently, however, the concept of the substrate envelope hypothesis was revisited. The hypothesis was initially based on various crystal structures which represent an average image of the low-energy states of the PR. However, is it more likely that the activity of the PR is intimately related to its dynamics and, thus, re-evaluating the substrate envelope using MD studies was recently performed (Özen et al., 2011). Taking the flexibility of the system into account is essential to the understanding of how substrate specificity is determined. Analyses of the substrate-bound PR complexes as a dynamic ensemble allowed for the substrate envelope to be defined as a probability distribution as opposed to a static substrate envelope (Figure 4), which is essentially a step function (Özen et al., 2011). As such, the edges of the dynamic envelope are not defined by a sharp cut off but, rather, the envelope is more heterogeneous in nature with discrete boundaries. This depiction of the substrate envelope provides a more realistic representation of the substrate consensus volume and may prove to be a more powerful tool in structure-based drug design algorithms.

Drug resistance is a subtle change in the balance between two types of recognition: (i) the mutant PR will still recognise and process the natural substrates, but (ii) will not effectively recognise and bind, and be inhibited by, drug molecules. Many of the resistance-associated primary mutations are conservative substitutions of hydrophobic amino acids located in the active site such as leucine, isoleucine, valine, and methionine. Residues such as L24, A30, V32, M46, I47, G48, I50, V82, and I84 form interactions with substrates and inhibitors in the active site cavity and are mutated in drug resistant PR variants. Several of these primary mutations are a result of exposure to specific PIs which in turn can be a signature for particular drugs used in therapy (Table 1). It is evident that a variety of primary mutations are selected during PI therapy some of which are unique to a particular inhibitor, for example the G48V mutation in response to SQV exposure. In contrast, other mutations, such as I84V, can confer cross-resistance to multiple inhibitors such as IDV, FPV, ATV, DRV and (Table 1). It is clear, therefore, that in the case of primary mutations the mechanism by which resistance is conferred is shared between multiple PIs.

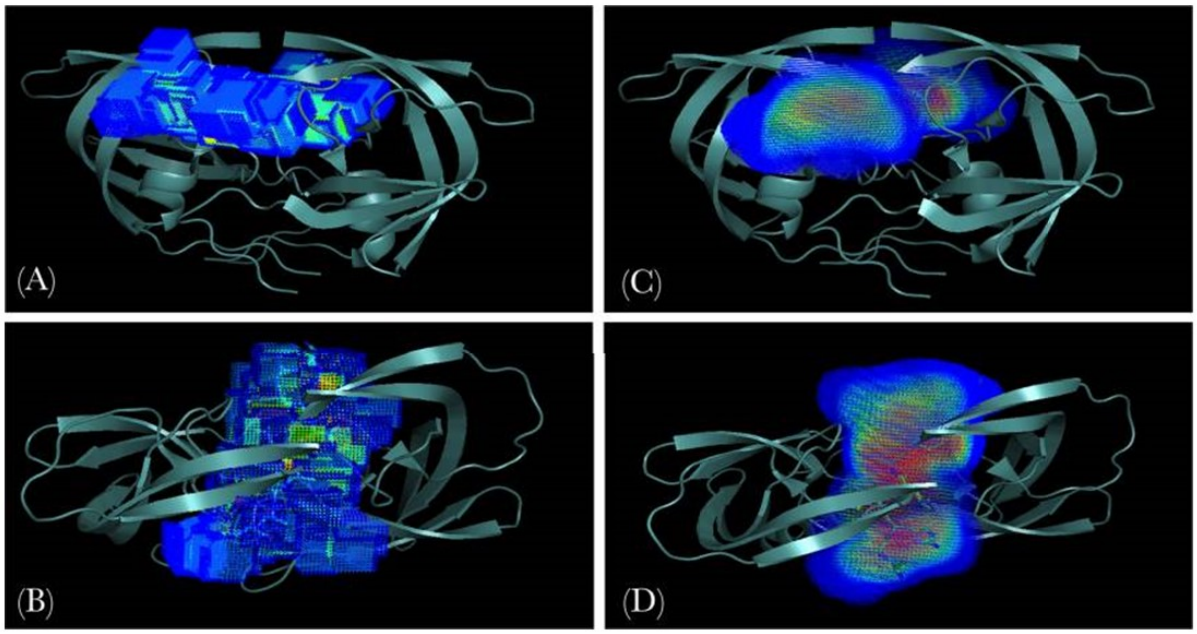


Figure 4. The dynamic substrate envelope is a probability distribution which is likely to be more accurate to incorporate into structure-based drug design (Özen et al., 2011). A) Side and, B) top views of the static substrate envelope in comparison to the respective views of dynamic substrate envelope (C, D). The cell centres of the dynamic envelope are coloured according to substrate occupancy from high (red) to low (blue). The envelopes are superimposed onto the structure of CA-SP1. Taken from Özen et al., 2011.

The mechanism by which active site mutations confer PI resistance has been clarified through structural analyses of inhibitor-bound PR complexes (King et al., 2004; Prabu-Jeyabalan et al., 2002; Tie et al., 2005). Their mode of action has been attributed to lost interactions and/or steric effects created by altered active site geometries that directly affect the drug/protein interface (Ohtaka et al., 2002). One straightforward example where the loss of interactions is evident is with the I84V mutation. In the crystal structure of the I84V variant in complex with DRV, as a result of the reduced bulk of the side chain of the valine residue in comparison to isoleucine this mutation causes a loss of two van der Waals contacts between residue 84 and DRV (Tie et al., 2004). Interestingly, nearly 45 amino acid positions of the 99 found in each monomer of the PR have been implicated in drug resistance (Baxter et al., 2016; Rhee et al., 2010) and 25 of these positions have been shown to contribute significantly to decreased inhibitor susceptibility. Ten amino acid positions associated with reduced inhibitor susceptibility are located within the active site (Table 2) (Baxter et al., 2016; Rhee et al., 2010; Wu et al., 2003), which implicates the remaining 15 positions as non-active site mutations. Therefore, approximately 15% of the entire PR amino acid sequence is associated with drug resistance as well as being found outside of the active site region (Baxter et al., 2016; Rhee et al., 2010; Wu et al., 2003). As such, the importance of resistance-associated non-active site mutations cannot be overlooked.

Table 1. Resistance mutations of HIV-1 PR in response to PI exposure (Baxter et al., 2016).

PIs	Primary resistance mutations	Secondary resistance mutations
SQV/r	G L 48 90 V M	L L I I A G V V I 10 24 54 62 71 73 77 82 84 F I V V V S I A V
IDV/r	M V I 46 82 84 I A V	L K L V M I A G L V L 10 20 24 32 36 54 71 73 76 77 90 I M I I I V V S V I M
NFV/r	D L 30 90 N M	L M M A V V I N 10 36 46 71 77 82 84 88 F I I V I A V D
FPV/r	I I 50 84 V V	L V M I I G L V L 10 32 46 47 54 73 76 82 90 F I I V L S V A M
LPV/r	V I L V 32 47 76 82 I V V A	L K L L M I F I L A G I L 10 20 24 33 46 50 53 54 63 71 73 84 90 F M I F I V L V P V S V M

ATV/r	I	I	N				L	G	K	L	V	L	E	M	G	F	I	D	I	I	A	G	V	L	I
	50	84	88				10	16	20	24	32	33	34	36	48	53	54	60	62	64	71	73	82	90	93
	L	V	S				F	E	R	I	I	F	Q	I	V	L	L	E	V	L	V	C	A	M	M
TPV/r	I	Q	T	V	N	I	L	L	M	K	M	I	H	L											
	47	58	74	82	83	84	10	33	36	43	46	54	69	89											
	V	E	P	L	D	V	V	F	I	T	L	A	K	I											
DRV/r	I	I	L	I			V	V	L	T	L														
	50	54	76	84			11	32	33	74	89														
	V	M	V	V			I	I	F	P	V														
RTV	V	I					L	K	V	L	M	M	I	I	A	V	L								
	82	84					10	20	32	33	36	46	50	54	71	77	90								
	A	V					F	M	I	F	I	I	I	V	V	I	M								

Table 2. The locations of PR mutations associated with significantly reduced PI binding (Rhee et al., 2010).

Mutation location	Mutation														
Active site	D	V	M	I	G	I	F	V	N	I					
	30	32	46	47	48	50	53	82	83	84					
	N	I	I	A	M	L	L	A	D	A					
Outside the active site	L	V	K	L	L	E	K	I	Q	G	T	L	N	L	L
	10	11	20	24	33	35	43	54	58	73	74	76	88	89	90
	F	L	T	F	F	G	T	A	E	C	P	V	D	V	M

Non-active site mutations associated with drug resistance

Since primary resistance mutations possess the potential to reduce the affinity of the PR for its natural substrates (Borman et al., 1996), compensatory secondary mutations have been seen to accumulate within the PR sequence. Secondary mutations are hypothesised to improve catalytic efficiency and viral replication of mutated PR variants, which are diminished as a result of primary mutations (Boden and Markowitz, 1998; Borman et al., 1996; Gulnik et al., 1995; Kozal, 2004). While secondary mutations are seen to accumulate in drug-resistant isolates, a relatively small number are located within the active site region (Table 1). Interestingly, in a way that is similar to primary mutations, certain patterns of secondary mutations associated with drug resistance emerge. Secondary mutations accumulate in a relatively predictable manner, based on the associated timing and combinations of inhibitor exposure (Kim and Baxter, 2008; Kozal, 2004; Winters et al., 2000). *In vitro* analyses involving serial passages of HIV strains in the presence of individual inhibitors have demonstrated that these mutations accumulate in particular patterns that are specific to each drug (Baxter et al., 2016). Furthermore, as primary and secondary mutations begin to accumulate within the PR sequence, it is likely that the PR will exhibit decreased susceptibility to more than one PI as

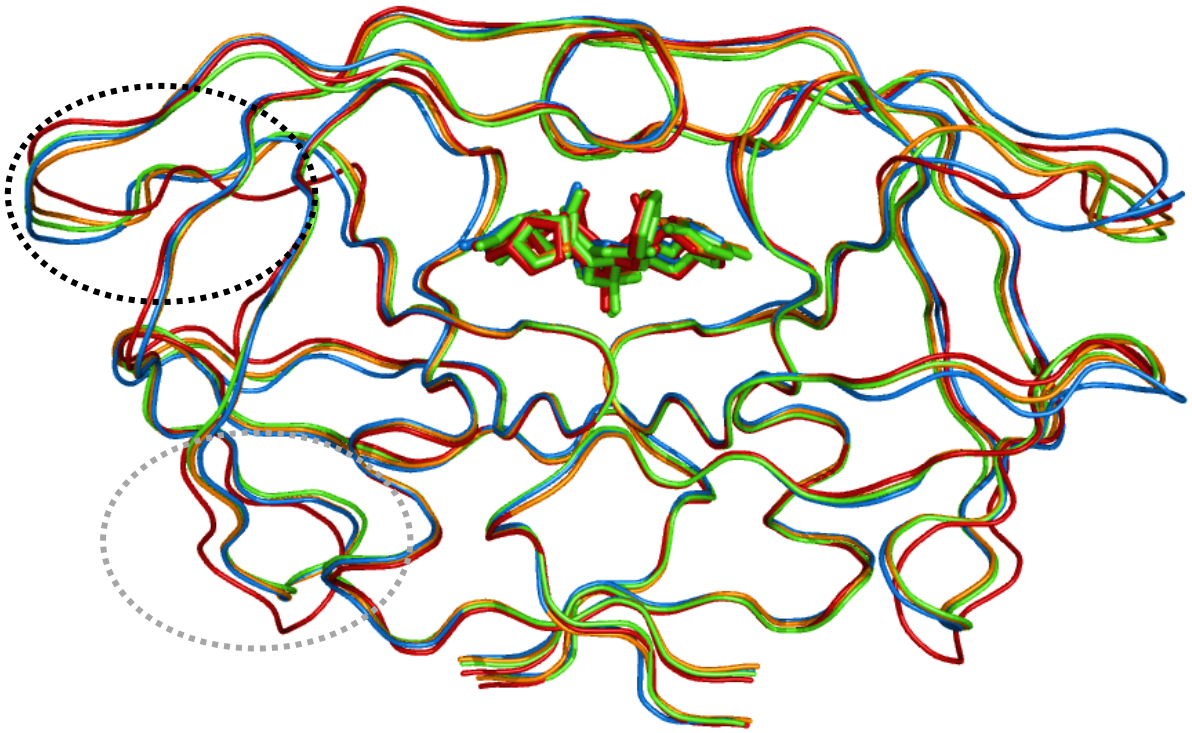
certain mutations are associated with broad resistance to PIs (Wensing et al., 2017). Interestingly, one analysis found that positions 10, 54, 71, 77, 82, 84, and 90 were the locations for the most frequently occurring mutations in multi-PI resistant isolates (Hertogs et al., 2000). The I84V mutation, for example, is not considered a primary mutation for some PIs, but it is associated with resistance to all currently available PIs (Table 1) (Wensing et al., 2017).

Non-active site mutations are known to contribute to drug resistance significantly (Clemente et al., 2004; Foulkes-Murzycki et al., 2007; Muzammil et al., 2003; Rhee et al., 2010; Weber et al., 2015). However, how these mutations impart their effect is yet to be fully understood. Some non-active site mutations are primary resistance mutations while others are suggested to contribute to drug resistance when found alongside other mutations. The cooperativity of non-active site mutations has been suggested previously (Foulkes-Murzycki et al., 2007; Goldfarb et al., 2015; Olsen et al., 1999). A recent study (Sherry et al., 2021; manuscript under review) focused on four cooperative non-active site mutations; namely, I13V, L33F, I62V and V77I. In this instance, the presence of the L33F mutation in PR isolates increases the occurrence of I13V from 34% to 51%, I62V from 16% to 48%, and V77I from 18% to 35% relative to the general population (Foulkes-Murzycki et al., 2007). This observed trend suggests a cooperativity in the mechanism of action of these mutations. The I13V/I62V/V77I mutations increase the dynamics of the fulcrum, cantilever, hinge, and flap regions of the mutated protease while in complex with SQV, DRV and ATV (Sherry et al., 2021 manuscript under review). Specifically, these mutations shift the conformational equilibrium of the flaps towards the semi-open conformation when bound to these inhibitors. Moreover, with the addition of the L33F mutation, the flap tips of the PR adopted curled conformers and coordinated several water molecules into the active site cavity in the presence of PIs.

The importance of non-active site mutations is emphasised by the observation that PR variants that exhibit cross-resistance to multiple inhibitors contain a constellation of non-active site mutations and only one or two active site mutations (Baxter et al., 2016; Clemente et al., 2004; Kozal, 2004; Ohtaka et al., 2002; Olsen et al., 1999). Moreover, it has been clarified that the effect of major primary drug resistance mutations is significantly diminished in the absence of associated secondary non-active site mutations (Ishima et al., 2001; Johnson et al., 2003). Although it remains relatively unknown as to how these mutations impart their effect, some modes of action of non-active site mutations can be ruled out. For example, secondary mutations rarely involve charged residues and, therefore, they cannot influence activity through long-range electrostatic interactions. Furthermore, the majority of compensatory mutations (i.e. secondary mutations) occur outside of the active site and have little influence on inhibitor binding directly (Schock et al., 1996). Moreover, these non-active site mutations

lead to minor geometric rearrangements in the active site which does not explain the noticeable increase in activity that is observed (Chen et al., 1995; Ridky et al., 1998; Schock et al., 1996). Based on crystallographic data of the HIV PR, non-active site mutations appear to not impart any major change in the PR backbone (Figure 5). That being said, the largest differences in the overlaid structure of a WT (4LL3) and multi-drug resistant (MDR) PR (3TTP), as well as two mutated PRs (3D1Z & 3S53) are seen in the more-flexible regions of the PR such as the hinge regions and the β-turn of the cantilever tip region (Figure 5). The large differences in these regions are likely attributed to the inherent flexibility of these regions given that their flexibility is essential to the functioning of the enzyme (Gustchina and Weber, 1990). Additionally, the flexible nature of these regions and their location on the surface of the PR has been linked to the differences in crystal packing, which in turn possibly accounts for the observed differences in the crystal structures.

Based on crystallographic evidence, it is evident that the global structure of the PR does not change when multiple mutations are present in the sequence; however, these mutations still have a considerable effect on the binding affinity of inhibitors (Foulkes-Murzycki et al., 2007; Olsen et al., 1999; Perno et al., 2002; Rhee et al., 2010; Weber et al., 2015). Although crystallographic studies provide valuable insight into how a point mutation in the active site can directly affect inhibitor binding, the true effect of non-active site mutations on the functional mechanism of the PR is more intricate. This stems from the fact globular enzymes are dynamic systems that are capable of experiencing large structural shifts over nanosecond timescales as has been seen during numerous MD simulations (Foulkes-Murzycki et al., 2007; Hornak et al., 2006; W. R. Scott and Schiffer, 2000). Therefore, while analyses of crystallographic data provide useful insight into the structure of a protein, these data only represent a static “snapshot” of a system which is inherently dynamic and, thus, most of the dynamic information about these system over time is absent. Crystallographic atomic displacement parameters, or B-factors, are available in all protein crystal structures and contain information about the fluctuations of atoms in the structure. These B-factors have previously been used to determine conformational entropies (Polyansky et al., 2012). However, extraction of conformational entropies in this way has been shown to be unreliable when compared to entropies measured by NMR relaxation or through MD simulations (Caldararu et al., 2019). Furthermore, most protein crystal data are collected at 100 K which might lead to an underestimation of conformational entropies when compared to data collected at room temperature or *in vivo*. Perhaps most importantly, standard B-factors do not provide an accurate reflection of the dynamics of a protein in a crystal structure (Kuriyan and Weis, 1991).



```
4LL3      PQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMNLPGRWKPKMIGGIGGFIKVRQYD  60
3D1Z      PQITLWKRPLVTIKIGGQLKEALLDTGADDTVIEEMSLPGRWKPKMIGGIGGFMKVRQYD  60
3S53      PQITLWKRPLVTIKIGGQLKEALLDTGADDTIIEEMSLPGRWKPKMVGGIGGFIKVRQYD  60
3TTP      PQITLWQRPLVTVKIGGQLREALLDTGADDTIFEDINLPGKWTPKMVGGIGGFMKVRQYD  60

4LL3      QILIEICGHKAIGTVLVGPTPVNIIGRNLLTQIGCTLNF      99
3D1Z      QIIIEIAGHKAIGTVLVGPTPVNIIGRNLLTQIGATLNF      99
3S53      QIIIEIAGHKAIGTVLVGPTPINIIGRNLLTQIGATLNF      99
3TTP      QVVIEICGHKVTSPVLVGPTPLNIIGRNVLTQLGCTLNF      99
```

Figure 5. A) Structural overlay of WT and mutated PR variants in complex with DRV and B) the corresponding amino acid sequence alignment of the variants. The structure and sequences represent the MDR (3TTP), WT (4LL3), and two mutant PR variants (3D1Z & 3S53). The MDR PR contains 19 amino acid polymorphisms with respect to the WT PR. The flexible hinge region and cantilever tip regions of the PR are highlighted in the black and grey dotted ovals, respectively. Image generated using PyMOL (Schrödinger, LLC), sequence alignment generated using CLUSTAL-Omega.

A possible explanation for the effect of non-active site mutations may lie in the difference in conformational flexibility between the WT and mutant PRs (Chen et al., 1995; Gulnik et al., 2000). There is evidence to indicate that these non-active site mutations would influence the drug/protein interface in a cooperative manner by changing the dynamics of regions within the PR (Nalam et al., 2010; Prabu-Jeyabalan et al., 2003). Therefore, these mutations alter drug binding without directly disrupting the drug/protein interface through steric and geometric changes as is the case with active site mutations. It is likely that the mechanism of action of these mutations is centred around changing the dynamics of the flap regions which are inherently flexible and, therefore, more liable to change. Since these regions play a crucial role in the exposure of the inhibitors to the active site, as well as a significant role in binding, alterations in their dynamics could have considerable effects on drug binding. Therefore, mutations outside of the active site may have significant implications on the extent to which the flaps open or close and, as a result, change the nature of the network of hydrophobic interactions thereby modulating access to the active site cavity. Consequently, by indirectly influencing the dynamics of the flap regions, these non-active site mutations may reduce the ability of the inhibitor to bind to or access the active site cavity (Goldfarb et al., 2015).

Notably, primary mutations have the potential to diminish the binding affinity of natural substrates; Gag and Gag-Pol (Borman et al., 1996) and, thus, subsequent mutations have been seen to accumulate outside of the active site of the PR. These so-called secondary mutations have been hypothesised to compensate for variations generated by primary mutations in ways that improve catalytic efficiency and viral replication (Boden and Markowitz, 1998; Borman et al., 1996; Gulnik et al., 1995; Kozal, 2004). In summation, the combined effect of primary and secondary mutations alters the viral phenotype to decrease susceptibility of the PR to inhibition by one or more PIs while maintaining selectivity for natural substrates. It is also noteworthy that mutations outside of the viral PR sequence have been known to accumulate in the processing sites of the Gag/Gag-Pol polyprotein, which further improves substrate binding and cleavage, and, therefore, viral fitness (Doyon et al., 1996; Williams et al., 2019; Zondagh et al., 2019). As such, it is necessary for the co-evolution of mutations within the PR as well as the Gag/Gag-Pol cleavage sites to maintain and/or improve the viral fitness of HIV.

The role of Gag mutations in resistance

As mentioned previously, the Gag polyprotein and the PR play key roles in the process of viral maturation. Specifically, the PR cleaves Gag precursor (Figure 6) into structural proteins and functional enzymes which results in the production of mature and infectious viral progeny (Göttlinger et al., 1989). The Gag cleavage sites do not share a consensus sequence (Figure 6) but, rather, these sites occupy a conserved shape within the active site cavity which is

proposed to determine substrate specificity (Prabu-Jeyabalan et al., 2002). To inhibit processing of Gag and, therefore, prevent viral maturation, PIs are designed to block the activity of the PR in a competitive manner with Gag substrate. Mutations associated with PIs have been reported to occur concurrently in both the PR and Gag which indicates that a synergy exists between both in order to overcome PIs (Codoñer et al., 2017). Therefore, it is understood that mutations in the Gag cleavage sites directly contribute to PI resistance (Clavel and Mammano, 2010). Moreover, mutations in the Gag that occur outside of the cleavage sites contribute to drug resistance in a manner that compensates for the loss of viral fitness which arise from primary mutations in the PR (Özen et al., 2014). It is interesting to note the emergence of new clinical resistance mutations in the PR have been decreasingly reported, which suggests that there is a limit to the number of mutations that are tolerable within the PR sequence. In contrast, Gag is approximately 500 amino acids in length and, likely, has more tolerance for compensatory mutation in the PR-Gag synergy relating to drug resistance.

The mechanisms through which Gag cleavage site mutations contribute to PI resistance and improve viral fitness have yet to be completely understood. Some insight into these mechanisms has been elucidated from investigations into the co-evolution of PR and Gag mutations. In particular, the PR mutation I84V which is associated with mutations in Gag cleavage sites; SP2/p6 (L449F) and NC/SP2 (L429F-Q430R-A431V) and, similarly, the PR mutation V82A which is associated with Gag mutations A431V and I437V (Clavel and Mammano, 2010). For instance, the A431V mutation improves cleavage of the NC/SP2 peptide. In this instance, the smaller alanine residue at position 82 in the PR results in lost contacts between the active site and PIs, thereby, reducing their binding affinity. However, this change also affects the extent to which NC/SP2 makes contact with the PR. The A431V mutation results in a novel point of contact with the PR in the active site cavity at a location that is not naturally occupied by the substrate. Interestingly, this structural change does not involve residues of the substrate that make contact with the PR at position 82 but, rather, compensates for lost contacts in an alternative location and produces an increase in enzyme activity of the mutated PR (Prabu-Jeyabalan et al., 2004). As such, these mutations work synergistically to simultaneously minimise inhibitor contacts while maintaining substrate contacts. This model of co-evolution is likely to extend to other cleavage site mutations and PR mutations. An unanswered question involving Gag mutations revolves around how such mutations can produce resistance to PIs in a manner that is independent of changes in fitness. This may be explained by the substrate binding with better efficacy to the PR in the presence of PIs.

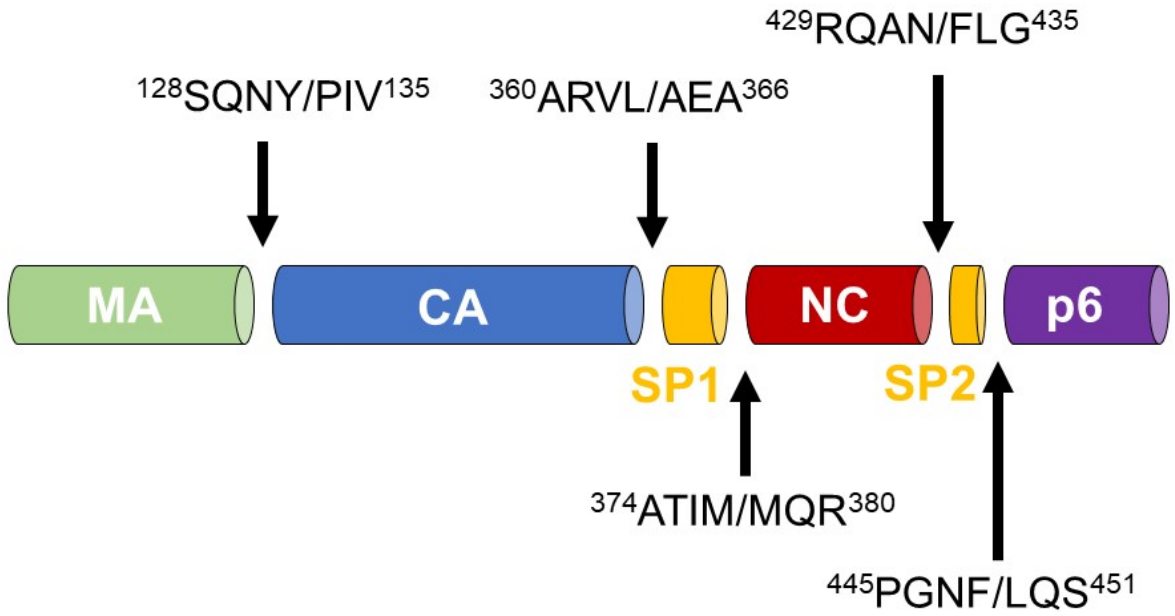


Figure 6. Diagrammatic representation of the Gag polyprotein. The Gag polyprotein is comprised of the matrix (MA) protein, capsid (CA) protein, nucleocapsid (NC) protein and p6 protein. Two additional linker proteins, SP1 and SP2, are present in the polyprotein. The SP1 protein separates the CA and NC proteins while the SP2 protein is located between the NC and p6 proteins. The black arrows indicate the position of PR cleavage sites. The sequences of the seven residues which comprise the cleavage sites are labelled, with a slash to indicate the position of the scissile bond. Adapted from Ohishi et al., 2011.

This explanation may seem unlikely given that the binding affinity of PIs for the PR is several orders of magnitude greater than the affinity for the Gag cleavage sites (picomolar vs micromolar) (Moore et al., 1989). However, the intra-virion concentration of PIs is largely unknown and, thus, it is possible that the local concentration of the competing binders in the virion is such that the concentration of Gag cleavage sites can be considered as being very high in comparison to that of the PR and PIs (Briggs et al., 2004). Therefore, enhanced Gag processing could shift the binding equilibrium in favour of Gag substrate which would confer resistance to inhibitors. Clearly, more work is required to evaluate the intracellular concentrations of PIs at the site of assembly of the virion in order to understand how the substrates can out-compete the binding of the PIs.

Flexibility-assisted functionality of the PR

A variety of mutations can occur over a wide range of locations within the PR sequence, but outside of the active site, and still impart a similar effect to active site mutations in terms of reducing drug susceptibility. Since structural changes as a result of non-active site mutations are minimal in the PR, a common mechanism unrelated to structural changes could be responsible for the observed effect these mutations have on drug binding. Additionally, as stated previously, the PR is seen to accumulate multiple non-active site mutations which suggests cooperativity in the nature of their mode of action towards drug resistance. Upon drug binding, the structure of the PR adjusts to accommodate the inhibitor through intricate rearrangements of residues throughout the protein sequence, including those that are mutated (Prabu-Jeyabalan et al., 2003; Shen et al., 2010). Largescale motions of the PR have been found to be crucial to direct the substrate into an optimal position in which successful catalysis can occur, a process that is largely dependent on the overall plasticity of the enzyme/substrate complex (Piana et al., 2002). Similar conclusions have been drawn for other enzymatic systems (Kahn and Bruice, 2000; Radkiewicz and Brooks, 2000) and, thus, it has been hypothesised that conformational flexibility plays a role in substrate and inhibitor binding as well as in changes in the catalytic efficiency of mutated PR variants.

A flexibility-assisted catalysis mechanism has been suggested as an explanation for the changes observed in mutated PR variants. The mechanism suggests that the enzyme/substrate complexes sample an ensemble of conformations when in complex with one another. Some of the conformations in this ensemble, referred to as “reactive conformations”, have the chemical groups of the enzyme and substrate arranged in the correct location and orientation that is necessary for the cleavage reaction to occur. Furthermore, building on this notion, some mutation-containing complexes might not sample reactive conformations as often as WT-containing complexes because of changes in the global flexibility of the complex. As such, even some peptide substrates could act as inhibitors if,

when complexed to the PR, the complex is unable to achieve the proper reactive conformations (Piana et al., 2002). This could be due to the structural rearrangements required of the complex acting as an energetic barrier that is unable to be overcome and, therefore, catalysis would be hindered. This hypothesis suggests a more intricate and delicate system that is influenced through minor changes in the overall plasticity of the complex as opposed to local or global structural changes when mutations occur. The importance of the plasticity of the hydrophobic core in the functionality of the PR has been illustrated using site-directed cysteine cross-linking (Mittal et al., 2012). This study provided evidence for the requirement of conformational changes in the hydrophobic core of the PR in order for successful catalysis to occur. It was found that when crosslinks were formed within the double cysteine containing mutant (G16C/L38C), the conformational restraints in the core of the PR compromised its activity. However, the activity of the PR was largely restored upon reduction of the disulfide bridges, which implicated the importance of core rearrangements and flexibility in the function of the PR. The oxidation and reduction of these restrictive intramolecular crosslinks acted as a molecular switch which could essentially control the activity of the protease and turn the PR "off" and "on", respectively (Mittal et al., 2012).

Small changes imposed by non-active site mutations can induce significant differences in the dynamics of the protein without modifying the global structure of the enzyme. For example, an MD study of G48T/L89M PR revealed that the variants exhibited wider flap conformations when in complex with SQV which reduced the van der Waals interactions between the enzyme and the inhibitor which in turn reduces the binding affinity (Goldfarb et al., 2015). Interestingly, fluorescence-based enzyme assays performed on WT, G48V, L90M, and G48V/L90M PRs revealed that the association rates (k_{on}) for the proteins in complex with SQV were not significantly different. While, in contrast, the dissociation rates (k_{off}) were increased for the mutated PRs in comparison to the WT. Therefore, the reduced affinity observed for these mutant PRs was primarily the result of increased dissociation rates which could be attributed to increased flap opening rates, decreased flap closing rates, or both (Maschera et al., 1996). Similarly, the E35D mutation is located in the hinge region of the PR, yet it is associated with increased resistance. This mutation is seen to disrupt the water-mediated interaction between the amide proton of I50 and the sulfonyl-oxygen of ATV which is purported to be crucial for tight ATV binding (Kim et al., 1995). Interestingly, this loss of interaction is not accompanied by opening of the flaps, rather, this disruption is caused by a rotation of the φ-angle of I50 from -73° to approximately +60° (Meiselbach et al., 2007). The E35D mutation, therefore, altered the conformation of the bound state of the PR in a manner that destabilised the complex to make inhibitor binding less favourable.

Interestingly, in a similar manner, a variant containing I13V/L33F/I62V/V77I mutations exhibited a reduction in the number of conserved hydrogen bonds with the active site Asp25 residues (manuscript in preparation), which are key in the formation of the inhibitor/protein complex for ATV, DRV and SQV. Furthermore, this variant exhibited a tendency to adopt wider flap conformations and to facilitate the coordination of several water molecules (between 2 and 3) in the active site in order to compensate for the lost hydrogen bonds, which collectively modulated the interactions between the PR and PIs. The greater propensity for water-bridge formation was hypothesised to increase the mobility of the PIs in the active site because these water molecules can rotate and slide across each other. As a result, the hydrogen bonds formed between the enzyme and inhibitor would be more transient than those that would form directly between the inhibitor and the protein. Therefore, mutated PRs may increase flexibility of the complexed enzyme in a manner that minimises interactions with PIs and destabilises the complex so as to increase the dissociation rate of the inhibitors. MD simulations of WT and V82F/I84V mutant PRs revealed that the average backbone RMSF value per residue of the WT was larger than that of the mutant PR (Hou and Yu, 2007). Thus, the mutated PR demonstrated a higher degree of stability while the WT PR had a larger overall flexibility in the apo state. These findings were echoed in differential scanning calorimetry experiments where the mutated PR underwent thermal denaturation at 4 °C higher than the WT, meaning that the mutated PR was more stable (Todd et al., 2000). It is evident from MD simulations that non-active site mutations may serve to increase the rigidity of the regions within the PR, including the hinges and flaps (Goldfarb et al., 2015). This observation has further been revealed by MD simulations of PR containing flap region mutations, such as M46I and K45I, which cause an increase in the stability of the closed flap conformation and less flap flexibility in comparison to the WT (Mahalingam et al., 2001; Piana et al., 2002).

What is even more intriguing to note is that this phenomenon is evident when comparing the efficacy of PIs between different HIV-1 subtypes. For instance, the subtype C PR differs from the subtype B PR at eight amino acid positions outside of the active site cavity; namely, T12S, I15V, L19I, M36I, R41K, H69K, L89M and I93L (Velázquez-Campoy et al., 2003). These naturally occurring polymorphisms (NOPs) do not affect the overall structure of the subtype C PR in comparison to the subtype B PR; however, their effect on the efficacy with which the PR binds to PIs is significant (Naicker and Sayed, 2014). This is due to the fact that all of the FDA-approved PIs that are currently available were designed using the subtype B PR as a target. Thermodynamic analyses of PIs tested against the subtype C PR indicated that, in general, the subtype C PR exhibited 3- to 24- fold weaker binding affinity when compared to subtype B PR (Velázquez-Campoy et al., 2003). Moreover, viral phenotypic assays performed on the subtype C PR revealed that IC_{90} values were 0.16- to 3- fold greater in comparison to the

control subtype B PR strains (R. W. Shafer et al., 1999). In addition, MD studies of the subtype C PR indicated that the hinge, fulcrum, and cantilever regions of PR display increased stability in comparison to the subtype B PR (Naicker et al., 2014). An increase in the stability of these regions may stabilise the core of the PR, which may explain the wide range of open conformers the subtype C PR exhibits in comparison to the subtype B PR. The greater propensity that the C-SA PR exhibits for open conformers may explain the decreased binding affinity of PIs (Naicker et al., 2014). Furthermore, differential scanning calorimetry (DSC) experiments investigating the stability of the native conformation of the A, B, C and G subtype PRs revealed that the A, C and G subtype PRs appeared to exhibit higher structural stability than the B subtype PR (Velázquez-Campoy et al., 2001). The DSC technique measures the magnitude of the forces that stabilise the native structure of proteins (Todd et al., 1998), and alterations in structural stability as a result of NOPs may affect binding affinity due to increases in the energy needed to induce the fit between the enzyme and inhibitor. A similar conclusion could, therefore, be drawn with regard to the impact of non-active site mutations on the structural stability of the PR. Since the closing and rearrangement of the flaps during ligand binding is associated with an unfavourable enthalpic cost (Luque et al., 1998), if mutated PRs prefer closed or semi-open conformations more so than the WT, then this feature could represent a mechanism for drug resistance since all drugs would pay a larger enthalpic cost to alter the conformation of the mutants flaps. In summation, the role that non-active site mutations play in resistance may be two-fold. Firstly, non-active site mutations may serve to increase the rigidity of the apo PR which would potentially impose a larger energetic penalty for the conformational changes necessary for induced-fit inhibitor binding; therefore, decreasing the energetic favourability of PI binding (Figure 7, D). Secondly, these mutations may increase the flexibility of the complexed PR and, specifically, alter the flexibility and mobility of the flaps. In this way, these mutations can alter or even eliminate intermolecular interactions that facilitate inhibition and, in doing so, decrease the likelihood of the complex assuming and maintaining a conformation that results in inhibition (Figure 7, E). The non-active site mutations in the core of the PR may result in novel conformational heterogeneity that selects for conformers that preferentially bind to a substrate in reactive conformations and simultaneously diminish inhibitor binding, perhaps by excluding inhibition conformations. Furthermore, non-active site mutations change the dynamics of the hydrophobic core which, in turn, alters the functional properties of the active site thus shifting the relative binding affinities for inhibitors and substrates in favour of inhibitor resistance and greater substrate binding. It is apparent, therefore, that the non-active site mutations may alter the flexibility of the enzyme in a manner that decreases the association rate of drugs with the mutated enzyme or, conversely, increases the dissociation rate of the drugs from the complex (Figure 7).

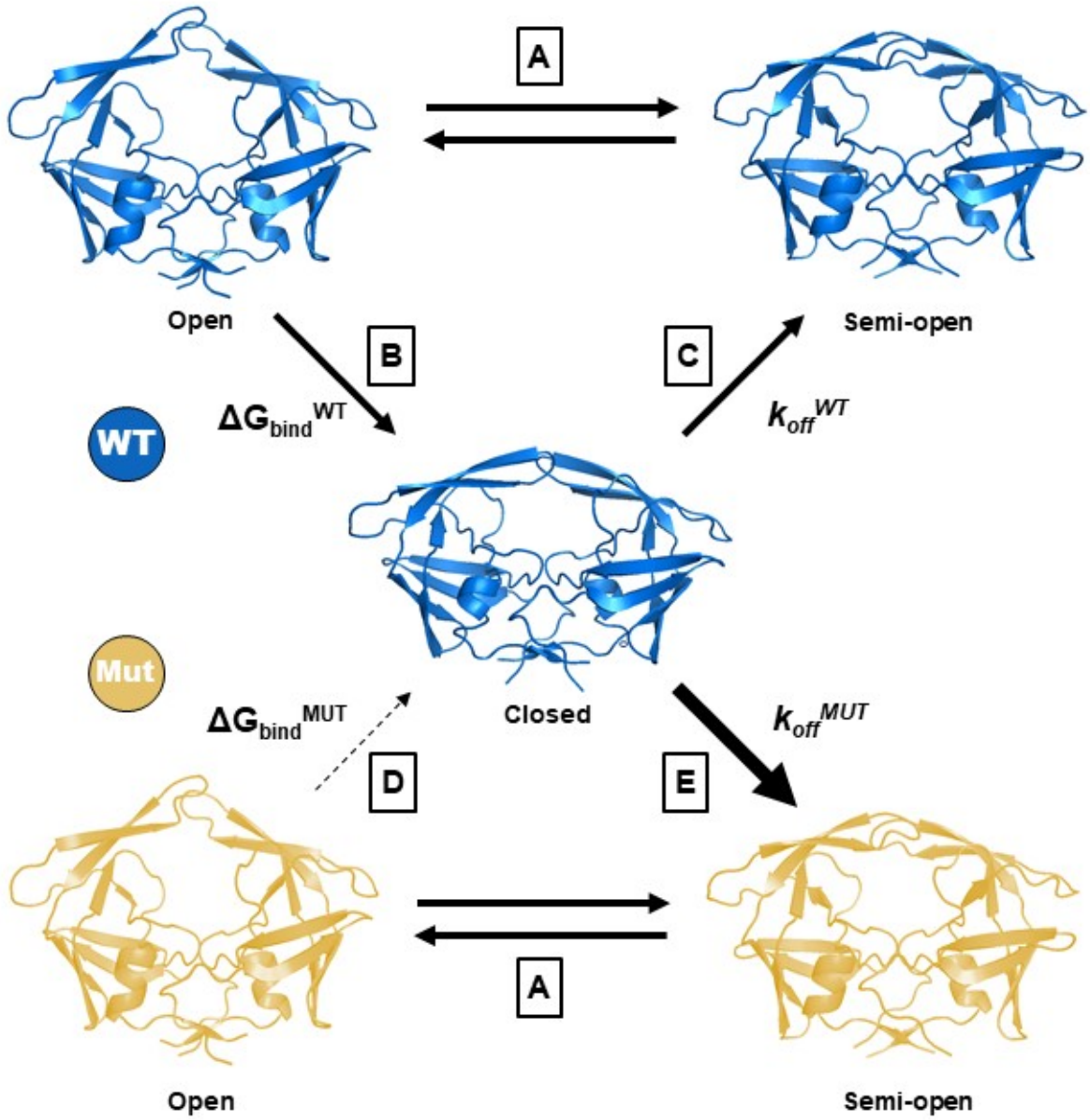


Figure 7. Non-active site mutations alter PR dynamics in favour of resistance. A) The apo PR exists largely in equilibrium between the open and semi-open conformations B) PIs bind to and bring the WT PR into the closed conformation C) PIs dissociate from the active site cavity of the PR D) Non-active site mutations may serve to decrease the energetic favourability (ΔG_{bind}) of the inhibitor binding to the mutated PR or E) Non-active site mutations could increase the off rate (k_{off}) of inhibitors bound to the mutated PR.

Allosteric inhibition of the PR

Considering that the PR exists in an equilibrium between closed, semi-open, and open conformations, designing inhibitors that perturb this equilibrium may present a promising avenue for the next generation of PIs. Allosteric inhibition of the PR has been suggested previously since such inhibitors would not directly compete with the substrate for the same binding site. Assuming that the dynamics of the enzyme must be conserved, it may be more difficult for the virus to develop mutations that are able to maintain the functional dynamics of the enzyme while simultaneously evading inhibitors that target the PR dynamics. MD simulations have revealed that the mobility of the flaps is closely correlated with movements in the hinge region (Foulkes-Murzycki et al., 2007; Gustchina and Weber, 1990; Hornak et al., 2006) and, as such, allosteric inhibitors that target these regions have been suggested. Four allosteric binding sites have been identified to date, summarised in Figure 8.

Fragment-based crystallographic screening is increasingly popular as a technique used to design novel therapeutics. Such studies have revealed a novel allosteric binding site, termed the ‘exo site’, on the surface of the PR near the β-turn of the fulcrum region formed by residues Lys14, Gly16, Gly17, Gln18 and Leu63. The ‘exo site’ has been shown to bind the small molecule, 4D9 (2-methylcyclohexanol) (Figure 9) (Perryman et al., 2010). Similarly, a study making use of a library of brominated fragments found that a number of these fragments bind in a region beyond the original description of the ‘exo-site’ based on the crystal structures with 4D9 (PDB ID: 3KF0, 3KFN)(Tiefenbrunn et al., 2014). In this study, a number of brominate compounds extend beyond the ‘exo site’ and bind between Leu63 and Glu65 (Figure 8) (Tiefenbrunn et al., 2014). The most promising compound from this study, Br27 (1-bromo-2-naphthoic acid) (Figure 9), binds at the base of the ‘exo site’ and is a much larger molecule than 4D9 and shows that stacking interactions with Leu63, Glu65, and Lys70 can be exploited in future drug development. Additionally, a new interaction was identified in this Br27-bound complex in the form of an intermolecular salt bridge with Lys14 (Tiefenbrunn et al., 2014). Further investigations into possible allosteric binding sites of the PR have revealed a targetable pocket on the top of the flap region. This ‘flap top’ pocket is formed by residues Trp42, Pro44, Met46, Lys55, Arg57 and Pro79 (Figure 10). Compounds such as 1F1 (indole-6-carboxylic-acid), 1F1-N (3-indolepropionic acid), AK-2097 and 2-acetylbenzothiophene have been shown to bind to this site on the PR (PDB ID: 4EJK)(Figure 10) (Perryman et al., 2010; Tiefenbrunn et al., 2014). Notably, ligand binding at this site has been seen to cause the PR to favour the closed flap conformation (Tiefenbrunn et al., 2013).

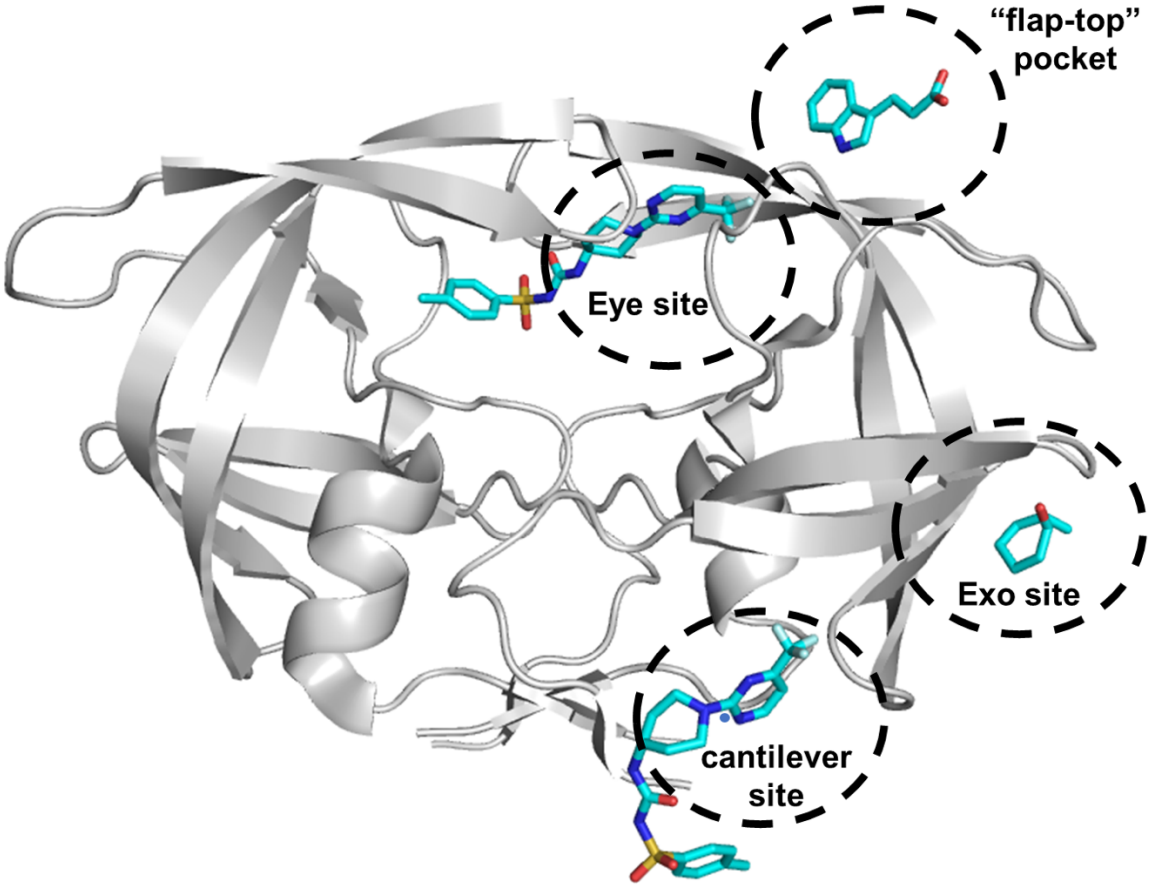


Figure 8. Allosteric binding sites identified for the HIV PR. Four binding sites for small molecules have been identified on the surface of the PR; namely, the exo site, the "flap top" pocket, the Eye site and the cantilever tip-region site.

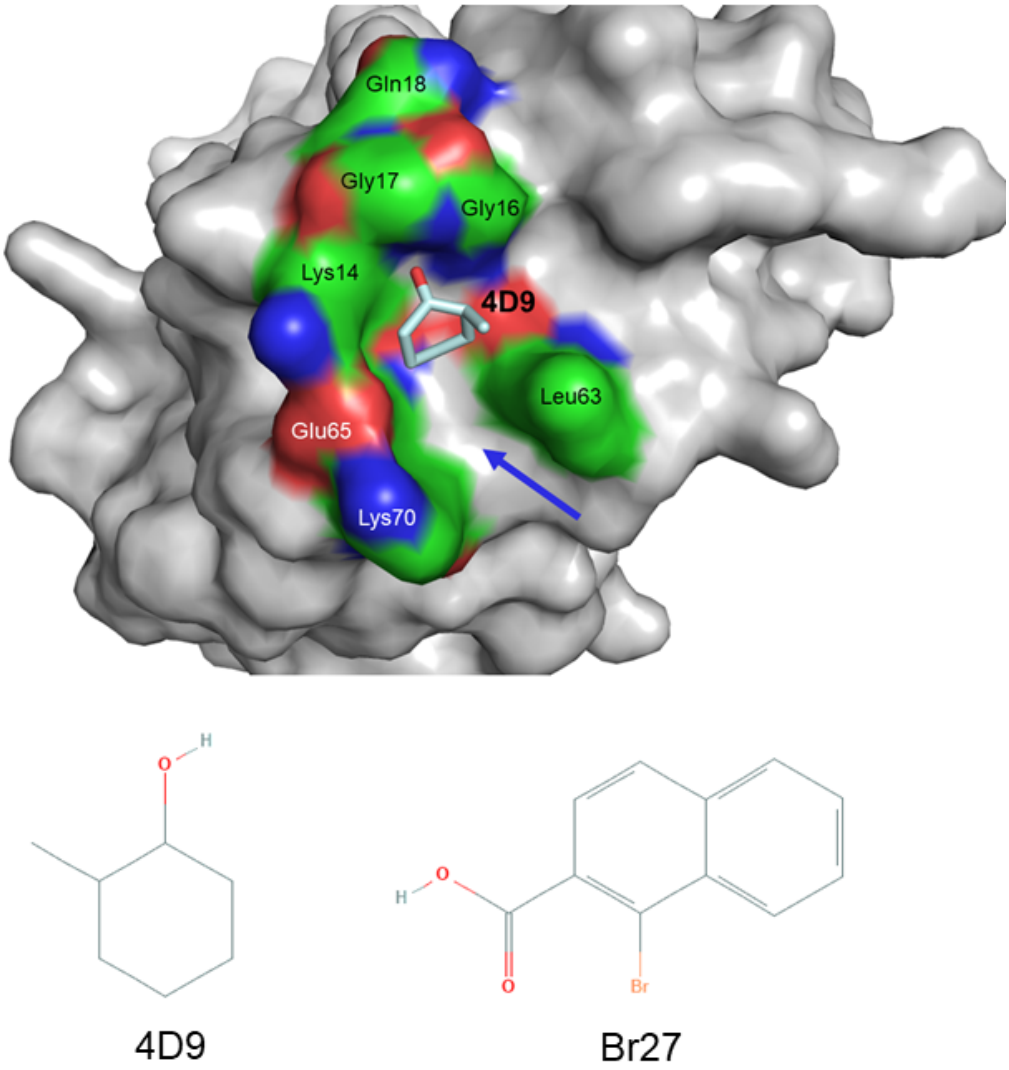


Figure 9. The 'exo site' of the PR occupied by the small fragment 4D9 (PDB ID: 3KF0). The exo site is primarily comprised of residues Lys14, Gly16, Gly17, Gln18 and Leu63. An extension of the exo site which could be exploited to accommodate larger ligands is indicated by the blue arrow. The structures of fragments that bind to this site (4D9 and Br27) are shown.

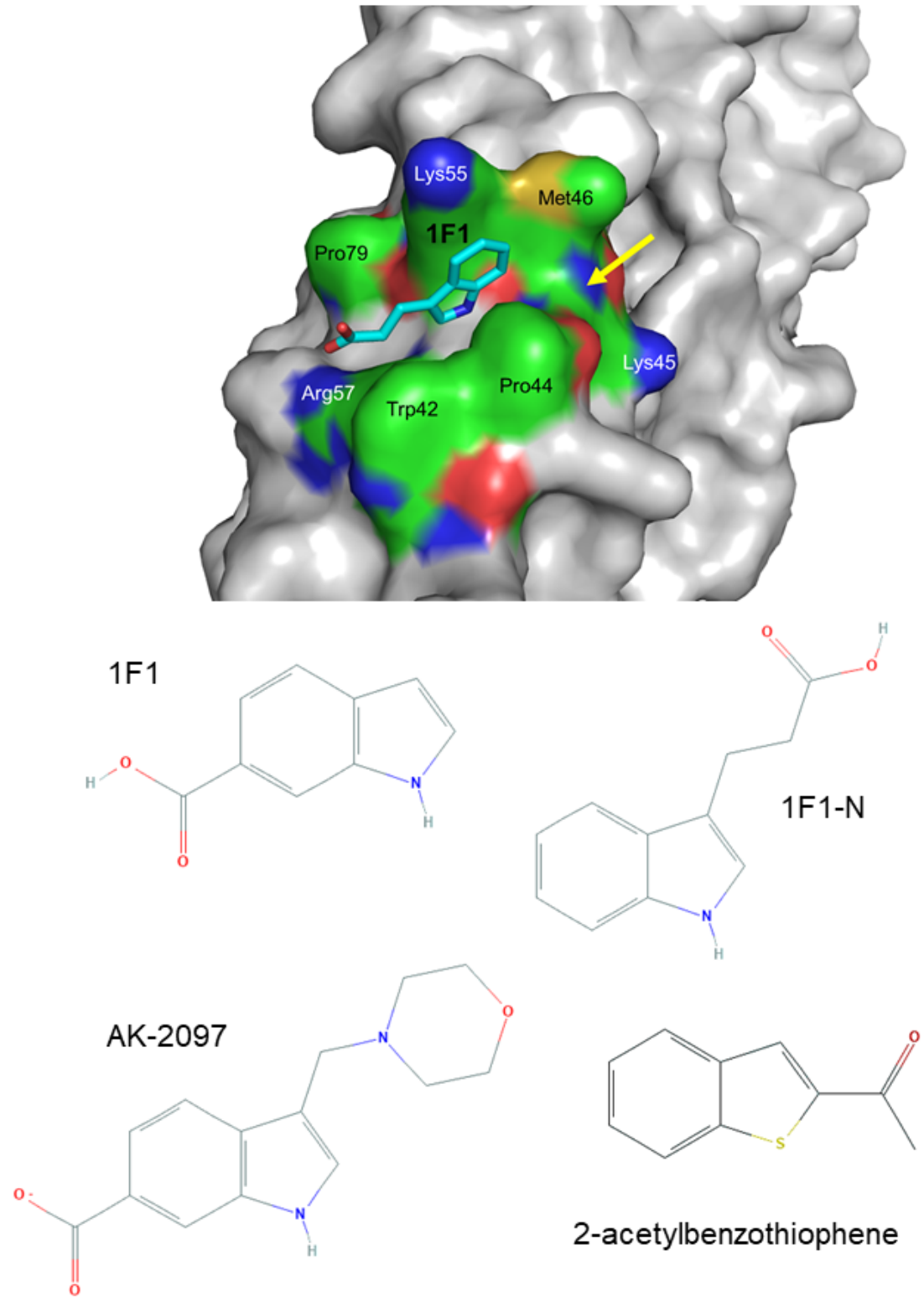


Figure 10. The 'flap top' pocket located on the surface of the PR flaps with 1F1 bound (PDB ID: 4EJK). The flap top pocket is formed by Trp42, Pro44, Met46, Lys55, Arg57 and Pro79. The structures of allosteric binders at this binding pocket; namely, 1F1, 1F1-N, AK-2097 and 2-acetylbenzothiophene are represented. The P2 site which extends from the flap top pocket is indicated by the yellow arrow.

Free energy calculations revealed that the largest contribution to binding affinity for 1F1 and 1F1-N binding to the 'flap top' pocket relies on an intermolecular salt bridge formed between Arg57 and the carboxylate ion of the ligand (Tiefenbrunn et al., 2014). It is understood that the disruption of the Glu35-Arg57 salt bridge results in the flaps adopting an open conformation (Swairjo et al., 1998; Zondagh et al., 2019) and, furthermore, the absence of this salt bridge has been associated with drug resistance in the subtype C PR (Naicker et al., 2013). As such, drugs that are able to form a stable ionic interaction with Arg57 could lock the flaps in the closed conformation since the absence of this salt bridge is necessary for flap opening. Therefore, the formation of the stable salt bridge with Arg57 may provide a possible explanation for the mechanism of action of binders at this site. Further investigation into the efficacy of Br6, revealed that the presence of this fragment at the flap top pocket causes an MDR PR (PDB ID: 2AZC) mutant to behave much more like the WT PR when in complex with an active site inhibitor, TL-3 (Luchi et al., 2018). In this study, the presence of the MDR mutations shift the equilibrium of the flap conformations from closed in the WT to semi-open in the MDR variant. When Br6 was placed at the flap top pocket, the equilibrium of the MDR PR was restored back to the closed conformation of the flaps (Luchi et al., 2018). Structural analyses revealed that the L63P and V77I mutations, located in the cantilever region, increase the strength of the backbone interactions of the MDR PR, thus, preventing the clockwise/anti-clockwise movement of the hinges which is required for keeping the flaps closed. When Br6 is present, however, it perturbs the Glu35-Arg57 salt bridge which results in restoration of clockwise/anti-clockwise motion of the hinges in the resistant form of the PR, returning the dynamics of the MDR PR to that of the WT (Luchi et al., 2018). The manner in which ligand binding to this site occurs is further revealed by AK-2097. This ligand buries its indole ring in the pocket and simultaneously forms a hydrogen bond with the carbonyl oxygen of Val56 via the NH located in the indole group of AK-2097.

A second binding site (P2) which is an extension of the 'flap top' pocket has also been seen to bind a number of brominated fragments (Tiefenbrunn et al., 2014). This P2 site is formed by the side chains of Glu35, Lys55, Arg57, and Pro79 and is located adjacent to the 'flap top' site but towards Met46 (Figure 10). This second pocket shows the potential for designing larger, more potent ligands starting from 1F1 and 1F1-N with the intention of exploiting this second adjacent pocket while simultaneously preserving interactions within the first pocket and, most importantly, the intermolecular salt bridge with Arg57. Biochemical assays of 1F1 and 1F1-N revealed K_d values in the low μM range (Tiefenbrunn et al., 2013). Since the currently confirmed ligands exhibit weak binding, expanding these fragments to form more intermolecular interactions with the PR can increase their binding affinity and, therefore, their efficacy.

The existence of these binding sites outside of the active site of the PR can establish a starting point for the development of larger drug molecules, < 500 Da considering Lipinski's rule of five, with higher binding affinities for these sites which is not unreasonable since certain PIs have already been seen to bind to the 'flap top' pocket. Crystal structures of the PR have been found to bind DRV and APV at this alternative binding pocket (Figure 11). In the case of these crystal structures, the inhibitors form two hydrogen bonds with the PR, specifically with the carbonyl oxygen of Lys45 and the guanidino group of Arg57 (Figure 11).

Further structural investigations of the PR have revealed that the enzyme possesses hydrophobic flap-tip recognition pockets, or 'Eye sites', which are comprised of residues Val32, Ile47, Gly48, Gly49, Ile50, Ile54, Val56, Gly78, Pro79, Thr80, Pro81 and Ile84 (Figure 12). Upon substrate binding, residues of the flap tip region (49-GGIG-52) interact with this highly conserved region on the opposite monomer. The Eye sites are not accessible when the PR is in the closed conformation as the flap tip of the opposite monomer covers each site. However, in the open conformation the flap tip undocks which opens up the Eye sites. One study modelled a potential lead compound capable of binding to the Eye sites; 2,2,4-trimethyl-1,2-dihydroquinolin-6-yl benzoate and, subsequently, tested a derivative of this compound (2,2,4-trimethyl-1,2-dihydroquinolin-6-yl-4-methoxy benzoate) experimentally (Damm et al., 2008). The derivative of the lead compound was found to inhibit the PR with an IC_{50} value of ~18 uM (Damm et al., 2008). Further confirmation of the possibility of targeting the Eye site was confirmed through fragment-based crystallography, wherein 5NI (5-nitroindole) was found to bind to the Eye site with the PR in the semi-open conformation (Perryman et al., 2010). The power of allosteric inhibition at this site was exemplified in the study of a novel compound, named NIT, which was equipotent against an MDR and WT PR (Ung et al., 2014). NIT was shown to exhibit mixed competitive inhibition, demonstrating its ability to modulate activity of the PR through mechanisms other than competing for the active site (Ung et al., 2014). Furthermore, MD simulations of the PR in complex with NIT illustrated that NIT remains stably bound to the Eye site and affects the dynamics of the flaps (Ung et al., 2014). Interestingly there is a crystal structure of the PR (PDB ID: 5VJ3) with a novel compound (HV9) bound to the Eye site. However, to date this structure has no publication associated with the deposition.

Targeting the Eye site could pose interesting mechanisms of inhibition for both the closed and open conformations of the PR. Curling of the flap tip regions into the Eye site is thought to drive the conformational change into the open conformation (Perryman et al., 2004; S. Piana et al., 2002; W. R. Scott and Schiffer, 2000). Then, blocking the Eye site may block this interaction in a manner that prevents the flaps from being able to open fully and accommodate substrate. Alternatively, blocking the Eye site could prevent the flaps from closing properly

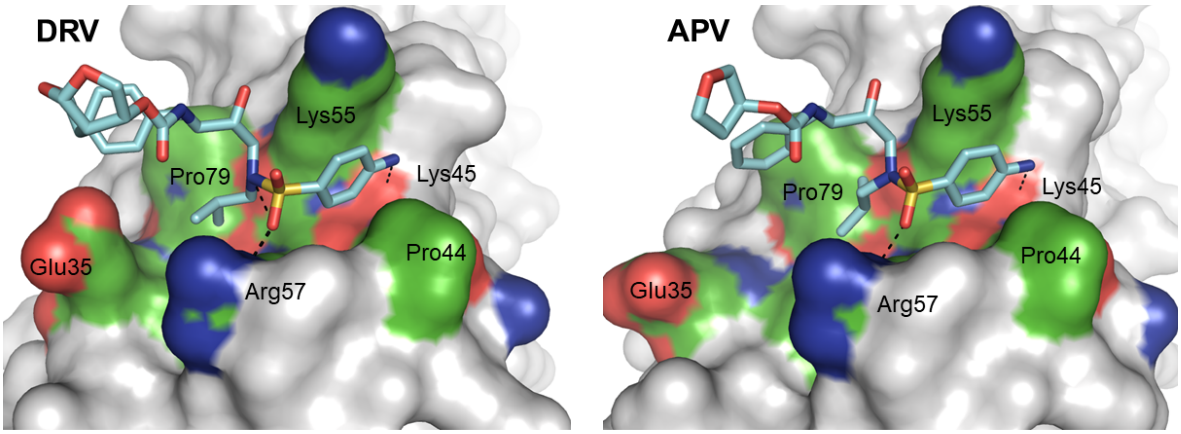


Figure 11. PIs bound to the 'flap top' binding pocket located on the surface of the flaps. Two PIs, APV and DRV, bind to the flap top pocket located on the surface of the flaps of the PR.

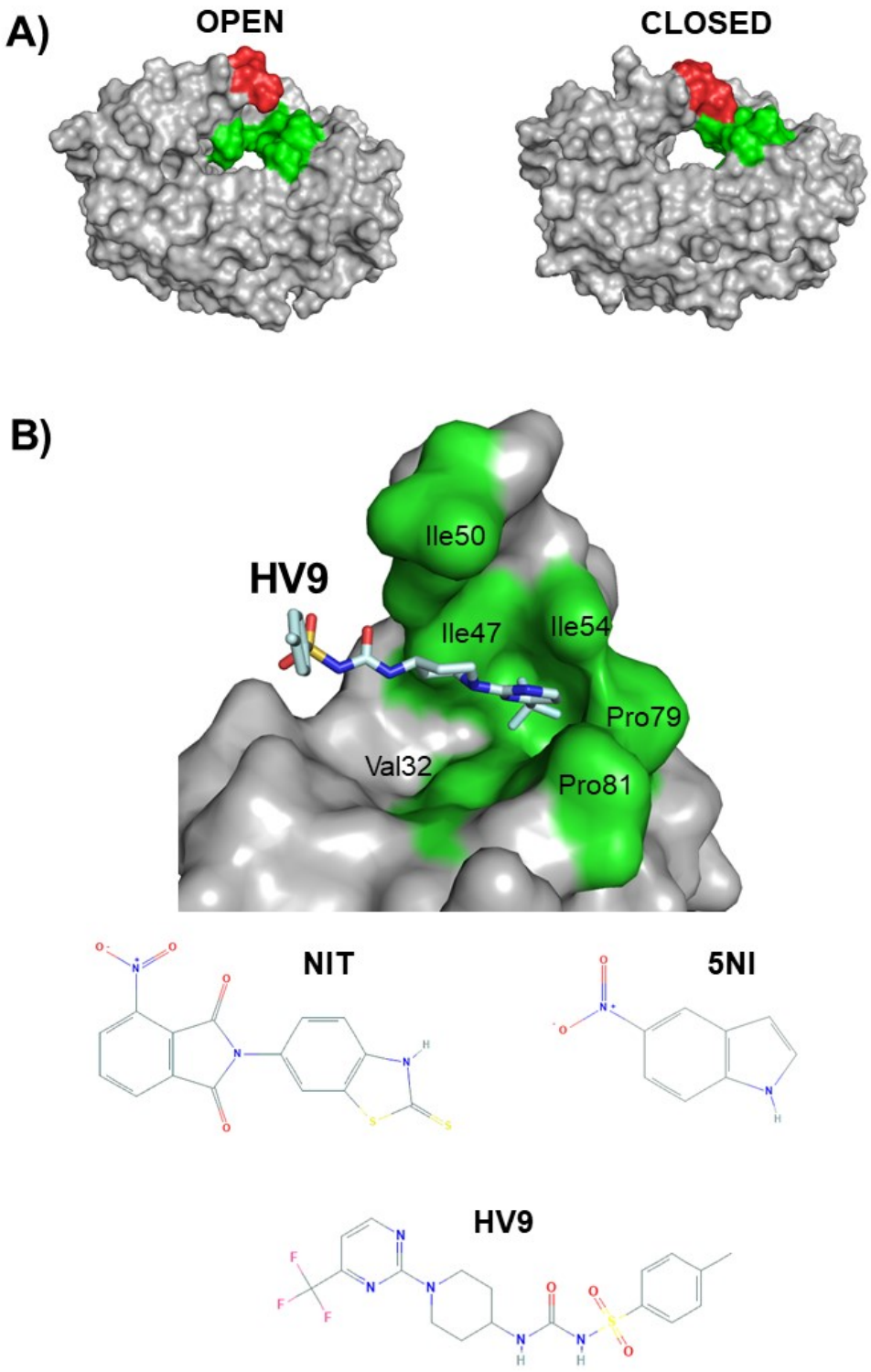


Figure 12. The Eye site allosteric binding pocket of the HIV PR (PDB ID: 5VJ3). A) The Eye site (green) is covered by the flaps tips (49-GGIG-52, red) when the flaps are in the closed conformation B) HV9 bound to the Eye site allosteric binding pocket. The Eye sites are conserved flap tip recognition sites comprised of the following residues: Val32, Ile47, Gly48, Gly49, Ile50, Ile54, Val56, Gly78, Pro79, Thr80, Pro81 and Ile84. The structures of ligands that have been found to bind to this location (NIT, 5NI and HV9) are depicted.

around the substrate which would negatively impact the catalytic efficacy of the PR. If either, or both, of these mechanisms are possible then inhibiting the Eye site could alter the dynamic equilibrium of the flaps in a manner that diminishes substrate binding and catalysis. Further, since changes in the flexibility of the system are thought to be linked to reduced binding affinity, modulating the flap dynamics may prove to be a potential mechanism for eluding inhibitor resistance.

More recently it was shown that curling-in and -out of the residues of the tip of the cantilever region (residues Ile66, Cys67 and Gly68) may be correlated with opening and closing of the flaps of the apo PR, respectively (Sherry et al., 2021; manuscript under review). As mentioned previously, curling of the flap tips has already been associated with opening the flap regions and, so, it is not beyond possibility that curling of the cantilever tips, although distal to the active site, may have some role in the dynamic movements of the protein. A surface pocket, not unlike the ‘exo site’ or the ‘flap top’ pocket, is located between residues near to the dimer interface; namely, Pro1, Gln2, Ile3, Lys7, Thr12 and Cys67 (Figure 13). Notably, a crystal structure of HV9 bound to this site on the PR (PDB ID: 5W5W) has been deposited into the PDB. However, as with the crystal structure with HV9 bound to the Eye site, this structure also has no associated publication. As is the case with similar allosteric sites that have been identified, an adjacent pocket poses the potential for designing larger allosteric binders that further exploit these pockets on the surface of the protein. Here, the secondary binding pocket is largely formed by Lys7 (Figure 13) which could provide a method of forming intermolecular salt bridges with ligands. Such salt bridges have been shown to contribute significantly to the binding favourability of fragments such as 1F1 and 1F1-N which bind to the flap top pocket (Tiefenbrunn et al., 2014) and, thus, exploiting this mechanism of binding could improve the binding affinity of drugs at this location. The mobility of the tip regions of the cantilever may play a significant role in the functional dynamics of the PR (Sherry et al., 2021; manuscript under review) and, therefore, developing allosteric inhibitors that bind to this site and modulate the mobility of the cantilever tips could prove to be a novel avenue of interest for the development of inhibitors.

In conjunction with our understanding of the role that the hinges play in flap dynamics as well as the hydrophobic sliding mechanism it is evident that the global flexibility of the PR is necessary for its function. As such, combination therapy that makes use of allosteric inhibitors could complement the ‘locking down’ effect of active site inhibitors and, in doing so, restore potency against MDR HIV mutants.

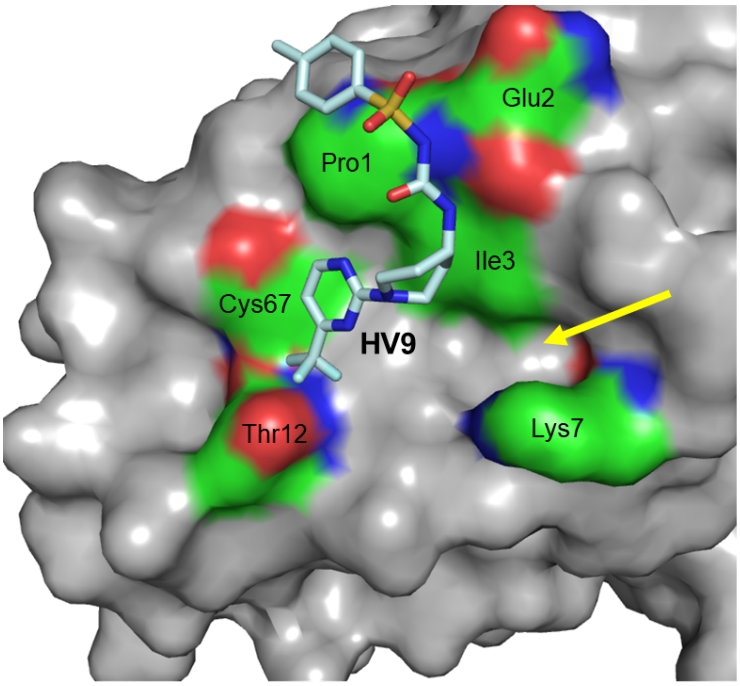


Figure 13. HV9 bound to a probable allosteric binding site located near the N-terminus of the PR. This binding pocket is formed by residues Pro1, Gln2, Ile3, Thr12 and Cys67 and an extension of this pocket (yellow arrow) is created by Lys7 which could be exploited to develop higher affinity allosteric binders.

Conclusions

The role that non-active site mutations play in drug resistance mechanisms is not fully established. The overall changes to the structure of the PR, including the flexible flaps, are minimal in the presence of a wide range and variety of mutations. As such, in order for the PR to be catalytically viable and maintain viral fitness, the backbone conformation of the active site must remain relatively well-conserved. This has interesting implications for drug resistant PR variants, because the presence of mutations that maintain the function of the PR cannot significantly alter the active site backbone conformation. Evidently, a better understanding of the mechanistic events associated with the binding of inhibitors to the HIV PR is important for the design of novel inhibitors. There is evidence that implicates the importance of flexibility of many enzymes in the role of inhibition and, thus, when considering the flexibility of the PR, a different inhibition paradigm arises in which inhibition does not depend on accommodating changes in the active site shape or the affinity of the binding site, but rather, interfering with the global dynamics of the protein. The current paradigm for non-active site mutations describes their cooperative role in resistance by relaying changes to the active site through distal perturbations that are transmitted through the backbone to the active site. Secondary mutations are said to compensate for the deleterious effects of primary mutations, but this concept is poorly understood and not well investigated. Perhaps, the role that these mutations play in increasing drug resistance, or viral fitness, lies in their effect on the dynamic movement of the PR. It can be hypothesised that non-active site mutations may generate novel conformational heterogeneity in the PR, in a manner that selects for conformers that exclude inhibitor binding while still being able to process substrates. Therefore, intrinsic protein dynamics can mechanistically relate protein function molecular recognition, and drug resistance. That is, the effect that non-active site mutations have on the ability of the PR to sample subtly different reactive conformations that are exclusive to PI binding while maintaining substrate specificity, could explain the increase in off-rates of the inhibitors. Furthermore, these mutations might decrease the flexibility of the flaps of the apo PR in a manner that resists the conformational changes necessary to allow entry of the inhibitor into the active site as well as the formation of intermolecular interactions with the flaps. We hypothesise that variations in the dynamics of the flaps opening and closing would have a greater effect on the stable binding of rigid inhibitors than the transient binding of flexible substrates. As a result, these non-active site mutations are likely to shift the equilibrium of conformational states of the PR in a direction that is largely exclusive of inhibitor-bound states in favour of conformations that bind the substrate. Finally, it has been suggested that allosteric sites are an intrinsic property of all dynamic proteins. Consequently, disrupting the flexibility of the PR in a manner that shifts the conformational equilibrium of the PR to a behaviour that

resembles WT PR could prove to be an effective course of inhibition since differences in flap mobility are likely a contributing factor of multi-drug resistance in the PR. Therefore, the coadministration of PIs with supplementary ligands that bind to sites outside of the active site and supplement the action of active site inhibitors would prove to be a promising model of inhibition. As such, investigating compounds that alter the dynamic nature of the flap regions of the PR are essential to the pursuit of future drug compounds.

References

Baxter, J.D., Chasanov, W.M., Adams, J., 2016. An Update on HIV-1 Protease Inhibitor Resistance. J AIDS Clin Res 7. https://doi.org/10.4172/2155-6113.1000581

Bloom, J.D., Gong, L.I., Baltimore, D., 2010. Permissive Secondary Mutations Enable the Evolution of Influenza Oseltamivir Resistance. Science 328, 1272–1275. https://doi.org/10.1126/science.1187816

Boden, D., Markowitz, M., 1998. Resistance to Human Immunodeficiency Virus Type 1 Protease Inhibitors. Antimicrob Agents Chemother 42, 2775–2783.

Borman, A.M., Paulous, S., Clavel, F., 1996. Resistance of human immunodeficiency virus type 1 to protease inhibitors: selection of resistance mutations in the presence and absence of the drug. J. Gen. Virol. 77 (Pt 3), 419–426. https://doi.org/10.1099/0022-1317-77-3-419

Briggs, J.A.G., Simon, M.N., Gross, I., Kräusslich, H.-G., Fuller, S.D., Vogt, V.M., Johnson, M.C., 2004. The stoichiometry of Gag protein in HIV-1. Nat. Struct. Mol. Biol. 11, 672–675. https://doi.org/10.1038/nsmb785

Caldararu, O., Kumar, R., Oksanen, E., Logan, D.T., Ryde, U., 2019. Are crystallographic B-factors suitable for calculating protein conformational entropy? Phys. Chem. Chem. Phys. 21, 18149–18160. https://doi.org/10.1039/C9CP02504A

Chang, M.W., Torbett, B.E., 2011. Accessory mutations maintain stability in drug-resistant HIV-1 protease. J Mol Biol 410, 756–760. https://doi.org/10.1016/j.jmb.2011.03.038

Chen, Z., Li, Y., Schock, H.B., Hall, D., Chen, E., Kuo, L.C., 1995. Three-dimensional structure of a mutant HIV-1 protease displaying cross-resistance to all protease inhibitors in clinical trials. J. Biol. Chem. 270, 21433–21436. https://doi.org/10.1074/jbc.270.37.21433

Clavel, F., Mammano, F., 2010. Role of Gag in HIV Resistance to Protease Inhibitors. Viruses 2, 1411–1426. https://doi.org/10.3390/v2071411

Clemente, J.C., Moose, R.E., Hemrajani, R., Whitford, L.R.S., Govindasamy, L., Reutzel, R., McKenna, R., Agbandje-McKenna, M., Goodenow, M.M., Dunn, B.M., 2004. Comparing the accumulation of active- and nonactive-site mutations in the HIV-1 protease. Biochemistry 43, 12141–12151. https://doi.org/10.1021/bi049459m

Codoñer, F.M., Peña, R., Blanch-Lombarte, O., Jimenez-Moyano, E., Pino, M., Vollbrecht, T., Clotet, B., Martinez-Picado, J., Draenert, R., Prado, J.G., 2017. Gag-protease coevolution analyses define novel structural surfaces in the HIV-1 matrix and capsid involved in resistance to Protease Inhibitors. Sci Rep 7, 3717. https://doi.org/10.1038/s41598-017-03260-4

Damm, K.L., Ung, P.M.U., Quintero, J.J., Gestwicki, J.E., Carlson, H.A., 2008. A poke in the eye: inhibiting HIV-1 protease through its flap-recognition pocket. Biopolymers 89, 643–652. https://doi.org/10.1002/bip.20993

Doyon, L., Croteau, G., Thibeault, D., Poulin, F., Pilote, L., Lamarre, D., 1996. Second locus involved in human immunodeficiency virus type 1 resistance to protease inhibitors. J Virol 70, 3763–3769.

Foulkes-Murzycki, J.E., Scott, W.R.P., Schiffer, C.A., 2007. Hydrophobic Sliding: A Possible Mechanism for Drug Resistance in Human Immunodeficiency Virus Type 1 Protease. Structure 15, 225–233. https://doi.org/10.1016/j.str.2007.01.006

Freedberg, D.I., Ishima, R., Jacob, J., Wang, Y.-X., Kustanovich, I., Louis, J.M., Torchia, D.A., 2002. Rapid structural fluctuations of the free HIV protease flaps in solution: Relationship to crystal structures and comparison with predictions of dynamics calculations. Protein Sci 11, 221–232.

Ghosh, A.K., Osswald, H.L., Prato, G., 2016. Recent Progress in the Development of HIV-1 Protease Inhibitors for the Treatment of HIV/AIDS. J Med Chem 59, 5172–5208. https://doi.org/10.1021/acs.jmedchem.5b01697

Goldfarb, N.E., Ohanessian, M., Biswas, S., McGee, T.D., Mahon, B.P., Ostrov, D.A., Garcia, J., Tang, Y., McKenna, R., Roitberg, A., Dunn, B.M., 2015. Defective Hydrophobic Sliding Mechanism and Active Site Expansion in HIV-1 Protease Drug Resistant Variant Gly48Thr/Leu89Met: Mechanisms for the Loss of Saquinavir Binding Potency. Biochemistry 54, 422–433. https://doi.org/10.1021/bi501088e

Göttlinger, H.G., Sodroski, J.G., Haseltine, W.A., 1989. Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. Proc Natl Acad Sci U S A 86, 5781–5785.

Gulnik, S., Erickson, J.W., Xie, D., 2000. HIV protease: enzyme function and drug resistance. Vitam. Horm. 58, 213–256.

Gulnik, S.V., Suvorov, L.I., Liu, B., Yu, B., Anderson, B., Mitsuya, H., Erickson, J.W., 1995. Kinetic characterization and cross-resistance patterns of HIV-1 protease mutants selected under drug pressure. Biochemistry 34, 9282–9287.

Gustchina, A., Weber, I.T., 1991. Comparative analysis of the sequences and structures of HIV-1 and HIV-2 proteases. Proteins 10, 325–339. https://doi.org/10.1002/prot.340100406

Gustchina, A., Weber, I.T., 1990. Comparison of inhibitor binding in HIV-1 protease and in non-viral aspartic proteases: the role of the flap. FEBS Lett 269, 269–272. https://doi.org/10.1016/0014-5793(90)81171-J

Haynes, B.F., Burton, D.R., Mascola, J.R., 2019. Multiple roles for HIV broadly neutralizing antibodies. Science Translational Medicine 11. https://doi.org/10.1126/scitranslmed.aaz2686

Hertogs, K., Bloor, S., Kemp, S.D., Van den Eynde, C., Alcorn, T.M., Pauwels, R., Van Houtte, M., Staszewski, S., Miller, V., Larder, B.A., 2000. Phenotypic and genotypic analysis of clinical HIV-1 isolates reveals extensive protease inhibitor cross-resistance: a survey of over 6000 samples. AIDS 14, 1203–1210. https://doi.org/10.1097/00002030-200006160-00018

Hornak, V., Okur, A., Rizzo, R.C., Simmerling, C., 2006. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. PNAS 103, 915–920. https://doi.org/10.1073/pnas.0508452103

Hou, T., Yu, R., 2007. Molecular Dynamics and Free Energy Studies on the Wild-type and Double Mutant HIV-1 Protease Complexed with Amprenavir and Two Amprenavir-Related Inhibitors: Mechanism for Binding and Drug Resistance. Journal of Medicinal Chemistry 50, 1177–1188. https://doi.org/10.1021/jm0609162

Hsu, D.C., O'Connell, R.J., 2017. Progress in HIV vaccine development. Hum Vaccin Immunother 13, 1018–1030. https://doi.org/10.1080/21645515.2016.1276138

Ishima, R., Freedberg, D.I., Wang, Y.X., Louis, J.M., Torchia, D.A., 1999. Flap opening and dimer-interface flexibility in the free and inhibitor-bound HIV protease, and their implications for function. Structure 7, 1047–1055.

Ishima, R., Louis, J.M., Torchia, D.A., 2001. Characterization of two hydrophobic methyl clusters in HIV-1 protease by NMR spin relaxation in solution. J Mol Biol 305, 515–521. https://doi.org/10.1006/jmbi.2000.4321

Johnson, V.A., Brun-Vézinet, F., Clotet, B., Conway, B., D'Aquila, R.T., Demeter, L.M., Kuritzkes, D.R., Pillay, D., Schapiro, J.M., Telenti, A., Richman, D.D., International

AIDS Society-USA Drug Resistance Mutations Group, 2003. Drug resistance mutations in HIV-1. Top HIV Med 11, 215–221.
Kahn, K., Bruice, T.C., 2000. Transition-State and Ground-State Structures and Their Interaction with the Active-Site Residues in Catechol O-Methyltransferase. J. Am. Chem. Soc. 122, 46–51. https://doi.org/10.1021/ja993084o
Kim, E.E., Baker, C.T., Dwyer, M.D., Murcko, M.A., Rao, B.G., Tung, R.D., Navia, M.A., 1995. Crystal structure of HIV-1 protease in complex with VX-478, a potent and orally bioavailable inhibitor of the enzyme. J. Am. Chem. Soc. 117, 1181–1182. https://doi.org/10.1021/ja00108a056
Kim, R., Baxter, J.D., 2008. Protease inhibitor resistance update: where are we now? AIDS Patient Care STDS 22, 267–277. https://doi.org/10.1089/apc.2007.0099
King, N.M., Prabu-Jeyabalan, M., Nalivaika, E.A., Schiffer, C.A., 2004. Combating susceptibility to drug resistance: lessons from HIV-1 protease. Chem. Biol. 11, 1333–1338. https://doi.org/10.1016/j.chembiol.2004.08.010
Kozal, M., 2004. Review: Cross-Resistance Patterns Among HIV Protease Inhibitors. AIDS Patient Care and STDs 18, 199–208. https://doi.org/10.1089/108729104323038874
Kuriyan, J., Weis, W.I., 1991. Rigid protein motion as a model for crystallographic temperature factors. Proc. Natl. Acad. Sci. U.S.A. 88, 2773–2777. https://doi.org/10.1073/pnas.88.7.2773
Kwong, P.D., Mascola, J.R., 2018. HIV-1 Vaccines Based on Antibody Identification, B Cell Ontogeny, and Epitope Structure. Immunity 48, 855–871. https://doi.org/10.1016/j.immuni.2018.04.029
Landais, E., Moore, P.L., 2018. Development of broadly neutralizing antibodies in HIV-1 infected elite neutralizers. Retrovirology 15, 61. https://doi.org/10.1186/s12977-018-0443-0
Leaver-Fay, A., Tyka, M., Lewis, S.M., Lange, O.F., Thompson, J., Jacak, R., Kaufman, K., Renfrew, P.D., Smith, C.A., Sheffler, W., Davis, I.W., Cooper, S., Treuille, A., Mandell, D.J., Richter, F., Ban, Y.-E.A., Fleishman, S.J., Corn, J.E., Kim, D.E., Lyskov, S., Berrondo, M., Mentzer, S., Popović, Z., Havranek, J.J., Karanicolas, J., Das, R., Meiler, J., Kortemme, T., Gray, J.J., Kuhlman, B., Baker, D., Bradley, P., 2011. Rosetta3: An Object-Oriented Software Suite for the Simulation and Design of Macromolecules. Methods Enzymol 487, 545–574. https://doi.org/10.1016/B978-0-12-381270-4.00019-6
Lefebvre, E., Schiffer, C.A., 2008. Resilience to resistance of HIV-1 protease inhibitors: profile of darunavir. AIDS Rev 10, 131–142.
Li, H., Omange, R.W., Plummer, F.A., Luo, M., 2017. A novel HIV vaccine targeting the protease cleavage sites. AIDS Research and Therapy 14, 51. https://doi.org/10.1186/s12981-017-0174-7
Louis, J.M., Ishima, R., Torchia, D.A., Weber, I.T., 2007. HIV-1 protease: structure, dynamics, and inhibition. Adv. Pharmacol. 55, 261–298. https://doi.org/10.1016/S1054-3589(07)55008-8
Luchi, A., Angelina, E., Bogado, L., Forli, S., Olson, A., Peruchena, N., 2018. Flap-site fragment restores back wild-type behaviour in resistant form of HIV protease. Mol Inform 37, e1800053. https://doi.org/10.1002/minf.201800053
Luque, I., Todd, M.J., Gómez, J., Semo, N., Freire, E., 1998. Molecular basis of resistance to HIV-1 protease inhibition: a plausible hypothesis. Biochemistry 37, 5791–5797. https://doi.org/10.1021/bi9802521
Mager, P.P., 2001. The active site of HIV-1 protease. Med Res Rev 21, 348–353.
Mahalingam, B., Louis, J.M., Hung, J., Harrison, R.W., Weber, I.T., 2001. Structural implications of drug-resistant mutants of HIV-1 protease: high-resolution crystal structures of the mutant protease/substrate analogue complexes. Proteins 43, 455–464. https://doi.org/10.1002/prot.1057
Markowitz, M., Mo, H., Kempf, D.J., Norbeck, D.W., Bhat, T.N., Erickson, J.W., Ho, D.D., 1995. Selection and analysis of human immunodeficiency virus type 1 variants with increased resistance to ABT-538, a novel protease inhibitor. J. Virol. 69, 701–706.

Maschera, B., Darby, G., Palú, G., Wright, L.L., Tisdale, M., Myers, R., Blair, E.D., Furfine, E.S., 1996. 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. 271, 33231–33235. https://doi.org/10.1074/jbc.271.52.33231

Meiselbach, H., Horn, A.H.C., Harrer, T., Sticht, H., 2007. Insights into amprenavir resistance in E35D HIV-1 protease mutation from molecular dynamics and binding free-energy calculations. Journal of Molecular Modeling 13, 297–304. https://doi.org/10.1007/s00894-006-0121-3

Menéndez-Arias, L., 2010. Molecular basis of human immunodeficiency virus drug resistance: an update. Antiviral Res. 85, 210–231. https://doi.org/10.1016/j.antiviral.2009.07.006

Mittal, S., Cai, Y., Nalam, M.N., Bolon, D.N.A., Schiffer, C.A., 2012. Hydrophobic core flexibility modulates enzyme activity in HIV-1 protease. J Am Chem Soc 134, 4163–4168. https://doi.org/10.1021/ja2095766

Moore, M.L., Bryan, W.M., Fakhoury, S.A., Magaard, V.W., Huffman, W.F., Dayton, B.D., Meek, T.D., Hyland, L., Dreyer, G.B., Metcalf, B.W., 1989. Peptide substrates and inhibitors of the HIV-1 protease. Biochem. Biophys. Res. Commun. 159, 420–425. https://doi.org/10.1016/0006-291x(89)90008-9

Muzammil, S., Ross, P., Freire, E., 2003. A major role for a set of non-active site mutations in the development of HIV-1 protease drug resistance. Biochemistry 42, 631–638. https://doi.org/10.1021/bi027019u

Naicker, P., Achilonu, I., Fanucchi, S., Fernandes, M., Ibrahim, M.A.A., Dirr, H.W., Soliman, M.E.S., Sayed, Y., 2013. Structural insights into the South African HIV-1 subtype C protease: impact of hinge region dynamics and flap flexibility in drug resistance. J. Biomol. Struct. Dyn. 31, 1370–1380. https://doi.org/10.1080/07391102.2012.736774

Naicker, P., Sayed, Y., 2014. Non-B HIV-1 subtypes in sub-Saharan Africa: impact of subtype on protease inhibitor efficacy. Biol. Chem. 395, 1151–1161. https://doi.org/10.1515/hsz-2014-0162

Naicker, P., Stoychev, S., Dirr, H.W., Sayed, Y., 2014. Amide hydrogen exchange in HIV-1 subtype B and C proteases--insights into reduced drug susceptibility and dimer stability. FEBS J. 281, 5395–5410. https://doi.org/10.1111/febs.13084

Nalam, M.N.L., Ali, A., Altman, M.D., Reddy, G.S.K.K., Chellappan, S., Kairys, V., Ozen, A., Cao, H., Gilson, M.K., Tidor, B., Rana, T.M., Schiffer, C.A., 2010. Evaluating the substrate-envelope hypothesis: structural analysis of novel HIV-1 protease inhibitors designed to be robust against drug resistance. J. Virol. 84, 5368–5378. https://doi.org/10.1128/JVI.02531-09

Navia, M.A., Fitzgerald, P.M., McKeever, B.M., Leu, C.T., Heimbach, J.C., Herber, W.K., Sigal, I.S., Darke, P.L., Springer, J.P., 1989. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. Nature 337, 615–620. https://doi.org/10.1038/337615a0

Nicholson, L.K., Yamazaki, T., Torchia, D.A., Grzesiek, S., Bax, A., Stahl, S.J., Kaufman, J.D., Wingfield, P.T., Lam, P.Y., Jadhav, P.K., 1995. Flexibility and function in HIV-1 protease. Nat. Struct. Biol. 2, 274–280. https://doi.org/10.1038/nsb0495-274

Nishimura, Y., Martin, M.A., 2017. Of Mice, Macaques, and Men: Broadly Neutralizing Antibody Immunotherapy for HIV-1. Cell Host & Microbe 22, 207–216. https://doi.org/10.1016/j.chom.2017.07.010

Ohishi, M., Nakano, T., Sakuragi, S., Shioda, T., Sano, K., Sakuragi, J., 2011. The relationship between HIV-1 genome RNA dimerization, virion maturation and infectivity †. Nucleic Acids Research 39, 3404–3417. https://doi.org/10.1093/nar/gkq1314

Ohtaka, H., Freire, E., 2005. Adaptive inhibitors of the HIV-1 protease. Prog. Biophys. Mol. Biol. 88, 193–208. https://doi.org/10.1016/j.pbiomolbio.2004.07.005

Ohtaka, H., Velázquez-Campoy, A., Xie, D., Freire, E., 2002. Overcoming drug resistance in HIV-1 chemotherapy: the binding thermodynamics of Amprenavir and TMC-126 to wild-type and drug-resistant mutants of the HIV-1 protease. Protein Sci. 11, 1908–1916. https://doi.org/10.1110/ps.0206402

Olsen, D.B., Stahlhut, M.W., Rutkowski, C.A., Schock, H.B., vanOlden, A.L., Kuo, L.C., 1999. Non-active Site Changes Elicit Broad-based Cross-resistance of the HIV-1 Protease to Inhibitors. Journal of Biological Chemistry 274, 23699–23701. https://doi.org/10.1074/jbc.274.34.23699

Özen, A., Haliloğlu, T., Schiffer, C.A., 2011. Dynamics of Preferential Substrate Recognition in HIV-1 Protease: Redefining the Substrate Envelope. J Mol Biol 410, 726–744. https://doi.org/10.1016/j.jmb.2011.03.053

Özen, A., Lin, K.-H., Kurt Yilmaz, N., Schiffer, C.A., 2014. Structural basis and distal effects of Gag substrate coevolution in drug resistance to HIV-1 protease. Proc. Natl. Acad. Sci. U.S.A. 111, 15993–15998. https://doi.org/10.1073/pnas.1414063111

Perno, C.F., Cozzi-Lepri, A., Balotta, C., Bertoli, A., Violin, M., Monno, L., Zauli, T., Montroni, M., Ippolito, G., d'Arminio-Monforte, A., 2002. Low prevalence of primary mutations associated with drug resistance in antiviral-naive patients at therapy initiation. AIDS 16, 619–624. https://doi.org/10.1097/00002030-200203080-00014

Perryman, A.L., Lin, J.-H., McCammon, J.A., 2004. HIV-1 protease molecular dynamics of a wild-type and of the V82F/I84V mutant: possible contributions to drug resistance and a potential new target site for drugs. Protein Sci. 13, 1108–1123. https://doi.org/10.1110/ps.03468904

Perryman, A.L., Zhang, Q., Soutter, H.H., Rosenfeld, R., McRee, D.E., Olson, A.J., Elder, J.E., Stout, C.D., 2010. Fragment-Based Screen against HIV Protease. Chem Biol Drug Des 75, 257–268. https://doi.org/10.1111/j.1747-0285.2009.00943.x

Piana, S., Carloni, P., Parrinello, M., 2002. Role of conformational fluctuations in the enzymatic reaction of HIV-1 protease. J Mol Biol 319, 567–583. https://doi.org/10.1016/S0022-2836(02)00301-7

Piana, Stefano, Carloni, P., Rothlisberger, U., 2002. Drug resistance in HIV-1 protease: Flexibility-assisted mechanism of compensatory mutations. Protein Sci. 11, 2393–2402. https://doi.org/10.1110/ps.0206702

Polyansky, A.A., Kuzmanic, A., Hlevnjak, M., Zagrovic, B., 2012. On the Contribution of Linear Correlations to Quasi-harmonic Conformational Entropy in Proteins. J. Chem. Theory Comput. 8, 3820–3829. https://doi.org/10.1021/ct300082q

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

Prabu-Jeyabalan, M., Nalivaika, E.A., King, N.M., Schiffer, C.A., 2004. Structural Basis for Coevolution of a Human Immunodeficiency Virus Type 1 Nucleocapsid-p1 Cleavage Site with a V82A Drug-Resistant Mutation in Viral Protease. J Virol 78, 12446–12454. https://doi.org/10.1128/JVI.78.22.12446-12454.2004

Prabu-Jeyabalan, M., Nalivaika, E.A., King, N.M., Schiffer, C.A., 2003. Viability of a Drug-Resistant Human Immunodeficiency Virus Type 1 Protease Variant: Structural Insights for Better Antiviral Therapy. J Virol 77, 1306–1315. https://doi.org/10.1128/JVI.77.2.1306-1315.2003

Radkiewicz, J.L., Brooks, C.L., 2000. Protein Dynamics in Enzymatic Catalysis: Exploration of Dihydrofolate Reductase. J. Am. Chem. Soc. 122, 225–231. https://doi.org/10.1021/ja9913838

Ragland, D.A., Nalivaika, E.A., Nalam, M.N.L., Prachanronarong, K.L., Cao, H., Bandaranayake, R.M., Cai, Y., Kurt-Yilmaz, N., Schiffer, C.A., 2014. Drug Resistance Conferred by Mutations Outside the Active Site through Alterations in the Dynamic and Structural Ensemble of HIV-1 Protease. Journal of the American Chemical Society 136, 11956–11963. https://doi.org/10.1021/ja504096m

Rhee, S.-Y., Taylor, J., Fessel, W.J., Kaufman, D., Towner, W., Troia, P., Ruane, P., Hellinger, J., Shirvani, V., Zolopa, A., Shafer, R.W., 2010. HIV-1 protease mutations and protease inhibitor cross-resistance. Antimicrob. Agents Chemother. 54, 4253–4261. https://doi.org/10.1128/AAC.00574-10

Rick, S.W., Erickson, J.W., Burt, S.K., 1998. Reaction path and free energy calculations of the transition between alternate conformations of HIV-1 protease. Proteins 32, 7–16.

Ridky, T.W., Kikonyogo, A., Leis, J., Gulnik, S., Copeland, T., Erickson, J., Wlodawer, A., Kurinov, I., Harrison, R.W., Weber, I.T., 1998. Drug-resistant HIV-1 proteases identify enzyme residues important for substrate selection and catalytic rate. Biochemistry 37, 13835–13845. https://doi.org/10.1021/bi980612k

Roberts, N.A., Martin, J.A., Kinchington, D., Broadhurst, A.V., Craig, J.C., Duncan, I.B., Galpin, S.A., Handa, B.K., Kay, J., Kröhn, A., 1990. Rational design of peptide-based HIV proteinase inhibitors. Science 248, 358–361.

Rose, R.B., Craik, C.S., Stroud, R.M., 1998. Domain Flexibility in Retroviral Proteases: Structural Implications for Drug Resistant Mutations † , ‡. Biochemistry 37, 2607–2621. https://doi.org/10.1021/bi9716074

Schaffer, L., Verkhivker, G.M., 1998. Predicting structural effects in HIV-1 protease mutant complexes with flexible ligand docking and protein side-chain optimization. Proteins 33, 295–310.

Schock, H.B., Garsky, V.M., Kuo, L.C., 1996. Mutational anatomy of an HIV-1 protease variant conferring cross-resistance to protease inhibitors in clinical trials. Compensatory modulations of binding and activity. J. Biol. Chem. 271, 31957–31963. https://doi.org/10.1074/jbc.271.50.31957

Scott, W.R., Schiffer, C.A., 2000. Curling of flap tips in HIV-1 protease as a mechanism for substrate entry and tolerance of drug resistance. Structure 8, 1259–1265.

Shafer, R.W., Chuang, T.K., Hsu, P., White, C.B., Katzenstein, D.A., 1999. Sequence and drug susceptibility of subtype C protease from human immunodeficiency virus type 1 seroconverters in Zimbabwe. AIDS Res. Hum. Retroviruses 15, 65–69. https://doi.org/10.1089/088922299311727

Shen, C.-H., Wang, Y.-F., Kovalevsky, A.Y., Harrison, R.W., Weber, I.T., 2010. Amprenavir complexes with HIV-1 protease and its drug-resistant mutants altering hydrophobic clusters. FEBS J. 277, 3699–3714. https://doi.org/10.1111/j.1742-4658.2010.07771.x

Sok, D., Burton, D.R., 2018. Recent progress in broadly neutralizing antibodies to HIV. Nature Immunology 19, 1179–1188. https://doi.org/10.1038/s41590-018-0235-7

Swairjo, M.A., Towler, E.M., Debouck, C., Abdel-Meguid, S.S., 1998. Structural role of the 30's loop in determining the ligand specificity of the human immunodeficiency virus protease. Biochemistry 37, 10928–10936. https://doi.org/10.1021/bi980784h

Thompson, M.A., Aberg, J.A., Cahn, P., Montaner, J.S.G., Rizzardini, G., Telenti, A., Gatell, J.M., Günthard, H.F., Hammer, S.M., Hirsch, M.S., Jacobsen, D.M., Reiss, P., Richman, D.D., Volberding, P.A., Yeni, P., Schooley, R.T., International AIDS Society-USA, 2010. Antiretroviral treatment of adult HIV infection: 2010 recommendations of the International AIDS Society-USA panel. JAMA 304, 321–333. https://doi.org/10.1001/jama.2010.1004

Tie, Y., Boross, P.I., Wang, Y.-F., Gaddis, L., Hussain, A.K., Leshchenko, S., Ghosh, A.K., Louis, J.M., Harrison, R.W., Weber, I.T., 2004. High resolution crystal structures of HIV-1 protease with a potent non-peptide inhibitor (UIC-94017) active against multi-drug-resistant clinical strains. J. Mol. Biol. 338, 341–352. https://doi.org/10.1016/j.jmb.2004.02.052

Tie, Y., Boross, P.I., Wang, Y.-F., Gaddis, L., Liu, F., Chen, X., Tozser, J., Harrison, R.W., Weber, I.T., 2005. Molecular basis for substrate recognition and drug resistance from 1.1 to 1.6 angstroms resolution crystal structures of HIV-1 protease mutants with substrate analogs. FEBS J. 272, 5265–5277. https://doi.org/10.1111/j.1742-4658.2005.04923.x

Tiefenbrunn, T., Forli, S., Baksh, M.M., Chang, M.W., Happer, M., Lin, Y.-C., Perryman, A.L., Rhee, J.-K., Torbett, B.E., Olson, A.J., Elder, J.H., Finn, M.G., Stout, C.D., 2013. Small molecule regulation of protein conformation by binding in the Flap of HIV protease. ACS Chem Biol 8, 1223–1231. https://doi.org/10.1021/cb300611p

Tiefenbrunn, T., Forli, S., Happer, M., Gonzalez, A., Tsai, Y., Soltis, M., Elder, J.H., Olson, A.J., Stout, C.D., 2014. Crystallographic Fragment Based Drug Discovery: Use of a Brominated Fragment Library Targeting HIV Protease. Chem Biol Drug Des 83, 141–148. https://doi.org/10.1111/cbdd.12227

Todd, M.J., Luque, I., Velázquez-Campoy, A., Freire, E., 2000. Thermodynamic basis of resistance to HIV-1 protease inhibition: calorimetric analysis of the V82F/I84V active site resistant mutant. Biochemistry 39, 11876–11883. https://doi.org/10.1021/bi001013s

Todd, M.J., Semo, N., Freire, E., 1998. The structural stability of the HIV-1 protease. J. Mol. Biol. 283, 475–488. https://doi.org/10.1006/jmbi.1998.2090

Tözsér, J., 2001. HIV inhibitors: problems and reality. Ann. N. Y. Acad. Sci. 946, 145–159.

Ung, P.M.-U., Dunbar, J.B., Gestwicki, J.E., Carlson, H.A., 2014. An allosteric modulator of HIV-1 protease shows equipotent inhibition of wild-type and drug-resistant proteases. J Med Chem 57, 6468–6478. https://doi.org/10.1021/jm5008352

Velázquez-Campoy, A., Todd, M.J., Vega, S., Freire, E., 2001. Catalytic efficiency and vitality of HIV-1 proteases from African viral subtypes. Proceedings of the National Academy of Sciences 98, 6062–6067. https://doi.org/10.1073/pnas.111152698

Velázquez-Campoy, A., Vega, S., Fleming, E., Bacha, U., Sayed, Y., Dirr, H.W., Freire, E., 2003. Protease inhibition in African subtypes of HIV-1. AIDS Rev 5, 165–171.

Volberding, P.A., Deeks, S.G., 2010. Antiretroviral therapy and management of HIV infection. Lancet 376, 49–62. https://doi.org/10.1016/S0140-6736(10)60676-9

Wang, X., Minasov, G., Shoichet, B.K., 2002. Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. J Mol Biol 320, 85–95. https://doi.org/10.1016/S0022-2836(02)00400-X

Weber, I.T., Kneller, D.W., Wong-Sam, A., 2015. Highly resistant HIV-1 proteases and strategies for their inhibition. Future Med Chem 7, 1023–1038. https://doi.org/10.4155/fmc.15.44

Weikl, T.R., Hemmateenejad, B., 2020. Accessory mutations balance the marginal stability of the HIV-1 protease in drug resistance. Proteins 88, 476–484. https://doi.org/10.1002/prot.25826

Wensing, A.M., Calvez, V., Günthard, H.F., Johnson, V.A., Paredes, R., Pillay, D., Shafer, R.W., Richman, D.D., 2017. 2017 Update of the Drug Resistance Mutations in HIV-1. Top Antivir Med 24, 132–133.

Wensing, A.M.J., van Maarseveen, N.M., Nijhuis, M., 2010. Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance. Antiviral Res. 85, 59–74. https://doi.org/10.1016/j.antiviral.2009.10.003

Williams, A., Basson, A., Achilonu, I., Dirr, H.W., Morris, L., Sayed, Y., 2019. Double trouble? Gag in conjunction with double insert in HIV protease contributes to reduced DRV susceptibility. Biochem. J. 476, 375–384. https://doi.org/10.1042/BCJ20180692

Winters, M.A., Baxter, J.D., Mayers, D.L., Wentworth, D.N., Hoover, M.L., Neaton, J.D., Merigan, T.C., 2000. Frequency of antiretroviral drug resistance mutations in HIV-1 strains from patients failing triple drug regimens. The Terry Beirn Community Programs for Clinical Research on AIDS. Antivir. Ther. (Lond.) 5, 57–63.

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

Wu, T.D., Schiffer, C.A., Gonzales, M.J., Taylor, J., Kantor, R., Chou, S., Israelski, D., Zolopa, A.R., Fessel, W.J., Shafer, R.W., 2003. Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments. J. Virol. 77, 4836–4847. https://doi.org/10.1128/jvi.77.8.4836-4847.2003

Zondagh, J., Basson, A.E., Achilonu, I., Morris, L., Dirr, H.W., Sayed, Y., 2019. Drug susceptibility and replication capacity of a rare HIV-1 subtype C protease hinge region variant. Antivir. Ther. (Lond.). https://doi.org/10.3851/IMP3308

Chapter 6 – General Discussion and Conclusions

The mutations found in the HP3 PR, namely I13V, I62V and V77I, are found in the hydrophobic core and are likely to alter the internal hydrophobic surface area of the PR. Since the hydrophobic sliding mechanism plays an important role in the function of the PR, subtle alterations in the hydrophobic core could alter the functional properties of the enzyme, in particular, the flexible flaps. As the flaps play a crucial role in drug binding, changes in the nature of their dynamic behaviour could modulate drug binding in a manner that favours drug resistance.

The content of this chapter deals with the following objectives as stated previously. For clarification purposes, the objectives are listed below:

- Evaluate the energetics of binding between the PR variants and PIs via displacement titration calorimetry using selected PIs which cause the mutations at the focus of this study; namely, atazanavir, darunavir, ritonavir, and saquinavir.
- Determine the effect that the mutations present in the HP3 PR have on the overall stability of the protease in comparison to the wild type through differential scanning calorimetry.
- Investigate the effect of these mutations on the dynamic movement of the core regions and flaps of the protease through in silico molecular dynamic simulations.

Altered apo PR stability

MD simulations have revealed that mutations in the PR, such as V82F/I84V or M46I/K45I, result in a decrease in the average backbone RMSF per residue of the apo PR in comparison to the WT (Hou and Yu, 2007). In particular, the closed flap conformation is stabilised and the flexibility of the flaps is diminished in comparison to the apo WT PR when such mutations are present (Goldfarb et al., 2015; Mahalingam et al., 2001; Stefano Piana et al., 2002). The results of these computational studies are reiterated in DSC experiments where the V82F/I84V PR exhibited a transition temperature (T_m) that was 4 °C higher than that of the WT (Todd et al., 2000). Furthermore, DSC has revealed that the subtype C PR exhibits higher structural stability than the subtype B PR (Velázquez-Campoy et al., 2001) and, notably, the subtype C PR exhibits 3- to 24-fold lower binding affinity for inhibitors in comparison to the subtype B PR (Velázquez-Campoy et al., 2003). Here, the three-dimensional structure of the HIV subtype C PR is unchanged in comparison to the subtype B, but eight NOPs are found in the amino acid sequence of the subtype C outside of the active site. This implicates these naturally occurring non-active site mutations (between subtypes) in diminished inhibitor binding and,

simultaneously, increased dimer stability. As such, it is reasonable to conclude that these NOPs may affect binding affinities of inhibitors by increasing the energy required to induce the fit between the PR and the inhibitor. Indeed, changes in T_m of DSC profiles for proteins as a result of mutations can provide an understanding of the effect that the mutations have.

The HP3 PR mutations lead to differences in the nature of the hydrophobic interactions in the core of the PR, which could lead to altered dynamics and stability of the apo enzyme. In the apo state, the T_m of the HP3 PR was relatively unaffected when compared to the corresponding WT PR. Interestingly, however, the MD data of these proteins indicate that the average RMSF values of distinct regions of the PR, such as the fulcrum, cantilever, hinge and flaps, were lower in the HP3 and HP4 PR in comparison to the WT. The HP3 PR remained in the closed flap conformation for the majority of the simulations and did not sample semi-open or open flap conformations. These lower average RMSF values in comparison to the WT PR alongside the observation that the flaps remained closed during the simulations suggest that these mutations may stabilise the apo PR in the closed conformation. Interestingly, the addition of the L33F mutation in the HP4 PR alters the flap conformation of the PR in another manner. The L33F mutation appears to promote the opening of the flaps into the semi-open conformation for the entirety of the apo simulation. Together, the four mutations present in the HP4 PR appear to simultaneously increase the rigidity of the flaps while also allowing the flaps to be sustained in a wider, semi-open, conformation. Therefore, the HP4 PR sustains wider flap conformations than the WT PR and sustains these conformations in a more stable manner than the WT PR is able to do.

The similar T_m values of the HP3 and WT PRs may indicate that the global stability of the PR in largely unaffected by these mutations, but DSC does not provide an indication of the dynamic conformations of the enzyme. Therefore, it is plausible that these mutations change the predominant apo flap conformation that is the most favourable but do not affect the global stability of the PR homodimer. A stabilised closed conformation of the apo PR would impose a greater energetic penalty for the conformation changes associated with opening the flaps in order to allow for inhibitor entry into the active site cavity as well as subsequent closing of the flaps. The implications of this are readily demonstrated when considering that drug binding is a process in dynamic equilibrium described by:

$$K_d = \frac{[PR][PI]}{[PR \cdot PI]} = \frac{k_{off}}{k_{on}} \qquad \text{[equation i]}$$

Here, the rate of drug binding (k_{on}) would be lowered as a result of the increased energetic penalty of changing the conformations of the flaps. In general, the closing and rearrangement of the flaps during drug binding is associated with an unfavourable enthalpic and entropic cost.

If mutations in the PR result in the closed or semi-open conformation being stabilised and favoured more so that the WT PR, then this could present a mechanism for drug resistance through decreasing the k_{on} since all drugs would pay a larger energetic penalty in order to induce binding and subsequent inhibition (Figure 9). While some mutations have been shown to increase the global stability of the PR, through increases in their T_m, the subtilties of the alterations may be more intricate than simply observing the global stability of the protein. Local structural changes in stability, such as those in the flaps, may not translate into global changes in stability however their role in drug resistance could still be evident and explicable.

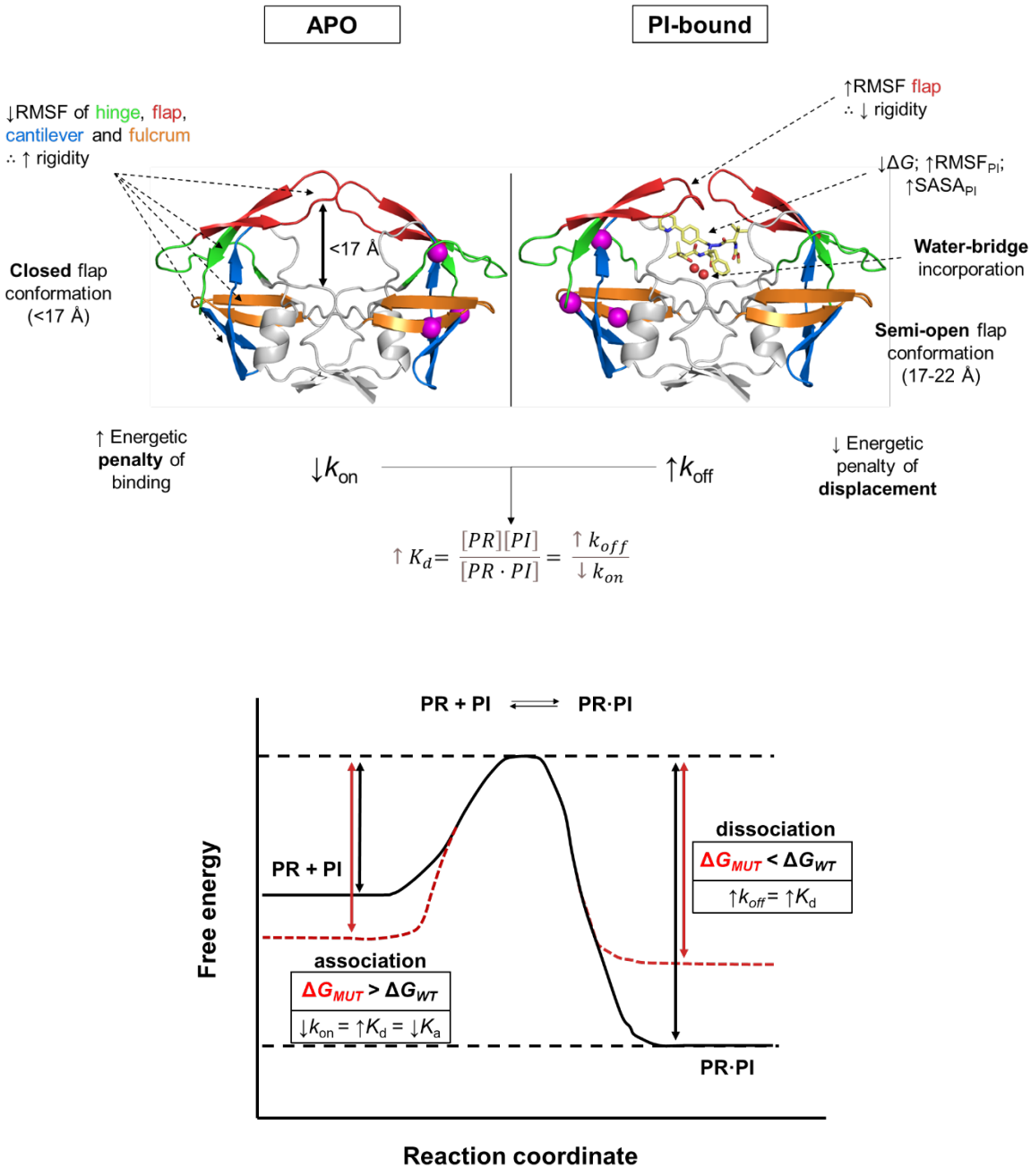


Figure 1. The influence of the HP3 mutations on the apo and inhibitor-bound HIV PR. The effect of the I13V/I62V/V77I mutations have on the PR are two-fold, affecting both the apo and the inhibitor-bound state. First, the rigidity of the apo PR is increased, which results in a greater energetic penalty of drug binding. The flaps of the apo PR are stabilised in the closed conformation and, therefore, have a diminished on-rate which leads to an increased dissociation constant. Secondly, the rigidity of the flaps of the PR is decreased in the presence of an inhibitor and the flaps exist in a semi-open conformation. As a result, the inhibitor is more mobile in the active site and is not desolvated effectively upon binding. The inability of the inhibitor to bring the flaps into the closed conformation necessitates water-bridge incorporation in order to form a quasi-stable complex. These effects cause a decrease in the favourable Δ*G* of binding, which correlates to an increased dissociation constant for the complex.

Diminished inhibitor-bound complex stability

Analysis of the energetics of inhibitor binding provides further insight into the binding characteristics displayed by HP3 PR and how these characteristics differ from those of the WT PR. Since inhibitors compete with substrates for binding to the active site cavity, PIs should be able to bind with significantly high affinity and a slow off-rate (k_{off}). With this in mind, any mutations in the PR that decrease the energetic penalty associated with the dissociation of inhibitors (thereby increasing the k_{off}) would lead to inhibitor resistance.

The change in Gibbs free energy (ΔG) of the inhibitor binding event provides a measurement of the favourability of the protein-inhibitor interaction. A negative ΔG is indicative of an energetically favourable binding interaction and can be defined by:

$$\Delta G = \Delta H - T\Delta S = RT \ln K_d \qquad \text{[equation ii]}$$

Where the magnitude of the favourable, negative, ΔG is inversely proportional to the dissociation constant (K_d) of the ligand (i.e., an increased K_d is associated with a less favourable ΔG). Therefore, it is noteworthy that the ΔG for the binding of ATV, DRV, RTV and SQV is less negative when binding to the HP3 PR than the WT PR which is indicative of an increased dissociation constant for these inhibitors. The most significant change in these binding affinities is evident for the binding of ATV and DRV which were reduced 64- and 16-fold, respectively, by the presence of the HP3 mutations. Interestingly, the HP3·ATV and HP3·DRV complexes were less thermally stable than the corresponding WT PR complexes. The T_m for these complexes were 4 and 7 °C lower, respectively, than the major transition peaks observed for the WT complexes. The implications of this are further elucidated from equation [i], where the lower stability of the complex would result in a higher k_{off} for inhibitors, since the energy required to separate the enzyme and inhibitor is lowered, thus, shifting the equilibrium in favour of dissociation (Figure 9).

The favourability of the thermodynamics of protein-ligand interactions is dependent on both the change in enthalpy (ΔH) and entropy (ΔS) of the system. Briefly, the change in enthalpy is representative of breaking and/or formation of bonds, indicated by a positive or negative change, respectively. Entropy is representative of the dynamics or disorder of the system. Ligand binding brings the binding partners into more conformationally retrained states and the loss of entropy is unfavourable. However, this is largely compensated for by the increase in the entropy of the surrounding solvent as a result of desolvation of the hydrophobic surface areas of the interacting surfaces. The hydrophobic effect is the major contributor to the favourable increase in the solvent entropy (S_{solv}) upon ligand binding.

This was evident from the ITC analyses, where the entropy change was favourable when PIs bound to the WT PR, with the exceptions of RTV which was slightly unfavourable. A large negative Δ*H* was observed for ATV, DRV and RTV when binding to the WT PR, making these interactions enthalpically driven while SQV binding was entropically driven. The effect of the mutations which are present in the HP3 PR resulted in a decrease in the favourability of Δ*G* for all of the inhibitors tested when compared to the WT PR. As is evident from equation [ii], this change can be correlated with an increased dissociation constant and, therefore, decreased inhibitor susceptibility. Binding of ATV and DRV was enthalpically driven for the WT and HP3 PRs, however, the HP3 PR displayed an unfavourable entropy change for these interactions. Which is to say, the binding of ATV and DRV to the HP3 PR resulted in a decrease in entropy of the system.

Further insight into how these mutations alter inhibitor binding for the HP3 PR can be clarified through MD simulations of these complexes. The unfavourable change in entropy may be explained by the calculated solvent accessible surface area (SASA) of the PIs when in complex with the PR variants. The SASA of the PIs were consistently larger in the mutant complexes in comparison to the corresponding WT complexes. Since the major contribution to the favourable entropy change arises from an increase in S_{solv} due to the desolvation of the ligand and protein surfaces (Kožíšek et al., 2014), an increased SASA suggests that the binding mechanism is altered in a manner that increases the exposure of the PI to the surrounding solvent, thereby diminishing the entropic favourability of binding. The flap regions play an important role in inhibitor binding and, thus, the dynamics of these regions is important to consider. The HP3 PR exhibited a wider, semi-open, flap conformation when in complex with ATV, DRV and SQV in comparison to the corresponding WT PR complexes which remained closed. This observation was further enhanced with the addition of the L33F mutation, in the HP4 PR complexes. Furthermore, analysis of the average RMSF of the flap regions of the mutant complexes was greater than those of the WT complexes in the majority of cases investigated. Strikingly, the average RMSF for the flap regions of the HP3·ATV and HP3·DRV complexes were increased by 154% and 56% in comparison to the corresponding WT complexes. Furthermore, the HP4 PR in complex with ATV, DRV and SQV exhibited higher average flap region fluctuations of 81%, 116% and 30%, respectively. In addition, it was seen that these PIs exhibited higher average RMSF values per atom when in complex with the mutated PRs in comparison to the WT PR, suggesting higher mobility of the drugs and decreased stability of the bound state. A noteworthy change in the binding characteristics of the HP3 PR is the inclusion of water molecules into the bottom of the active site of the PR when in complex with PIs. In these complexes, the PIs form intermolecular interactions with the flap regions but, the inhibitors are unable to induce the closed conformation of the flaps.

As a result, in order to simultaneously maintain contact with the active site residues and the flap regions, water bridges must be formed between moieties of the PI and residues at the bottom of the active site such at Asp25, Asp29 and Asp30. These coordinated water molecules possess the ability to rotate and slide across each when forming part of the PR·PI complex. The implication of this in drug resistance lies in the fact that hydrogen bonds formed in the water bridges would be more transient than those that would form directly between the rigid inhibitor and the closed conformation of the PR. Such changes in the dynamics of the PR·PI complexes would diminish, or even eliminate, these intermolecular interactions which could decrease inhibitor efficacy since the strong binding of PIs is mainly driven by a favourable enthalpy change (Velazquez-Campoy et al., 2001a; Velázquez-Campoy et al., 2001b).

Collectively, these findings suggest that these mutations in the PR decrease the rigidity of the flap regions of the PR when in complex with PIs and, simultaneously, increase the average flap distance from closed to semi-open when bound to PIs. As a result, the PIs fit less securely into the active site cavity and, perhaps, protrude into and interact with the surrounding solvent which decreases the entropic favourability of binding. The PIs do not appear to bring the mutated flaps into the closed conformation which is illustrated by the necessity for water molecules to bridge the gap between the PI and the active site residues. These alterations are likely to contribute to the decrease in the observed T_m for the mutated PR·PI complexes which reflects a decrease in the stability due to diminished binding affinity of the inhibitors. Thus, the PIs exhibit increased K_d values in comparison to the WT, which is indicative of reduced inhibitor susceptibility, or drug resistance (Figure 9).

Conclusion

Mutations located outside of the active site are thought to play a compensatory role in the PR, by increasing viral fitness in response to the potentially deleterious effects of active site mutations. However, there is evidence which demonstrates that such non-active site mutations are implicated in inhibitor resistance, although their mechanism of action in resistance-specific mechanisms has yet to be fully elucidated. The I13V/I62V/V77I set of mutations occur concurrently and have been correlated with decreased drug susceptibility but, hitherto, their mechanism has not been explained.

The flap regions of the HIV PR establish a dynamic equilibrium of conformations based on their relationship with distinct regions of the enzyme which facilitate their mobility. These regions include the hinges and, through more recent understanding, the fulcrum and cantilever regions. The ability of the cantilever and fulcrum regions to slide across one another is crucial to the conformational changes associated with flap movement of the PR and this ability is reliant on the near-isoenergetic transfer of intramolecular contacts within the core of the

enzyme. Thus, mutations that occur in these core regions of the PR are likely to affect the dynamic equilibrium of the PR flaps without imparting direct geometric or steric effects at the active site residues. Non-active site mutations, such as I13V/I62V/V77I are likely to accumulate in a manner that shifts the dynamic equilibrium of the flaps so as to diminish drug binding in a manner that single non-active site mutations are unable to convey. Due to the rigid nature of PIs, their efficacy is highly susceptible to alterations in the active site, as is seen with active site mutations, but these alterations need not be limited to the “static” geometric changes imparted by polymorphisms which directly change the active site geometry. The mechanism of non-active site mutations likely lies in altering the dynamics of the flap regions of the PR which affects the ability of the rigid inhibitors to lockdown the active site into the closed conformation. On the other hand, the natural substrates of the enzyme are able to compensate for such changes in the dynamic ensemble of the active site, through their own inherent flexibility as well as through mutations in their own sequences.

The findings of the current study illustrate the salience of non-active site mutations in drug resistance mechanisms, alongside their previously proposed effects on viral fitness. Through considering the importance of the dynamic ensemble of the enzyme during inhibition, a slightly different inhibition paradigm arises in which inhibition depends on interfering with the global dynamics of the enzyme and/or through returning the dynamics to a more “wild-type like” state in which standard inhibitors are still effective. Inhibitors or small molecules which can bind to allosteric sites on the PR and disrupt the effect of non-active site mutations so as to shift the equilibrium towards a behaviour that resembles the WT PR could prove to be a more effective course of inhibition. Further research is required in order to develop supplementary ligands which can be co-administered with PIs in order to target MDR PR variants which are likely to have altered flap dynamics.

Therefore, the intrinsic dynamics of the protein ensemble can mechanistically relate enzyme function, molecular recognition and drug resistance to each other. Which is to say, non-active site mutations can subtly change the dynamic ensemble of the PR towards conformers which are exclusive of inhibitors but still maintain substrate binding and, therefore, functionality of the enzyme.

References

Ahmed, S.M., Kruger, H.G., Govender, T., Maguire, G.E.M., Sayed, Y., Ibrahim, M.A.A., Naicker, P., Soliman, M.E.S., 2013. Comparison of the molecular dynamics and calculated binding free energies for nine FDA-approved HIV-1 PR drugs against subtype B and C-SA HIV PR. Chem Biol Drug Des 81, 208–218. https://doi.org/10.1111/cbdd.12063

Bebenek, K., Abbotts, J., Roberts, J.D., Wilson, S.H., Kunkel, T.A., 1989. Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase. J Biol Chem 264, 16948–16956.

Borman, A.M., Paulous, S., Clavel, F., 1996. Resistance of human immunodeficiency virus type 1 to protease inhibitors: selection of resistance mutations in the presence and absence of the drug. J. Gen. Virol. 77 (Pt 3), 419–426. https://doi.org/10.1099/0022-1317-77-3-419

Briggs, J.A.G., Kräusslich, H.-G., 2011. The molecular architecture of HIV. J Mol Biol 410, 491–500. https://doi.org/10.1016/j.jmb.2011.04.021

Brik, A., Wong, C.-H., 2003. HIV-1 protease: mechanism and drug discovery. Org. Biomol. Chem. 1, 5–14.

Chan, D.C., Kim, P.S., 1998. HIV entry and its inhibition. Cell 93, 681–684. https://doi.org/10.1016/s0092-8674(00)81430-0

Chen, Z., Li, Y., Schock, H.B., Hall, D., Chen, E., Kuo, L.C., 1995. Three-dimensional structure of a mutant HIV-1 protease displaying cross-resistance to all protease inhibitors in clinical trials. J. Biol. Chem. 270, 21433–21436. https://doi.org/10.1074/jbc.270.37.21433

Clavel, F., Guyader, M., Guétard, D., Sallé, M., Montagnier, L., Alizon, M., 1986. Molecular cloning and polymorphism of the human immune deficiency virus type 2. Nature 324, 691–695. https://doi.org/10.1038/324691a0

Clavel, F., Mammano, F., 2010. Role of Gag in HIV Resistance to Protease Inhibitors. Viruses 2, 1411–1426. https://doi.org/10.3390/v2071411

Clemente, J.C., Hemrajani, R., Blum, L.E., Goodenow, M.M., Dunn, B.M., 2003. Secondary Mutations M36I and A71V in the Human Immunodeficiency Virus Type 1 Protease Can Provide an Advantage for the Emergence of the Primary Mutation D30N. Biochemistry 42, 15029–15035. https://doi.org/10.1021/bi035701y

Codoñer, F.M., Peña, R., Blanch-Lombarte, O., Jimenez-Moyano, E., Pino, M., Vollbrecht, T., Clotet, B., Martinez-Picado, J., Draenert, R., Prado, J.G., 2017. Gag-protease coevolution

analyses define novel structural surfaces in the HIV-1 matrix and capsid involved in resistance to Protease Inhibitors. Sci Rep 7, 3717. https://doi.org/10.1038/s41598-017-03260-4

Darke, P.L., Nutt, R.F., Brady, S.F., Garsky, V.M., Ciccarone, T.M., Leu, C.T., Lumma, P.K., Freidinger, R.M., Veber, D.F., Sigal, I.S., 1988. HIV-1 protease specificity of peptide cleavage is sufficient for processing of gag and pol polyproteins. Biochem Biophys Res Commun 156, 297–303. https://doi.org/10.1016/s0006-291x(88)80839-8

Debouck, C., 1992. The HIV-1 protease as a therapeutic target for AIDS. AIDS Res Hum Retroviruses 8, 153–164. https://doi.org/10.1089/aid.1992.8.153

Foulkes-Murzycki, J.E., Scott, W.R.P., Schiffer, C.A., 2007. Hydrophobic Sliding: A Possible Mechanism for Drug Resistance in Human Immunodeficiency Virus Type 1 Protease. Structure 15, 225–233. https://doi.org/10.1016/j.str.2007.01.006

Freed, E.O., 1998. HIV-1 gag proteins: diverse functions in the virus life cycle. Virology 251, 1–15. https://doi.org/10.1006/viro.1998.9398

Freedberg, D.I., Ishima, R., Jacob, J., Wang, Y.-X., Kustanovich, I., Louis, J.M., Torchia, D.A., 2002. Rapid structural fluctuations of the free HIV protease flaps in solution: Relationship to crystal structures and comparison with predictions of dynamics calculations. Protein Science 11, 221–232. https://doi.org/10.1110/ps.33202

Galiano, L., Ding, F., Veloro, A.M., Blackburn, M.E., Simmerling, C., Fanucci, G.E., 2009. Drug pressure selected mutations in HIV-1 protease alter flap conformations. J. Am. Chem. Soc. 131, 430–431. https://doi.org/10.1021/ja807531v

Goldfarb, N.E., Ohanessian, M., Biswas, S., McGee, T.D., Mahon, B.P., Ostrov, D.A., Garcia, J., Tang, Y., McKenna, R., Roitberg, A., Dunn, B.M., 2015. Defective Hydrophobic Sliding Mechanism and Active Site Expansion in HIV-1 Protease Drug Resistant Variant Gly48Thr/Leu89Met: Mechanisms for the Loss of Saquinavir Binding Potency. Biochemistry 54, 422–433. https://doi.org/10.1021/bi501088e

Gustchina, A., Weber, I.T., 1990. Comparison of inhibitor binding in HIV-1 protease and in non-viral aspartic proteases: the role of the flap. FEBS Lett 269, 269–272. https://doi.org/10.1016/0014-5793(90)81171-J

Henderson, L.E., Krutzsch, H.C., Oroszlan, S., 1983. Myristyl amino-terminal acylation of murine retrovirus proteins: An unusual post-translational protein modification. Proc Natl Acad Sci U S A 80, 339–343.

Hofmann, T., Hodges, R.S., James, M.N.G., 2002. Effect of pH on the activities of penicillopepsin and Rhizopus pepsin and a proposal for the productive substrate binding mode in penicillopepsin [WWW Document]. https://doi.org/10.1021/bi00299a008

Hornak, V., Okur, A., Rizzo, R.C., Simmerling, C., 2006. HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. PNAS 103, 915–920. https://doi.org/10.1073/pnas.0508452103

Hornak, V., Simmerling, C., 2007. Targeting structural flexibility in HIV-1 protease inhibitor binding. Drug Discov Today 12, 132–138. https://doi.org/10.1016/j.drudis.2006.12.011

Hou, T., Yu, R., 2007. Molecular Dynamics and Free Energy Studies on the Wild-type and Double Mutant HIV-1 Protease Complexed with Amprenavir and Two Amprenavir-Related Inhibitors: Mechanism for Binding and Drug Resistance. Journal of Medicinal Chemistry 50, 1177–1188. https://doi.org/10.1021/jm0609162

Ishima, R., Freedberg, D.I., Wang, Y.X., Louis, J.M., Torchia, D.A., 1999. Flap opening and dimer-interface flexibility in the free and inhibitor-bound HIV protease, and their implications for function. Structure 7, 1047–1055. https://doi.org/10.1016/s0969-2126(99)80172-5

Jacobo-Molina, A., Arnold, E., 1991. HIV reverse transcriptase structure-function relationships. Biochemistry 30, 6351–6361. https://doi.org/10.1021/bi00240a001

Johnson, V.A., Brun-Vézinet, F., Clotet, B., Conway, B., D'Aquila, R.T., Demeter, L.M., Kuritzkes, D.R., Pillay, D., Schapiro, J.M., Telenti, A., Richman, D.D., International AIDS Society-USA Drug Resistance Mutations Group, 2003. Drug resistance mutations in HIV-1. Top HIV Med 11, 215–221.

Kear, J.L., Blackburn, M.E., Veloro, A.M., Dunn, B.M., Fanucci., G.E., 2009. Subtype Polymorphisms Among HIV-1 Protease Variants Confer Altered Flap Conformations and Flexibility. J Am Chem Soc 131, 14650–14651. https://doi.org/10.1021/ja907088a

Kohl, N.E., Emini, E.A., Schleif, W.A., Davis, L.J., Heimbach, J.C., Dixon, R.A., Scolnick, E.M., Sigal, I.S., 1988. Active human immunodeficiency virus protease is required for viral infectivity. Proc Natl Acad Sci U S A 85, 4686–4690. https://doi.org/10.1073/pnas.85.13.4686

Kožíšek, M., Lepšík, M., Grantz Šašková, K., Brynda, J., Konvalinka, J., Rezáčová, P., 2014. Thermodynamic and structural analysis of HIV protease resistance to darunavir - analysis of heavily mutated patient-derived HIV-1 proteases. FEBS J. 281, 1834–1847. https://doi.org/10.1111/febs.12743

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

Lin, L.N., Mason, A.B., Woodworth, R.C., Brandts, J.F., 1994. Calorimetric Studies of Serum Transferrin and Ovotransferrin. Estimates of Domain Interactions, and Study of the Kinetic Complexities of Ferric Ion Binding. Biochemistry 33, 1881–1888. https://doi.org/10.1021/bi00173a035

Mager, P.P., 2001. The active site of HIV-1 protease. Med Res Rev 21, 348–353.

Mahalingam, B., Louis, J.M., Hung, J., Harrison, R.W., Weber, I.T., 2001. Structural implications of drug-resistant mutants of HIV-1 protease: high-resolution crystal structures of the mutant protease/substrate analogue complexes. Proteins 43, 455–464. https://doi.org/10.1002/prot.1057

Markowitz, M., Mo, H., Kempf, D.J., Norbeck, D.W., Bhat, T.N., Erickson, J.W., Ho, D.D., 1995. Selection and analysis of human immunodeficiency virus type 1 variants with increased resistance to ABT-538, a novel protease inhibitor. J. Virol. 69, 701–706.

Mata-Munguía, C., Escoto-Delgadillo, M., Torres-Mendoza, B., Flores-Soto, M., Vázquez-Torres, M., Gálvez-Gastelum, F., Viniegra-Osorio, A., Castillero-Manzano, M., Vázquez-Valls, E., 2014. Natural polymorphisms and unusual mutations in HIV-1 protease with potential antiretroviral resistance: a bioinformatic analysis. BMC Bioinformatics 15, 72. https://doi.org/10.1186/1471-2105-15-72

Mittal, S., Cai, Y., Nalam, M.N., Bolon, D.N.A., Schiffer, C.A., 2012. Hydrophobic core flexibility modulates enzyme activity in HIV-1 protease. J Am Chem Soc 134, 4163–4168. https://doi.org/10.1021/ja2095766

Muzammil, S., Ross, P., Freire, E., 2003. A major role for a set of non-active site mutations in the development of HIV-1 protease drug resistance. Biochemistry 42, 631–638. https://doi.org/10.1021/bi027019u

Naicker, P., Achilonu, I., Fanucchi, S., Fernandes, M., Ibrahim, M.A.A., Dirr, H.W., Soliman, M.E.S., Sayed, Y., 2013a. Structural insights into the South African HIV-1 subtype C protease: impact of hinge region dynamics and flap flexibility in drug resistance. J. Biomol. Struct. Dyn. 31, 1370–1380. https://doi.org/10.1080/07391102.2012.736774

Naicker, P., Achilonu, I., Fanucchi, S., Fernandes, M., Ibrahim, M.A.A., Dirr, H.W., Soliman, M.E.S., Sayed, Y., 2013b. Structural insights into the South African HIV-1 subtype C protease: impact of hinge region dynamics and flap flexibility in drug resistance. J Biomol Struct Dyn 31, 1370–1380. https://doi.org/10.1080/07391102.2012.736774

Navia, M.A., Fitzgerald, P.M., McKeever, B.M., Leu, C.T., Heimbach, J.C., Herber, W.K., Sigal, I.S., Darke, P.L., Springer, J.P., 1989. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. Nature 337, 615–620. https://doi.org/10.1038/337615a0

Nicholson, L.K., Yamazaki, T., Torchia, D.A., Grzesiek, S., Bax, A., Stahl, S.J., Kaufman, J.D., Wingfield, P.T., Lam, P.Y., Jadhav, P.K., 1995. Flexibility and function in HIV-1 protease. Nat. Struct. Biol. 2, 274–280. https://doi.org/10.1038/nsb0495-274

Nijhuis, M., Schuurman, R., de Jong, D., Erickson, J., Gustchina, E., Albert, J., Schipper, P., Gulnik, S., Boucher, C.A., 1999. Increased fitness of drug resistant HIV-1 protease as a result of acquisition of compensatory mutations during suboptimal therapy. AIDS 13, 2349–2359.

Partaledis, J.A., Yamaguchi, K., Tisdale, M., Blair, E.E., Falcione, C., Maschera, B., Myers, R.E., Pazhanisamy, S., Futer, O., Cullinan, A.B., 1995. In vitro selection and characterization of human immunodeficiency virus type 1 (HIV-1) isolates with reduced sensitivity to hydroxyethylamino sulfonamide inhibitors of HIV-1 aspartyl protease. J. Virol. 69, 5228–5235.

Perryman, A.L., Lin, J.-H., McCammon, J.A., 2006. Restrained molecular dynamics simulations of HIV-1 protease: the first step in validating a new target for drug design. Biopolymers 82, 272–284. https://doi.org/10.1002/bip.20497

Piana, S., Carloni, P., Rothlisberger, U., 2002. Drug resistance in HIV-1 protease: Flexibility-assisted mechanism of compensatory mutations. Protein Sci. 11, 2393–2402. https://doi.org/10.1110/ps.0206702

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

Ragland, D.A., Nalivaika, E.A., Nalam, M.N.L., Prachanronarong, K.L., Cao, H., Bandaranayake, R.M., Cai, Y., Kurt-Yilmaz, N., Schiffer, C.A., 2014. Drug Resistance Conferred by Mutations Outside the Active Site through Alterations in the Dynamic and Structural Ensemble of HIV-1 Protease. Journal of the American Chemical Society 136, 11956–11963. https://doi.org/10.1021/ja504096m

Rhee, S.-Y., Fessel, W.J., Zolopa, A.R., Hurley, L., Liu, T., Taylor, J., Nguyen, D.P., Slome, S., Klein, D., Horberg, M., Flamm, J., Follansbee, S., Schapiro, J.M., Shafer, R.W., 2005. HIV-1 Protease and reverse-transcriptase mutations: correlations with antiretroviral therapy in subtype B isolates and implications for drug-resistance surveillance. J. Infect. Dis. 192, 456–465. https://doi.org/10.1086/431601

Rhee, S.-Y., Taylor, J., Fessel, W.J., Kaufman, D., Towner, W., Troia, P., Ruane, P., Hellinger, J., Shirvani, V., Zolopa, A., Shafer, R.W., 2010. HIV-1 protease mutations and protease inhibitor cross-resistance. Antimicrob. Agents Chemother. 54, 4253–4261. https://doi.org/10.1128/AAC.00574-10

Roberts, J.D., Bebenek, K., Kunkel, T.A., 1988. The accuracy of reverse transcriptase from HIV-1. Science 242, 1171–1173.

Robins, T., Plattner, J., 1993. HIV protease inhibitors: their anti-HIV activity and potential role in treatment. J. Acquir. Immune Defic. Syndr. 6, 162–170.

Rose, R.B., Craik, C.S., Stroud, R.M., 1998. Domain Flexibility in Retroviral Proteases: Structural Implications for Drug Resistant Mutations. Biochemistry 37, 2607–2621. https://doi.org/10.1021/bi9716074

Schaffer, L., Verkhivker, G.M., 1998. Predicting structural effects in HIV-1 protease mutant complexes with flexible ligand docking and protein side-chain optimization. Proteins 33, 295–310.

Schechter, I., Berger, A., 1967. On the size of the active site in proteases. I. Papain. Biochem Biophys Res Commun 27, 157–162. https://doi.org/10.1016/s0006-291x(67)80055-x

Schramm, V.L., 2011. Enzymatic Transition States, Transition-State Analogs, Dynamics, Thermodynamics, and Lifetimes. Annu. Rev. Biochem. 80, 703–732. https://doi.org/10.1146/annurev-biochem-061809-100742

Shafer, R.W., Hsu, P., Patick, A.K., Craig, C., Brendel, V., 1999. Identification of Biased Amino Acid Substitution Patterns in Human Immunodeficiency Virus Type 1 Isolates from Patients Treated with Protease Inhibitors. J Virol 73, 6197–6202.

Shrake, A., Ross, P.D., 1990. Ligand-induced biphasic protein denaturation. J. Biol. Chem. 265, 5055–5059.

Suguna, K., Padlan, E.A., Smith, C.W., Carlson, W.D., Davies, D.R., 1987. Binding of a reduced peptide inhibitor to the aspartic proteinase from Rhizopus chinensis: implications for a mechanism of action. Proc Natl Acad Sci U S A 84, 7009–7013. https://doi.org/10.1073/pnas.84.20.7009

Takahashi, K., Fukada, H., 1985. Calorimetric studies of the binding of Streptomyces subtilisin inhibitor to subtilisin of Bacillus subtilis strain N'. Biochemistry 24, 297–300.

Taylor, B.S., Sobieszczyk, M.E., McCutchan, F.E., Hammer, S.M., 2008. The challenge of HIV-1 subtype diversity. N. Engl. J. Med. 358, 1590–1602. https://doi.org/10.1056/NEJMra0706737

Thompson, M.A., Aberg, J.A., Cahn, P., Montaner, J.S.G., Rizzardini, G., Telenti, A., Gatell, J.M., Günthard, H.F., Hammer, S.M., Hirsch, M.S., Jacobsen, D.M., Reiss, P., Richman, D.D., Volberding, P.A., Yeni, P., Schooley, R.T., International AIDS Society-USA, 2010. Antiretroviral treatment of adult HIV infection: 2010 recommendations of the International AIDS Society-USA panel. JAMA 304, 321–333. https://doi.org/10.1001/jama.2010.1004

Todd, M.J., Luque, I., Velázquez-Campoy, A., Freire, E., 2000. Thermodynamic basis of resistance to HIV-1 protease inhibition: calorimetric analysis of the V82F/I84V active site resistant mutant. Biochemistry 39, 11876–11883. https://doi.org/10.1021/bi001013s

Tözsér, J., 2001. HIV inhibitors: problems and reality. Ann. N. Y. Acad. Sci. 946, 145–159.

UNAIDS, 2020. Global HIV & AIDS statistics - 2020 fact sheet.

Velázquez-Campoy, A., Kiso, Y., Freire, E., 2001a. The binding energetics of first- and second-generation HIV-1 protease inhibitors: implications for drug design. Arch Biochem Biophys 390, 169–175. https://doi.org/10.1006/abbi.2001.2333

Velázquez-Campoy, A., Todd, M.J., Vega, S., Freire, E., 2001b. Catalytic efficiency and vitality of HIV-1 proteases from African viral subtypes. Proceedings of the National Academy of Sciences 98, 6062–6067. https://doi.org/10.1073/pnas.111152698

Velázquez-Campoy, A., Vega, S., Fleming, E., Bacha, U., Sayed, Y., Dirr, H.W., Freire, E., 2003. Protease inhibition in African subtypes of HIV-1. AIDS Rev 5, 165–171.

Volberding, P.A., Deeks, S.G., 2010. Antiretroviral therapy and management of HIV infection. Lancet 376, 49–62. https://doi.org/10.1016/S0140-6736(10)60676-9

Weber, I.T., Kneller, D.W., Wong-Sam, A., 2015. Highly resistant HIV-1 proteases and strategies for their inhibition. Future Med Chem 7, 1023–1038. https://doi.org/10.4155/fmc.15.44

Williams, A., Basson, A., Achilonu, I., Dirr, H.W., Morris, L., Sayed, Y., 2019. Double trouble? Gag in conjunction with double insert in HIV protease contributes to reduced DRV susceptibility. Biochem. J. 476, 375–384. https://doi.org/10.1042/BCJ20180692

Wlodawer, A., Erickson, J.W., 1993. Structure-based inhibitors of HIV-1 protease. Annu Rev Biochem 62, 543–585. https://doi.org/10.1146/annurev.bi.62.070193.002551

Wu, T.D., Schiffer, C.A., Gonzales, M.J., Taylor, J., Kantor, R., Chou, S., Israelski, D., Zolopa, A.R., Fessel, W.J., Shafer, R.W., 2003. Mutation patterns and structural correlates in human immunodeficiency virus type 1 protease following different protease inhibitor treatments. J. Virol. 77, 4836–4847. https://doi.org/10.1128/jvi.77.8.4836-4847.2003

Zondagh, J., Balakrishnan, V., Achilonu, I., Dirr, H.W., Sayed, Y., 2018. Molecular dynamics and ligand docking of a hinge region variant of South African HIV-1 subtype C protease. J. Mol. Graph. Model. 82, 1–11. https://doi.org/10.1016/j.jmgm.2018.03.006